The London School of Economics and Political Science

Regulating Asset Ownership: Capabilities and Market Failures in Infrastructures

Enrico Rossi

A thesis submitted to the Department of Management of the London School of Economics for the degree of Doctor of Philosophy, London, July 2017

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 107,568 words.

Abstract

The regulation of assets and infrastructures has always been a central problem of economic theory. The first approach of regulatory and antitrust authorities was to regard asset ownership as a source of welfare inefficiency. This was the "monopoly explanation" that was contested by Coase. Developing from Coase's original intuitions, a contractual approach emerged in regulatory economics. This second tradition relies on the concept of transaction costs, or market failures, to justify the role of ownership for welfare and regulatory purposes. Yet, even though in the contractual approach ownership is not regarded as a necessary source of welfare inefficiencies, it still remains the necessary consequence of some dysfunctional characteristic of the market mechanism.

This study addresses the role of asset ownership understood in terms of the owner's subjective use value. The aim is thus to provide a theory of ownership based on a reinterpretation of the concept of value, relying on the classical dualism between value-in-exchange and value-in-use. Private ownership of a good or resource becomes relevant for welfare and efficiency purposes whenever assets can be redeployed internally across alternative subjective opportunities by the owner to satisfy their private, subjective, requirements. Through in-house redeployment, the asset owner becomes independent of the performances and requirements of external market mechanisms.

If auto-employment is allowed, then two conditions are satisfied. First, the performance of the market mechanism becomes unnecessary to understand how assets are allocated among alternative uses. This makes market failures and transaction costs a non-necessary requirement to justify asset ownership and to understand its welfare implications. Second, the knowledge of an actor's subjective capabilities in the use and employment of the asset (knowing how an actor prefers to privately use and consume an asset) becomes necessary in order to understand how different ownership patterns affect the set of idiosyncratic opportunities perceived by different potential assets owners.

If auto-redeployment (in-house enjoyment or auto-consumption) is allowed, then we can see that the idiosyncratic opportunities perceived by the actors are not necessarily driven by external market mechanisms, nor by its performance. Yet, they remain relevant in order to derive normative conclusions on the allocative outcomes. This can be seen as a make-or-buy problem where the trade-off between "make" and "buy" can be reinterpreted as a

trade-off in value terms, between subjective value-in-use and objective value-in-exchange. The different interpretation of the make option marks the difference between the make-or-buy problem modelled in the contract-based theory of ownership versus a capability-based theory of ownership.

The work argues that, whenever physical assets are privately owned and can be employed "in-house", in order to legitimately derive normative conclusions on how privately owned assets ought to be employed in a society, some form of public regulation is always needed in order to overcome the inherent presence of subjective (actor-specific) valuations. For this reason, the work concludes that whenever value in not a monism, the legal framework should always have logical and temporal priority over the competitive mechanism of the market, independently from the performance of the latter.

Acknowledgments

This work could not be concluded without the help of multiple people that have supported me along this long journey.

My gratitude goes first and foremost to my PhD supervisor, Jonathan Liebenau that has always encouraged me. His constant support through the entire process, since the years of my Master degree has been invaluable and contributed in a decisive way to the thoughts here developed. It is hard to overstate the many opportunities that he has offered to me through these years. Moreover his patience, even in the hardest periods, has been fundamental in order to get through this long journey. My gratitude also goes to my second supervisor Saul Estrin. Although he stepped in at a later stage, his suggestions and recommendations in the later stages of the PhD were of great importance. The discussions with him have always been very fruitful and greatly appreciated.

It is not possible to overlook the support that I got from many friends that have surrounded me during the last five years. A special thanks goes to Samer Abdelnour whose friendly suggestions and moral support have been invaluable. This thesis is also the result of some of his teaching and tips. Adeline Pelletier has been a loyal travel companion, and an incredible source of idea exchange. Her effort in the reading of much of the material here presented has been just incredible. I have been very fortunate to share some of my years with Koen Rutten. Although short, our time spent together at the LSE has been great and enriching for me. A special thanks goes also to Nadia Millington, who has always been available, and has spent time in reading some of the material of the thesis. Nico Tempini and Martha Poon were great friends and wise advisers. I thank them both for the great time spent together while they were at LSE. Further thanks go to all other friends that shared this experience with me, although it is impossible to list them all: Laura Zimmerman, Ceren Erdem, Nofie Iman, “Charly” Chen, Laura Schang, and many others.

Finally, the most important thanks should be reserved to the people that more than any others have shared the passion, the struggle, the achievements and the delusions of this thesis with me. First, my lovely girlfriend Roberta. It would have been literally impossible to keep going and to finish this thesis without her daily patience, her support and her love.

Finally, I have to thank my parents for all they have done for me, and for what they are still doing.

Table of Contents

CHAPTER 2

A reconsideration of Value theory in classical political economy: use value, exchange value, and implications for an new economic theory of institutions and asset of ownership 45

CHAPTER 4

CHAPTER 5

Table of Figures

INTRODUCTION

Preliminary overview of the work

1. Background of the research and motivations

The recent deregulatory wave starting in the late 1970s has radically modified our perception of so-called natural monopolies or "essential facilities". For more than a century, the standard way to deal with assets and infrastructures that were considered to be of essential importance for the public, but that were not convenient, or not feasible, to duplicate (due to increasing returns in production and/or in demand) was through public regulation or control (mainly U.S.), or public ownership (mainly Europe). Among them, the institution of common carriage has represented a particularly powerful regulatory tool in the United States. Both the regulatory institution of common carriage and public ownership has substantially had the same aim: to inhibit discretionary control over the employment of the essential infrastructure by the part of a private owner. The aim of both these regulatory measures (and of regulation more generally) was, precisely, to subsume *jus privatum* to the *public juris*.

Starting in late 1970s, the approach towards essential facilities radically changed. On the one hand, regulation became a subsidiary of antitrust law: the institution of common carriage was voided of all meaning, and started to be regarded as useless as long as competition could be "somehow" created and enforced. This implied a radically different interpretation of the nature of regulation as an institution: it ceased to deal with private and public rights (in ownership and control) associated to the ownership of a certain property, and became an enforcer of competition "of last resort" instead. On the other hand, competition has been introduced into former monopolised industries by unbundling (separating) the physical underlying asset (the infrastructure) from the services delivered "over-the-top": service-based competition has been introduced next, and in replacement, to asset-based competition.

This new course, which reached its peak with the 1996 Telecommunication Act in the US, and with the 2002-03 EU directive, posed brand new problems to policy makers.

In the first place, the way in which the decoupling between the ownership of the underlying asset and the services delivered over it should be treated and interpreted became unclear. The standard way in which the economic literature interprets vertical integration does not fit the specific case of an unbundled asset: integration or separation in this case are not between two distinct technological stages, but they are defined looking at the relationship between the ownership of an asset (the physical infrastructure) and the control of the services delivered *over-the-top*. In general, two brand-new issues have captured the attention of policy makers and of the general public: non-price discrimination and neutrality. The peculiarity of these new problems derives from the fact that neither can be addressed nor understood through the standard (static) economic variables, such as price and quantity. Non-neutral and discriminatory actions are not pricing (nor quantity) issues, but they involve the particular (dynamic) relationship between the ownership of the asset, and the control of the services provided over-the-top. These problems have been particularly magnified in the Information & Communication (IC) sector where the switch from analogic to digital transmissions of data has posed particularly difficult questions to regulators.

Secondly, it became unclear the way in which the regulation of the access to the infrastructure should be linked to the enforcement of antitrust law with respect to the services delivered over it: what is the difference between antitrust and regulation, and how should they be treated? What should the relationship between regulatory and antitrust authorities be in an unbounded environment?

These two aspects are two sides of the same coin (deregulation through unbundling) and they are strictly linked. The unclear relationship between the ownership of an asset and its control (the way in which the asset ought to be employed among the various services) is at the heart of the relationship between antitrust law and regulation: can we deregulate by merely severing asset ownership from its control, thus implementing an institutional shift from regulation to competition law? Under which circumstances is this legitimate?

These questions call for a better understanding of the role of institutions in economic theory. Ownership and control (use) rights are two distinct, but strictly linked, institutions (and certainly the most fundamental ones). In the same vain, regulation and antitrust law are two different institutional arrangements substituting, or complementing, the operations of the “free” market.

2. Brief overview of the essence of the work: aim and approach

2.1. General argument

This work provides a reconsideration of the role and the nature of institutions in general, and asset ownership in particular, in normative economics.

The basic aim of the work is to better understand *under which conditions the identity of the actor owning an asset, and having certain control rights over it, matters in order to derive legitimate normative conclusions on how the asset ought to be employed among alternative services*. Stated differently, the aim is to understand under which conditions it becomes legitimate to deregulate through the enforcement of competition (antitrust) law over an "unbundled network" and, vice versa, under which conditions deregulating by decoupling the ownership of a privately owned asset from the control of its employment generates unmanageable problems for the public authorities from an economic point of view in order to derive sound policy prescriptions.

The work redefines institutions as "shared mental and social frameworks enabling value judgments on what ought to be done" and asserts that the only way to understand the role of ownership and control rights in economics is to reconsider the way in which "value" is treated in economic theory. For this reason, the work asserts that in order to develop any normative economic theory of institution, it is first necessary to develop a theory of value: the nature and the role of institutions in economic theory will depend on the specific theory of value adopted.

The work illustrates how the institution of ownership (and the relative identity of the asset owner) becomes relevant in order to understand how assets or resources ought to be employed whenever value becomes subjective. The subjectivity of value derives from the fact that that there can be no objective agreement on the way in which a certain asset or resource ought to be allocated: different actors will solve the valuation problem differently, and therefore will allocate the asset differently, depending on their specific identity.

In the case of infrastructures and productive assets, the subjectivity of value derives from the fact that there can be no clear agreement on the way in which the specific infrastructure ought to be allocated among different services, and therefore *among alternative markets*. The work points out that whenever the allocation problem depends on the specific opportunity set held by an actor, then the valuation problem becomes subjective, or identity-dependent. Obviously, the way in which an asset ought to be allocated among alternative services becomes a subjective problem only if different asset owners would employ the same asset in order to serve different markets. Whenever the market is specific, or dedicated to one specific service/market, no asymmetries in the opportunity sets can be generated, and therefore no subjectivity in the valuation problem arises.

2.2. Detailed overview: the two dimensions of value that confer economic relevance to asset ownership

The work identifies the two fundamental drivers that make non-market institutions, such as asset ownership, economically relevant independently form the competitive performance of the market mechanism: the *exogeneity* of value (in use) and the *subjective* nature of value (in use).

On the one hand, the work asserts that the institution of property becomes economically relevant independently from the institution of exchange whenever value becomes a dualism, meaning that there is a source of value deriving from the employment of the property "in use" that is independent from the value generated from the employment of the property "in exchange." The *exogeneity* of use value with respect to exchange value is therefore a precondition for the institution of property to acquire the same legitimacy of the institution of exchange.

The most general case in which the value of any good or asset is a dualism corresponds to the situation where actors can privately employ assets for their own private purposes, in auto-consumption. This is a case when the asset owner can self-employ his own property "in house". Whenever production and consumption are not sharply severed, but production for self-consumption is allowed, there are the preconditions for dualism in value, and for the exogeneity of use value with respect to exchange value. The work suggests that production for auto-consumption is also the way in which the concept of vertical integration should be intended.

On the other hand, the work suggests that the *subjectivity* of value is necessary in order to justify the relevance of actors' identities (private and social). Use value can be exogenous, yet objective. While the independence of use value with respect to exchange value ensures that the market mechanism does not have an exclusive normative role in determining the way in which assets and resources ought to be employed, the subjectivity of use value is what ensures that the alternative institutional arrangements derive from the social and institutional sphere, and not from mere material properties of the physical assets and resources in question.

The fundamental argument of the work can therefore be summarised as follows: in order to derive normative conclusions concerning the remedies (prescriptions) that should be adopted with respect to the way in which assets or resources ought to be employed and used, the conditions affecting the subjectivity and the exogeneity of use value (as opposed to exchange value) must be verified first. Thus, if (i) assets can be redeployed "in-house", and (ii) the redeployment is guided and constrained by subjective, yet institutionally-embedded, opportunities and motives, then normative prescriptions on how assets ought to be employed must necessarily rely on some prior regulatory and legal institution, as markets cannot resolve the inherent dualism in value.

If this is the case, then the legal environment always has priority over the competitive mechanism, and deregulation through the mere enforcement of the competitive mechanism cannot be legitimate.

3. Brief overview of the structure and organization of the work

This work is structured in two main parts.

Chapters two, three and four represent the theoretical and conceptual foundations of the argument. Chapters five, six and seven (together with chapter eight that however also constitutes a concluding section), develop the conceptual implications for modern theory of institutions, and for modern theory of regulation.

3.1. Value theory in economic thought: from the classical dualism in value to the neoclassical monism in value and implications for the interaction between the private and the public domain

The first three chapters provide an historical inquiry on the history of value theory in economics, and on the meaning of value, from classical economics until the present day. The work shows that the contemporary approach to institutions, that interpret asset ownership as the mere outcome of some behavioural or allocative failure relies on two very specific conditions, which can be summarised as follows.

One, even though private ownership of physical assets and resources is allowed, the first necessary condition of the modern neoclassical framework is that only flowing services can be transacted in the market, and priced by the market. The first condition is therefore an ontological condition, and introduces a sharp decoupling between the materiality of the physical objects, such as assets, and the immateriality of service. Only if the underlying physical support is severed from the services operated “over the top”, can the standard general competitive equilibrium be computed.

The second condition is a behavioural one and provides a normative content to the competitive framework. This second condition states that actors that own material assets can never employ their own properties in order to satisfy their own value in use, but they necessarily have to “hold them out” to serve the public.

This condition introduces the decoupling between the private and the public dimension, which is the most fundamental distinction in any work in welfare, or institutional economics. The present work defines the “private sphere” as that institutional domain characterised by the personal employment of a physical asset “in-house”, thus in auto-consumption. Conversely, the public, or social, sphere is defined based on the employment of physical properties in order to serve third parties or, the public.

This second condition, also relabelled the “rationality assumption” ensures that any property owner will always treat his own material property (physical capital) as if it was mere financial capital or, money. Only when material wealth is employed as if it was mere monetary wealth, can the general competitive framework have normative content.

The work then develops along a double dualism, which constitutes the backbone of the present argument. Two main dualities can be identified: the one between physical assets and services, and the one between the private sphere and the public sphere. While the unit of analysis of the private sphere is the material property accumulated, the unit of analysis of the public sphere is the service transacted on the market.

The private sphere is characterised by the subjective use value generated by the internal auto-consumption of the privately accumulated material wealth. Conversely, the public sphere is characterised by the objective exchange value generated by the social consumption of the services provided by the underlying property by the part of the members of the public. The physical asset in this second case constitutes the physical support enabling the transaction with the public.

This approach to the public-private dichotomy very clearly shows why whenever private self-employment of private properties (in use) is ruled out, the public and the private domains conflate into an “institutional singularity”, where the public and the private spheres become indistinguishable, and value becomes a monism, meaning that use value conflates into exchange value, and the two becomes the same, unique, source of value.

The irrelevance of private ownership of assets and infrastructures in modern "failure-based approaches" derives from the fact that the classical dualism in value is ruled out by assumption, by imposing a univocal mapping between the *social monetary wealth* that any actor can generate through the sale of services on the market, and the *private material wealth* that any asset owner can privately accumulate in-house, and (potentially) use for his own self-consumption.

As a result, the first important consideration that emerges from the work is that the role of institutions in economic analysis depends on the way in which value theory is approached: monism in value theory implies monism in institutional theory, and the equivalence of the private and social domains. Conversely, dualism in value theory implies dualism in institutional theory, and the sharp independence of the private sphere from the social sphere.

3.2. The meaning and origins of regulation in common law: *juris privati*, *juris publici* and common law duties

After a lengthy discussion on the three main dualities (use value vs. exchange value, physical asset vs. flowing service, private domain vs. social/public domain), the work turns to a critical re-examination of the foundations of public regulation, public control of private assets.

Chapters 6 and 7 provide an in-depth reconceptualization of the role of regulation, of the meaning of *juris privati* and *juris publici* in property law, and of the institution of common carriage, which represents the key institution for the public regulation of privately owned infrastructures, especially, in modern times, railways and telecommunications.

On the one hand, the work shows that the dichotomy between private and public in common law has to be intended in value terms: as a dichotomy between self-employment of a property, and "social employment" of a property in order to serve third parties (the public). The concept of "affectation with public interest" represents the conceptual cornerstone justifying modern public control of private assets. The work shows that the fundamental distinction between *juris privati* and *juris publici* has to be intended in property law, thus should be applied to the concept of property, not to the nature of the service. The distinction between the private and the public status of a property depends precisely on the problematic interactions between the services operated by the property, and the property itself. Thus, even in common law the public-private dichotomy relies on the service-property dichotomy that was illustrated in the first part of the work.

On the other hand, the wok discusses the two social duties of a common carriage at length: the duty to hold out to the public, and duty to serve with due care. The work shows that these duties were originally conceived not in order to regulate competition, but rather in order to impose legal liabilities over private parties. The ultimate role of the two duties was precisely to establish a univocal mapping between the private sphere where the property was operated, and the public sphere characterised by the provision of a service.

As a result, the work shows that two medieval social duties of a public calling reflect precisely the two constraints of modern competitive general equilibrium: the duty to serve with due care corresponds to the ontological constraint: it ensures that a given property is employed in a specific way, in order to provide a specific service, so that only a specific type of service, provided in a specific way, can be associated to a given underlying property, or asset. On the other hand, the duty to "hold out to the public" corresponds to the behavioural constraint: it ensures that actors employ the property as a means for exchange, and not as a means for private self-subsistence.

In other words, the common law origin of public regulation shows that the irreducible dichotomy between the private domain where properties are owned, and the public domain where services are transacted can only be solved if some social constraints are devised in the first place, in order to regulate the way in which private owners ought to employ their own assets in the social domain.

4. Final considerations

In light of the considerations developed above, the work argues that the contemporary faith in deregulation through the substitution of public regulation with competition law is misplaced: no matter how competitive the market is, some regulatory framework will always be required in order to legitimise social normative conclusions on how privately owned assets ought to be employed.

Stated differently, the work shows that when it comes to regulating asset ownership, a "failure-based" approach to inefficiencies is not sufficient. The legal framework will always have to have logical and temporal priority over the competitive framework whenever physical assets and infrastructures can be privately appropriated.

CHAPTER 1

Prelude to the work: background, motivations, main problematics, and relevance of the research

Chapter synopsis

This chapter represents a very preliminary overview of some of the key concepts, and offers a preliminary overview of the main problems that will be developed in the reminder of the work.

1. A brief review of some of the basic economic approaches to asset ownership: from industrial organization to the theories of incentives

1.1. Asset ownership and integration in the "monopoly approach": antitrust, regulation and merger guidelines

The private ownership of assets has always been regarded as the fundamental reason generating natural monopolies or concentrated market structures. The role of antitrust and regulatory authorities have always been to prevent the leverage of monopoly power deriving from firms' ownership of assets. On the one hand, the private ownership of large (indivisible) facilities, represents a problem as it often leads to *naturally concentrated markets*, and to natural monopolies in the extreme case. On the other hand, firms can use their private ownership of assets in one stage of the chain to "*leverage their monopoly power*" in other stages of the chain. In this second case, by acquiring other assets from firms in contiguous stages firms can leverage their power and foreclose competition in otherwise competitive segments of the chain.

In any case, by acquiring assets and expanding their private property firms integrate either vertically or horizontally.

This *integration* "through asset ownership" (vertical and horizontal integrations) has always been regarded as a practice undertaken by firms in order to leverage their monopoly power either internally to a market (horizontal integration), or across layers of the supply chain (vertical integration). The antitrust merger guidelines (vertical and horizontal) have been developed precisely in order to deal with the monopolistic behaviour deriving from the leverage (vertical), or the concentration (horizontal) of market power deriving from private ownership of assets. In the extreme case where ex-post antitrust could not cope with an excessive concentration of power adopting ex-post "behavioural" remedies, ex-ante regulation has usually been regarded as the ultimate (structural) remedy of last resort to constrain the distortive power deriving from the private ownership of assets.

This has been regarded as the "monopoly" approach to asset ownership: asset ownership intended as a way to concentrate the market and increase, or leverage, market power.

1.2. Asset ownership and integration in the shift from monopoly failures to market failures: the "efficiency approach", rights and incentives to invest

Not only contemporary institutional economics (NIE) in the form of Transaction-Cost Economics (TCE), but also contemporary property-right theory (here labelled New-PRT, or NPRT) and organizational economics (OE) developed, as disciplines, from the unsolved problems posed by the ownership of assets.

Starting from the 1970s, asset ownership started to be seen not only as a "problem" for social welfare (as the driver generating and enabling market power), but also as a *positive remedy* for potential (dynamic) inefficiencies. This was, for instance, Williamson's thesis: TCE interpreted as an "efficiency theory" of asset ownership as opposed to the monopoly approach of standard industrial organization adopted by antitrust authorities. The new focus of these new theories (especially TCE and NPRT) was on *incentives*. The fundamental step was the recognition that ownership was not incentive-neutral, and that the identity of the actor owning the assets does matter as asset ownership changes the relative incentives of the parties involved in the economic system. No theory of asset ownership could be developed without a proper theory of incentives.

The various theories differ precisely in their way to treat and interpret incentives, and how they model the way in which *ownership affects incentives* (and vice versa).

On the one hand, the "*efficiency approach*" of TCE focusses on the way in which ownership affects *ex-post* incentives in adaptation: asset owners have the faculty to decide how to redeploy their owned property when facing an unforeseen circumstance. This unexpected or unforeseen change in situation might generate hold-ups and opportunism, and this is why vertical integration becomes economically valuable: the joint ownership of the various assets aligns the incentives of the various parties on a single "best response" to the adaptation process. This alignment of incentives is cost-efficient as it saves all the potential inefficiencies deriving from a costly and time-consuming bargaining between two (disintegrated) parties. This is, in a nutshell, the essence of the “efficiency” approach to asset ownership.

On the other hand, the “*rights approach*” of PRT focusses on *ex-ante* investment incentives: asset owners have the faculty to enjoy the returns generated by the employment of their own assets in production. As a result, ownership affects the (ex-ante) incentives that a party preliminary perceives with respect to the level of investment, due to the fact that a higher share of ownership implies a higher share of monetary returns that can be ex-post appropriated by the part of the respective asset’s owner(s).

The two approaches clearly do not share the same definition and interpretation of ownership. Their different interpretations of ownership reflect their different treatments of incentives as the way in which incentives are defined directly affect the way in which the institution of ownership is defined in turn.

Yet, these approaches do have common points. In particular, they share two common characteristics.

First, they all start from the assumption that extra-market institutions, such as the institution of ownership, arise as a consequence of market failures or, to use a widely known (but very generic) expression, transaction costs. These failures are usually regarded either as a break-down of the allocative system, or as behavioural limitations deriving from the “rationality” and knowledge of actors or, more commonly, from a combination of the two. Usually, the “key failure” in both TCE and NPRT is the incompleteness of contracts deriving from uncertainty and the inability to properly forecast future circumstances.

In a nutshell, both TCE and NPRT share a “market-failure” interpretation of asset ownership in the sense that they both interpret asset ownership as a second-best type of institution, to be regarded as a necessary response to behavioural or allocative failures. It is the presence of these behavioural and allocative failures that rehabilitates the institution of ownership and highlights its incentives-based benefits vis-à-vis its monopolistic incentive-free shortcomings.

Second, they both start their analysis *focussing on the given market* they are interested into, thus focussing on the specific service(s) or good(s) to be provided by the asset. As a result, both TCE and NPRT interpret asset ownership as an *investment* in order to *enter a* *specific market*, so to produce a *given service or good* (or set of services or goods). In both cases, ownership becomes instrumental for the provision of a given service or good (or set or services). In sum, *the given good/service market is the relevant unit of analysis, and ownership is regarded as the necessary investment in order to enter that specific market, and start production*.

This is perfectly consistent with the way in which the field of industrial organization adopted by antitrust authorities interpret the institution of ownership: as the preliminary fixed, or sunk, investment required in the long-run, in order to *enter a given market* and *produce a given product* (or set of products) in the short run.

This approach, which treats asset ownership as an institution instrumental, and subordinated, to the production of a given service, in order to serve a given market, will be here defined as the "*antitrust approach" to ownership*.

For expositive reasons, in what follows, we first focus on this second common point (the focus of the antitrust approach on the single market intended as the ultimate unit of analysis), and we will subsequently come back to the first aforementioned common point (point one above), namely the common focus on market failures in order to explain and understand the role and the nature of the institution of ownership.

2. Antitrust approach and regulatory approach: causality issues between production, allocation, ownership and control

2.1. The antitrust approach: asset ownership causally subordinated to the production problem given a certain (set of) market(s)

The key feature of what we here call an "antitrust approach" is the fact that the focus is not really on the way in which ownership shapes the incentives of the asset owner with respect to the way in which assets should be allocated, or employed, across *alternative services* or, in other words, across *alternative markets*. The antitrust approach does not perceive ownership in standard microeconomic terms, as a precondition to the constrained allocative problem (describing how a certain asset ought to be employed). Conversely, the antirust approach reverses the logical and conceptual causality between the institution of ownership, the production problem and the allocation problem: the latter occurs in the background, and represents an autonomous mechanism independent form the other two, while the production problem constitutes the preliminary precondition to the former.

Consistent with the approach adopted by antitrust authorities, an antitrust approach first defines (the boundaries of) the market. Then, once that the market (and the competitive arena) is clearly defined, the institution of ownership is approached in standard industrial organization terms and it is regarded as the necessary investment required in order to enter that single*, given* market. In other words, the market (characterised by its own degree of competition) is the unit of analysis and should be the entity defined first. The market then is (exogenously) given to the subsequent production problem, and drives any subsequent economic problem. Asset ownership then becomes instrumental for undertaking production in that specific market (or set of markets).

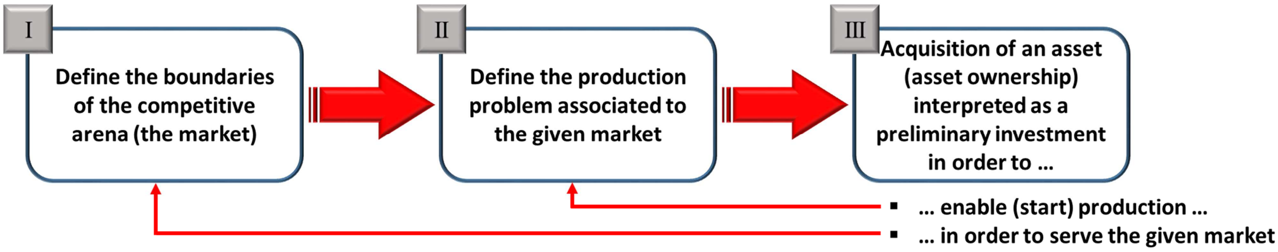

Figure 1.1 – A graphical representation of the conceptual logic behind the antitrust approach: ownership instrumental to production and the production problem causally determined by the preliminary definition of the competitive arenas (the markets)

For these reasons, the problem examined by any antitrust approach is not really *how ownership affects the way in which an asset should be used* or, in other words, which market (or set of markets) a certain asset owner should serve. Conversely, the antitrust approach already assumes that the market in which an asset should be employed is given and is not part of the economic problem at stake: *the economic problem is not given by the choice of the set of alternative markets that should be served*. The economic problems to be analysed then derive from the fact that (consistent with the approach adopted by antitrust authorities) the private acquisition of an asset by the part of any firm entails fixed or sunk costs in the long-run, and these fixed or sunk costs necessarily have implications for the short-run competitive environment of the *given market*, either with respect to the level of prices (p), or with respect to the volume of production of a given product (q) (price and quantity being the two short-run marginal variables).

The way in which this investment in sunk of fixed assets (in order to enter a certain market) should be treated and interpreted depends on the specific branch of the “antitrust approach” that is adopted, as recalled in section 1.1 above. Thus, the monopoly approach perceives any investment as an entry barrier to the specific market. Asset ownership in this case becomes the precondition to the concentration of a certain given market, and to the presence of some degree of monopoly power in that final market. Conversely, both the efficiency and property-right approaches treat investment in incentives terms.

As a result, the difference between the new "incentive-based" theories dealing with asset ownership (mainly TCE and PRT) and the original monopoly approach does not derive from the way in which the two branches interpreted asset ownership. In both cases ownership is intended as the fixed or sunk investment required in order to enter a certain market and this aspect of the standard monopoly approach was retained and shared by all theories belonging to the so-called "antitrust approach".

Conversely (as already remarked), the key difference between the new theories dealing with asset ownership and the mainstream "monopoly approach" of industrial organization, derives from the way in which they interpreted the *incentives shaping firms' motivations to acquire assets*, and therefore *to invest* in fixed or sunk assets in the first place: not as a mere detrimental barrier to competition in the given market any more, but rather as a "defence" against socially detrimental market failures. These failures are interpreted by these "incomplete contract theories" mainly as possible unforeseen situations that would otherwise restrain firms' commitment to invest in the acquisition of certain assets and produce certain outcomes. As a result these failures are detrimental to social welfare as they would undermine a healthy and (socially) needed level of investments, and therefore of volume of production, in certain sectors of the economy.

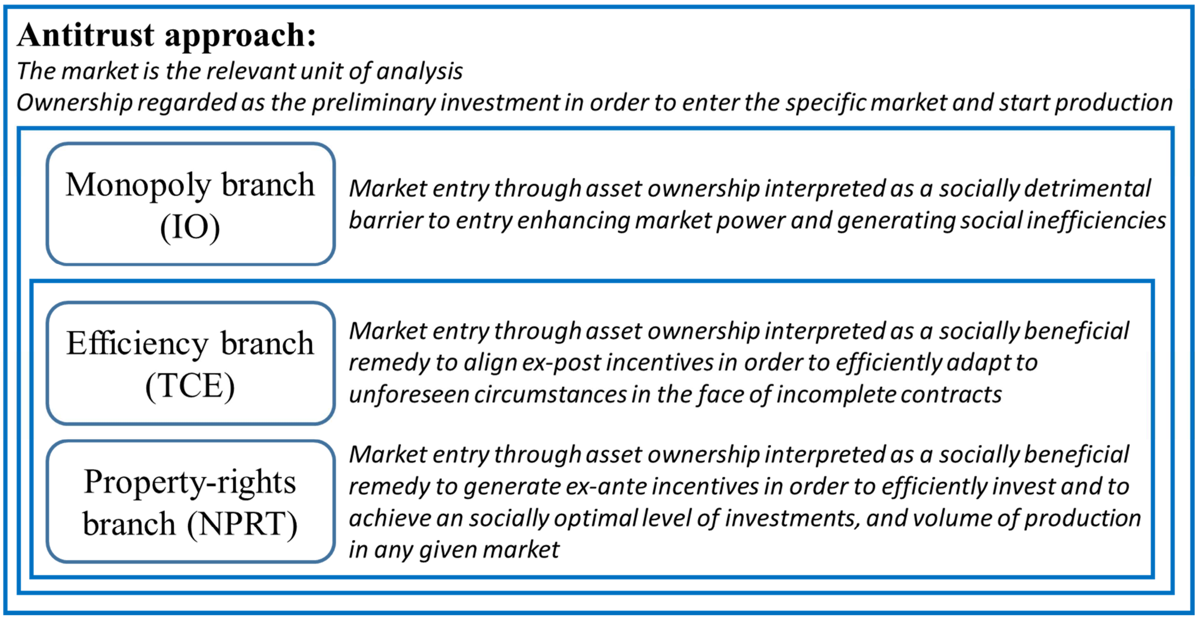

Figure 1.2 – The various branches of the antitrust approach. While their interpretation of the institution of ownership differ, together with their analysis of the implications of ownership, they all share the key characteristics of the antitrust approach: assets are acquired given a certain employment is decided first, meaning that ownership is undertaken in order to (conditional to) serve a certain given market (or given sets of

markets). The choice of the market (or set of markets) to serve, and therefore of the production process to undertake is never dependent on, or constrained by, the institution of ownership in the first place.

In brief, in the new incentive-based approaches, the “sunk” commitment deriving from a firm’s decision to acquire an asset, and therefore to own it, is not regarded as a mere beneficial fact for the asset owner any more (as a generator of entry barriers into the industry, thus of monopoly power), but also as a possible source of inefficiencies and threats for the asset owner. While asset ownership has to be sanctioned and discouraged in the former (monopoly) approach, it has to be encouraged and safeguarded in the latter (efficiency, or rights-based) approach. This key difference has been already highlighted above, and is well described by the relevant literature.

Yet, the analytical and epistemological approach towards the institution of ownership adopted by these theories is substantially identical, and it corresponds to the logic of the antitrust approach: ownership is instrumental to the production problem, and both are subsumed, and causally subordinated, to the allocation process. And this simply means that the decision concerning how an asset should be employed across alternative uses (or markets) is not delegated to the specific asset owner, but already solved ex-ante, when the various competitive arenas (markets) are defined.

2.2. The regulatory approach and the 1980s deregulatory wave: unbundling between asset ownership and control (of services)

In the meanwhile, things changed not only in the realm of abstract theory, but also in the realm of practical policy making. The deregulatory wave that began in the late 1970s, and fully matured in the 1980s, tackled the problem of asset ownership from a different perspective. This different perspective can be renamed as the *regulatory approach*.

The aim of regulatory authorities and policy makers starting from the late 1970s has been to introduce competition in those industries where the private ownership of an “essential facility” prevented the development of competition. The standard solution usually adopted by regulators (especially in the US) in order to deal with the private ownership of monopolistic (essential) facilities was the institution of common carriage: owners of essential facilities were obliged to provide services on a non-discriminatory basis.

Conversely, starting from the late 1970s, a new solution was adopted: the separation of asset ownership from the services implemented over it. In other words, the essence of the deregulatory wave has been to sever the ownership of the underlying asset from its control: essential facilities can still be owned by private actors, but the latter do not have full control over their usage in the sense that infrastructure owners cannot fully restrain or exclude others from the access and the use of their own property. Moreover, owners cannot even independently decide how their own privately owned infrastructure should be used, which services should be provided and therefore which markets should be served.

In the remainder of the work "control" is defined as the faculty to decide how an asset, good or resource ought to be employed and allocated among alternative uses. Thus, the control of an asset confers the faculty to decide *which* services should be delivered or performed, and *how*.

In sum, the separation of the ownership of the underlying physical infrastructure from its control represents the essence of the regulatory approach.

It should be noted that any regulatory measure deals with, and affects, the control of an asset. In fact, this is precisely the purpose of regulation itself, which can be defined (and is here defined) as *the imposition of socially-defined constraints over the private control of a (privately owned) asset*. In other words, regulation is the primacy and the super-imposition of *public juris* over the *jus privatum* of the asset owner. For this reason, the regulatory approach does not focus on the economic implications of the mere *ownership of an asset* (like the antitrust approach does), but it rather focusses on its control or, in other words, on the way in which it ought to be employment and used.

Thus, the institution of common carriage (as any other regulatory measure) also represented a decoupling between ownership and control. Yet, the new regulatory wave was substantially different from the good-old institution of common carriage as control was not decoupled from ownership vis-à-vis final consumers, but the decoupling was mandated vis-à-vis other competitors through open access, sharing, leasing and resale. In the new regulatory framework, competition has to be implemented "over the top" of assets and infrastructures, at the service level. Wholesale markets (the market for the access of the infrastructure) are created next to retail markets (markets to sell the final service to consumers).

Since the late 1970s, a *duality asset-service* is therefore introduced: service-based competition is introduced beside (over-the-top of) facility-based competition. The two conceptual and legal layers are closely linked as one relies on, and is constrained by, the other and vice versa. Yet, they refer to different steps of the economic analysis. More precisely, service-based competition has usually been perceived as a preliminary alternative to the most sustainable facility-based competition.

The economic term used to describe this new process was “unbundling” (or “separation”), coupled with other brand new concepts such as the ones of “open access”, “sharing”, and “non-discriminatory access”. This new (de)-regulatory legacy raised new economic problems such as price and, especially, non-price *discrimination*.

The previous “simple” regulatory regime based on the institution of common carriage starting from the 1980s was substantially abandoned, and has been gradually replaced by a new “dual framework” characterised by a cohabitation of regulation and antitrust at the same time: while the former ensures open and non-discriminatory access at the wholesale level, and sets pricing rules to access the underlying assets, the latter ensures fair and non-monopolized competition at the retail level (final consumption market).

This cohabitation is not simple to implement: the specific way in which antitrust and regulatory authorities have to interact, and the scope of their respective spheres of influence is still unsolved. After the “Trinko” decision in 2004 by the US Supreme Court, antitrust has been subsumed to regulation, while the latter has been given legal and conceptual priority. Moreover, the way in which the new concepts of “open access”, “non-discriminatory treatment”, together with the even newer one of “neutrality”, have to be interpreted, modelled, and implemented is also still very confusing. What does “non-price discrimination” actually mean? What about “neutrality”? How should they be conceptualised and treated? There is still quite a big confusion in the relevant literature.

2.3. The different meanings of asset ownership and vertical integration: the antitrust approach vs. the regulatory approach and implications for the competition law-deregulation debate

A clear problem emerges from the previous discussion: the approach towards asset ownership (and private property) adopted by the aforementioned economic theories developed between the 1960s and the 1990s is of a completely different nature from the approach towards asset ownership (and private property) adopted by regulators in general, and by the deregulatory wave of the 1980s and 1990s in particular.

The so-called “antitrust” approach takes the single market as its fundamental unit of analysis, and looks at the leverage of monopoly power through vertical integration across different technological interfaces. On the other hand, the “regulatory” approach takes the single asset as its unit of analysis, and focuses on the specific way in which the asset ought to be used and employed, and therefore on the way in which the owner of the asset implements services over-the-top of it. In this second case, vertical integration is regarded as the conflation of the ownership of an asset with its control (the control of the services and functionalities provided by it). Thus, an integrated supplier for a regulatory point of view not only owns the underlying asset or infrastructure, but it also retains control over its employment, meaning that it can also supply its own services, at its own terms.

Stated differently, the antitrust approach embedded in standard IO and employed by antitrust authorities, is interested in the implications of the *merger of two different separated technological modules (the two assets) resulting in the common ownership of two different assets given a certain final market*. Vertical integration is therefore intended as the common (joint) ownership of two separated assets given a single final market.

As already mentioned, on the one hand, from a "monopoly perspective", the main problem deriving from this common ownership is represented by the leverage of monopoly power across stages. On the other hand, in contrast with the standard "monopoly approach" adopted by antitrust authorities, new theories starting from the 1970s (NIE, PRT and OE) highlighted that the main benefit of common ownership derives from the alignment of incentives, and from the efficient coordination of the different stages towards a same objective and a common goal.

In a nutshell, in the antirust approach (no matter whether it is a monopoly-approach, or an efficiency-approach) *to own a certain asset is beneficial or detrimental in function of the ownership of some other contiguous asset*. Optimality and efficiency (what is beneficial and what is detrimental) is computed according to the implications that (i) leverage of monopoly power, (ii) coordination and (iii) incentives' alignment *across these stages* (modules) have for the level of investment and, ultimately, for the volume (and price) of production of a given output in a given market. Ultimately, the optimal set of assets grouped under the same ownership defines the "boundaries of the firm". This is the essence of the antitrust approach, and this approach is genuinely in partial equilibrium (as the focus of the analysis is on the market, not on the constrained allocative problem).

This is clearly *not* the meaning of "vertical integration" provided by the regulatory approach, and also adopted by regulatory authorities starting from the 1980s by unbundling and mandating open access to infrastructures. In this second case, vertical integration is intended as the merger of the wholesale stage (the ownership of the asset) together with the retail stage (the implementation of the services over the asset). In this case the asset owner is vertically integrated if she can also provide services over-the-top. The rupture is not between two distinct technological stages, represented by the two owned assets. Conversely, the rupture is between the *ownership of an asset* and its relative *control*, where the latter is intended as *the power to decide how the asset should, or ought to, be employed across alternative services or uses* or, in other words, *across alternative markets*.

As a result, in the "regulatory" case, "to vertically integrate" does not mean to acquire a further (contiguous) asset and to sum the ownership of two assets. Nor is the focus of integration and asset ownership in this second case on the efficient coordination of two autonomous (but interdependent) stages. Conversely, "to vertically integrate" in this second case implies that *the asset owner is also able to fully control the employment of his own asset*; this means to have the power to use or employ the asset for his own private purposes, to decide and select which markets to serve and which services to provide, to control the way in which services are delivered to the final consumers, over-the-top of his own asset, and to also control the way in which others might be able to use it. The focus is on the decoupling between the *ownership* of the asset (the underlying wholesale stage) and its *employment* (the services implemented over-the-top of it). Stated differently, it is not an *asset-asset* integration, it is an *asset-service* integration.

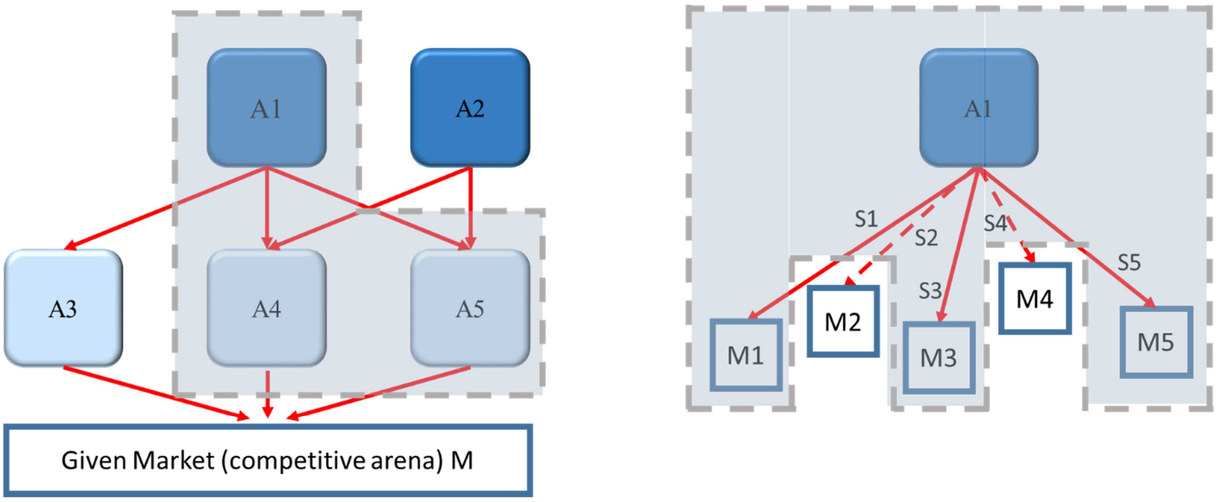

Figure 1.3 – The left panel retraces Williamson's TCE framework and shows the classic make-or-buy problem in an antitrust approach, where boundaries and integration (horizontal and vertical) are defined based on the set of assets (A1, A4, A5) jointly owned within the same propriety given a certain final market: in this case integration is defined as an asset-asset merge; the right panel shows the meaning of integration in a regulatory framework: in this case firm 1, owning asset A1 vertically integrates based on the set of services that it can provide (S1, S3, S5), in order to serve alternative final markets (M1, M3, M5). A regulatory approach to integration is defined as an asset-service merger. The dotted arrows corresponding to S2 and S4 mean that those services could have been potentially provided by the firm, and the corresponding markets potentially served, but it has been decided not to. Note that vertical integration in a regulatory approach in principle does

not limit itself to a mere decision of which markets to serve, but it also implies some form of control over the owned asset, meaning that the firm should also be able to decide (to a certain degree) how these various services should be provided.

Obviously, this second (regulatory) approach acquires relevance and makes sense only if there is no unanimous agreement on the way an asset ought to be employed and, therefore, on the market (or set of markets) that ought to be served, given the asset itself. This means that the regulatory approach becomes meaningful only if assets are not service or market specific, meaning that they are not completely dedicated to one single service, or employment. Here, assets that can only be employed in one single market will be called *mono-service assets*, while assets that, once acquired, can be employed in various ways, therefore across various markets, will be referred to as multi-services assets. Oliver Williamson's specific assets, in the sense that they are completely dedicated to one single use, are clearly mono-service assets.

Clearly, no control issue arises if an asset can only be employed in one specific way, and can only serve one single market, in the sense that every single actor would agree on the best way in which that asset should be employed. In this case of completely dedicated assets, it becomes clear that there is no real difference between an antitrust approach (where the given final market is defined first) and a regulatory approach (where the choice of the specific employment, or market, has to be decided ex-post, by the ones controlling the asset) when assets are mono-service. For this reason, Williamson's case of complete specificity can be correctly treated in monopoly terms (as TCE does), thus consistent with an antitrust approach to asset ownership.

As a result, there is another key difference implicitly deriving from this distinction between the antitrust and the regulatory approach. As already mentioned, in the antitrust approach the unit of analysis is the single final market in consumption, while the institution of asset ownership is analysed by looking at the implications that this has for the competitive level of this certain single market in consumption. Moreover, the definition of the market (the boundaries of the competitive arena) *causally precedes the institution of ownership*, meaning that the latter is analysed in function of the former (given the former).

Conversely, in the regulatory approach the unit of analysis is not the single, *given* market (or given set of markets). To control an asset precisely means *to decide which markets should be served* given a certain owned asset. Thus, in this case th*e institution of ownership causally precedes the institution of competition*, meaning that "boundaries of the analysis" are not provided by *the boundaries of the competitive arena* (i.e. the market), but *by the* *boundaries of the opportunity set of the asset owner*. In this case, the opportunity set is given by the scope of markets that the owner of the specific asset is potentially able to serve, thus from the services that he is potentially able to provide.

In this second case, the economic analysis of production (which depends on the specific market/s eventually decided to serve) should be subordinated (logically, temporally and analytically) to the preliminary acquisition of the asset by the relative owner, and therefore to the institution of asset ownership.

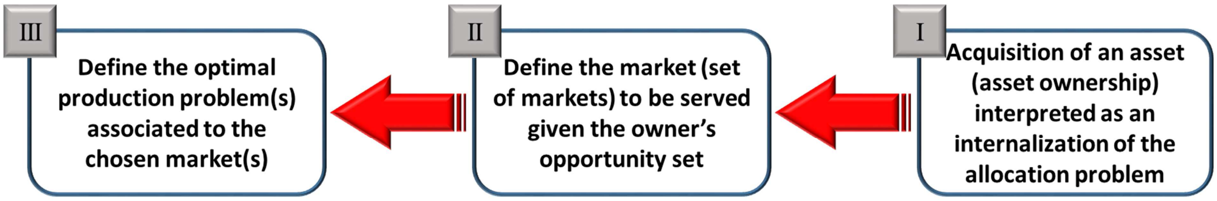

Figure 1.4 – In the regulatory approach the logical, analytical and temporal causality is reversed, vis-à-vis the antitrust approach. Actors first acquire an asset (they internalise a private property), and only subsequently, given the characteristics of their private property, and given their own (subjective) opportunity set, decide how to employ the asset or, in other words, how the productive capacity of the asset should be allocated among alternative final markets. Once the set of markets is defined, the optimal production problem is set-up and solved, in function of the specific services (outputs) to be provided.

In a nutshell, while in the antitrust approach the institution of ownership follows the institution of market, the regulatory approach reverses the causality, making the institution of market logically subordinate to the institution of ownership: actors first own (appropriate) an asset, then they decide how the latter should be employed or allocated (and not vice versa, where the specific given employment determines which asset should be acquired/invested into). Only lastly the specific production problem to implement (or resolve) is devised *in function of the specific market* (or set of markets) that the asset owner has originally decided to serve (or compete into).

The causality (logical, temporal and analytical) in the regulatory approach is completely reversed, and this derives from the fact that it is the starting point, the unit of analysis, to be reversed: from a given market, identifying a specific outcome, or service, to be provided or sold, we switch to its dual, the constrained allocation problem identifying a potential and alternative set of outputs, or services, to be provided or sold, given a certain resource. This reversed causality is graphically represented by figures 1.1 and 1.4: while the flow of causality runs from left to right (thus, from the given output market to be served, to the acquisition of the input asset) in figure 1.1, it flows from right to left in figure 1.4 (from the given asset as an input for production, to the choice of the set of outputs to serve).

3. Wrapping up: Incentives-based theories of asset ownership are unfit to deal with the regulatory approach

Neither of the aforementioned theories adopting the “antitrust approach” provides a convincing theory of asset ownership grounded on the dichotomy between ownership and control. Nor do they provide a satisfactory interpretative framework explaining how the ownership of the supporting infrastructure (the underlying asset) affects *how* the asset is employed, the *set of markets* served, the *types* of services provided, and *the ways* in which these services are provided (control of the way in which the asset is employed among alternative services).

In a nutshell, all theories rooted in the antitrust approach assume a *univocal relationship* between the (objective) characteristics of an asset, and its ultimate employment, or the market (or set of markets) the assets should serve.

On the one hand, NPRT is not interested in the way in which assets might be employed across alternative markets. As already recalled, the goal is a completely different one: to understand the optimal ownership arrangement (who should own what) in function of ex-ante investment incentives *given a certain market to be served*. The way in which an asset ought to be employed cannot be (endogenously) decided by its respective owner, but it is (exogenously) given to him.

On the other hand, when it comes to the relationship between the property of an asset and its employment, Oliver Williamson’s TCE represents a partial exception to the previous critique, while probably representing the closest framework to the present work. In effect, Williamson’s focus on actors’ ex-post incentives in adaptation can be intended as a very rough and primitive interpretation of the relationship between asset ownership and control. Not surprisingly, the main critiques to the open access policy introduced by the 1996 US Telecommunication Act have heavily relied on the concepts and the intuitions developed by TCE in order to criticise the liberalization of the retail market through asset unbundling. Moreover, many of the intuitions developed by Williamson have a strong regulatory flavour, and appear to have been influenced greatly by the debate on the deregulation of common carriage that ensued between the late 1970s and the early 1980s. More precisely, the concept of asset specificity was clearly influenced by the common carriage debate in the late 1970s[1] and is, at least in principle, a genuine regulatory concept, not an antitrust one.

Yet, it is undeniable that TCE adopts a clear antitrust approach to its analysis. As already mentioned, already asset specificity is, in principle, a concept developed in the regulatory literature, it is redefined by Williamson according to clear antitrust categories, such as the extent of scale and scope economies for the production of a given output (or set of outputs), and the degree of monopoly power within a given market (generating hold-up and bargaining power). In the same vain, the private acquisition of assets is regarded as a form of entry barrier, or sunk cost investment, influencing the relative monopoly power of two parties within a certain given market the asset naturally refers to, as already stated.

Consistent with these features, the "fundamental transformation" (which probably constitutes the key concept of TCE) is defined in function of the (dynamic) concentration of a certain market (the given competitive arena) following a certain (transaction specific) investment in production, consistent with a genuine antitrust approach. This is consistent with the fact that, despite the commendable intentions, TCE substantially always works with completely dedicated (mono-service) assets, which are always independent from the problem of control, and can always be analysed adopting a pure antitrust approach, as already remarked.

Moreover, TCE has a fundamental shortcoming that will be briefly discussed later in the section and that highlights its genuine "antitrust" nature: it merely focusses on underlying "material" support (the asset) and on the relative *ownership rights*, while it completely neglects the use rights held by different parties over the services provided by the same assets. This last point is consistent with what has been already stated above, namely that TCE is a theory of asset ownership, but is substantially silent on the way in which those assets ought to be used and employed by different actors. In other words, TCE is a theory of asset ownership, but does not provide any theory of control rights (or use rights). The most evident proof of the "antitrust nature" of TCE is therefore the way in which firms (and their boundaries) are defined and conceptualised: based on the *set of assets owned by the firm*, and not on *the set of activities or services under direct control of the firm*.

In conclusion, not even one of the economic theories that have been developed starting from the 1970s appears suitable to properly understand what the unbundling (or

[1] Not surprisingly Williamson introduces the concept in 1979 (Williamson, 1979), only after Klein et al. (1978) paper.

separation) between the underlying asset and the services provided over it actually means, how open access and sharing should be interpreted, how these phenomena should be treated, and what their economic and welfare implications are.

In order to understand how ownership and control interact, and how regulatory concepts such as the ones of unbundling, open access, and non-price discrimination should be interpreted, the work suggests that it is necessary to shift from a market-failure approach to the institution of asset ownership (suitable only in an antitrust perspective), to an opportunity-based approach to the institution of asset ownership (necessary when a genuine regulatory approach is required).

In fact, there is another common shortcoming linking all theories developed along the "antitrust" tradition, as already remarked in section 1.1.2 above: they all share a *market-failure approach* to the institution of ownership, and to institution in general. This is a further feature of TCE that renders the theory profoundly rooted in the antitrust tradition.

In a nutshell, the argument of the present work is that it is not possible to correctly conceptualise the *dichotomy* between an *antitrust approach* to ownership rooted in the enforcement of perfect competition tradition, and the *regulatory approach* to asset ownership, rooted in the distinction between public and private control, if a market-failure theory of institutions is not first contrasted and replaced by a new opportunity-based theory of institutions.

For this reason, the above conclusion can be generalised by saying that the preliminary aim of the present work is to trace the foundations for a new economic theory of institutions (and of asset ownership) that replaces the (objective) concept of market failure with the (subjective) one of opportunities and opportunity set.

To this is dedicated the bulk of the present work, and to this last point we now turn, before illustrating how this aim can be achieved, and subsequently applied to the (de)-regulation of assets and infrastructures.

CHAPTER 2

A reconsideration of Value theory in classical political economy: use value, exchange value, and implications for an new economic theory of institutions and asset of ownership

Chapter synopsis

The chapter is dedicated to the topic of value. The concept of value has been dismissed from the realm of economics since the neoclassical "solution" of the classical paradox of value, and its concomitant replacement with the concept of price. In fact, one way (but far from the only way) in which neoclassical economics can be defined vis-à-vis classical economics is by using the distinction between value and price: while the two are interchangeable in the former, they remain (at least in principle) two distinct concepts in the latter. This is especially true for the Ricardian legacy in value theory (as opposed to the Malthus-Mill approach).

As a result, the natural starting point of any work willing to enquire about the origin of the "neoclassical" (market-failure) approach to the theory of institutions and asset ownership must be the classical, and especially Ricardian, theory of value. More generally, any theory of ownership and institutions disjointed form the concept of market failures, and independent from the performance of the pricing mechanism, must start from a reconsideration of the theory of value, and from a separation between the concept of value and price. For this reason, a proper understanding of the characteristics of classical value theory is in order.

The chapter begins by establishing that the way in which institutions are modelled in economic analysis directly depends on the way in which value is modelled. It will be possible to identify as many (heterogeneous) sources of institutional arrangements as many (heterogeneous) sources of value are identified. Thus, if value is treated as a monism, it will be possible to identify one single source of value, meaning that institutions can be explained and justified by one single mechanism. This is the case of neoclassical economics where a monism in value theory (market price) corresponds to a monism in institutional sources: the market mechanism (and its failures). Conversely, an economic theory postulating a dual source of value will necessarily recognise two concurrent, heterogeneous and completely independent, institutional layers. This is the case of classical economics, where value can acquire two distinct natures: "*in use*" and "*in exchange*".

After a brief overview of the meaning of value in Ricardian economics in section one, followed by a detailed discussion of the classical paradox of value in section two, the main argument of the chapter is developed in section three. In a nutshell, two logical steps can be identified, with the second providing the justification to the first one.

First of all, the chapter explores whether it is possible to identify two distinct "institutional layers": the market and the ownership of property, if the classical paradox of value is retained,. The latter represents private endowment of material wealth in classical economics. The two are independent and heterogeneous, thus they exist, and have meaning, independently from each other. While the market is the institutional layer where value-in-exchange originates, the institution of asset ownership is the layer where value-in-use originates. In sum, it is sufficient to assume a dualism in value theory, rather than the neoclassical monism, in order to appreciate the nature of the institution of asset ownership, and its role in economic affairs, independently from the institution of exchange.

Secondly, the chapter develops the conditions ensuring that the classical dualism in value theory holds, and cannot be reduced to the standard monism: owners (actors endowed with ownership rights) must be able to employ their own properties "in-house". This means that the dualism in value holds whenever property owners are able to use their property as a means of addressing their own private consumption or, vice versa, whenever they are able to auto-consume the services produced by their own property.

The section argues that as soon as any good or resource can be indifferently interpreted *either* as *means for self-subsistence and auto-consumption* (use value), *or* a *means for market exchange* (exchange value), value acquires a dual nature. This leads to a new definition of the concept of asset ownership independently from the market institution, as that institution that confers to the owner of any good the faculty (rights) to satisfy their own needs by means of self-subsistence (through the private auto-consumption of the good), while disentangling them from the external social context.

The same concept can be expressed also in other, more meaningful words: asset ownership acquires an autonomous economic role whenever the sharp duality in the status of economic actors, severed between "consumers" and "producers" is discarded, while consumers are allowed to produce their own consumption, and producers are allowed to consume their own production.

The section reinterprets the duality of value in light of Adam Smith's division of labour, by showing that the two natures of value dynamically interact in function of the extent of the division of labour: the more primitive a society, the more relevant the relevance of the "use part" of value, and therefore the role of private property in shaping economic affairs. Conversely, in extremely developed societies where the division of labour is pushed to the extremes, satisfaction of the personal needs "in house" through the use of privately owned property becomes less and less likely, leading to a much higher prevalence of the share of value generated "in exchange". Use value and exchange value should therefore be intended as complementary, where their relative role is shaped by the characteristics of the external context. The section argues that the most appropriate way to interpret Smith's description of the division of labour is in "value terms", by looking at the implications of the dual nature of value, rather than in "production or efficiency terms", by employing genuine neoclassical concepts, such as the ones of production (scale and scope) economies.

Building upon this, the section also argues that the institution "firm" can be understood exclusively in function of the institution of private property (as opposed to the institution of market failures) by simply interpreting the make-or-buy problem in value terms, distinguishing between two different situations: one characterised by the fact that the status of producer and the one of consumer are clearly severed ("buy"), and one where a producer also becomes a consumer of his own produce and vice versa ("make"). While the source of value in the former is "in exchange", the source of value in the latter is "in use". "The nature of the firm" will therefore depend on the relative role of use-value vis-à-vis exchange value. The peculiarity of this approach derives from the fact that it is possible to explain the "make-or-buy" problem independently from the performance of the exchange (pricing) mechanism, while only focussing on the value-in-use deriving from the private ownership (and employment) of a certain asset.

Section four generalizes the above considerations and identifies three different cases that ensure that the two statuses of producer and consumer remain severed, thus guaranteeing monism in value theory: (i) absence of private ownership of resources so that any need and necessity must necessarily be satisfied by the purchase of services from third parties (the public authority); (ii) completely specialised assets, deriving from the perfect division of labour, so that the share of private needs that the property owner can satisfy "in-house" through his own assets, as compared to the share of needs satisfied for third parties, tends to zero; (iii) public regulation and social constraints on the way in which private owners *ought to* employ their own property, so that private owners are prohibited from employing their own property "in-house", for their own private self-subsistence, and are mandated to exclusively "hold it out" to the public.

The section briefly discusses how any normative approach to economic theory must necessarily always relies on at least one of the aforementioned cases in order to guarantee monism in value, and therefore justify and legitimize normative judgements on how actors ought to behave.

The section concludes by anticipating two fundamental themes developed in the next chapters.

The first is an institutional monism, following a monism in value, also implies an ontological monism: whenever the dualism between value-in-use and value-in-exchange is resolved, the distinction between material wealth in the form of physical assets, and immaterial wealth in the form of transacted services also disappears and services can be taken as the new unit of analysis, by replacing the classical notion of "factor". The duality between assets and services will represent a key aspect of the remainder of the work.

Secondly, the chapter introduces the very important concept whereby any social normative prescription must overcome the problems deriving from the multiplicity in valuations, and should ensure a collective and unique valuation framework. In other words, monism in value is needed for any normative judgment to be justified and legitimised, and the same holds for a normative theory of regulation. The chapter already briefly anticipates that, differently from the case of dedicated (specific) assets, the separation of consumption from consumption is alone, not sufficient to ensure monism in valuations whenever assets can be employed across multiple services. In order to ensure public control over the employment of assets a second obligation is needed also defining *how* a certain asset or resource *ought to be employed*. The sum of the two normative constraints constitutes the essence of public regulation as developed in common law, as chapters 6 and 7 will discuss later on.

1. Foundations of classical value theory: from Adam Smith to David Ricardo

The present section discusses the epistemological origins of the classical labour theory of value, and briefly illustrates Ricardian value theory, thus providing the foundations for the subsequent discussion. It especially clarifies the meaning of "classical value theory", and justifies the approach here adopted from an epistemological and historical point of view.

1.1. Is there a "classical theory of value"? Epistemological problems behind Smith and Ricardo, and the meaning of classical value theory

First of all, it is important to remark that the Ricardian framework as described in the *Principles* is probably the first clear example of a theory of value that is completely organic to, and consistent with, the overall framework it is placed into: Ricardian value theory is an integral part of his broader construction and gives substance to it.

Differently from Ricardo, Adam Smith, the only other main classical economist that wrote a full comprehensive treatise in political economy before Ricardo (at least in England),[2] did not have a clearly developed theory of value. At least, he did not have it in the sense that in Smith's *Wealth* there is no trace of a proper economic discussion explaining *how value theory and distribution theory are consistently and organically interwoven*.[3] As summarised by Sinha (2010: 48; emphasis in the original):

> "The epistemological background of Smith's theory is not the same as the modern epistemological background of unidirectional *cause and effect* relationship. Smith simply does not recognise our persistent question: is the distributional variables that determine the value of a commodity or it is the value of the commodity that determines the distributional variables?"

[2] "[B]etween 1776 and 1817 'not a single complete treatise on political economy appeared in England. Adam Smith remained the only authority, and he was little heeded'" (Elie Halévy, cited in Dobb, 1973: 65). The sole exception certainly worth mentioning is Jean-Baptiste Say who published his *Traité* in 1803. Yet the economics of Say is very different from the economics of English classical economists of the first decades of the 19th century, especially with respect to value theory. The ideas developed by Say are paradoxically much closer to the ones developed in England by economists two generations younger than him such as J.S. Mill, who was greatly influenced by Say.

[3] Thus, Peach (2009: 405) in his recent review concludes with the remark that "it must be granted that Smith's labor theory approach was extremely rudimentary. At least for the commercial economy, the focus was exclusively on differences and changes in exchangeable value, apparently with no recognition of the theory's implication for distribution …" However, O'Brien (2004: 92) gives a different interpretation of Smith's value theory, sustaining that "there was no very obvious way of linking distribution to value except by a cost of production theory of value"; thus implying that Smith's value theory has been developed precisely in order to respond to the complications arising in distribution theory.

So, for example, Blaug (1996: 37-38) goes as far as to deny that Smith has "any theory of value whatever", while O'Brien notes that there is a conceptual circularity in Adam Smith's treatment of value and distribution, implying that Smith "is not really offering a theory of value." (O'Brien, 2004: 92). Thus, if these interpretations are accepted, certainly there could be no other early economist "more classical" than Ricardo with respect to value theory, at least from the mere temporal dimension: classic political economy (the theory of value and distribution) has to start from Ricardo, if the criteria of judgement is the degree to which a theory of value is organic to a theory of distribution, *and vice versa*.

On the other hand, many concepts in value theory developed by Ricardo have been, in fact, borrowed from Adam Smith. Ricardo himself conceived of his work just as a refinement of Smith's ideas and this is especially true for the theory of value.[4] For example, the fundamental distinction between use value and exchange value, that will play such an important role in the reminder of the present work, was borrowed by Ricardo from Smith's *Wealth* (fully acknowledging the source and quoting the passages). This is also true for Ricardo's refinement of Smith's interpretation of labour as the ultimate source of (exchangeable) *real value* in a barter economy: both Smith and Ricardo distinguish between an "ultimate" source of value (real or absolute value) and a contingent "fluctuating" nominal value. The former is exogenously given and price-determining, while the latter is endogenously determined by the economic system, and corresponds to the actual, contingent, market price.[5]

Thus, it is certainly true that the *Wealth of Nation* does not develop an organic framework where value theory has its own distinct role in explaining, or complementing, all the other aspects of the economic system (and particularly distribution theory). Yet, it is undeniable that Smith did develop some theory of value.[6] As succinctly put by O'Brien (2004: 98), Adam Smith's theory of value "provided the framework for all the developments in Classical value theory", so that few would deny the status of (first) classical political

[4] For instance, this is how Schumpeter (1954: 472) summarises this issue: "Ricardo's theoretical structure represents a particular way of reconing the *Wealth*; Malthus theoretical structure represents another way of doing this." And later he adds further (p. 482) that "Ricardo's work, so far as general theory is concerned, started from the Wealth of Nations and recoined the latter's theoretical contents by a method that centered in the concept of value. Exactly the same thing is evidently true of the work of Malthus as presented in his Principles."

[5] For a synthetic account of Smith's distinction between real and nominal value the reader is referred to O'Brien (2004: 93-98).

[6] The literature on Adam Smith is of course immense, and is outside my capacity to present any representative list of works. With respect to Smith value theory, the reader is referred to the recent, fairly comprehensive, works of O'Brien (2004: 91-98) and Sinha (2010: chapter 1), and references therein.

economist to Adam Smith. But even in this case, the problem with Adam Smith's theory of value derives from the fact that Smith appears to be inconsistent throughout *Wealth* when he deals with the concept of value. It was on Smith's fundamental inconsistencies, or unsolved problems, that Ricardo (reinterpreting Smith's teachings) built his own framework (similar inconsistencies were however later assigned to Ricardo by Malthus but also Marx).

Smith's inconsistencies with respect to the subject of value are numerous and of different natures.[7] Probably the most relevant inconsistency is represented by the fact that while Smith (nearly everywhere) adopts examples and explanations solely based on labour quantity (making value an extra-economic and non-monetary entity), he also makes (albeit less often) references to the role of money-wages and to the three monetary components (wages plus rents and profits) in determining and explaining exchange value, and therefore prices (and it is at this point that Smith loses himself, and generates a circular reasoning between value and distribution theory). While the first approach was the one followed by Ricardo and the "Ricardians", the second was taken by all other classical and neoclassical economics (starting from Malthus and Torrens).

Thus, some scholars have interpreted Smith's approach as a cost-of-production theory of value that only becomes a labour theory of value in some special cases (the "early and rude societies") when only labour is employed (Blaug, 1996: 38; Dobb, 1973: 43-47; O'Brien, 2004: 91-92; Skinner, 2008). This is what Adam Smith sometimes declares through his *Wealth*, and this is also the starting point of Ricardo's critique. Others have been more dubious and have interpreted Smith's treatment much closer to a labour-theory of value *à la* Ricardo.[8] Peach (2009), for example, represents a recent reinterpretation of Smith's value theory in labour terms,[9] and even the recent treatise of Sinha (2010) seems to point

[7] There are many other confusions emerging from Smith's treatment of value, beyond the labour-cost vs. production-cost theory of value. For instance, it is not clear whether Smith deals with labour-value in absolute terms, as an analytical numérarire, or in relative terms. Also, it is not clear whether Smith computes value from the producer (labourer) perspective, or from the purchaser's perspective in a bargaining situation (thus whether the "quantity of labour" is considered in order to produce a commodity, or in order to purchase it). (O'Brien, 2004: 95-98; Schumpeter, 1954: 188-89; Sinha, 2010: 3-19).

[8] The difference between a cost-of-production theory of value and a labour-theory of value derives from the fact that all distributives shares cooperate for the determination of market prices (all distributive shares are costs, thus price-determining) in the former, while labour is the only source of value and cost in the latter. In this second case, all factors contributing to production can be reduced to, and expressed in, labour terms. Moreover, this makes all other distributive shares price-determined residuals, while only labour remains the sole price-determining (thus cost-determining) component.

[9] See also the study of Douglas (1927: 62; emphasis added) where it is observed that it is the introduction of money wages that "paves the way for a value theory based on the money cost of

in this direction.[10] In conclusion, the nature of Smith's value theory is far from clear and universally accepted.

Whether Smith's value theory was a labour theory of value, or just a cost-of-production theory of value does not interest us here, and we can leave the dispute to the historians of economic thought. More interesting for the present purposes are the epistemological and methodological implications that this confusion has generated through the decades (and still generates). In fact, depending on how classical political economy is defined, and depending on how Smith's value theory is interpreted, Smith can be grouped together with Ricardo, or he can be conceived as substantially different from Ricardo, and more closely associated to other economists such as, for instance, J.S. Mill.

Still, regardless of the disputes over the correct interpretation of Smith's theory, and no matter which definition of "classical political economy" (as opposed to the so-called "neoclassical" economics) is adopted, it is certainly possible to assert that both Ricardo and Smith's theories of value share some peculiar and paradigmatic aspects.

More precisely, two characteristics of both Ricardo's and Smith's value theory shall be highlighted. On the one hand, both Smith and Ricardo accept the paradox of value: they both decouple use value from exchange value, thus ruling out the role of utility (or usefulness) in determining "natural" or "real" value. On the other hand, they both distinguish a "real" (absolute or natural) price-determining *non-monetary* source value in the long-run, from a "nominal" monetary expression of value in the short-run: the market price.[11] Both authors (not only Ricardo) regard the former as solely determined by labour, thus substantially adopting a labour-theory of "real" value (even though this point is quite confusing and contradictory in Smith, as remarked above). To these, a third common aspect should be added, which mainly derives from the second point: both Smith and

---

production, which can *also be found lurking* in Smith's discussion." Schumpeter (1954: 307-311) highlights the confusion emerging from Smith's treatment of value, but it seems (at least to me) that he recognizes a preponderant labour-approach embedded in Smith's treatment (see also his remarks at p. 302 and pp. 472-75; but see pp. 188-89, especially footnote 20 for a different analysis).

[10] Sinha (2010: 10) for instance asserts the following: "Though it may be helpful for a modern reader of Smith to separate the notion of 'real value' from the notion of 'labour commanded' as a *numéraire* for price determination, Adam Smith apparently did not see it in such terms. … for Smith 'real value' is represented by the 'command of labour' of a commodity is at the same time represented by its 'real value'. The two notions cannot be separated."

[11] This is relabelled by Bharadwaj (1978: 255) as "the Smith-Ricardo distinction between 'natural price' and 'market price' (the former explained in terms of the conditions of production the 'ultimate regulator of value' and the latter viewed as arising out of deviations around natural prices brought about by temporary and accidental fluctuations in conditions of demand and supply) …"

Ricardo saw value theory (in the long-run) as an analytical prerequisite to distribution theory.

Thus, with respect to all the relevant aspects of value theory that shall interest us here, Ricardian and "Smithian" value theories are in complete agreement, the former following and building on the latter. For this reason it also becomes completely legitimate from an epistemological and methodological point of view to start the enquiry from Ricardo's value theory, and to develop the subsequent analysis from that point onwards.

1.2. Ricardian labour theory of exchange value in brief: usefulness and quantities of labour[12]

Among classical economists, value does not have a univocal meaning. As recalled by Ricardo at the outset of his *Principles*, while citing Adam Smith (Ricardo, 1911: 5; emphasis in the original):

> "The word Value has two different meanings, and sometimes expresses the utility of some particular object, and sometimes the power of purchasing other goods which the possession of that object conveys. The one may be called *value in use*; the other *value in exchange*."

While value in use depends on the usefulness (or utility) deriving from the *possession and employment of a good*, value in exchange expresses the ratio in which *two goods* can be exchanged for each other. In other words, value in exchange denotes the market price (the exchangeable ratio) of a good. Usefulness (or use value) is a prerequisite to exchange value, but it does not completely explain the latter, nor is usefulness identical to exchange value. As explained by Ricardo (1911: 5):

> "Possessing utility, commodities derive their exchangeable value from two sources: from their scarcity, and from the quantity of labour required to obtain them."

In truth, "scarcity" and "quantity of labour" are not complementary conditions, but alternative ones. "Scarcity" only applies to commodities or goods that cannot be

[12] The literature on Ricardian value theory is of course immense. Here a selection of the references used in the present section it is listed. For in in-depth overview of Ricardian value theory, see Blaug (1958: 15-37; 1996: chapter 4), Cassels (1935), Dobb (1973: 73-92); Hollander (1904); Meek (1973; esp. chapter 3) O'Brien (2004: 98-106); Pasinetti (1960), Peach (1988: 114-121; 1993), Schumpeter (154: 588- 605), Sinha (2010: chapter 2), Sraffa (1951: xiii – lxii); Stigler (1952; 1958). Note that Ricardo's interpretation of value changed through the years and along with his three editions of the *Principles* (see Peach, 1993: chapters 4 and 5).

reproduced or augmented in supply through labour expenditures.[13] This creates a sharp distinction in Ricardo between those goods that can be increased in supply (labour-goods) and those goods whose supply is given (it is scarce) and cannot be modified (scarcity-goods). The theory of exchange value only applies to the former (Ricardo, 1911: 6):

> "In speaking, then, of commodities, of their exchangeable value, and of the laws which regulate their relative prices, we mean always such commodities only as can be increased in quantity by the exertion of human industry, and on the production of which competition operates without restraint."

The distinction between "scarcity goods" (scarce resources) and "production goods" (goods produced through a technological process) has been retained in economics until our days and becomes very important especially in the field of regulation, as we shall discuss much later in this work.[14]

Thus (abstracting from the role of scarcity), a good possess value in exchange in the Ricardian framework if and only if two necessary conditions are fulfilled: (a) the good has a certain utility in use, or usefulness (use value); (b) the good is produced in a "laborious" way (i.e. using labour in the production process). Neither of them is, alone, sufficient.

This leads to the very well-known conclusion that Ricardian theory of exchange value is a *labour-theory of value*. This means that, *provided that goods are useful*, then two goods exchange for each other according to the relative quantities of labour expended in order to produce and sell them or, as summarised by Ricardo (1911: 29-30) himself, "their relative values will be governed by the relative quantities of labour bestowed on their production."

It should be noted that the generic expression "value" in the previous statement should be intended as "*exchange value*", as this other passage (among others) of the *Principles* (Ricardo, 1911: 14; emphasis added) makes clear:

> "in estimating the *exchangeable value* of stockings, for example, we shall find that their value, comparatively with other things, depends on the total quantity of labour necessary to manufacture them and bring them to the market."

[13] "There are some commodities, the value of which is determined by their scarcity alone. No labour can increase the quantity of such goods … Their value is wholly independent of the quantity of labour originally necessary to produce them, and varies with the varying wealth and inclinations of those who are desirous to possess them." (Ricardo, 1911: 6).

[14] Thus, while the radio-frequency spectrum is a "scarce resource" (meaning that its production capacity is given, and it is not subject to any economic calculation), any other communication asset is considered a component of a production process (whose supply, or production capacity, can be modified and it is part of the economic problem).

Yet, as use value has usually been considered outside the realm of political economy (at least since Adam Smith), classical economists usually intend the generic term *value* as a synonym of *exchange value*. This is not a light assumption and its full implications will be developed in great detail in later sections.

The pure labour theory of value is the interpretation of Ricardian value theory retained by Ricardo's closest disciples and contemporary followers (the "Ricardian" classical economists).[15]

Having briefly discussed the key aspects of Ricardian value theory, it is now necessary to go deeper than that and to highlight and clarify some key aspects of the theory.

Two aspects in particular of Ricardian value theory deserve further considerations. One, the *separation of use value from exchange value* and, two, the *exogeneity of value* (not only of use value, but also of exchangeable value itself). Both aspects of Ricardian value theory are key for understanding the notions of *property* (private ownership) and *asset* (or capital), together with the relationship between the latter and the services operated over it, which represent the real focus of the present work. We firstly tackle the distinction between use and exchange value in the section below, while we switch to the exogeneity of value in the following chapter, together with a general overview of the meaning of value in early neoclassical economics.

2. Use value and exchange value: preliminary considerations and critical reassessment of the literature

Before beginning the critical re-examination of the two classical components of the well-known "paradox of value", it is important to first highlight that Ricardo inherits Smith's analysis and approach to the topic (Ricardo directly quotes Smith, as highlighted above). Adam Smith formalizes the separation between use and exchange value and Ricardo simply accepts it with no further enquiry.[16] For this reason, the following discussion will often refer

[15] Thus, his most faithful disciple McCulloch summarizes the essence of the Ricaridian framework as follows: "Mr. Ricardo maintains, in this work, the fundamental principle, that the exchangeable value of commodities or their relative worth, as compared with each other, depends *exclusively on the quantities of labour* necessarily required to produce them and bring them on the market" (McCulloch, 1952; cited in Blaug, 1958: 33). For a discussion of the "Ricardians" (James Mill, McCulloch, West and De Quincey), see also Schumpeter (1954: 469-480).

[16] Note that, although the separation between use value and exchange value was not new in the intellectual panorama of that time, it is Smith to officialise and "institutionalise" the irreconcilable nature of the two. As summarised by Schumpeter (1954: 188): "In chapter 4 [of the *Wealth*] A. Smith … falling far below the level reached by many older authors and particularly Galiani, severs 'value in exchange' completely from 'value in use'."

to the Smith-Ricardo approach, although it is sufficient to just focus on Adam Smith's treatment, confident that this is the view shared by Ricardo.

As already highlighted, use value under Smith and Ricardo is a completely different concept from exchange value. The two are not comparable, nor commensurable. The present chapter will especially focus on the concept of use value, providing just a very brief overview of exchange value, while abstaining from a detailed discussion of the complex issues deriving from the nature of exchange value among classical economists.

Last, but not least, a warning is in order: neither has a clear and univocal meaning among classical economists (even though Ricardo is much clearer and less inconsistent than Smith).

2.1. Some preliminary considerations on the nature of exchange value

On the one hand, the exchange value of a good or commodity is *usually* regarded as just another term to define the (relative) market price of the good with respect to the other goods or commodities exchanged in the market. It should be noted that this is not completely correct, nor necessarily true. First of all, exchange value in Ricardo is not a *monetary price*, but a *quantity* of a certain resource employed in production (labour). More specifically, whether exchange value can be regarded as a monetary measure of purchasing power (in exchange) also depends on how exchange value itself is in turn defined: whether it is a "natural" or "real" exchange value, or a nominal (monetary) one. Only the latter can be intended as a pure monetary measure, but usually not the former.

These remarks are certainly true in Ricardo, while Smith is more confusing and contradictory on this point although he often seems to be in line with the Ricardian interpretation (1991: book 1, chapters 5,7). This distinction between "natural" and "real value" is an important one as it uncovers the nature (the origin) of value, although this point will bring us too far away in the discussion, and will not be developed further.

Moreover, it should also be remarked that Ricardo does not always interpret exchange value in relative terms, thus as a *relative* measure of exchange (as an exchange ratio). Conversely, exchange value is very often an absolute (not relative) measurement of value in Ricardo and the same hopeless confusion is present in Smith (even though the latter usually interprets exchange value in relative terms).

Nevertheless, abstracting from all these complications for the moment, with some approximation it is possible to conclude that the theory of exchange value *can* be interpreted as the theory of prices (usually "natural" prices, rather than nominal market prices). As a result, the important implication is that *exchange value can be measured*, and *it is measured in monetary terms*. Moreover, the relative exchange values of two commodities can be quantitatively computed and objectively compared.

2.2. Use value in the literature: material wealth, property, and the role of consumption

On the other hand, the concept of use value among classical economists has remained even more obscure. Classical authors have usually overlooked a detailed discussion of the concept so that, as remarked by Schefold (1999: 122), "[t]he analysis of use value … seems to have remained entirely outside the sphere of traditional political economy". On the other hand, later neoclassical economists have usually regarded the classical concept of use value either as incomprehensible, or as irrelevant, while even the more recent secondary literature seems to have largely ignored the concept, with only a limited number of exceptions (Aspromourgos, 2009: 115-118; Bharadwaj, 1978; O'Brien, 2004: 93; Schefold, 1999; Schumpeter, 1954: 300-311). The outstanding exception to this paucity of contributions is Marx (1995: book 1, section 1, chapter 1), the only classical economist that has discussed the meaning and the nature of the concept in some detail. Much of the received classical interpretation of use can be attributed to Marx'.

The following relevant characteristics of use value have been underlined by the literature: objectivity, qualitative incommensurability, and exogeneity (independence from exchange value). The first and the third in particular will only be introduced, as their meaning and implications will be developed at length in the following chapter.

2.2.1. Use value is an objective property of the physical "thing"

First of all, the use value of a commodity in Smith-Ricardo refers to the *objective and physical property* of a commodity, or resource, if employed by its respective owner. Use value is therefore objective usefulness (deriving from the physical properties of the material commodity) and not subjective utility. Thus, even though classical economists sometimes employ the term "utility", this has to be intended as the physical property of a material thing, and not as subjective gratification or feelings *à la* Jevons.[17] As summarised by Bharadwaj (1978: 256; emphasis in the original) use value in Smith is "not dependent

[17] Thus, Marx (1995: 27) introduces use value precisely in utility terms: "The utility of a thing makes it a use value." But he then goes on clarifying the meaning as follows: "But this utility is not a thing of air. Being limited by the physical properties of the commodity, it has no existence apart from that commodity. A commodity, such as iron, corn, or a diamond, is therefore, so far as it is a material thing, a use value, something useful."

upon the *individual's estimation* to gratify subjective inclinations, measured in quantitative terms." Use value is therefore inherent in the materiality of the possessed goods, in their underlying material substratum supporting their physical employment. This interpretation is unanimously shared by all main classical economists, from Smith to Marx.

This objective usefulness of a material good derives from its ability to serve some of the owner's purposes, wants or needs. As remarked by Aspromourgos (2009: 117), this physical and objective interpretation of usefulness derives directly from Aristotle:[18]

> "He [Aristotle] has an objective conception of use: the use-value of a chair is sitting, the use-value of a hat is covering the head, and so on. Use-values are heterogeneous, each specific to qualitatively distinct human purposes that things are fitted to serve."

2.2.2. Use value is a qualitative and incommensurable magnitude

Secondly, consistent with the above remarks, use value is a qualitative and incommensurable property of a commodity, or good and cannot be measured in monetary terms. In fact, use value is not subject to quantitative measurement at all. This point has a couple of interesting implications.

On the one hand, this is the reason why use value and exchange value remain two incomparable and incommensurable entities, as already remarked. This point is made particularly explicit in Marx (1995; 28):[19]

> "As use values, commodities are, above all, of different qualities, but as exchange values, they are merely different quantities, and consequently do not contain an atom of use value."

[18] See also Schefold (1999: 123) for similar remarks.

[19] Marx is particularly careful in insisting on the fact that use value is a qualitative, thus discretionary, measure of the usefulness of a commodity, differently from exchange value (1995: 27-28): "Every useful thing, as iron, paper, &c., may be looked at from the two points of view of quality and quantity. It is an assemblage of many properties, and may therefore be of use in various ways. To discover the various uses of things is the work of history. So also is the establishment of socially-recognized standards of measure for the quantities of these useful objects. The diversity of these measures has its origin partly in the diverse nature of the objects to be measured, partly in convention … Exchange value, at first sight, presents itself as a quantitative relation, as the proportion in which values in use of one sort are exchanged for those of another sort, a relation constantly changing with time and place … But the exchange of commodities is evidently an act characterised by a total abstraction from use value. Then one use value is just as good as another, provided only it be present in sufficient quantity."

On the other hand, it is important to note that use value in classical political economy is associated to the *accumulation and the appropriation of material wealth*, and can only be referred to (is only a property of) material wealth. As succinctly put by Marx (1995: 27):

> "Use values become a reality only by use or consumption: they also constitute the substance of all wealth, whatever may be the social form of that wealth. In the form of society we are about to consider, they are, in addition, the material depositories of exchange value."

So that (p. 33):

> "An increase in the quantity of use values is an increase in material wealth."

This is consistent with the above considerations concerning the "physical nature" of use value in classical political economy (and Aristotle).

Usefulness and use value are therefore an intrinsic characteristic of one actor's physical property (their wealth) and they are usually a function of the quantity of items possessed by the latter. The degree of use value experienced by an individual derives from the personal *accumulation of material goods* under their own property, and by *their employment* "in consumption" or, in other words, by their *wealth*. This is consistent with the fact that wealth (or "riches") among classical economists is always intended as the quantity of material things possessed. Thus, the higher the material wealth possessed by an actor, the higher the value in use available to them. As summarised by Schefold, (1999: 125; emphasis in the original):

> "Smith regarded riches partly as power, without reducing them to utility … [A diamond] use value is to reflect the power of its owner by impressing the observer. Its utility remains to be discovered, says Smith … but it can be used to display wealth, therefore the ability to *command* labour."

In the previous passage, the term "riches" represents the standard classical way to define wealth. The same interpretation of value in use as a measure of the amount of wealth *privately owned* (possessed) by an individual is shared by Ricardo (1911: 191; emphasis added):

> "Adam Smith … very justly distinguishes the nature of value which [natural agents] add to commodities – they are serviceable to us, by increasing the abundance of productions*, by making men richer*, by *adding value in use*; but as they perform

> their work gratuitously … the assistance which they afford us adds nothing to value in exchange."

In sum, value in use is a (physical) property of the tangible, material objects (goods, assets or resources) *possessed* by an individual. Yet, as highlighted by Ricardo in the quote above, this implies that use value is a characteristic inherent in the private ownership of goods and resources and is independent form the "act of exchange": no transaction is needed in order to generate use value as privately possessed things "perform their work gratuitously." According to the classical interpretation of value in use, if a man does not possess any good, resource or asset (capital), he will not experience any value in use.

This holds true regardless of whether the same individual is the recipient of other economic activities (or services), or if he can command some flow of income deriving from his labour: use value in classical economics derives from cumulated wealth, and the latter derives from the extent of cumulated material "things" (goods and assets) that an individual can employ for his own benefit. Monetary flows of incomes generated by exchanges are not regarded as material wealth by classical economists, and therefore they cannot be repositories of any use value.

2.2.3. Use value is independent from the production (and exchange) process

Thirdly, from the above quotation of Ricardo it is possible to infer that use value is independent from the characteristics of the production process and, more specifically, from the quantity of labour expended for the production of a certain good (or commodity). Going back to the aforementioned Ricardian distinction between "scarcity" and "labour expenditure", while exchange value only derives from the latter, use value is *mainly* a property of the former: *appropriable scarce resources* usually have use value (provided they have some intrinsic usefulness).

The same dichotomy between *labour* (repository of exchange value) and *materiality* (repository of use value), is retained by Marx (1995: 31):

> "The use values, coat, linen, &c., *i.e.*, the bodies of commodities, are combinations of two elements – matter and labour. If we take away the useful labour expended upon them, a material substratum is always left, which is furnished by Nature without the help of man. The latter can work only as Nature does, that is by changing the form of matter. Nay more, in this work of changing the form he is constantly helped by natural forces. We see, then, that labour is not the only

> source of material wealth, of use values produced by labour. As William Petty puts it, labour is its father and the earth its mother."

On the one hand, this remark not only re-establishes the separation of use value from exchange value (classical price theory), and the independence of the former from the latter: two types of commodities are postulated and only one (labour-goods) is *also* endowed with exchange value, while the other (scarcity goods) can only have use value.

On the other hand, Marx's passage highlights very clearly the fact that use-value, being a feature of the martial world, and a characteristics of the material things used in order to satisfy and address personal needs, is independent from the production process. This is especially evident for those classical economists, such as Marx, that, following Ricardo, adopted a pure labour-theory of value.

2.2.4. Use value and the theory of consumption: some remarks

On the other hand, it is clear that the way in which classical economists approach use value represents the fundamental reason why consumption (or demand) has no place in classical political economy. The theory of consumption according to classical economics is, simply, the theory of use value and of the physical usefulness of material goods and commodities. However, as use value is an objective and exogenous property of the physical (material) world that remains independent from the characteristics of production (thus from exchange value and price), then it was already regarded as a useless and irrelevant category in political economy. For this reason use value was unsurprisingly discarded by classical economists and for this reason the theory of consumption remains beyond the scope of classical political economy, while the theory of price is only developed by looking at the characteristics of the supply (and thus independently from the characteristics of the demand).

Consumption is indeed absent in classical price theory (but for some fundamentally irrelevant exceptions deriving from short-run fluctuations, of the market price). For these reasons the vestiges of the so-called "paradox of value" are usually regarded as the single most important factor delaying the ultimate advent of the neoclassical "subjectivist" revolution in value theory (which occurred in the 1870s), which is usually regarded as the official origin of the neoclassical framework.

In this respect, it should be remarked that the sharp distinction between use and exchange value proposed by Smith, and retained by Ricardo, has remained one of the most incomprehensible points of classical economics in the eyes of later economists. This also derives from the fact that all most prominent economists until Smith's time (including Smith's teachers) were more inclined towards a "utility" theory of exchange value not dissimilar from those of early marginalists *à la* Jevons, where use value represents the logical prerequisite for, and determinant of, exchange value (O'Brien, 2004: 91-93). As succinctly summarised by Schumpeter (1954: 302): "let us bear in mind that it was the 'subjective' or 'utility' theory of price that had the wind until the influence of the *Wealth of Nations* –and especially of Ricardo's *Principles*– asserted itself." Thus, Jevons (1957: 79) attributes a great deal of confusion on the subject to Smith, and reinterprets use value as the overall utility obtained in the consumption of a good, while exchange value remains as the exchange ratio of the good in the market.[20] Although this interpretation was not identical to the ones provided by the two other early marginalists (as we shall see in the next chapter Walras, and especially Menger, who interpret use value differently and more in line with Smith's original treatment), this is how use value has been interpreted since.

Obviously, once that value is explained in terms of subjective utility (as all economists starting from the second half of the 19th century gradually started to acknowledge) utility in use becomes the *origin* of market prices, and value in use and value in exchange eventually conflate into a unique (homogeneous) theory of value: price theory. Yet, if use value and exchange value remain two distinct and incompatible entities, there is no hope of tracing exchange value (thus market prices) back to subjective utility (or, usefulness) and therefore to make consumption a branch of economic theory.

2.2.5. *Some new perspectives on use value: self-employment of property and the conflation of production with consumption*

In conclusion, use value in classical political economy (a) characterizes the physical properties of a *material good*, (b) derives from the material *wealth accumulated and possessed* by an individual, and (c) is associated with the *consumption* of these goods (or wealth), independently from the production and the exchange aspect (the two being intimately related in classical economics where exchange-value is only shaped by the "supply side"). It is thus possible to summarise the above discussion by stating that use

[20] "It is sufficiently plain that, when Smith speaks of water as being highly useful and yet devoid of purchasing power, he means *water in abundance*, that is to say, water so abundantly supplied that it has exerted its full useful effect, or its *total utility*. Water, when it becomes very scarce, as in a dry desert, acquires exceedingly great purchasing power. Thus Smith evidently means by value in use, *the total utility of a substance of which the degree of utility has sunk very low, because the want of such substance has been well nigh satisfied*. By purchasing power he clearly means the ratio of exchange for other commodities." (Jevons, 1954: 79; emphasis in the original).

value corresponds to *the usefulness that someone receives, or can dispose of, by using and employing their own property in consumption*.

The above three characteristics of use value are the ones usually associated with the Smith-Ricardo approach to the concept. This is especially true for the association of use value with the "theory of consumption". Yet, this does not exhaust all the features of use value. There is a further (fourth) aspect of use value, mainly *but not exclusively* stressed by Marx, that is worthy of mention, that is its relationship with the (classical) theory of production.

Point three above asserted that in classical political economy, use value was a concept only valid in consumption. In fact, it can be regarded as the classical substitute for the theory of consumption. In truth, Marx, merging Ricardian labour-theory, with the ideas first developed by Aristotle, provides a more multifaceted interpretation of the concept (Schefold, 1999: 123). As described by Montani (1987), Marx identifies *two* conditions according to which a good can have use value, even though it may not have any exchange value. The first case is the one already identified by Smith-Ricardo: the existence of resources or goods that are not produced by labour. In this case, use value is completely independent from any production process, and can only become meaningful in consumption. The second case "regards a good produced with labour, but without any social utility, i.e. *subjectively useful but not exchangeable on the market*" (Montani, 1987; emphasis added). In this case the use value of a good is associated to (it derives from) production. Quoting directly from Marx (1995: 30; emphasis added), this is how the two cases are summarised by the author:

> "A thing can be a use value, without having value. This is the case whenever its utility to man is not due to labour. Such are air, virgin soil, natural meadows, &c. A thing can be useful, and the product of human labour, without being a commodity. *Whoever directly satisfies his wants with the produce of his own labour, creates, indeed, use values, but not commodities*. In order to produce the latter, he must not only produce use values, but use values for others, social use values."

Thus, Marx operates a very important extension of the concept of use value where he not only associates it with the realm of consumption, but also interprets use value as a peculiar characteristic of production: production for own consumption or, stated differently, production for auto-consumption. Moreover, a *subjective type* of use value is distinguished from "*social* use value" or value in exchange. The former derives from the employment of a one's *own labour* in order to satisfy *their own wants,* while the latter derives from the employment of labour for the satisfaction of others' wants.

In this second case use value is therefore associated to consumption *and* the production process *at the same time*.

This conflation between production and consumption is a fundamental aspect of the present work and it is important to develop it further. Although the recent literature on the Smith-Ricardo approach to use value denies that this was also the original interpretation of the forefathers of classical economics, the remainder of this section will argue that this conclusion is however unwarranted.[21] This interpretation probably derives from a biased "neoclassical reconstruction" of use value as a concept only applicable to consumption that is however not supported in Smith's *Wealth*, as we shall argue in the following section.

Thus, a closer look at the concept of use value is still necessary in order to fully grasp the essence of this distinction that has been progressively abandoned since J.S. Mill (contextually with the abandonment of Ricardian value theory by the late 1830s) and will be ultimately rejected in a definitive manner, under Marshall, Jevons and all later neoclassical economists following them (see Bharadwaj, 1978), but *not* however by the early Austrians, and Menger in particular. As we shall see, the latter will also retain Marx's distinction between *subjective* and *objective* (or social) value.

As the reminder of the section will argue, a proper understanding of the distinction between the act of consumption (linked to use value) and the act of production (associated to exchange value), becomes imperative for a new (non-neoclassical) theory of property and asset ownership to be developed.

3. Use value and the paradox of value: implications for asset ownership and the "make-or-buy problem"

The conflation of production and consumption is the necessary precondition in order to guarantee duality of value. For this reason, it is import to better develop this assertion, and to derive implications for the role and the nature of institutions in economic analysis.

[21] Thus, Schefold (1999: 141) concludes by saying that in "Adam Smith … the use value of the commodity is not reduced to its being an object of immediate consumption or of domestic production but it is used to impress others."

More specifically, the section focusses on the role that the duality of value deriving from the conflation of production with (auto)-consumption has for the institutions of asset ownership and for the make-or-buy problem or, in other words, the "nature of the firm".

3.1. Use value and exchange value: a reinterpretation of the paradox in light of the division of labour

In order to clarify the real meanings of use and exchange value, it is necessary to understand the way in which Smith introduces the two concepts in the *Wealth*. Adam Smith is the forefather of the so-called "paradox of value" as no other economist before him (not even his mentor and teacher), developed a clear dualistic approach to value theory.

The distinction between use and exchange value is discussed in the chapter on "*the origin and use of money*" (chapter 4; book 1), before discussing the mechanisms regulating exchange ("*of the real and nominal price of commodities*": chapter 5) and just after the chapters on the famous "*division of labour*" (chapters 1 to 3 in book 1). Thus, this is how Smith (1991: chapter 4, book 1; emphasis added) opens the chapter, just a couple of pages before distinguishing between the two forms of value:

> "When the division of labour has been once thoroughly established, it is but a very small part of a man's wants which the produce of his own labour can supply. He supplies the far greater part of them *by exchanging that surplus part of the produce of his own labour*, *which is over and above his own consumption*, for such parts of the produce of other men's labour as he has occasion for. Every man thus lives by exchanging, or becomes in some measure a merchant, and the society itself grows to be what is properly a commercial society."
>
> "One man, we shall suppose, *has more of a certain commodity than he himself has occasion for*, while another has less. The former consequently would be glad to dispose of, and the latter to purchase, *a part of this superfluity* … The butcher *has more meat in his shop than he himself can consume*, and the brewer and the baker would each of them be willing to purchase a part of it."
>
> "… every prudent man in every period of society, after the first establishment of the division of labour, must naturally have endeavoured to manage his affairs in such a manner as to have at alltimes by him, *besides the peculiar produce of his own industry*, a certain quantity of some one commodity or other, such as he

> imagined few people would be likely to refuse in exchange for the produce of their industry."

The essence of the above extracts is incredibly similar, if not fundamentally identical, to the description of use value provided by Marx (quoted above).

From the above excerpts, two important peculiarities of use value follow. Firstly, use value (but also exchange value) becomes meaningful only under the possession (or ownership) of the good. Ownership is the preliminary condition for value, and goods or resources have to be owned in order to have use value. Consistent with what has been said above, use value is a feature of the accumulated property of an individual or, in other words, of his personal wealth (2.2.2). Two, the use value, or usefulness, of the owned commodity, resource or good, derives from *its direct employment in consumption*. This is consistent with the third aspect of use value discussed above (2.2.3 and 2.2.4).

However, from Smith's descriptions it clearly emerges that it is not just consumption that generates use value. Conversely, it is the *auto-consumption* of the possessed commodity that generates its use value; it is the usefulness in auto-consumption from which use value derives. It is this aspect that places Smith's treatment of the topic next to Marx's or, more accurately, it is this interpretation of Smith's *Wealth* that renders Marx's reinterpretation of it much more faithful and correct than is usually acknowledged. Differently from the Ricardian treatment, use value in Smith is not severed from the act of production. Conversely, it complements it.

What really distinguishes Smith's treatment of the concept from Marx's (and, of course, Ricardo's), is the fact that the concept is unavoidably linked to the dynamic process of the division of labour. In the *Wealth*, value in use can be interpreted as the original value held by any useful commodity in those stages of the society that precede the division of labour, thus barter and market exchange. Use value is the only form of value in an "autarkic" society entirely based on the self-subsistence of its citizens, where actors can only use and consume those goods that they already own, or possess. In this case, when only auto-consumption is practiced, there is a univocal correspondence between use value and value in general: use value *is* value in its entirety and private possession is the preliminary (necessary and sufficient) condition to achieve it.

As soon as exchange (or barter) among individuals becomes possible, value assumes its twofold nature: in exchange and/or in use. Yet (and this is the fundamental point) *even* *when exchange appears*, it is only when the "produce of a man's own industry" generates a surplus that exceeds ("over and above") *his own consumption*, that the unconsumed portion of a commodity (or set of commodities) can be exchanged in the market, *thus* acquiring a certain exchange value.

It is important to appreciate the inherent consistency that runs through Smith's development of the concept: the distinction between the two types of value is introduced just after the discussion on the division of labour, but just before the chapter on the nature of price (expressed in monetary terms). With no division of labour, only self-sufficiency and auto-consumption occurs: actors satisfy their own necessities and wants by solely relying on their own means, provided either that (i) *they possess or own something in the first place,* or that (ii) *they can produce something for themselves using "their own industry"* (roughly reflecting Marx's two cases). It is the division of labour, following the expansion of the market, that triggers the switch from a primitive society merely based on self-sufficiency and the auto-consumption of one's own property (wealth) and own production, to a society that is *also* relying on the exchange of the actors' relative surpluses in consumption and/or production. In other words, *it is the division of labour that introduces value in exchange besides value in use*.

Two important points should be noted at this point.

Note that the term "surplus" should not be intended in monetary or profitability terms, as a reward exceeding costs, as it is standard in a pure theory of production. In this case, surplus should be interpreted as that part (spare capacity) of a person's private wealth that exceeds the private needs of an individual, and that is left idle, not being employed to address one's self-subsistence. In other words, "surplus" is that part of private possession of an individual that *exceeds auto-consumption*, and that can be "held out" for *third-parties' consumption*. It is the specific identity of the actor enjoying the consumption of a certain private wealth that characterises what is surplus, and what is not. In this respect, the expansion of the market and the appearance of trade becomes necessary in order to fruitfully employ this (otherwise idle) surplus as some form of trade between the property owner, and third parties is needed.

Moreover, it should also be further noted that even when the market enlarges, and trade begins, exchange value does not *replace* the original use value generated through auto-consumption. It *complements* it, in the same way in which the act of consumption complements the act of production in both Smith's and Marx's treatment. This derives from the fact that only that part of private wealth that is not privately employed in order to satisfy *personal needs*, is actually held out to third parties, and sold on the market to satisfy *someone else's needs*. A market transaction is then created.

Consistent with this interpretation of Smith's *Wealth*, the following definitions of use and exchange value are provided: while use value denotes the usefulness of a commodity *qua* means of self-sufficiency and (auto)-consumption, exchange value quantifies the "usefulness" of a commodity *qua* means of purchase and exchange.

Alternatively, the same definitions can be reframed in the following identical way: while use value denotes the *degree of usefulness* that can be directly obtained from a commodity by employing it in consumption, exchange value denotes the degree of usefulness that can be *indirectly* obtained from a commodity by employing it in order to buy some other commodities.

It should be noted that this definition of use value is different from Jevons' reinterpretation of use value intended as the cumulative (overall) utility generated by the consumption process. According to the definition here provided, use value should be intended as the instantaneous usefulness derived by *retaining the good*, and by employing it for private and personal purposes. In other words, value in use can be interpreted as the (instantaneous) usefulness deriving from using and consuming a person's private property "*in-house*". If it is more useful to employ some assets or goods "in-house" rather than to employ them in exchange (as means of exchange) then this would mean that the value-in-use that can be derived from the *act of employing the private property internally* (directly for internal purposes) is higher than the value that can be derived from the *act of exchanging it*.

3.2. Use value and the division of labour in Smith's *Wealth*: preliminary implications for a theory of asset ownership and for the "make-or-buy" problem

At this point, some important implications of use value have to be discussed. One set of implications concerns the relationship between value in use and the (neoclassical) dichotomy between production and distribution. On the other hand, a second set of implications highlights the relationship between value in use and ownership, especially with respect to the role and the nature of the latter in classical economics

3.2.1. *Use value and its implications for ownership: a new perspective on the role of asset ownership in economics*

Let us begin with some considerations concerning the relationship between value in use and ownership. In order to do that, it is important to underline, once again, the core message deriving from Smith's paradox of value: the fact that someone is either endowed with some commodities ("one man … has more of a certain commodity"), or is producing a certain quantity of these commodities ("beside the peculiar produce of his own industry"), *does not imply that he necessarily has to sell them all in the first place*, nor does it mean that these commodities cannot be *useful* for him in the first place.[22] In principle, it is only the surplus left "over and above his own consumption" that, once that private consumption is preliminary addressed, can then be sold into the market.

In Smith (and Marx) the *personal use* of one's private possession seems to have *temporal*,[23] *conceptual*[24] and *analytical*[25] *priority* over the act of exchange: exchange is "relegated" to a subaltern position and becomes a "residual" of the economic activity, and not the primary activity.

In other words, the fact that one's private property (his material wealth composed by productive assets and means of subsistence) *could* be employed as means of exchange, does not mean that it *should* necessarily be employed as such. Alternatively, the fact that an actor *might* use its property as a means of exchange does not necessarily means that they *will not, or should not*, use it as means of auto-subsistence, or self-sufficiency, in consumption.

These considerations highlight a very important property of ownership, namely that the possession of any useful good enables its owner to satisfy some of their needs *without being reliant on others*. In other words, private possession (ownership) generates a decoupling between the owner and the rest of the social system, conferring independence, or self-sufficiency, to the former.

---

[22] Remember that "usefulness" in classical economics is "in use", and is an intrinsic property of the physical matter employed by its respective owner in order to address some need "in consumption" (section 2.2).

[23] Personal use of a property "in auto-consumption" has *temporal priority* over exchange not only because the satisfaction of personal needs is always a preliminary necessity to address first, but also because it reflects the natural development of any society through time: wile early forms of society are founded on self-subsistence first, exchange develops only later on.

[24] The *conceptual priority* derives from the fact that, as mentioned above, classical economists always start from the assumption that people address their needs first through the property that they own, and that is available to them.

[25] Use value has *analytical priority* over exchange value as the latter is computed as a surplus, over and above the former.

Thus, a possible economic definition of ownership can be formulated based on the previous properties: ownership or possession of a good (private property) is that particular economic institution that confers the faculty of self-sufficiency, meaning that the owner of the goods or resources can have the discretion to rely on their own property in order to disentangle himself from the external economic environment, while still being able to address (part of) their needs and requirements.

The same concept can also be expressed differently: ownership is that institution that enables the person endowed with it to enjoy some value from his property independently from the value conferred to their property by the external economic environment (i.e. the social exchange). This is Marx's distinction between subjective and social value, as mentioned above.

It should be remarked that the above is by no means the sole possible definition of ownership. On the one hand, the above differs from the standard definition of ownership provided by cotemporary theory of incentives, where ownership confers the *ex-ante* rights and incentives to appropriate the "left-overs" (in monetary terms, as appropriation of the residual intended as profit). On the other hand, the above definition of ownership differs from the definition of ownership provided by the contemporary theory of incomplete contracts, where ownership confers the right to decide how a property should adapt *ex-post* to changed circumstances and contingencies.

Differently from the above modern characterizations of ownership, this other definition of ownership highlights a new and very relevant economic aspect of ownership: ownership confers emancipation or the discretion of not being reliant on others for the satisfaction of some needs. Obviously, whether this conclusion is true or not depends heavily on the concept of "usefulness", a much more complex aspect that will be discussed later.

The above definition is much closer to the legal definition of ownership, intended as that institution conferring the "right to exclude." The latter is the more fundamental and profound interpretation of ownership used by regulatory (rather than antitrust) authorities. In fact, the conceptual relationship between regulation and antitrust, and their relative spheres of intervention, are founded on the correct interpretation of the right to exclude, as economists have started to understand in the last quarter of century.[26]

[26] For a (non-exhaustive) selection of some key articles on the right to exclude, and on the fundamental problems deriving from it in the regulatory debate sees the following: Brennan (1993),

Leaving this point for later discussion, it is interesting to note that the "classical" reinterpretation of ownership appears to be different even from this third (legal) interpretation of ownership: the focus is not on the competitive *implications* of exclusion, but rather on its original *motivations* and economic foundations. In this case exclusion does not derive from a willingness to increase monopoly power, as it is usually (but not always) intended in regulatory law, but always in antitrust law. Conversely, exclusion in this classical reinterpretation should be rather intended as a willingness not to treat a certain property as a means for trade and exchange in the first place, which is something still preliminary to the competitive motivations adopted by regulators.

In conclusion, two remarks are in order.

First, it should be noted that no classical economist (certainly not Smith, Ricardo nor Marx) offer an interpretation of use value and private ownership much different from the one just presented. The only reason why neither of them developed the concept further is that they (arbitrarily) limited the scope of political economy to one twofold aim: to understand how the private wealth possessed by the various individuals is valued through an exchange (or barter) mechanism (value theory), so to ultimately understand how those contributing to its creation are rewarded, *given* this market valuation preliminary provided (distribution theory). Distribution theory was the real focus of classical economists, not value theory; the latter was just instrumental to the former.[27]

For this reason, it was of no interest to classical economists to study how private wealth is valued if used and employed internally *by their own possessor*, and how the latter *rewards themselves* for the value generated by "their own industry". In fact, distribution theory collapses in this latter case where one "rewards himself" for the auto-consumption of his own produce. As distribution theory was the real focus of classical political economy, then any situation falling outside the realm of the former would automatically also fall outside the realm of the latter. This is just another way to understand why use value was rapidly

---

Carlton (2001), Hovenkamp (2005), Meese (2005), Popovsky (2006), Schon (2005), Singer (1996), Vickers (2010).

[27] It is the prominence of the law of returns in distribution theory that distinguishes classical political economy under and after Ricardo (and Malthus) from classical political economy before them (namely, Adam Smith political economy). As summarised by Peach (1993: 31): "The labour theory had supplied Ricardo with a graphic means of demonstrating that worsened conditions of agricultural production (more labour expenditure per unit of output) would 'permanently' depress general profitability. His appeal for the theory undoubtedly stemmed, in part, from its suitability to his specific requirements in the realms of distribution theory."

discarded by all classical economists (but Marx), and the theory of consumption with it, as already discussed above.

However, it is fundamental to underline an important implication deriving from the above considerations: the fact that this case cannot be studied by distribution theory does not mean that it was *ruled out* by classical political economists. It was accepted and contemplated, but simply not developed further. It was not labelled as either unrealistic, or unfeasible, or illogical, or meaningless. The approach will be very different in neoclassical economics where the same situation will *be ruled out by assumption* (as illogical, meaningless or, in other words, irrational).

Second, it should be noted that another way to reframe the previous property (and relative definition) of ownership is the following one: ownership potentially enables the owner of a certain good (provided usefulness) to be the supplier of his own needs in consumption in the same way in which it enables the same owner to be the consumer of his own property employed in production.

This last definition of ownership leads the discussion to the second set of considerations concerning the implications that use value has for the conceptual categories of production and consumption.

3.2.2. Use value and some further considerations on the division of labour: reinterpreting the "make-or-buy" problem and new perspectives on "the nature of the firm"

As already remarked above, the possession of goods enables the owner to get things done *in-house*, without entering in exchange transactions and without relying on external third parties. This is precisely the situation characterising a pre-market (pre-social) world with "no division of labour", as discussed by Adam Smith. It should be noted, that this feature of ownership is also consistent with the essence of any theory of the firm. According to the standard theory of the firm, firms exist precisely because "to get things done in-house" is more convenient (efficient) than to rely on third parties by means of market exchanges. The former represents the "make" option, *vis-à-vis* the "buy" decision represented by the set-up of a market transaction.

Not surprisingly, Adam Smith's discussion of the factors generating the division of labour in a society has long been regarded as one of the fundamental insights in order to develop an economic theory of the firm (Coase, 1988b: 10; Stigler, 1951; Williamson, 1980).[28] These

[28] See West (1988: 30-33) for a (not-so-recent) review on the influence of Adam Smith on the theory of the firm and organization.

authors have all stressed the role of scale economies arising from the division of labour (Williamson, 1985b) and speed economies (Chandler, 1977: 235-239) and rightfully so: Smith does describe the advantages deriving from the division of labour mainly in efficiency and productivity terms.[29]

What all these authors seem to overlook is the fundamental role played by use value in justifying and supporting Adam Smith's discussion of the division of labour. This might derive from the fact that, already starting from the mid-19th century, use value has usually been reinterpreted by *all* economists (with, to my knowledge, no exception) as a very rough substitute for subjective utility *in consumption*, as already remarked above (2.2.4). In other words, the concept of use value has usually been regarded as those part of classical value theory belonging to the (missing) second leg of the standard Marshallian framework: the demand curve. The demand curve, of course, fulfils the role of consumption in the determination of market prices and there is no doubt that there is no demand curve in classical economics (in fact, the lack of the demand curve is precisely what characterises classical economics in the eyes of later neoclassical economists).

Yet, it might be the case that this interpretation reflects what Blaug (1990) calls a "rational" reconstruction of the early classical theory of value through neoclassical lenses. A careful analysis of the first chapters of Smith's *Wealth* would suggest that it is not completely clear whether it is possible to frame Smith's analysis according to the standard categories of "production" (supply) and "consumption" (demand), by clearly distinguishing one from the other in a clear-cut way.

As already remarked, all classical economists (with no real distinction between the various classical economists) perceived economic analysis in dynamic terms, as a study of the implications of growth, overpopulation and economic development for the relative distributive shares (wage, profits and rent) accruing to the various actors (or classes) of the society (Schumpeter, 1954: 554-574). Differently from Ricardo, where the fundamental mechanism driving growth is given by Malthus' law of diminishing returns for land, the engine driving growth in Smith's *Wealth* is the division of labour, as already discussed. As summarised by O'Brien (2004: 40): "Economics is seen as providing the statesman's guide to economic growth and in all this division of labour is given a central role as the

[29] According to Smith, the advantages coming about from the division of labour deriving from the more "intensive" employment of the latter are mainly three: dexterity, reduced waste of time, and the aid of specialised machinery. See Cannan (1903: chapter 3) for an in-depth discussion on the topic.

mechanism of growth." And, as Smith titles in chapter 3 of the *Wealth*, and as recalled by Stigler (1951), "the division of labour is limited by the extent of the market."

However, there are two distinct ways to reinterpret this dynamic mechanism of growth enquired and analysed by Adam Smith. And it is at this point that the neoclassical refutation of use value plays a fundamental role. The first reinterpretation is the mainstream neoclassical "rational" reconstruction. According to this first reinterpretation, as the market for a certain good expands, scale economies acquire more prominence and suddenly mass production for that good becomes more and more efficient, and convenient. Specialization in production, supported by the investment in highly dedicated (or specific) assets, ensue. This is the standard neoclassical story as evoked by Stigler (1951), and supported by Chandler (1977) and subsequently also retained and accepted by all the prominent exponents of NIE and TCE, as recalled above.

The reason why this interpretation might be a "rational reconstruction" (or "Marshallian bias") of Smith's division of labour derives from the underlying assumption that there is a clear-cut distinction between the role of the producer (shaping the supply curve) and the role of the consumer (shaping the demand curve). In other words, it assumes that Smith was already thinking in terms of production and consumption, or in terms of (decreasing) marginal cost curves (scale economies in supply) and (decreasing) demand curves. In this (Marshallian or "neoclassical") framework producers are pure producers: their only task is to supply and sell the goods they produce *to third parties*. Conversely, consumers are pure consumers: they spend their available income in order to purchase and consume the goods supplied to them *by third parties*.

Yet, this interpretation of Smith's division of labour is not supported by Smith's treatment in *Wealth*. On the one hand, Smith, and all other English classical economists, not only did not have a theory of consumption based on such a concept such as the "demand curve", but he did not even have any theory of production, let alone a theory of marginal cost curves and scale economies. This second point is shared by Ricardo and by all other classical economists up until Torrens (1921) and Senior (1936), the first classical authors to put forward a proper economic theory of production based on the combination of production factors (i.e. labour and capital).[30]

[30] See Robbins (1958) on the key early contribution of Torrens to classical production and distribution theory. See also Blaug (1958: chapter 3) and O'Brien (1988).

The lack of a theory of production is not a secondary detail. Conversely, it is probably one of the most relevant aspects of all early classical economists (until Ricardo and his close disciples after him). A proper economic theory of production intended as the combination of variable shares of factors, having (marginal) decreasing returns if substituted among them (the fundamental principle on which the concept of "economies in production" relies on) will only appear on the scene at the very end of the 19th century, following the second marginalist revolution in (not surprisingly) production and distribution theory.[31] Yet, this event occurred not only much after the end of classical economics, but even a quarter of a century after the first marginalist revolution in value theory was developed.

For this reason, to base Adam Smith's concept of the division of labour on contemporary concepts in production theory (such as that of "economies in production") seems unwarranted from both an epistemological, but also a conceptual point of view.

On the other hand, the way in which Smith seems to interpret the mechanism of the expansion of the market, and of the division of labour is one where *any actor is at the same time producer and consumer* of the commodities that they *possesses or produces*. The extent to which each actor is a consumer rather than a producer of their own commodities (and vice versa) depends, precisely, on the development stage of the society they operates in. In other words, the status of "consumer" and "producer" for the *same* actor are themselves endogenously dependent on the extent of the market and therefore, ultimately, on the degree of the division of labour.

Once again, a brief look at Smith's *Wealth* reveals that no distinction is assumed between consumption and production, and that the division of labour should be interpreted precisely as that mechanism that (endogenously) generates and creates the conceptual distinction between production and consumption. The two would otherwise be indistinguishable in principle. Thus, in chapter 1, after having illustrated the benefits of the division of labour, Smith (1991: 13; emphasis added) asserts the following:

> "It is the great multiplication of the productions of all the different arts, in consequence of the division of labour, which occasions, in a well-governed society, that universal opulence which extends itself to the lowest ranks of the people. Every workman has a great quantity *of his own work* to dispose of *beyond what he himself has occasion for* …"

[31] This was due to Barone (1896), Clark (1899), and Wicksteed (1996).

And again (p. 18; emphasis added):

> "As it is by treaty, by barter, and by purchase that we obtain from one another *the greater part* of those mutual good offices which we stand in need of … And thus the certainty of being able *to exchange all that surplus part of the produce of his own labour, which is over and above his own consumption*, for such parts of the produce of other men's labour as he may have occasion for …"

While chapter 3 on the "division of labour limited by the extent of the market" is opened with the following clear statement (p. 20; emphasis added):

> "When the market is very small, no person can have any encouragement to dedicate himself entirely to one employment, *for want of the power to exchange all that surplus part of the produce of his own labour, which is over and above his own consumption*, for such parts of the produce of other men's labour as he has occasion for."

Smith's assertions discussed in chapter 4 of the *Wealth* have been already presented above, and will not be repeated.

From the above extracts one thing becomes clear: in Smith's framework production and consumption conflate and become indistinguishable; any actor can indifferently be consumer or producer of the same commodity. The only mechanism that allows us to discern the two is the extent to which he can sell the goods that he is already producing "with his own labour" to other third parties. However, as already recalled above, these goods would be originally produced and used *for his own consumption* anyway in the first place.

In sum, on the one hand, it is possible to interpret Smith's division of labour as a dynamic mechanism that magnifies productivity and efficiency in production deriving from the exploitation of scale economies. This approach however is based on the (neoclassical) static dichotomy between consumption and production. On the other hand, it is possible to interpret Smith's division of labour as a dynamic mechanism whose main implication is precisely to dynamically shift the conceptual boundary separating the act of consumption from the act of production.

Note that the fundamental assumption behind the neoclassical approach does not imply that producers cannot also be consumers of some other production stage in turn. In other words, it does not prevent producers from also being consumers *of someone else's* *production*, or consumers from also being producers *for someone else's consumption*. Of course, in the neoclassical framework any producers is *also* a consumer and any consumer is *also* a supplier. However, this is not the relevant problem at stake here. The fundamental assumption behind the standard "supply and demand" framework is much more subtle: producers cannot be also consumers *of their own production*, *or wealth*, while consumers cannot be also producers *of their own consumption, or wealth*.

In conclusion, the present section argues the following: it is probably hopeless to properly understand the underlying nature of Smith's "division of labour limited by the extent of the market" if the sharp neoclassical distinction between production and consumption is retained.

Conversely, if the second interpretative approach is adopted, the two conflate and become, at least in principle, indistinguishable. The division of labour is precisely that mechanism that resolves this confusion and induces the members of a society to move from production for self-sufficiency and *self-consumption*, to production for sale to third parties in order to address *somebody else's consumption*.

After all, what does the "make" option in the theory of the firm actually mean if not "production for auto-consumption"? "To make" simply means "to produce for yourself", or to dispose of the faculty of being self-sufficient and autonomous in the satisfaction of a certain necessity by self-employing the available property.

The above discussion suggests that an economic theory of organizations (or firm) can only be correctly formulated if firms are not merely conceived as producers (or suppliers) but as consumers of their own outputs as well. In other words, it is important to start to develop *a theory of production as if it was a theory of consumption.*

The above considerations suggest that any theory of the firm, as well as of asset ownership, must rely on the original (Smithian) duality of value, and should see the latter as an outcome of the dynamic process of the expansion of the market, and the division of labour. For this reason, in order to develop a new theory of the firm where production and consumption conflate into each other, a new theory of property based on the value-in-use concept is required. This is why the concept of use value is so relevant and its distinction from exchange value so important.

4. Private property, exchange and the source of value: further considerations on the production-consumption duality and definition of a general taxonomy for the study of institutions in economic analysis

Adam Smith highlights only one possible case in which the complete distinction of production and consumption undermines the original duality of value: the "extent of the market".

In truth, two other interesting cases can be identified: absence of private ownership of assets and public control on the way in which assets ought to be used. The interesting aspect of these other two cases derive from the fact that, differently from the expansion of the market, they are strictly linked with the external legal (rather than socio-economic) context. The section highlights some interesting implications of the monism in value theory, and introduces some important concepts and aspects that will become the key topic of later chapters on regulation.

4.1. In classical economics the institution of exchange is causally determined by the institution of private property (asset ownership) as long as value remains a dualism

Above, it was noted that use value depends on the presence of private property. It is private ownership that allows a certain good or asset to be used for auto-consumption. Yet, this means that, provided that these goods possess some use value for their owner, *value is inherent in ownership*. In other words, *ownership is itself a source of economic value*, and this is true independently from the characteristics and the properties of the external context (ownership was defined precisely as that property that renders the external environment irrelevant).

These remarks are consistent with the way in which ownership has been defined above, namely as "that institution that enables the person endowed with it to enjoy some value from his property independently from the value conferred to his property by the external economic environment." It is property itself that is inherently generating value for its user.

Conversely, this also means that *it is possible to find some economic value outside the exchange process*, thus independently from the characteristics of the external economic context (provided that the owned goods have some usefulness). However, this second assertion is just another way to reframe the fundamental feature of classical "value theory" already highlighted above (which neoclassical economists have regarded as the fundamental "flaw" of classical value theory); namely, that use value in consumption is independent from the value generated in exchange.

What is more important, this interpretation can also provide a much more profound understanding of Smith's "much cited but little used" mechanisms behind the division of labour: while in the "early and rude states of society" private property is *the only source of value*, as markets develop and exchange is established, individuals can *also* rely on their own peers to address their needs and requirements. In this situation *a second potential source of value develops beside the institution of private property*: the institution of exchange. *These two institutions are the two complementary and independent sources of value in classical political economy*.

The dual nature of the institutional system in political economy is the most important consequence of the dual nature of value. However, the respective role of the two institutional layers (private property and exchange) is not equivalent. In the same way in which use value comes first and precedes exchange value (it is a necessary precondition for it), the institution of private property comes first and precedes the institution of exchange. Consistent with what has been asserted above, it is possible to summarise this by asserting that in classical economics, the institution of private property has *temporal*, *conceptual* and *analytical priority* over the institution of exchange, where the latter becomes a "residual" of, and derives from, the preliminary possession, and employment of property (intended as cumulated wealth available for use).

In conclusion, if the works of "early" classical economics are taken literally, private property (ownership) is the first and original source of economic value. It is the institution of exchange that derives from the institution of private property and not vice versa. This approach is pretty consistent across all classical economists and is a peculiar characteristic of political economy in general. As recalled by O'Brien (2004: 41), "[e]ssential to [classical economics] was security of property, an emphasis that stems from the British natural-law tradition, especially Locke and Hutcheson. Smith has emphasized this in his Lectures, and it remains a permanent assumption of his work – so permanent and so generally accepted that he felt little need to support his presumption with lengthy argument."

Figure 2.1 provides a summary of the above conclusions.

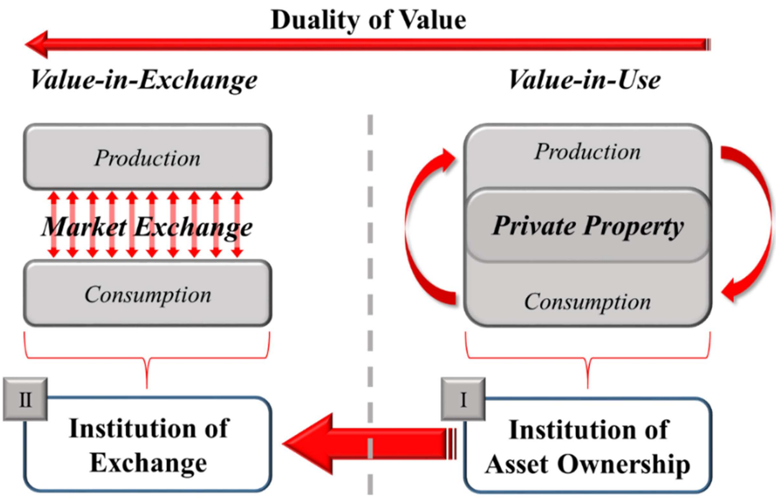

Figure 2. 1 – In a framework where production and consumption conflate, the institution of private property is legitimised by the presence of use-value and has causal priority over the institution of exchange, based on exchange-value

The logical precondition on which this specific temporal, logical and analytical causality rests is the duality of value, and the existence of use value, next to, and independent from, the presence of any value-in-exchange.

4.2. A recap of the three important cases in which the classical duality in value theory becomes a monism: a prelude to new epistemological foundations of a normative theory of regulation and public control

Even when the classical approach is adopted, use value is not necessarily always complementing exchange value, meaning that value cannot always be divided into its two independent components.

More specifically, three situations can still be identified where, granted usefulness, even according to the Smith-Ricardo legacy, only exchange value exists, while use value disappears. The three cases represent the three possible sources of monism in value theory.

The first case is obvious: use value will always be zero for those actors who do not own any property. This reason for the absence of use value can be defined as “structural” as there is no use value outside private property as already remarked. By denying actors the faculty to privately own assets (“material wealth”), the possibility to derive any usefulness in consumption that is also independent from exchange is also denied. This situation corresponds to the one where the right of private ownership (sometimes also referred to as “property rights”: see chapter 5) is absent, or denied. According to the framework developed above, the lack of private property of assets and resources has the considerable advantage of overcoming the difficulties generated by the classical paradox of value, and to transform value into a monism (which is much easier to handle, especially conceptually and analytically).

The second case is an implicit consequence of the dynamic process of the division of labour: whenever actors are perfectly specialised into one single task, they necessarily can only use their property (and specialised productive capacity) only and exclusively "as means of exchange", so to sell their services to third parties. In this case ownership is not ruled out. Yet, with perfect division of the tasks it is possible to assume that everyone is specialised to the extent that they will only be able to satisfy their own needs by purchasing services form other "hyper-specialised" actors, and vice versa. This is the last stage of society in Adam Smith's model, and it is consistent with the reinterpretation of Smith's dynamic model as developed above. Note that in this case use value might not go all the way to zero, but it might approach it "infinitesimally."

The third case is much more subtle: use value is always zero every time that the owners of certain properties cannot potentially *also employ them in consumption* (thus using these properties as means for self-subsistence). This can have two main causes.

On the one hand, the owner of a property will not perceive any use value simply because the good has no "usefulness" to him whatever. Hopefully for the owner, the good will have some usefulness for someone else, and this asymmetry in "usefulness" will trigger an exchange, and therefore generate value exclusively in the form of exchange value. Although this case is conceptually close (although not identical) to the case of perfect division of labour as described above, it still cannot be considered a classical approach to use value as usefulness among classical economists is not a subjective feature associated to the actor, but an objective physical property of the material objects, as discussed above. On the other hand it can also be the case that the owner of a property will not perceive any use value because they are limited and constrained with respect to the use that they can have of their own property.

Differently from the first case, this second case can support the classical approach to objective usefulness. In this case, the absence of use value derives from the fact that there is some external regulatory (or institutional) constraint limiting the control that an owner can exercise over their own property. In this specific case, the limitation concerns the potential option "use for self-sufficiency", or "use for personal purposes". In this case, use value goes to zero because the external institutional environment rules out the "use option", obliging the owner to necessarily transact with other (potentially interested) parties, independently from the perceived "usefulness" of the good (for private consumption).

It should be noted that, in all three cases, the common condition ensuring that use value goes to zero is that consumption is strictly separated from production so that no ambiguity in the respective roles of the various actors emerges.

The first situation, where private ownership is absent, ensures that consumers will never have the option to also become producers of the goods that they employ in consumption in order to satisfy their own personal necessities and needs. The condition of not owning any goods ensures that those who are property-less can only satisfy their needs by putting their productive power at the service of third parties (the producers) and by purchasing someone else's production with the monetary returns that they obtain in exchange. In this case, the total value perceived by property-less actors will correspond to the monetary income accruing to them in return for their productive service.

In this specific case, the (use) value of the inputs purchased in return will always equalise their (monetary) exchange value. Moreover, it is easy to see that what characterises "pure" consumers is the fact that their only endowment is in the form of monetary income, and not in the form of material wealth (which is the precondition for use value).

This first situation of no ownership is the one occurring in the case of firms in the Walrasian von Neumann-Arrow-Debreu (vNAD) general competitive model (Arrow and Debreu, 1953, Arrow and Hahn, 1971, Debreu, 1959, von Neumann, 1945):[32] All (private) wealth in the

[32] The contemporary framework in microeconomics is usually defined as "Arrow-Debreu". This is how all scholars usually denominate the standard framework in general competitive analysis. However, this expression is quite reductive. The mathematical bearing-load pillars sustaining, not only of the "Arrow-Debreu" framework, but also all contemporary normative microeconomics in general, were provided by John von Neumann's achievements in mathematics and logic during the 1920s and 1930s. On the one hand, the economic conception of rationality relies on von Neumann theory of ordinal numbers and finite or bounded sets developed in the late 1920s. On the other hand, the same paper from Arrow and Debreu (1954) could not be written if von Neumann had not provided a first complete solution of the Walrasian framework substituting equalities with inequalities, and using fixed-point theorems and the duality concept (von Neumann [1945]). For these reasons, here the contribution of von Neumann is explicitly acknowledged and the expression "Arrow-Debreu" is replaced with the more correct "von Neumann-Arrow-Debreu" (vNAD). For similar considerations concerning the role of von Neumann, see the preface of Debreu *Theory of Value* (Debreu, 1959: ix), and the historical (mainly introductory) chapters of the following monographs: Arrow and Hahn (1971; esp. p. 10), Border (1985: chapter 1), Dantzing (1963: chapter 2) Dorfman et al. (1958: chapter 1), Koopmans (1951: 1-12), and the introductory chapter in Mas-Colell (1985). For a historical assessment of the figure of John von Neumann in the history of economics, the reader is referred to Mirowski (2002: chapter 3) who went as far as stating that "he

standard competitive model is initially endowed to consumers, which are also the ones supplying productive services to firms. For this reason, although the Arrow-Debreu model is usually regarded as a model of "private ownership", this is only half-true as the adjective "private" only applies to consumers, never to producers (i.e. firms).

Conversely, the second situation of perfect division of labour corresponds to the one usually portrayed by industrial organization, and competitive models in partial equilibrium *à la* Marshall.[33] This second case of pure specialization also corresponds to the case of completely dedicated assets where any property can only render one specific service to the public, as described by TCE (Williamson, 1979; 1985; 1988). As already argued above, there is little chance of self-subsistence through auto-consumption if the privately owned property is dedicated and can only be employed in one specific way. In this case, the "relative share" of value-in-use with respect to the exchange counterpart tends to be zero as every need but one should be purchased from third parties. Under perfect specialization, the same holds for every economics actor.

More generally, both cases of absence of private ownership and of dedicated assets share the same peculiarity: the behaving actor does not face any subjective economic constraint and this is responsible for the disappearance of use value. This point will be briefly developed in the subsequent chapter.

On the other hand, the third situation ensures that those that do own some property, and that could potentially enjoy its inherent use value (at the expenses of exchange value), will *not be allowed* to do so. In this second case, property owners are "forced" to only employ

---

will come to be regarded as the single most important figure in the development of economics in the twentieth century" (p. 94).

[33] A note is in order here. An economic framework is in partial equilibrium if the economic decisions undertaken by an actor with respect to a certain transaction depend only and exclusively on the economic terms of that transaction. For informal definitions and comparison of the two the reader is referred to Mas-Colell et al. (1996: 538-540), Arrow and Hahn (1971: 6-8), or Tirole (1988: 7-12). For instance, Arrow and Hahn (1971: 6) use the following expression: "The demand and supply of a single commodity are conceived of as functions of the price of that commodity alone; the equilibrium price is that for which demand and supply are equal." The Marshallian framework is in general equilibrium for a given consumer's market as demand and supply for a certain consumer's good endogenously interact. However, the Marshallian framework is in partial equilibrium in the market of productive factors and resources, where productive services are exchanged between property holders and firms, as the purchasing prices of a given resource's service faced by firms do not depend on what other firms in other industries do (which is the whole essence of the industry-based focus of industrial organization). The same way to see this is to note that what distinguishes a partial from a general competitive equilibrium is the presence of a constraint linking the different choices. In other words, whether an analysis is in partial or in general equilibrium can be appreciated by looking at how actors' decisions are linked to a common capacity constraint. A more formal treatment is provided in Laffont (1988: 132-152) and Vives (1999: chapter 3), among others.

their property to satisfy the necessities and needs of others. This situation is identical to that of perfect division of labour as private property is allowed, yet it is the auto-consumption of this private property that is "artificially" ruled out. Thus, identically to the previous situation of "natural" perfect division of labour, even in this case, their perceived value of their property will merely and exclusively be its exchange value (the value others are willing to pay for it in an exchange relationship), and this derives from the fact that the owned property can only be used as a pure means of exchange for the indirect satisfaction of the owner's wants.

This is the situation that interests us the most here. The imposition of a "fictitious" or "artificial" constraint on actors' behaviour that impedes them to leverage their own property "internally" for their own private needs (in-use) is the essence of any normative constraint. Normative constraints can be defined precisely as those exogenously-defined impositions that purposefully constrain the behaviours of private actors.

This normative (super-imposed) divorce of production from consumption characterises two situations. One, it is the essence of the normative assumptions constraining the behaviour of consumers in the vNAD general competitive model (remember that firms are characterised by the absence of property). As the work will discuss in chapter 4 below, property owners in vNAD are constrained by very specific normative assumptions forcing them to necessarily sell the services of their private properties to others. Although these normative constraints are usually referred to as "rational behaviour", the situation is much more complex than that. Two, this third case represents the essence of public regulation in common law, and it will become fundamental once that the institution of public calling (or common carriage) is considered (see chapter 6 below).

In sum, it is now possible to have a complete overview of all the three cases in which production and consumption are sharply severed by ruling out the option of auto-consumption. Whenever autarkic employment of private properties for private self-subsistence is ruled out, consumption and production become two perfect complementary (but never overlapping) sides of the same coins, thus becoming the *dual* of each other's. In this case, value-in-use ceases to exist, and is completely taken over by value in exchange, as already mentioned above.

Figure 2.2 summarizes the argument, and complements figure 2.1 above.

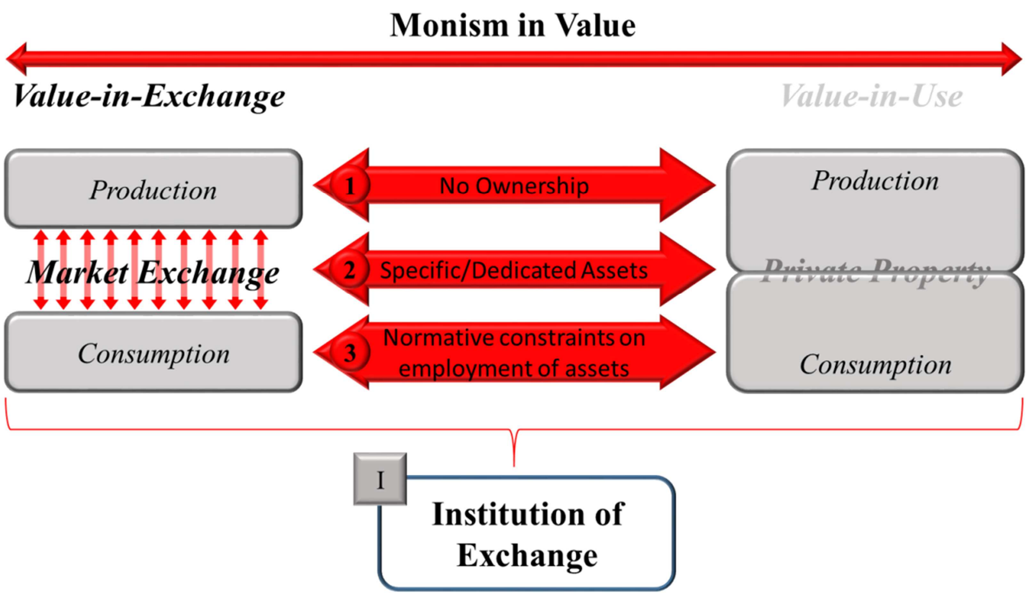

Figure 2. 2- The three specific cases in which production and consumption are severed, so to set the "use component" of value to zero. In this case the original classical duality of value is transformed into a (neoclassical) monism

4.3. From material property to immaterial (flowing) services: shift in the unit of analysis and in the meaning of "wealth"

The above discussion on the characteristics of a world where production and consumption conflate highlights a very important feature of all the cases characterised by a monistic value theory: the unit of analysis changes as the classical focus on material things (assets or resources) is replaced by a new (neoclassical) focus on (immaterial) services.

More specifically, whenever consumption and production can be clearly distinguished and the sole source of value becomes the market transaction, "wealth" ceases to be associated to the material support (the assets, or things, owned and cumulated by the various actors), and becomes a synonym of monetary income. If "something" can only be used as "means of exchange", its only source of value corresponds to the monetary rewards generated by the same thing on the market. In this case, wealth depends on the value of the services exchanged (and compensated) on the market. In other words, whenever use-value conflates into exchange-value, wealth ceases to have the original dual meaning as well, becoming a synonym of monetary income.

Thus, the duality in production and consumption, deriving from the duality in value theory, also generates another very important duality in the way in which wealth is intended: from the ownership of *assets* to the sale of (productive) *services*. As a result, the collapse of dualism in value theory also triggers the collapse of a second original dualism: the one between assets and services. In a world where only exchange value exists, material assets and flowing services (measured in monetary terms) suddenly corresponds, while the latter substitutes the former as a source of wealth.

The shift in the way in which property is intended, from the usefulness of possessed "*things*", to the rewards deriving from productive *services* sold on the market probably represents one of the key conceptual revolutions that occurred in the switch from classical to neoclassical economics. In fact, the switch from the materiality of assets and resources to the "immateriality" of services can be probably regarded as the essence of the second marginalist revolution in production and distribution theory.[34] But while the literature has usually only focussed on the role of assets and services in production and distribution, very few have noticed that this also has implications for the way in which the concept of "wealth" should be intended (Heilbroner, 2008). Similarly to value, "wealth" in a monistic world is only in exchange, never in use, as the institution of private property (on which it originally relies) disappears.

In conclusion to the section, an important remark must be made with respect to the dichotomy between assets and services. It should be noted that the case of dedicate assets (second case of monism in value) is an even stronger condition if compared to the case of external behavioural constraints (third case). As the reminder of the work will show, the former does not implement a switch of unit of analysis from assets to services, but it also automatically establishes a univocal mapping between the two: in the case of a completely dedicated (industry or service-specific) asset, there is an intrinsic univocal mapping between the service performed for the public, and the underlying asset used. This is not necessarily the case if behavioural constraint of "holding out" are imposed on the property owner.

In the case of public regulation, a second necessary duty is required to make sure that it is possible to establish a univocal mapping from the service to the underlying asset as private owner would still retain the freedom to decide which service ought to be served even

[34] Ricardian distribution theory, based on the concept of price-determined residuals to factors of production could only collapse once that the materiality of classical factors, assets or resources were replaced by the concept of services. On this, famous critiques are the ones of Knight (1935a), Robbins (1935: chapter 1), Stigler (1941), Wicksteed (1991). Thus, Robbins (1935: 4) opens his famous treatise with an enquiry on the meaning of wealth by noticing that "[i]n ordinary speech there is unquestionably a sense in which the word 'economic' is used as equivalent to 'material'", but then boldly asserts (p. 9) that "services … are wealth. Economics deals with the pricing of these services … Whatever economics is concerned with, it is *not* concerned with the causes of material wealth. The causes which have led to the persistence of this definition are mainly historical in character." This point will be the focus of chapter 4.

though they are bound to serve third parties (the "market") rather than themselves. These two behavioural constraints will correspond to the two duties of a public calling (common carriage) or to the two assumptions of welfare economics. Although this point will be developed at length in chapter 4 below, it was important to anticipate it at this stage.

In conclusion, figure 2.3 summarises the two new situations, where the duality between services and assets (and its implication for wealth) is now highlighted.

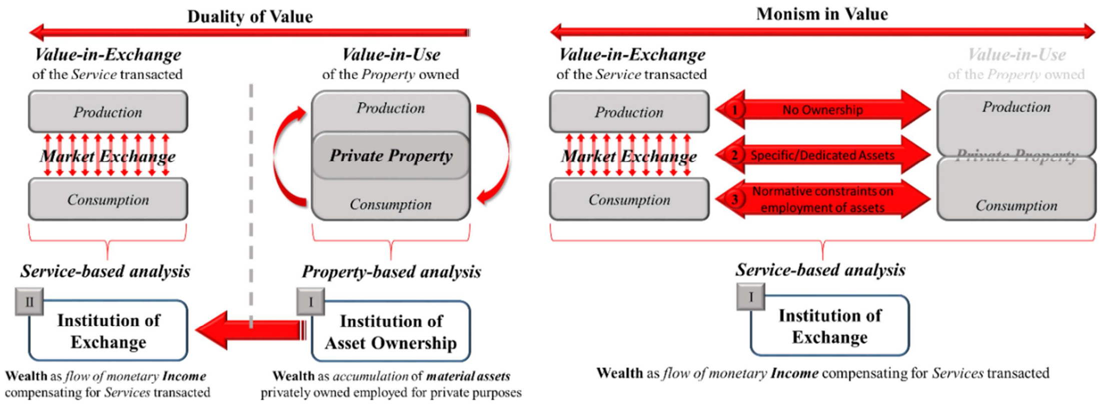

Figure 2. 3- The three key dualities summarised: production vs. consumption, use value vs. exchange value, property vs. service

4.4. Duality in value, duality of status, and the dynamic division of labour: concluding remarks

In conclusion it is possible to conclude this section by asserting the following: it *is the sharp distinction of the "consumer status" from the "producer status" that ensures "monism" in value theory.*

According to what Smith writes in *Wealth*, this strict separation between consumption and production would have been accepted by Adam Smith only for the very final stage of a society where the market has extended at its maximum and ultimate level. This extreme case corresponds to the situation of complete specialization, where value in use disappears, so that the "make" option is completely replaced by the "buy" option. Only in this extreme case, all possible activities can be provided by specialised (dedicated) producers. Only in this extreme case, the productive capacity of the property of any man will be fully dedicated to one single activity, and for the exclusive sale of its services in order to satisfy third parties' needs (in consumption). Vice versa, only in this case any need and necessity (in consumption) of any man will only be addressed "externally" (rather than internally) by purchasing the services provided by third parties' properties (in production). Last but not least, only in this extreme case does private property ceases to have an "intrinsic value" in itself, beside, and preliminary, to the institution of exchange. As a result, wealth conflates into monetary income.

In this very last stage of society the institution of exchange completely takes over the institution of asset ownership, making the latter an irrelevant sub-product of the former.

5. Asset ownership and institutions reinterpreted through the dualism of value: concluding remarks

The chapter has discussed the classical "paradox of value", and has argued that the precondition in order to have a duality in value is to allow the conflation of the activities of production and consumption.

A normative theory of asset ownership based on the concept of property can be developed only if the intrinsic value-in-use that the owner can derive from their property can be exercised beside, and independently from, value-in-exchange. Thus, the normative implications of private property in economic affairs can only be properly understood if consumers are treated as if they were producers of the services rendered by their own property and, similarly, only if producers are treated as if they were consumers of the services of their own property.

As a result, the chapter has argued that as long as any economic problem will be framed by clearly severing producers from consumers, no real theory of institution based on the concept of private property or asset ownership can be developed. In this case, the market (the exchange mechanism) can alone solve the social valuation problem. As a result, as long as the economic problem is framed under the assumption that actors can only satisfy their personal needs by *purchasing someone else's services* (through the market), and that assets or resources' owners can only employ their properties by selling their services in order to satisfy *somebody else's needs* (through the market), no economic theory of private property, or the firm, will be ever developed independently from the performance of the exchange mechanism and therefore *unless some failures in the exchange process are introduced*.

These points will be developed in later chapters of the work.

While section three has mainly reinterpreted the duality of value in light of Adam Smith's "division of labour", the last section has enlarged the picture and has highlighted the three key cases in which value ceases to be a dualism and becomes a monism instead: the absence of private ownership of material property, completely dedicated (i.e. specific) assets, and public control over the way in which private ownership ought to employ their own property. In neither of these three cases can private auto-consumption for self-subsistence occur, meaning that use value disappears, consistent with the treatment of value provided by classical economists.

The fundamental aspect that shall be noted is that nearly all branches of normative economics, and law and economics, adopt at least one of the above assumptions: while Paretian (ordinalist) welfare economics blends the first with the third case, industrial organization and its most notable application, antitrust law, is characterised by the second point, in the same way in which TCE is. This should not be surprising as TCE is a framework that originally developed in the bosom of antitrust law (Williamson, 1971; 1974; 1975). Conversely, the theory of public regulation relies on the latter case: the imposition of behavioural constraints on the way in which actors ought to employ their own properties and assets.

The three are intimately linked. Not only because they all rely on the same fundamental assumption of monistic value, but especially because they frame the economic problem by clearly identifying univocal *status* for the various actors: all frameworks (implicitly) rely on the verification that the necessary precondition in order to shift from a dualism to a monism in value theory is the clear-cut separation of economic roles into producers and consumers.

Thus, by anticipating a concept that will be developed later in the work (chapters 7), it is possible to reframe the above conclusions by saying that value theory becomes a monism whenever actors are endowed with a clear *social status*.

This chapter has highlighted one specific way to confer public status to the various actors: the distinction between the ones who only consume *from* third parties (consumers) and the ones who only produce *for* third parties. In between the two prepositions *for* and *from*, there is the market that mediates between the two, and solves the unique and univocal valuation problem. Clearly, in the case in which production (*for*) and consumption (*from*) belong to the same person (the asset owner), it is the property that mediates between the two sides, and that therefore solves the valuation problem that, in this second case, becomes "*in use*". In truth, something more is needed in order to univocally identify an actor through his social status, as we shall discuss later on. Yet, the separation of consumption and production is a first, necessary, step in that direction.

Among the three aforementioned cases, the present work will focus on the last two, and on the latter in particular. In order to understand the foundations of a normative theory of regulation, a normative theory of asset ownership and private property is first required. And, of the aforementioned three cases, the third case, dealing with super-imposed social (normative) behavioural constraints on the way in which private owners *ought to use* their properties, must necessarily represent the focus of any normative theory of regulation willing to understand the role of asset ownership in the theory of regulation, and in the antitrust-regulation debate.

CHAPTER 3

Explaining institutions through value: understanding the dismissal of classical value theory in order to generate a general taxonomy for the study of the theory of value and institutions

Chapter synopsis

While the previous chapter provided a first overview of the relationship between the classical theory of (use) value and the theory of institutions, and asset ownership in particular, the present chapter completes the picture and starts to offer some fundamental insights on the way in which value theory should be modelled in order to develop a correct taxonomy of institutions in economic analysis.

The chapter does so by resuming the discussion from where chapter 2 (section 1.2) left off, by critically examining those problems that were left open by classical value theory that convinced economists between the classical and the "neoclassical" period to overcome classical value theory and, ultimately, to dismiss it. Thus the initial focus of the chapter will be on the problems and inconsistencies emerging from the classical treatment of use value, and from the way in which exchange and use values have been linked in classical economics.

The general argument of the chapter can be summarised as follows.

The attempts made by economists over the decades to overcome the difficulties inherent in the classical theory of use value, so to provide a brand new theory of value where consumption and usefulness (utility) could finally have a new central role in economic analysis, have usually followed two lines of attack. On the one hand one line of attack focused on the problems and inconsistencies deriving from the objective treatment of usefulness and use value, and pointed out the necessity to reinterpret them in subjective, rather than objective terms. On the other hand, others focussed on the shortcomings of the theory deriving from the sharp separation of the two dimensions of value, which in classical economics remain independent and therefore mutually exogenous to each other.

Section one argues that these two lines of attack have not been clearly distinguished, nor clearly spelled out. Too often the two features of the classical theory of (use) value have been conflated and treated as if they were two alternative manifestations of the same underlying problem. Usually, the lack of subjectivity in the treatment of value has been treated as a signal of the lack of a clear analytical connection between use and exchange value.

The chapter argues that this failure to clearly distinguish the two features of classical value theory and to treat them as two different features is unfortunate, and it is at the root of many later misunderstandings concerning the nature of value in economics. More importantly, the confusion concerning the relative role of the two features of classical value theory that followed the dismissal of classical value theory (and the paradox of value) is the source of much later confusion and misunderstandings concerning the role of institutions (and asset ownership) in economic analysis. The essence of the introductory section can then be summarised as follows: a clear understanding of the way in which value shapes the nature and the role of institutions in economic analysis must rely on both dimensions, and must understand how the two fundamental characteristics of classical value theory interact and can be linked.

The goal of the chapter, developed in sections two and three, is twofold.

On the one hand (section two), it provides a new analytical framework in order to identify, characterise and understand the main features of any theory of value. On the other hand (section three), the chapter applies this framework to study of institutions (and asset ownership) in economic analysis.

The final aim of the chapter is to provide a general and reliable taxonomy of the various characteristics of value theory that allows to univocally understand the nature and the features of any economic theory of institutions by simply analysing the characteristics of the underlying theory of value on which the theory of institutions is built on. In other words, the aim of the chapter is to understand the nature and the role that institutions have in any economic theory, based on the way in which value is treated by each theory.

The detailed argument of the chapter goes as follows.

After a brief discussion on the reasons why classical economists decided to abandon the concept of use value (section one), the chapter first provides a general taxonomy for the study of value theory (section two). This general framework can be intended as the first main contribution of the chapter and it can be summarised as follows: in order to understand the nature of any theory of value, it is necessary to look at two independent (yet closely linked) dimensions.

The first dimension looks at the nature of the valuation problem, or the type of constraint shaping the valuation problem. This is represented by the dichotomy between the concept of subjectivity and the one of objectivity. Value is subjective whenever the valuation problem depends on the identity of the subject employing a certain object, and not on the intrinsic characteristics of the object, in which case value becomes subjective. More generally, the value of any entity can be regarded as objective whenever it does not depend on the identity of the specific subject involved in the valuation process.

The second dimension that characterises any theory of value looks at the way in which the two (potential) dimensions of value are linked to each other. The section distinguishes between the exogeneity and endogeneity. If the two components of value are endogenous to each other, this means that one can be expressed as a function of the other. In this case, the section argues that any theory of value can be expressed as a monism as given any (say) value-in-use, the latter can always be retrieved as a function from a given value-in-exchange and vice versa. Exogeneity characterises a situation where the two components of value have an autonomous and independent nature. If a certain dimension of value is exogenous to the other it means that it cannot be retrieved from the latter, nor can it be expressed as a (endogenous) function of the latter. In this case value remains a dualism.

Section two discusses at length the meanings of these two dichotomies, by showing that the classical treatment of use value (relying on objectivity and exogeneity of the two components of value) presents some unanswered points and is open to critiques. The section discusses the two main concerns of economists through the 19th century: the lack of a theory of usefulness and consumption based on the subjectivity of use values and the lack of a clear explanation of how use value and exchange value could interact and endogenously affect each other. The section concludes by noting that the incapacity to connect the intrinsic subjectivist nature of usefulness with the concept of value in exchange was the fundamental reason economists revised the assumptions behind classical value theory.

Section three represents the core of the chapter as it addresses the second aim: it shows how the previous two-dimensional taxonomy of value theory should be used in order to derive conclusions on the nature and the role of institutions in economic analysis.

The basic claim of section three can be summarised as follows: contrary to the usual interpretation of the "marginalist revolution", it is not subjectivity per se that ensures the relevance of (non-market) institutions in economic analysis. Although subjectivity in valuations becomes a necessary characteristics in order to study the role of institutions in economic analysis, it is irrelevant when it comes to the classical dualism in value theory. Rather, the dimension that constitutes the necessary precondition guaranteeing dualism in value theory, and therefore the relevance of the institution of ownership as opposed to the institution of market exchange, is the mutual independence (exogeneity) of the two sources of value: dualism in value is guaranteed as long as use value cannot be univocally retrieved as a function of exchange value.

The section argues that subjectivity becomes relevant in order to understand the role of asset ownership only subsequently, and only if dualism is ensured by the reciprocal independence of the two sources of value in the first place.

Thus, the first important point made by section three is that the focus that all early marginalists placed on the lack of subjective valuations in classical economics was misplaced: the introduction of subjectivity in valuations, alone, is irrelevant if not combined with the contextual independence of use value from exchange value. It is this last feature that confers an autonomous role to consumption, and can give full meaning to the concept of usefulness and utility, if the aim is to re-evaluate the consumption side that was completely neglected in classical economics.

The section develops this intuition further in order to link the characteristics of value with the study of ownerships and institutions. The reminder of the section introduces the second major point of the chapter: the dichotomy between the private and the public domain.

The section discusses the two concepts as follows. On the one hand, a valuation problem remains private whenever the only actor involved in the process of valuation is the one employing (consuming) the object. A social value is characterised by the fact of being the outcome of an interaction of individuals, meaning that there is at least a second actor beyond the one employing the good. It is this second dimension that guarantees the classical duality of value, and that distinguishes auto-consumption of a good (based on private valuations), from market exchange (based on social valuations).

The section discusses how only the mutual exogeneity of the two dimensions of value guarantees the separation between the private and the public domain, while monism in valuations also implies a monism in institutional domains. Only when the private sphere where auto-consumption occurs is distinct and separated from the social sphere where social exchange occurs, do subjective identities matter, and subjectivity in valuations matter in turn.

Thus, the second important point made by the section is that whenever value reduces to a monism, the social system becomes a “singularity”, where it is not possible to distinguish between the private and the public sphere as the two conflate: under monism in valuations what is public becomes private and what is private becomes public. This implication of value theory becomes fundamental in understanding why subjectivities in valuations become meaningless and irrelevant in a “monistic world”, while they only make sense, and become meaningful and relevant, if value remains a dualism so that the private and the social sphere can be kept separated.

In the last part of section three the chapter also defines two typologies of goods: money (or monetary goods) and “appropriable goods”. The former are those goods that are “intrinsically social” as they can be operated and used only and exclusively in the social dimension, as “money” does. These goods are always subject to monism in value theory, meaning that the private domain of value-in-use does not exist. Conversely, “appropriable goods” are those goods that can be potentially employed “in use” (even though they do not necessarily need to). For these goods, both institutional domains, the private (ownership) and the social (exchange) make sense. Consistent with the discussion developed through the chapter, subjectivity in valuation only makes sense, and becomes economically relevant, for “appropriable goods”, while it loses all meaning, becoming irrelevant, in the case of things that can be treated as if they are money.

The chapter concludes in section four by wrapping up the discussion and defining a taxonomy of economic approaches organised around the two key dimensions of value: the institutional locus of the valuation problem (private vs. public domain) and the nature of the valuation problem (subjective vs. objective).

A normative theory of asset ownership (or private property) willing to go beyond the concept of market failures must necessarily adopt a theory of value where (i) the valuation problem in the private domain is independent from the one in the social domain and where, at the same time, (ii) the former depends on the private identity of the asset owner (subjectivity in valuations).

1. Unsolved problems inherent in the "use value – exchange value" dualism: demise of the use value concept in Smith-Ricardo and in neoclassical economics

1.1. Dualism in value in classical economics: a recap of the approach towards use value by the part of classical economists

The sharp separation between consumption and production, leading to the disappearance of use value and, ultimately, to the irrelevance of private property (ownership) for the theory of value, will be the blueprint of neoclassical economics, as is well-known.

One of the important points emerging from the previous chapter is that the "paradox of value" should be contextualised, and understood in dynamic terms. This would simply reduce neoclassical economics to a very specific case of classical economics: in the same way in which monism in value is just a specific case of a more general situation allowing dualism in the nature of value, price theory simply becomes a very specific case of value theory, whenever use value goes to zero.

Consistent with what has been discussed above, the assumptions on which this monism relies can be of a different nature. On the one hand, price theory can simply be interpreted as that branch of economics studying economic interactions in the peculiar case of a society that has reached its very last stage of growth and development and that has achieved the most extreme form of division of labour. As already stated, this should be the stage where the institution of exchange (the market) remains the sole and undisputed source of value. On the other hand, nothing impedes the conception of price theory as the branch of economics that models a society where the employment of private assets or resources is publicly constrained by means of social duties and obligations.

While the first interpretation of price theory is consistent with the second case illustrated in chapter 2, the second one would correspond to either case three, or one, depending on the "degree of public control" on the employment of resources, public ownership being the ultimate, extreme remedy. Both interpretations have their merits and are legitimate from both a conceptual and epistemological point of view. The second, especially, will play a key role in the two debates that developed at the beginning of the 20th century concerning the nature of value on the one hand, and the role of social planning and welfare economics on the other. The reminder of the chapter will develop this very important point further, so we will discard it for the time being.

Although in Smith's *Wealth* the nature of value is quite explicitly put in relationship with the (dynamic) characteristics of the external context, this was not the way in which economists (not only neoclassicals, but also classical contemporaries of Ricardo) interpreted the Smith-Ricardo distinction between use value and exchange value. In fact, this is not even the way in which Ricardo himself interpreted Smith's concept of use value. The "dynamic" mechanism ensuring growth and development in Ricardo, is not represented by Smith's division of labour (triggering the endogens shift from use to exchange value), but by the Malthusian law of diminishing marginal fertility of land.

As a result, the distinction between use value and exchange value has usually not been regarded in dynamic terms, but as a fact that had to be accepted and that could not be subject to economic enquiry. Usefulness, or value in use, existed independently from exchange value. This was the very-well-known paradox of value.

For this reason, given the sharp division between the two natures of value, the concept of use value has never been really employed in classical political economy, and all classical economists dismissed the concept, and therefore ruled out the numerous complications deriving from the paradox of value.

Before moving on to discuss the role of value in the debate for socialist planning, that has greatly influenced the contemporary debate on the theory of institutions, it is therefore first necessary to provide an overview of the way in which early and late classical economists tried to address, or dismiss, the complications deriving from the concurrent presence of two forms of value.

1.2. The demise of use value on the part of classical economists: a brief review of the analytical, methodological and epistemological motivations

First of all, it is useful to better understand the reasons behind the dismissal of value-in-use by the part of classical economists. In sum, use value has been dismissed for a mix of methodological, analytical and epistemological reasons.

From an epistemological pint of view, a possible source of problem was represented by the fact that classical political economy was predominantly interested in the dynamic accumulation of wealth, and in the way in which the latter was divided into monetary shares reimbursing to various parties participating in the production process (distribution theory). As summarised by Ricardo in his very first page of the Principles (1911: 1): "To determine the laws which regulate this distribution is the principal problem of political economy." Value theory, and the origin of value, were not interesting questions *per se*. The theory of value was a pure instrumental tool serving the cause of distribution theory. This was particularly true for that aspect of value that does not even participate to productive and distributive process, such as use value.

Adam Smith's justification for the dismissal of use value is a mix of epistemological and methodological reasons: the underlying mechanism justifying the wealth of Nations is the increasing division of labour deriving from specialization. Yet, the underlying assumption running through the *Wealth* starting from chapter 5 onward, is that *the division of labour already occurred*, and that political economy is that branch of social sciences that only enquires about the origin and the nature of value for well-developed societies. "Exchange" then becomes the default option in which privately owned wealth can be employed, as direct self-subsistence becomes an irrelevant portion of the economic activity for highly specialised societies.[35]

On the other hand, Ricardo's justification is also a blend of epistemological and methodological reasons. On the one hand, Ricardo, after having introduced the Smithian distinction between use and exchange value, and having linked to former to "scarcity goods", and the latter to "labour-goods", he dismisses the former. Yet, it is important to underline the fact that this logical step is not justified on the grounds of the Smithian division of labour, but it is described as a mere analytical simplification, as follows (Ricardo, 1911: 6; emphasis added):

> "There are some commodities, the value of which is determined by their scarcity alone … Their value is wholly independent of the quantity of labour originally necessary to produce them, and varies with the varying wealth and inclinations of those who are desirous to possess them. *These commodities, however, form a very*

[35] This is how chapter 5 of the *Wealth* starts, with a clear and explicit abandonment of the self-consumption option, which has been nevertheless the core of the first four chapter of the work: "Every man is rich or poor according to the degree in which he can afford to enjoy the necessaries, conveniences, and amusements of human life. But after the division of labour has once thoroughly taken place, it is but a very small part of these with which a man's own labour can supply him. The far greater part of them he must derive from the labour of other people, and he must be rich or poor according to the quantity of that labour which he can command, or which he can afford to purchase. The value of any commodity, therefore, to the person who possesses it, and who means not to use or consume it himself, but to exchange it for other commodities, is equal to the quantity of labour which it enables him to purchase or command."

> *small part of the mass of commodities daily exchanged in the market*. *By far the greatest part of those goods which are the objects of desire are produced by labour* … In speaking, then, of commodities, of their exchangeable value, and of the laws which regulate their relative prices, we mean always such commodities only as can be increased in quantity by the exertion of human industry, and on the production of which competition operates without restraint."

Ricardo operates an arbitrary restriction of the field, justified by the scarce analytical relevance of those commodities that do not derive their value from the expenditure of labour in production, but by their mere abundance or scarcity, *while possessed and employed*.

In contrast to Smith, where the conceptual link between self-consumption, use value, and the division of labour is clearly drawn since the beginning of the *Wealth*, the conceptual link between use value, "scarcity goods", personal wealth (riches), and self-employment in consumption is not clearly stated in chapter 1 of the *Principles*, and Ricardo seems to be quite unclear and ambiguous about that. Yet, the similarity with the Smithian treatment becomes quite evident in chapter 20, when the distinction between value and wealth is spelled out.

At the same time, the dismissal of use value can also be justified on epistemological grounds: as the value of "scarcity resources" is independent, not only from labour, but from the production (and distributive) process in general, they should be left out of the analysis.

Skipping the faithful Ricardian disciples of the 1830s (especially James Mill and McCulloch), as they did not differ in any way from the Ricardian treatment of value, but also ignoring for the moment the anti-Ricardians of the 1830s (such Senior, De Quincey), that have never been fully convinced by the dual nature of value as described by Smith, we are left with J.S. Mill and Marx.

Like all other classicals, J.S. Mill's approach towards the rejection of use-value is also twofold: analytical and epistemological.

Firstly, what is interesting about J.S. Mill's treatment of value is his explicit connection of the concept with the Smithian division of labour. More specifically, from J.S. Mill treatment of value it appears that the author bases his refutation of the concept of use value on the same analytical considerations already advanced by Smith. This is how J.S. Mill (1948: book 3, chapter 1, section 1) justifies the treatment of the topic of value:

> "In a state of society, however, in which the industrial system is entirely founded on purchase and sale, each individual, for the most part, living not on things in the production of which he himself bears a part, but on things obtained by a double exchange, a sale followed by a purchase – the question of Value is fundamental …"

Where the term "Value" in J.S. Mill means "exchange value" only. This passage is very interesting as it is perfectly in line with the Smithian interpretation of the division of labour in value terms, as discussed in chapter 2 above. Thus, the above statement seems to suggest that only in an advanced society where self-consumption of own wealth is not a meaningful (or feasible) option anymore, exchange value becomes a fundamental concept in economics. As a result, the above assertion by J.S. Mill also seems to imply that in the *other* stages of society, when self-employment can be feasible, and is widely practised, it is the preponderant role of use value that makes (exchange) value of little relevance.

This is however not the case. The reason for this is epistemological and it derives from the fact that the epistemological assumption formulated by J.S. Mill is that political economy, as a discipline, "has nothing to do with the comparative estimation of the different uses" of a material good. More precisely, J.S. Mill agrees with the standard classical approach to use value, asserting that "the use of a thing, in political economy, means its capacity to satisfy a desire or serve a purpose" (Mill, 1948: book 1, chapter 1, section 2). However, he also disposes of use value quite abruptly and with few, clear-cut words:

> "political economy has nothing to do with the comparative estimation of different uses in the judgement of a philosopher or of a moralist."

In other words, economics does not deal with the problems deriving from the way in which resources or goods ought to be employed by their owners as this belongs to the realm of ethics and morality.

Thus, the essence of the epistemological approach to the demise of use value simply derives from the specific sphere of enquiry that classical economists thought political economy should have been circumscribed to. But these same authors do not deny the legitimate and independent role of use value in society. To assume that use value is not part of political economy does not mean that use value, by itself, does not exist, or does not have his own conceptual legitimacy.

What is interesting about J.S. Mill treatment is that the same exact reasons for which he discards use-value on an epistemological ground are precisely the ones that will foster the rise of marginalism in consumption and utility theory: the dependence of a certain good employment on the idiosyncratic will of its own administrator. We will come back on this later in the chapter.

Lastly, Karl Marx. Marx is the last classical economist to carefully develop the dichotomy between use and exchange value. Consistent with Smith and Ricardo, even Marx dismisses use value. Marx's motivations are yet different from both Smith's and Ricardo's. Marx does not dismiss use value on the grounds of some eschatological unfolding of a dynamic process (as in Smith's division of labour), nor on the ground of the "relevance of the analysis" (as in Ricardo). Conversely, Marx's concerns mainly derive from the fact that while use value is a qualitative measure of usefulness, exchange value can be quantitatively measured. It is the problem of measurement and relative comparison that bothers Marx.

Thus, this is how Marx disposes of the duality of value (1995: 28):

> "As use values, commodities are, above all, of different qualities, but as exchange values they are merely different quantities, and consequently do not contain an atom of use value. If then we leave out of consideration the use value of commodities, they have only one common property left, that of being products of labour …"

In other words, Marx recognises that material goods possess some inherent utility, to be intended as an absolute property of the materiality of the good. Yet, this level of utility is "hidden", making it "impossible to grasp." For this reason, the only possibility to grasp and measure this absolute level, is to generate a relationship with some other goods, and this makes the concept of exchange value much more useful and informative than the elusive one of use value. In his words (Marx, 1995: 33):

> "The value of commodities is the very opposite of the coarse materiality of their substance, not an atom of matter enters into its composition. Turn and examine a single commodity, by itself, as we will, yet in so far as it remains an object of value, it seems impossible to grasp it … In fact we started from exchange value, or the exchange relation of commodities, in order to get at the value that lies hidden behind it."

In sum, if we want to measure and compare the values of qualitatively heterogeneous goods, we must abandon the elusive concept of use value, and just focus on the second manifestation of value: exchange.

On the one hand, Marx's labour theory of value cannot accept a theory of (exchange) value based on subjective wants, even though Marx is well aware of the fact that use value can only derive from the satisfaction of actors' subjective wants in consumption. At the same time Marx knows that the concept of use value generates some insurmountable analytical difficulties deriving from its qualitative nature, from its hidden and unknowable nature, and from its incommensurability with exchange value.

For these reasons, Marx follows the path already taken by both Smith and Ricardo, and puts use value outside the realm of political economy. However, as use value is that aspect of value associated with the employment of material goods in consumption, this necessarily also places consumption outside the realm of classical political economy.

In conclusion, beyond the different motivations put forward by the various classical economists, one common aspect links them all: use value was simply discarded and brought outside the realm of political economy for analytical, methodological, or epistemological reasons. The concept was not rejected on logical grounds.

2. Some fundamental difficulties in the classical approach to use value: trying to overcome the objectivity and exogeneity of usefulness in consumption

If one reconsiders the two issues left open by the classical approach towards use value, it is easy to see that the germs for the marginalist revolution in value theory were already present since at least the 1840s. Thus, it shouldn't be surprising that the Smith-Ricardo-Marx paradox of value was already under attack (or already under reconsideration) well before the 1870s Jevons-Menger marginalist revolution in value theory (Blaug, 1958: chapter 3).

More precisely, the Smith-Ricardo-Marx approach to use value and the paradox of value leaves the theory exposed to two broad types of critiques.

First, a theory of use value based on the usefulness of the goods in satisfying wants in consumption necessarily calls for a reconsideration of the role of consumption in economics (beside production), and for a new economic theory explanting how subjective wants and subjective utility in consumption can shape and influence (use) value, and therefore the specific way in which the underlying material thing is used and employed.

Second, the strong dualism in value theory called for an attempt to solve the irreducibility and incommensurability of the two sources of value: "in use" and in "exchange."

These two weaknesses of classical economics can also be reinterpreted in function of the characteristics of use value as introduced in section 2.2 of chapter 2 above.

On the one hand the first problem derives from the *objectivity* of the concept of use value (section 2.2.1 of chapter 2 above), and from the materiality of the concept of usefulness. The latter in classical economics is hopelessly disjointed from the inherent *subjectivity* of the wants of the actors employing factors, resources and goods.

On the other hand, the second problem is linked to the mutual *exogeneity* of use-value (and usefulness) with respect to exchange-value (section 2.2.3). It should be noted that this point is conceptually linked to the first one: if use value originates from the intrinsic material properties of the physical things, it is necessarily also independent from, or *exogenous to*, exchange value. The opposite is also true in classical economics where exchange value is an outcome of the sole production process, thus it has nothing to do with the specific usefulness (in consumption) of goods and resources (section 2.2.4).

It is important to develop these two points more in depth, as subsequent developments in all fields of normative economics (microeconomics and the theory of consumer behaviour, Paretian welfare economics, institutional economics and regulatory economics) can all be traced back to the problems raised by the material *objectivity* and by the *exogeneity* of use value in classical economics, for reasons that will be clarified later on in the present chapter.

2.1. Difficulties in the classical approach to use value 1: objectivity of the matter and subjectivity of its employment

The first serious limitation inherent in the classical treatment of use value is represented by the notion of "usefulness".

As already discussed above, the usefulness of a commodity was intended as an objective property of the commodity by all major classical economists. In other words, in classical value theory usefulness is embedded in the materiality of the physical artefacts and resources, and for this reason it becomes an objective property of any good, resource, or artefact of the material world. On the one hand, this is consistent with the classical "material" interpretation of wealth, regarded as the accumulation and stockpiling of physical things. On the other hand, this peculiar interpretation of usefulness is what rendered value-in-use an *objective property of the physical good* employed in consumption.

It is important to understand what the objectivity of use value means and implies: Use value in classical economics is objective as the usefulness of any physical good or resource does not depend on the *subject* employing the good, but it is inherent and entrenched in the (objective) physical characteristics of the object. In other words, value-in-use in classical economics is a property of the material *object*, while it is independent from the acting *subject*. Yet, this also implies that the objective physical properties and characteristics of the good in question univocally define the function(s) that the good is supposed to fulfil, and therefore the way in which any subject ought to employ the given good or asset.

The objective interpretation of usefulness highlights a fundamental characteristic of classical political economy: *the primacy of the physical properties of the material world over the decisions that can be potentially taken by the (economic) actors*.

Stated differently, it is possible to say that the classical interpretation of use value implies *the primacy of the ontology of the material world over the economic dimension*. Not only does the (given) ontology of factors determine the usefulness of these factors, and therefore their inherent (use) value, but all actors are equally constrained in their economic decisions and behaviours by the unavoidable ontology of the material goods employed, and by their physical materiality. The (given) ontological properties of the material world are given to the subjects and constraint all actors' behaviours in the same way or, to use Marx's (1995; 27) words:

> "A commodity is, in the first place, an object outside us, a thing that by its properties satisfies human wants of some sort or another … The utility of a thing makes it a use value. But this utility is not a thing of air. Being limited by the physical properties of the commodity, it has no existence apart from that commodity. A commodity, such as iron, corn, or a diamond, is therefore, so far as it is a material thing, a use value, something useful."

Thus, all actors perceive the same (objective) physical properties of the material goods that are located "outside them". Moreover, these physical properties of the *material object* inform the *subject* on how the former *ought to be employed*.

In sum, the objective nature of classical value-in-use can be interpreted in two ways. On the one hand, usefulness (use value) is objective as the constraints to the employment of the material goods derive from the physical properties (the ontology) of the *objects*, rather than from the autonomous decisions of the behaving *subjects*. On the other hand, usefulness is also objective because these physical properties of the object are identically perceived and experienced by all subjects, and the former remain independent from the (economic) decisions taken by the latter.

It should be noted that this peculiar implication of use value, leading to the primacy of the ontology of the physical world over the economic dimension represented by the decisions of the actors is consistent with what has been discussed above (sections 3.2.1 and 4.1 of chapter 2) concerning the priority of private property and personal use (generating use value) over the institution of exchange and interpersonal use (generating exchange value); this is illustrated in figure 3.1 below.

For this reason, it is also possible to assert, along the same lines, that, among classical economists, *the ontology of the material objects seems to have temporal, conceptual, and analytical priority over the way in which the subjects employ and operate them*.

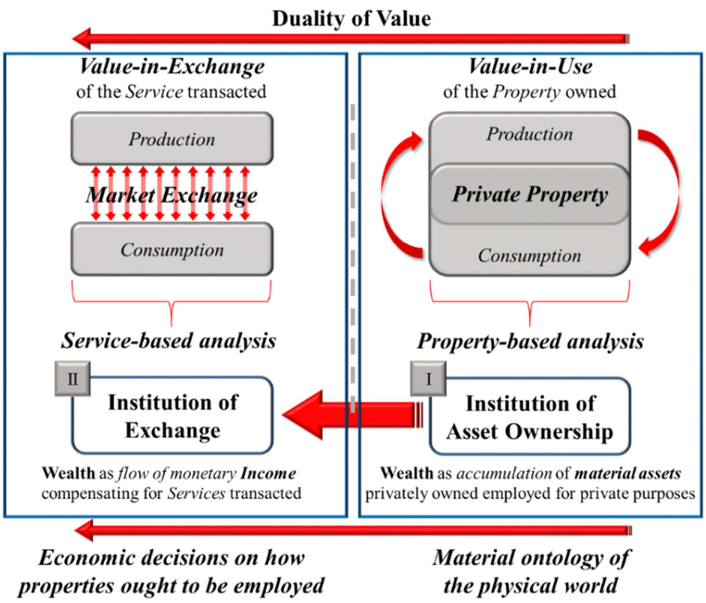

Figure 3. 1 – In classical economics, the ontological constraints posed by the materiality of objects is given to the actors, and is treated as an object constraints to actors' decisions

From figure 3.1 an interesting parallel should be noted: while actors' decisions on how things ought to be employed are in the camp of a service-based analysis, the material ontology of the physical world is linked to the camp of property. Leaving these considerations for the next chapter, it should be already noted that this retains logical consistency as the employment of assets or resources deals with services, while ownership deals with the physical support employed. The same dichotomy can also reinterpreted as one between, respectively, control and ownership.

The above considerations show a clear inconsistency in the classical approach towards usefulness: on the one hand, usefulness was regarded as an objective property deriving from the materiality of the underlying asset or resource; but on the other hand, it was clearly recognised that use value depends on the specific way in which the subject was employing the underlying "material support".

It is sufficient to go back to Marx or J.S. Mill discussion of use value (above) to realise that these authors consider use value as a function of the *different wants* (Marx), or of the *different uses* (J.S. Mill) things can be put to. How the objective materiality of the underlying asset could be reconciled with the variety of (qualitatively diverse) uses in which the same materiality could be employed was an open problem that was never satisfactory solved, nor addressed by classical economists. In fact, it is precisely because of this problem that the elusive concept of use value (and usefulness in consumption) was *purposefully* and *explicitly* discarded (this is, in a nutshell, the essence of both Marx and J.S. Mill's rejection of use-value). The classical silence on the role of consumption and utility is just the ultimate, more visible, consequence of this purposeful choice to discard use value from their analytical and theoretical inquiry.

It is this objective interpretation of usefulness, leading to evident conceptual inconsistencies between the object and the role of the subject, that has been regarded by later neoclassical economists as one of the most fundamental shortcomings of classical economics in general, and certainly as the most important shortcoming of classical value theory.

Early neoclassical economists reproached to classical political economy the complete demise of the fundamental role that the *subject* has in determining how *the object* (the material artefact) *ought to* be employed and used. According to neoclassicals it is the perception, or inclination, of the *subject* that defines usefulness, rather than the materiality of the *object*. This epistemological shift from the object to the subject as the ultimate source of usefulness is what will characterise the subjectivist revolution in value theory, starting with the first marginalist revolution of the 1870s.

Here it is important to clarify a potential source of confusion with respect to the way in which the concepts of objectivity and subjectivity are interpreted among classical economists.

The previous considerations might seem inconsistent with some claims developed in the previous chapter (section 2.2.5). Above, it was stated that the Marxian interpretation of the paradox of value could be intended as a dichotomy between subjective value-in-use and social value-in-exchange. However, the Marxian way to use the term "*subjective*" is different from the neoclassical interpretation of the term. In classical economics (mainly in Marx though) use value is subjective as it is the value generated by the physical usefulness *in the direct consumption of the good*, as illustrated in chapter 2. In this sense "subjective" should be intended as a synonym of "private", as opposed to "social", or collective. The act of consumption is subjective as it does not involve the participation of other people. In this sense, the employment of a good for self-consumption for the direct satisfaction of an actor's private needs (value-in-use), is not a social activity. In Marx's words, it is subjective. But this does not mean that the nature of value-in-use in classical economics is not objective, as it originates from the ontological properties *of the object* employed in consumption (the material goods).

More precisely, Marx distinguishes two different sources of value, respectively associated to use and exchange value: the *material substance* of value, generated by the physical properties of the material world, and the *social substance* of value, deriving from the social interactions of the various actors (volume 1 of *Capital*: Book 1, Chapter 1, Section 3). While the material substratum constraints the employment that any actor can derive from any material good *qua means of consumption*, the "social substratum" constraints the employment that any actor can derive from any material good *qua means of exchange*. In this in this terms that "subjective" has to be interpreted in classical economics, consistent with the discussion developed in chapter 1: as defining the *locus of the constraint* to actors' behaviour (their own private property vs. the social context), and not the *nature of the constraint* depending on the specific identity of the subject. .

It is important to remark this distinction as the Mengerian reinterpretation of Marxian value-in-use will derive from a conflation of the classical interpretation of subjectivity (denoting private self-consumption) with the neoclassical interpretation of subjectivity (denoting subjectivity in feelings and perceptions).

2.2. Difficulties in the classical approach to use value 2: independence and mutual endogeneity with exchange value

The second serious limitation inherent in the classical treatment of use value is represented by the incompatible natures of the two manifestations of value: "in use" and "in exchange". In classical economics, not only are use value and usefulness exogenous to exchange value, but also exchange value is exogenously given to use value. The two trace their origins back to two completely different, and incommensurable, mechanisms.

2.2.1. Some digression on the meaning of exogeneity: independence and functional relationships

It should be noted that the mutual independence (exogeneity) of the two dimensions of value is the essence of the well-known "paradox of value" itself. Yet, this same statement might also sound completely conventional. After all, use value and exchange value in classical economics are associated to, respectively, the act of consumption and the act of production and the two are treated as independent (exogenous) primitives in neoclassical economics as well. This confusion concerning the appropriate way to interpret the "mutual exogeneity" of use value with respect to exchange value depends on the specific way in which the terms "exogeneity", or "independence" are intended in turn.

The term "exogeneity" only makes sense if referring to something: something is exogenous with respect to something else. In this case, the exogeneity of some element with respect to another one simply means that the former *does not depend* on the latter, but that *it is given* to the latter. In classical political economy, the resulting value-in-use experienced by any actor employing a certain physical thing is exogenous and independent with respect to two different things: the production process *and* the exchange (market) mechanism (value-in-exchange). The two are quite different.

On the one hand, value-in-use is exogenous with respect to the production process due to the fact that usefulness is an inherent property of the material object, and of its physical features and properties, as discussed above. In this sense, the intrinsic level of usefulness that can be derived from any material object does not depend on the features of the production process shaping exchange value.

This way to interpret the exogeneity of use value in classical economics is perfectly consistent with the standard exogeneity of consumption's preferences with respect to the production process, normally assumed by neoclassical economics later on.

Yet, there is a fundamental difference between the exogenous source of value-in-use in neoclassical economics, and the exogenous source of value-in-use assumed by classical political economy: in neoclassical economics utility is not the exogenous source of value-in-use only, but it is also contributing to value-in-exchange (i.e. price in neoclassical terms). This means that, although the origin of usefulness (and utility) is exogenous, exchange value is in not independent from use value and *vice versa*, and this makes the latter just another *endogenous* aspect of the former *and vice versa*.

The previous point leads to the second interpretation of exogeneity: in classical value theory value-in-use is also exogenous with respect to value-in-exchange. This second point is much more powerful than the simple consideration that value-in-use is independent from the characteristics of the production process. The fact that use value is exogenous with respect to exchange value means that the physical usefulness of any material good does not depend on the valuations that actors formulate on the market, nor the degree of usefulness experienced by an actor employing the specific physical good will vary with the exchanges operated in the market. Thus, no matter what the valuation of a good generated by the market (its price, or value-in-exchange), the value-in-use experienced by any actor employing the physical good will remain unaffected.

The reason why it is important to stress this point is that the same does not hold in neoclassical economics where the level of usefulness (value-in-use) generated by the use of a (certain quantity of a) good does depend on its value in the market and can be retrieved from it.

The previous considerations can be better understood if the concept of “exogeneity” is redefined in a further, but very meaningful, way: a parameter is exogenous with respect to another one if the former cannot be univocally retrieved from the latter or, in other words, if the knowledge of the latter is not necessarily informative in order to understand the value of the former.

This definition of an “exogenous parameter” derives from the recognition that a variable is *endogenous* to another one only if the former can be expressed as a function of the latter or, in other words, only if the knowledge of the latter is sufficient in order to retrieve the value of the former. This is simply the definition of a function, meaning that a certain variable y is therefore endogenous to another variable x if it is possible to express the former in function of the latter ($y = f(x)$). A function, by definition, establishes an endogenous relationship between the independent (x) and the dependent (y) variable, meaning that the knowledge of the former is sufficient in order to also univocally retrieve the value of the latter.

2.2.2. *Mutual independence and mutual dependence of values: distinguishing classical form neoclassical economics*

The fact that value-in-use is independent from value-in-exchange means that the former behaves as an exogenous parameter with respect to the latter. For this reason, in classical economics, value-in-use cannot be univocally retrieved from value-in-exchange as it is not possible to establish a univocal mapping between exchange value and use value: value-in-use cannot be written as a function of value-in-exchange in classical economics. Moreover, the reverse is also true as value-in-exchange is also exogenous to value-in-use in classical economics.[36]

One of the main reasons why use-value remains exogenous and independent from exchange-value in classical economics certainly derives from the third characteristic of use value highlighted in chapter 2 above (section 2.2.2): its qualitative and incommensurable nature. The qualitative nature of use value makes it incommensurable, and the latter, in turn, implies that use value cannot be compared with any other magnitude, nor can it be quantitatively expressed in function of any other magnitude.

Conversely, any quantitative measure, in order to be defined as such, should be expressed relative to some other magnitude (unit of measure). In other words, comparability implies commensurability, and commensurability implies relativity. Yet, as the usefulness of different objects cannot be quantified, it cannot even be compared, and therefore it cannot be expressed as a relative relationship.

The independence of use value with respect of exchange value deriving from the incommensurability of the former was already clearly identified by Marx, as recalled in section 1 above. In fact, the measurement problem, and its implications for the quantification of use value, was precisely one of the two main motivations why Marx, last among classical economists, abandoned the concept of use value. In Marx's words (1995: 28):

> "We have seen that when commodities are exchanged, their exchange value manifests itself as something totally independent of their use value. But if we

[36] Yet, it should be remembered that this does not hold with respect to the relation between the two institutional layers: as discussed in chapter 2, the institution of market exchange always follows, and derives from, the institution of ownership in classical economics.

abstract from their use value, there remains their Value as defined above. Therefore, the common substance that manifests itself in the exchange value of commodities, whenever they are exchanged, is their value. The progress of our investigation will show that exchange value is the only form in which the value of commodities can manifest itself or be expressed."

The problem of the incommensurability, and therefore lack of comparability, of value in use and usefulness will represent a key topic for many decades in economics. On the one hand the problem of the lack of comparability of the intrinsic (ontological) properties of material goods represents the most important reason why the unit of analysis of the economic process shifted from the notion of material resources or factors, in favour of the concept of services during the second marginalist revolution in production and distribution through the 1890s (see chapter 4, section 3 below). On the other hand the incommensurability of use value and utility in the private employment of any good (consumption) will represent the main reason why, several decades later, microeconomics switched to the ordinalist framework of Paretian welfare economics.

Both these points will be developed in the remainder of the work. For the moment, it should be highlighted the fact that the exogeneity of value-in-use is the precondition for the paradox of value: it is precisely because value-in-use is exogenous with respect to (and independent from) value-in-exchange (and vice versa) that value has a dual nature in classical economics. Conversely, it is clear that the monistic nature of value derives precisely from the fact that the valuations formulated in the market by the part of the transacting parties can endogenously affect the level of use value experienced by each actor. In this case, the latter can be (univocally) expressed in function of the former.

Whenever the two magnitudes can be expressed in function of each other, by establishing a mutual endogeneity between the two natures of value, value-in-use can be univocally expressed in function of value-in-exchange, and vice versa, value-in-exchange can be univocally expressed in function of value-in-use. In this case, a duality between the two aspects of value can be established, and value can be treated as a monism, while solving the so-called "paradox of value".

In conclusion, it is certainly true that the exogeneity of use value with respect to exchange value is also a consequence of the fact that value-in-use in classical economics exclusively depends on the objective properties of the material goods, as discussed above (section 2.1). However, the previous consideration highlights a more general, and more fundamental, relationship: there is a clear linkage between the *monistic-dualistic* nature of value, and the *exogenous-endogenous relationship* between use value and exchange value.

The case of exogenous value-in-use, when value-in-exchange is "enacted" only whenever there is a surplus in consumption, as already remarked in chapter 2, can be represented as in figure 3.1. On the other hand, the mutual endogeneity between value-in-use and value-in-exchange establishes a univocal mapping between the two that is very close to the situation represented in figure 3.2.

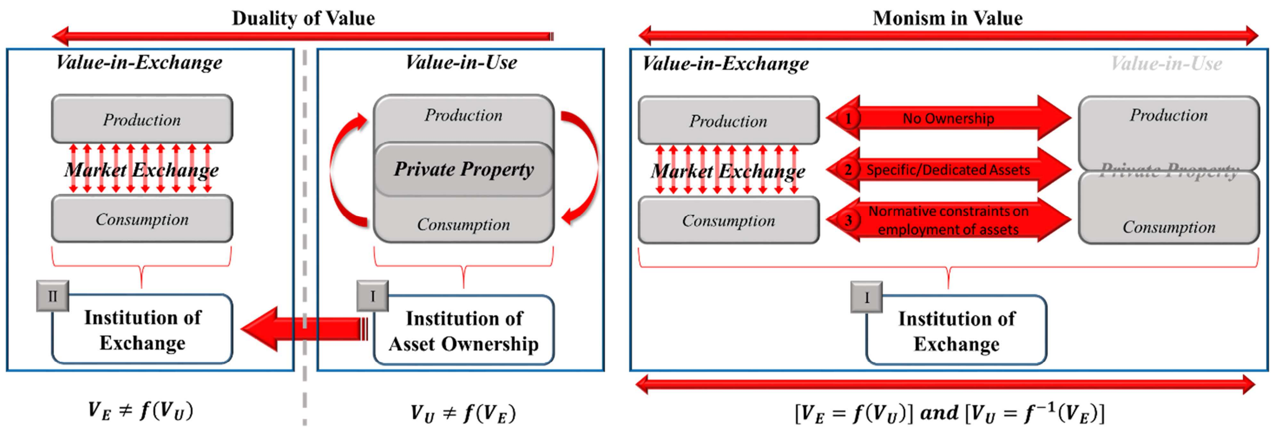

Figure 3. 2 – Monism in value is guaranteed by the fact that one dimension of value can be univocally retrieved from the other and vice versa; dualism in value is guaranteed by the fact that the two dimensions of value can exist independently and cannot be expressed one as the function of the other and vice versa

Figure 3.2 graphically represents the close relationship that can be established between the reciprocal independence of the two sources of value and the dualism in value theory on the one hand (left), and the mutual endogeneity of the two dimensions of value and the monism in value theory on the other hand (right).[37]

Here an important clarification is in order.

In the classical paradigm both use and exchange value are exogenous to each other. This means that neither analytically determines the magnitude and the level of the other as their respective magnitudes derive from two independent and irreducible sources and mechanism. The classical incommensurability of use value vis-à-vis exchange value fulfils precisely this aim. For this reason the mono-directional right-left arrows that appear in all the figures used through the work up to this point (i.e. the three arrows in figure 3.2), should not be represented as suggesting a functional causality of the form $[V_E = f(V_U)]$, meaning that it is possible to retrieve the magnitude of exchange value starting from use

[37] The expression f^{-1} simply denotes the inverse function of f, showing that value-in-use and value-in-exchange are the dual of each other's.

value. If this was actually the case, then we would not have a dualism in valuations anymore, but we would find ourselves back into a monism, where use value could be used as the sole origin of value, also determining exchange value (this was actually the –never-fulfilled– aim of early marginalists).

Conversely, the arrows represent a logical and temporal causality that establishes what should be considered first, and what causally triggers what. In other words, they define (logical and conceptual) *priorities over time* between incommensurable variables. It is this type of priority that gives substance to Smith's process of the "division of labour" (chapter 2). The incommensurability of the two dimensions makes sure not only the neither can be expressed in function of the other, but also that the principle according to which one (exchange value) is triggered by the other (use value) cannot be represented numerically (by a standard function), but must follow other logics. In chapter one, it was illustrated that the logic followed is the following: only when personal necessities are satisfied (self-subsistence), then exchange can be triggered beyond auto-consumption. "Incommensurability" is therefore a key element in order to understand the nature of classical value theory.

2.2.3. Preliminary attempts to solve the classical dualism of value: expressing use value in function of (monetary) value in exchange

The intimate relationship between the incommensurability of use value, its independence with respect to exchange-value, and the persistence of the classical paradox of value was starting to be understood already in Marx's times. As a result, the intellectual effort starting from the early 1840s was mainly devoted towards the solution to the paradox of value.

The focus was soon placed on the fact that the measurement and incommensurability problem of use value deriving from its objective (hidden) materiality, could be automatically solved whenever it could be expressed in function of exchange value. The latter, different from the former, could be quantified and measured, thus compared.

As a result, the first line of attack of these early critiques can be summarised as follows: even if it is accepted that use value and exchange value are "in principle" originally two different entities, nothing prevents the former from converging with, and eventually conflating into, the latter. In other words, it might very well be the case that a closer look to use value and to its nature will reveal that the initial distinction is only apparent and that the former will necessarily resolve itself into exchange value, rendering the two indistinguishable and equivalent components of the same, unique value concept.

Thus, the apparent shortcoming of the Smithian paradox at the eyes of late classical economists, (and early marginlists) derives from the fact that the possibility that use value can simply be interpreted as the “extreme limit” (or the upper bound) of value-in-exchange” was not considered as an acceptable option. This remark was already highlighted by Thomas De Quincey in the 1840s, and this critique has been retained since by J.S. Mill in his *Principles* and, especially, by Marshall (see Bharadway, 1978). Restated differently, under Marshall value-in-use becomes *the monetary compensation* of a foregone use-value.

The way in which Marshall (independently from, and probably before, Jevons himself) solves the paradox is by attributing a monetary measure to use value, thus making the latter quantifiable and measurable. Use value thus becomes perfectly commensurable with exchange value. As clearly explained by Bharadwaj (1978: 261):

> “[Marshall] … proceeded to define the value-in-use of a thing to a person as 'the value of the things which must be given him in order that he may be induced to give it up, or which he will give up rather than not obtain it'. By so defining it, Marshall was already moving on to a measure of use value in terms of the price that the individual is willing to pay for a certain unit of a commodity rather than forgo its possession.”

In other words, Marshall’s approach simply interprets the price formed in a market as the monetary counterpart of the various parties’ use values. Use value thus becomes the “reservation price”, or limit price, defining actors’ willingness to pay for a certain good.

The Marshallian reinterpretation of value in use anticipates (and will inspire) the “compensatory” approach that will be later formalised by Hicks (1939; 1956), which will later become the blueprint of contemporary demand theory in welfare economics following the ordinalist revolution of the 1930s. For the moment it will suffice to highlight the fact that whether this univocal mapping from exchange value to use value is possible depends on whether it is possible to express the former as a specific function of the latter.

3. The duality between use value and exchange value reinterpreted in function of the duality between the private and the social sphere

The two peculiarities of classical use value can be summarised as follows:

- The value-in-use of a material thing derives from the objective properties of the "thing", while it does not depend on the subjective identity of the actor employing and consuming it;
- The value-in-use of a material thing is incommensurable and independent from the relative value-in-exchange that could be generated employing it in exchange: the exogeneity of use value with respect to exchange value guarantees its independence.

The two points summarise the fundamental issues inherent in the value-in-use concept that induced classical economists to abandon the concept, and to place it outside the realm of political economy (together with consumption and utility theory).

Both points are closely related to the role of consumption, and to the meaning of "usefulness". In the same vain, both points stress the inadequate treatment that classical political economy reserves to use value, and to the role played by actors in deciding how the physical goods they possess ought to be employed. Yet, their essences, and their ultimate implications could not be more different. In fact, depending on the line of attack adopted, the role played by use value in the economic framework can either be greatly enhanced, or completely annihilated.

Ironically, although the neoclassical course in value theory emerged precisely as an attempt to confer to use value that analytical legitimacy that the concept could not have in classical political economy, the final outcome will be the complete opposite. The ultimate result of the marginalist revolution, far from being a rehabilitation of the classical interpretation of use value, will be to render use value irrelevant and to substantially take it out from the picture. This brings the neoclassical approach to value theory back to the same starting point where classical political economy originally left it.

The reasons behind this epistemological "short circuit" are found in the confusing treatment and interpretation of the two problems distinguished above. For this reason the aim of this section is to clearly understand what the two characteristics of classical value in use mean, how they should be interpreted, and, especially, what their overcoming implies for the theory of value and institutions. This will clarify some fundamental aspects before going on to examine the historical demi-tour that brought neoclassical value theory back to its starting point: to the irrelevance of use value.

3.1. The classical paradox of value derives from the dichotomy between the public and the private domain while subjectivity only denotes the nature of the institutional layers

The most important point to note with respect to the two aspects highlighted above is that they are two different characteristics of use value that can be addressed independently. The objectivity of use value is not a necessary condition for its exogeneity. The essence of the exogeneity condition lies in the fact that the inherent value that a person can derive from the employment of a good is independent and disconnected from the performances of the exchange mechanism and from its signals. Yet, this obviously does not impede value-in-use to be subjective. Nor does its exogeneity imply that the exogenous source of value-in-use must necessarily be the physical properties of the material goods. The obvious alternative is, for instance, the external cultural and social context.

More generally, the specific nature of use value (whether it is objective or subjective), is independent from its relationship with exchange value (whether it endogenously depends on it or not). In principle, one condition can be used to also address the other, but this does not need to be the case. For instance, there is no doubt that all early marginalists (Jevons, Menger, Walras) were genuine subjectivists, and that, among them, Jevons and Walras used their theory of subjective value in order to also resolve the classical paradox of value (even though only in a rough and preliminary way). Yet, this does not hold for Mengerian (and early Austrian) theory of subjective value, where the classical dualism in value is retained, and use value cannot be explained in function of, nor derived from, exchange value (Menger, 1981).[38] The same holds for all other possible combinations between the two dimensions.

For this reason, the roots of the classical paradox of value might be ambiguous as two dimensions play a role in it: the subjective-objective dimension and the endogenous-exogenous dimension. Moreover, a further source of ambiguity emerges when switching from the classical dualism in value to the marginalist monism in value: the meaning of the first dichotomy highlighted above (between subjectivism and objectivism) changes as we

[38] Differently from Walras and Jevons, the legacy of Menger has always attracted much attention and this is certainly due to the inherent classical exogenous approach towards use value. This topic will be developed at greater length in chapter 4 below. For studies on Menger, see Alter, 1990; Caldwell, 1990; Hicks and Weber, 1973.

switch from one paradigm (the classical) to another one (the neoclassical). This creates a second source of ambiguity.

More generally, the potential ambiguities emerging from the attempts to overcome the classical dualism in value have plagued the history of the economic discipline, and generated much confusion over the decades (and centuries). While the next chapter will address one of those in depth, it is now necessary to address these two sources of ambiguity.

First of all, it should be noted that the dichotomy between subjective and objective can be interpreted in two different ways depending on the way in which the concept "subjectivity" is interpreted.

On the one hand, in classical economics an economic action is private or subjective if it involves the personal (auto)-consumption of a resource for private purposes. Conversely, the act of exchange is social as it involves the participation of other people, as opposed to the act of consumption, which is inherently private.[39] Thus, in this case the adjective "subjective" means "belonging to the private sphere" dealing with one's personal use of a certain property, while at the same time "disentangled form the social sphere" of interpersonal exchange. Yet, the objectivity of use value in classical economics also belongs to the private sphere. Use value in classical economics is objective, yet it always arises from the auto-consumption of a certain property. This point has already been made in section 2.1.

On the other hand, "subjective" also implies "depending on the specific identity of the subject" rather than on the physical characteristics of the object. This is the essence of the new "subjectivist" approach of the first marginalist revolution, in direct opposition with the classical school. The private sphere in classical economics is shaped and constrained by the *objective* properties of the material thing (object) employed, rather than by the subjective identity of the acting subject, as extensively illustrated above.

The previous two interpretations highlight the common feature of the term "subjective" that links both interpretations: subjectivity is always associated with the private sphere of the acting individual employing his own property. No matter whether "subjective" is

[39] This is consistent with Ricardo's distinction between the wealth, which is always intended as private and derives from the employment of useful material things (*riches*, or value-in-use), and the value generated collectively in a society (see also O'Brien, 2004: 105). In the same fashion, Marx distinguishes between *utility* and *value*.

intended in the classical or in the marginalist way, the common ground of the two is represented by the verification that value is subjective whenever it belongs to the private, or personal, domain, and therefore whenever there exists an *exclusive* relationship between the property in question, and the individual with the faculty to employ and consume it. In this sense, the objectivist approach to value in use of classical economists was still a subjective one as it belonged to the private (personal) sphere of an individual.

Both approaches agree on one point already highlighted above: use value is always a subjective entity in the sense that it always characterises the private actions and decisions of an individual in their private domain, when these private actions never involve, nor interfere with, third parties lying outside the private sphere of the actor in question.

In sum, as highlighted in section 2.2, the paradox of value in classical economics is generated by the heterogeneous and independent nature of use value with respect to exchange value, not necessarily by its objectivism. With very few exceptions,[40] in classical economics, both use and exchange value are exogenous to each other *and objective*. This means that it is the mutual endogeneity of use value with respect to exchange value (and the possibility to univocally retrieve the former from the latter) that breaks-down the dual nature of value, rather than their respective natures. And what really distinguishes the two values in classical economics is the nature of the institutional layer the two types of value belong to: while value in exchange is a social value generated by *inter-personal employment* of a certain property, value in use is a private value generated by the *personal employment* of the property.

As a result, what guarantees the dualism in value in classical economics really is the dual nature of the institutional layers, and the (unsolved) dichotomy between the *private* and the *social* sphere, and not subjectivism *per se*. Subjectivism, as opposed to objectivism, defines the *nature* of value, and the type and nature of the institutional dimension from which value originates (thus, whether it depends on the materiality of the physical asset employed, or on the idiosyncratic identity of the actor employing it). In other words, subjectivism in use value ensures that the private identity of an actor does matter and affects the way in which he employs assets and resources at his disposal. But this alone

[40] Both labour-theorists and cost-of-production theorists share a pure objectivist approach to value-in-exchange deriving, respectively, from the conditions contributing to the subsistence of the labourer in the former, or from the characteristics of the technological production process for the latter. Say and Senior are the two notable exceptions, being the two classical economists close to a subjectivist theory of value (Bowley, 1972; deMarchi, 1972; O'Brien, 2004: 114-123; Schumpeter, 1954: 469-493 & 588-605; Sotiropoulos, 2009).

does not guarantee monism in value, and therefore the relevance of the institution of private property (asset ownership) vis-à-vis the institution of exchange. Conversely, what ensures that subjectivity in (use) value really remains meaningful is the impossibility to univocally retrieve the subjective dimension from the social dimension.

The two dichotomies highlighted above can therefore be reconceptualised according to the following taxonomy.

- The subjectivity-objectivity dichotomy defines the nature of the institutional layer value comes from: whether value depends on the specific identity/ies of the actor/s employing the property, or it is independent from it/them;
- The exogeneity-endogeneity dichotomy defines the logical, temporal and analytical causality among these institutional layers, and therefore it defines whether value and institutions shall be treated as a monism (mutual interdependence), or in dualistic terms.

It should be remarked that whenever use and exchange value exist independently, then the dichotomy between objectivity and subjectivity, intended as a dichotomy between, respectively, "something independent from the specific identity of the actor(s)", and "something dependent on the specific identity of the actor(s)" can apply to both exchange and use values.[41] For this reason, when dealing with the subjective-objective dichotomy the general expression "value" will be used, knowing that the mere fact that this distinction is meaningful already implies that value is a dualism.

The above discussion can be summarised as follows. On the one hand, the dichotomy between subjectivity and objectivity *of use value* identifies the nature and the locus of the institutional layer from which use value originates, given dualism in value and institutional layers. On the other hand, the dichotomy between exogeneity and endogeneity defines the logical, temporal and analytical causality between the various institutional layers and determines whether there is a monism (mutual endogeneity) or a dualism (exogeneity) in value and institutional layers.

Both subjectivism in use value and the independence of use value are needed in order to understand the economic role and nature of asset ownership independently from the

[41] The "s" denoting plural forms are there precisely to distinguish between subjectivity in use value, which only depends on the identity of the single specific actor, and subjectivity in exchange value, which derives from the inter-personal interaction of multiple identities operating within a social system.

performance of the market mechanism. On the one hand the subjectivity of value ensures that the private identity of the actor does matter in order to understand why and how a certain property has been employed in a certain way rather than in another way, while on the other hand the dualism in value (exogeneity) ensures that the market cannot be used as the necessary ultimate explanatory factor in order to understand, once again, why a certain property has been employed the way it has.

Once the key distinction between the specific nature of the institutional layers from which value originates is clarified, it is important to derive some interesting conclusions for both cases.

Two concluding remarks on the public-private dichotomy are in order at this point. Both points will become key when the nature of public control and regulation will be examined.

3.2. The relationship between private and public domain in case of dualism in value

First of all, figure 3.3 summarises the above considerations by showing that the paradox of value in classical economics can also be reinterpreted as a dichotomy between two independent institutional domains: the *private or subjective domain* of use value and the *public or social domain* of exchange value. While the former deals with the private self-employment of a property as means for auto-consumption, the latter deals with the employment of a property in the social sphere, thus a means for exchange and monetary rewards (wealth changes meaning between the two cases as well, as already remarked).

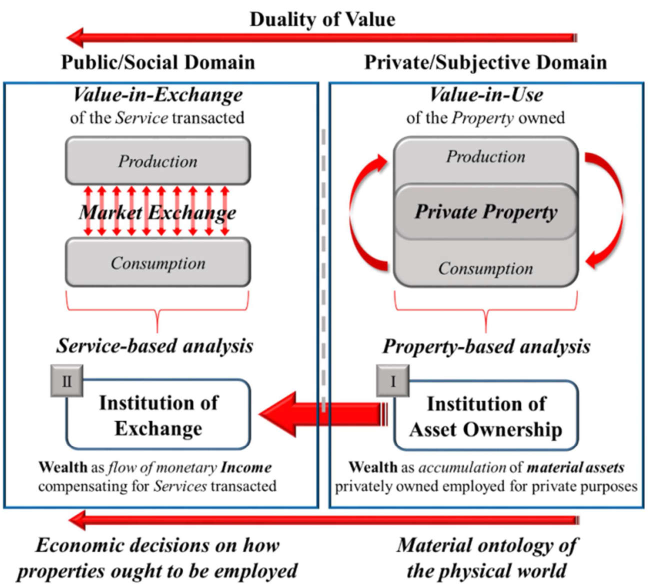

Figure 3. 3 - Dualism in value reinterpreted as a dichotomy between the subjective sphere of the individual, and the social sphere of public interaction

From figure 3.3 a very important point emerges: the “classical” causality from the institution of property (private auto-consumption) to the institution of exchange (social interaction), also implies a causality running *from the private domain to the public domain* and not vice versa.

This point deserves some further explanation. The fact that in classical economics the private sphere has priority over the social sphere should not necessarily be interpreted as a symptom of “methodological individualism”.

In order to better understand the essence of this causality it should be noted that the primacy of the private domain (characterised and shaped by the specific subjectivity of actors) over the social one simply reflects the primacy of the institution of private property over the institution of exchange. Moreover, this is also consistent with the key feature of classical economics, meaning that personal auto-consumption of the property has priority over the exchange process.

In general, in classical political economy, *social* institutions are always causally subordinated (from a logical and temporal point of view) to the institutions shaping and constraining the *private* sphere of the individuals and, more specifically, the preliminary private employment of available properties (“in-house”).

Moreover, this new type of causality also closely retraces the other typical causality of classical economics (represented in figure), namely the logical and temporal priority of the material object (the physical asset or resource) over its economic employment.

As a result, classical economics is characterised by the following logical chain.

First, physical things, such as assets and resources, exist independently from the actors and their employments (characterising the primacy of material ontology of the physical world). Second, given these material things, they are first appropriated, accumulated and held by the various actors (primacy of the private domain), with the preliminary and specific aim to first address some personal need and necessity (third: primacy of self-subsistence and auto-consumption). Fourth, and lastly, the properties can be employed in the social sphere, provided that the original private needs and requirements are addressed, and there is a “surplus” of (spare) productive capacity left.

In sum, the causal priority of the private sphere over the social sphere in classical economics is just the consistent outcome of the fact that in classical economics, once that a certain property is available, the employment of the property as means of self-subsistence in order to satisfy private needs has priority over the employment of the property as means of exchange *in order to derive monetary gains*.

This seems to be the way in which the primacy of the personal sphere over the social sphere should be interpreted in the classical framework: as the outcome of the primacy of the satisfaction of *personal (human) needs and necessities* over the accumulation of capital, based on the achievement of *monetary gains*.

In fact, it should be remembered that the private dimension is always "in use", or "in consumption", while exchange value is *never* subjective (private) and *always* social (public) by definition. Thus, any employment of a property in the social sphere is necessarily also an employment "for monetary profit" (this is the essence of the property used as means of exchange), while an employment of the property in the private sphere is necessarily an employment "for personal satisfaction", that has nothing to do with the monetary compensation that would always be generated through market transactions in the social sphere.[42] This is also consistent with the fact that wealth in classical economics is not regarded as a flow of monetary income, but rather as an intrinsic property deriving from the accumulation of *useful* physical "things" (which can be purposefully used in order to address private necessities).

In this sense, it is clear that the priority of the private domain over the social domain should not be necessarily regarded as symptomatic of methodological individualism, at least in the way in which the term is normally interpreted in neoclassical economics, thus linked to the maximization of *monetary* wealth.

The primacy of the private domain is a very important aspect. As chapter 6 will illustrate, it constitutes a key feature of the common law approach to regulation and to the concept of "public interest". Moreover, the above four-step procedure retraces very closely the steps undertaken to check the duties of a common carrier in common law.

3.3. The relationship between private and public domain in a monistic case: the role of subjective identities and time

Second, figure 3.4 shows the implications of monism in value theory for the private-social dichotomy.

[42] Note that this distinction between a "for-profit" use and a use for "personal satisfaction" is what makes the concept of usefulness (or utility) an incommensurable and a non-compensated one.

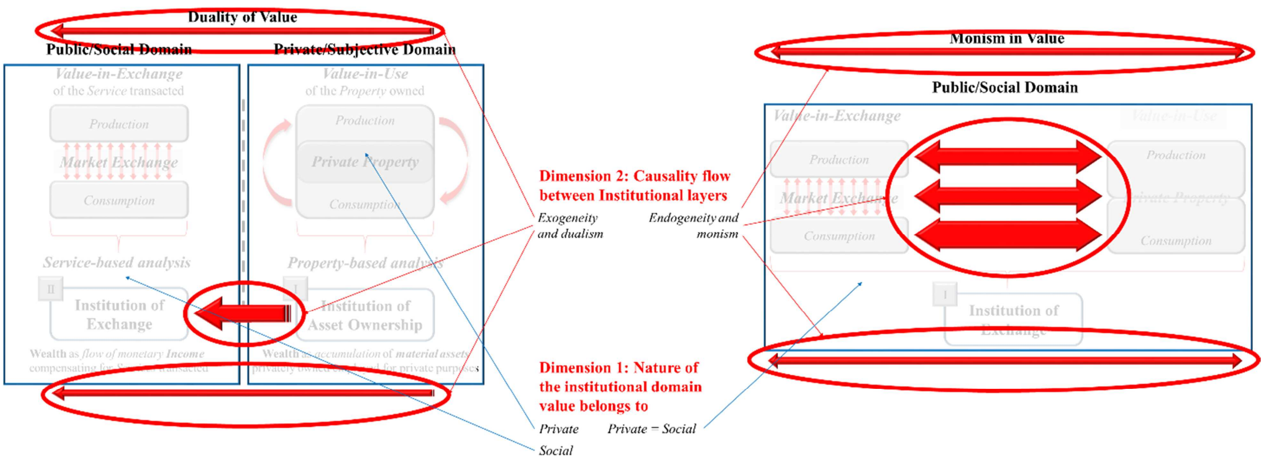

Figure 3. 4 - Institutional layers and causality flow between institutional layers in the case of dualism and monism in value

If the paradox of value is reinterpreted according to the private-public dichotomy, a very important implication follows: the two institutional domains (private and public) conflate and become indistinguishable. This means that the private and subjective sphere becomes public and, vice versa, the public and social sphere univocally informs and shapes the private and subjective sphere. Thus, in the case of mutual endogeneity between use and exchange value, what occurs in the private domain is also necessarily public, and what occurs in the public domain is also necessarily private. After all, this is the whole essence of a mutual endogenous relationship between two variables, where each one univocally informs the other one.

The reason why it is very informative to reframe the paradox of value by using the public-private dichotomy is that it clearly shows that whenever monism in value (and institutional layers) is guaranteed, there are no boundaries severing the private from the public domains. Later on in the work, this aspect will also be reinterpreted as a situation where an actor's private (subjective) identity can always be revealed, and substituted, by a public identity through the imposition of a social status. It is when the subjective identities of the asset owners become irrelevant, that the institution of private property and asset ownership become irrelevant in turn. This point will become fundamental in order to understand the essence of public control and regulation of private assets developed in chapters 6.

Moreover, there is a second important consideration to make with respect to the endogenous case.

Whenever there is mutual endogeneity between use and exchange value, it does not matter whether the causality flow with respect to value and institutional arrangements runs from the public domain to the private domain or vice versa. The mutual endogeneity ensures that the direction of the logical, analytical and temporal causality can be inverted at will with no loss of information accuracy, and with no distortion of the final (univocal) result. It does not matter whether value is generated in the private or in the public domains as, on the one hand, the two become conceptually identical and, on the other hand, time does not matter, meaning that it is not even possible to distinguish "before" from "after".

In other words, what disappears in a monistic world is the dimension of time, and its linked concepts, such as the one of causality (or path dependency). Overlooking the philosophical meaning and implication of this assertion, there is a very simple way to better understand why this is the case: in a monistic world, there are not two distinct layers anymore, but just one. As figure 3.4 shows, the two departments of value theory merge into one single department, and this is the whole essence of a monism. Yet, while it is possible to establish priorities and causality between two elements, it becomes impossible to establish priorities "between" one single element. It does not make sense to ask which institutional layer comes first, if there is only one.

It is important to summarise the important meanings and implications of these two features of a "monistic world".

First of all, whenever dualism in value collapses to a monism it is not possible to distinguish the subjectivities of actors, which means that the private identities disappear: all actors can be treated in the same way as they become indistinguishable from each other, as if they all are part of a "general public" with no distinctions nor exceptions. In other words, a world where value is a monism is a subjectivity-less world, and therefore an *identity-less world*.

Secondly, whenever dualism in value collapses to a monism, it is not possible to distinguish the *origin* of actors' decisions, choices and decisions. Monism in value can only be reached if the role of temporal and logical causation between the private and the public sphere is ruled out, so that each institutional layer can be at the same time the cause and the outcome of the other. In a situation where temporal causation does not play any role, the long-standing debate on whether actors' decisions should be considered as genuinely "private" or whether they are "socially-induced" (or, institutionally embedded) is simply meaningless.[43] A world where value is a monism is therefore a *causality-less world*.

[43] See for instance Swedberg (2005), Rutherford (1994: chapter 3), Hodgson, (1988), Granovetter, (1985), Burt, (1992)

In conclusion, a society where monism in valuations can be imposed becomes a sort of "institutional singularity" out of time and space: it is identity-less, and causality-less.

These conclusions are far to be new. The lack of subjective identities and time in neoclassical economics has been remarked on by a vast amount of scholars for many decades, if not centuries.[44] Yet, no-one to my knowledge seems to have clearly spelled out the precise analytical and logical mechanism that ensures this identity-less and causality-less situation holds. The reason here proposed is that whenever value becomes a monism, the private and the social spheres cease to be independent and separate institutional layers, but converge and collapse into each other, generating a world where what belongs to the public sphere also becomes private, and vice versa what belongs to the private sphere also becomes public. This is the most fundamental implication of monism in value, and it constitutes its deepest essence. This will become key when analysing the meaning of public regulation and its *normative* foundations in chapters 4.

For these reasons, from now on the dualism between private and public will be generally used to intend the dualism in value.

3.4. Some important remarks on the meanings of subjectivity and objectivity in light of the dichotomy between private and public

3.4.1. Clarifying the meaning of objectivity and its relationship with the social domain: are social magnitudes also "objective"?

One of the most important points highlighted in this section is that subjectivism in use value, alone, is useless and meaningless if there is no clear use value (as distinguished form exchange value) in the first place. Exchange value is by definition social, thus necessarily belongs to the social sphere, as all classical economists clearly understood. The rejection of use value in classical economics was justified precisely on the ground that "political economy has nothing to do with the comparative estimation of different uses in the judgement of a philosopher or of a moralist." Consumption remains a private act, and political economics can only deal with exchange for the simple reason that it can only deal with social matters.

Yet, the "social" character of exchange value is also supported by two other characteristics: the fact of being measurable and commensurable, which in turn also entail its "shared" nature. Social magnitudes are by definition *shared* among the various members of a

[44] For the role the reader is referred to the path-breaking work of Nelson and Winter (1982) and to the literature developing from there on path dependency and evolutionary economics.

society. Exchange value is shared in the sense that it is perceived equally by all members of a society: all actors share the same understanding of the magnitude of the exchange value. This is not the case for use value for the precise reason that it is subjective, thus it depends on the specific subject. Use value cannot be identically shared and perceived by various subjects.

This other way to interpret the dichotomy between private use value and social exchange value opens the door to a further interpretation of the dichotomy between subjectivity and objectivity, and to its relationship with the second dichotomy previously discussed, between monism and dualism in value theory. One possible reason why the two distinct dichotomies characterising classical value theory have been constantly conflated over the centuries might derive from the fact that, beside the concept of "subjectivity", even that of "objectivity" can be potentially interpreted in two slightly different ways.

The way in which the term "objectivity" has been used here in order to explain the classical approach to value theory is by highlighting that something is "objective" if it does not depend on the subject. In this case, the noun "subject" is singular, not plural. Thus, as the specific way in which a material thing ought to be employed in classical economics does not depend on the subjectivity of the specific actor actually employing the thing, but is intrinsic in the object itself, it was asserted that use value is "objective".

However, it is also possible to interpret the concept of objectivity in a slightly different way, substantially as a synonym of "social", or of "shared", "common", or "identical" to all subjects (plural).

A good whose employment and usefulness is independent from the specific identity of the actor employing it, and only depends on its exogenously-given material ontology (this is where exogeneity in use value derives from in classical economics) is necessarily also identically perceived and used by all actors in the same way. The same functionality or usefulness will apply to every single actor independently from the subjectivity of the specific actor.

Along these lines, any social, or shared magnitude can be regarded as "objective" in the sense that this magnitude is equally perceived by every single actor, and does not vary with the subjective identity of the person who is dealing with it (or perceiving it). In this sense, any price formed in a market is certainly objective. Yet, market price is not objective in the sense that it does not depend on the subjects (plural). Obviously a market price does depend on the various identities of the actors participating in the exchange: this is the basics of any elementary course in economics. In this second interpretation, market price is objective because it is a shared piece of information that is identical to every actor perceiving it. In this sense, value-in-exchange is objective *because it is social*. It is its social dimension that renders it objective.

Thus, both terms “subjective” and “objective” can be intended according to two slightly different interpretative approaches.

On the one hand something is subjective because it belongs to the private sphere of the subject, and not to the social domain (classical interpretation of subjectivity). At the same time, something can be regarded as depending on the subject employing the object, and not on the exclusive characteristics of the object employed (neoclassical critique to subjectivity). On the other hand, something can be objective because it only depends on the given ontological characteristics of the material object (classical interpretation of objectivity). At the same time something can be regarded as objective also because it belongs to the social domain, and therefore it is identically perceived by all actors.

This second interpretation of the “objectivity” concept in “social” terms is a bit narrow. Obviously not all social elements must be identically perceived, and therefore “objective”. But if an element can only exist in the social dimension and derives its existence precisely from its unambiguous meaning that is identically perceived and shared by all members of a society, then its social nature also implies its objective nature and vice versa.

These “intrinsically social” elements might derive their nature from their (objective) quantitative measurability and commensurability: to measure univocally is a precondition to comparison, and to compare is an act that generates “objectivity”.

Although this condition of quantitative “intrinsic sociality” might seem nebulous at first, *money* is a basic example: money is an “intrinsically social” entity (also) in virtue of its objectively quantifiable nature.

As a result, the two interpretations of “objectivity” actually conflate in the specific case of value-in-exchange or, in other words, price: a price is objective in the sense that it is “attached” to the specific object transacted: it is a characteristic that univocally defines the transacted object independently of the “subjectivities” of the various actors perceiving the object, purchasing, or using it. At the same time, a market price is by definition a social dimension as it can only make sense when “more than one person” interact: a transaction can never be defined in the private dimension, by definition (and this is precisely what generate the dualism in institutions between auto-consumption and market exchange).

As the topic of the present work is value, it is completely legitimate to use the adjective "objective" as a synonym of "social" and vice versa, provided this remains confined to the concept of price (value-in-exchange).

However the fact that it is possible to establish a univocal mapping between the institutional layer and the nature of the valuation process (objective vs. subjective) in the case of a very specific "commodity" such as money (in function of which value-in-exchange is expressed), does not carry over to the case of the overwhelming majority of objects and "things" that can very well belong to the private sphere even though their nature is an objective one. In general, whenever an object can be employed "in consumption", thus can belong to the private sphere, there is no meaningful univocal relationship between the institutional domain (private vs. social) and the nature and the source of the valuation process itself (subjective vs. objective).

The relationship between objectivity and subjectivity can therefore ne understood also in function of the ontology (the nature) of the object considered. This deserves some further considerations.

3.4.2. Monetary goods and Appropriable goods

Two types of goods can be identified based on their ontology. Goods that can only be employed in the social domain or, in exchange, and goods that can also be employed in-house, in consumption.

The first type of "intrinsically social" objects (goods, resources or assets) can be called "*monetary goods*", due to their similarity with the ontology of money. "Monetary goods" makes only sense in a monistic environment, and for this reason normative considerations apply to them. Monetary goods have the fundamental property of making ownership irrelevant, due to the fact that they are mono-dimensional (they can only be applied under monism in value and institutional domains). As subjective valuations do not apply to monetary goods, private identities become irrelevant as well (due to the disappearance of the private sphere).

As a result, it is possible to add monetary goods to the list of cases that generate the irrelevance of ownership (chapter 2: section 4). Whether monetary goods are a fourth, distinct case or just a particular manifestation of other cases is not clear. On the one hand, pure money income corresponds to the first case of no ownership: property-less actors can only rely on the monetary income that their services generate. On the other hand, monetary goods can also be regarded as the outcome of behavioural constraints on goods that would otherwise be employed "in-house" if their employment was not regulated. This second case would render monetary goods pure financial capital, and would make monetary goods a specific case of the third condition leading to the irrelevance of ownership: the presence of normative constraints on the employment of capital, supported by public regulation.

Secondly, here we define those things or objects that are not "intrinsically social", and for whom the univocal mapping between the two dimensions of classical value theory does not hold "appropriable".

"Appropriable objects" can be defined as those things that do not cease to be useful if employed in the private sphere on the part of the *single* actor. In other words, appropriable things can be defined as those things for whom dualism in value is valid and possible as they can generate value (in use) if employed in-house through auto-consumption in order to address the private needs of the actor employing them. Appropriable things are those that can potentially generate value (in use) for the actor employing them *independently from the social context*. All things that can be potentially used for personal self-subsistence are necessarily appropriable. In sum, appropriable things are simply those that are endowed with some degree of *usefulness* for the subject, intended in the classical way.

It should be noted that the concept of "appropriability" relies on the definition of ownership provided in chapter 2 above (section 1.3). As a result, it is possible to say that appropriable objects are those characterised by the fact that their private ownership is not only possible, but especially *meaningful* (as they can potentially decouple value and generate a source of value that can only come with private ownership: in use). For this reason, we should expect appropriable things to be the ones that, if privately owned, can generate the so-called "endowment effect" discussed by behavioural economics.[45]

As a result, monetary goods can never be employed as means for self-subsistence, and for this reason they necessarily support a situation of perfect division of labour, where

[45] The decoupling in the so-called "endowment effect" as described by Kahneman et al. (1990) and Kahneman et al. (1991) is between WTP and WTA. In this case, the subjectivity of value in use would be associated to the latter, while the objectivity of value in exchange is represented by the equivalent monetary compensation (WTP).

production and consumption are perfectly severed. Appropriable goods can be employed for private self-subsistence, and therefore might lead to a situation where production and consumption overlap, meaning that their ownership becomes economically meaningful.

In conclusion, two remarks are in order.

One, the distinction between monetary and appropriable goods reflects the long-standing confusion over the concept of capital in economics (Hodgson, 2014). Monetary goods are those that legitimise a pure financial interpretation of capital. In other words, even though monetary goods have a physical nature, they can be simply be considered as mere financial capital. Appropriable goods are resources and assets properly said, thus they can be considered as material (stocked) capital. The former can only be exchanged on the market, the second can be accumulated and stocked.

Two, the two types of goods also reflect the aforementioned distinction in the way in which wealth is treated. Monetary goods can be treated in the same way as monetary (financial) wealth, thus as mere flows of monetary income. This is the neoclassical definition of wealth that, not surprisingly, only deals with financial capital or, monetary goods. Appropriable goods can be treated as material wealth in the form of physical stocks of material things. These physical things, in virtue of their materiality, can also be employed in-house and used for self-subsistence purposes, conversely from financial capital that has to be necessarily spent, thus transacted. This is the classical definition of wealth, or riches in Ricardian terminology, that is defined in function of the level stocked by private actors for their personal purposes. Figure 3.5 summarises the discussion above.

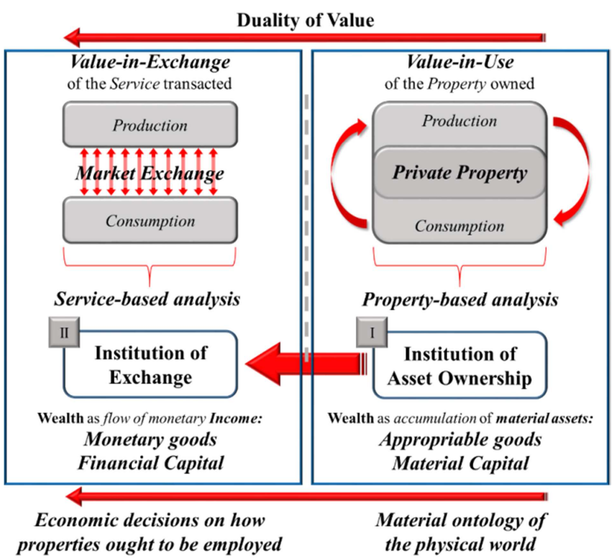

Figure 3. 5 - Monetary goods and Appropriable goods in the perspectives and implications for the meaning of capital and wealth

3.5. Wrapping up: final considerations on the relationship between public and private domains and subjective and objective valuations

In conclusion, it is certainly possible that the excessive focus on the means through which exchanges occur (money) has led to the constant confusing conflation of the two peculiarities (or dimensions) of classical value theory: its objectivity and its exogeneity (i.e. dualism). It is true that for the "intrinsically social" and objectively quantifiable object "money", it is indeed possible to overcome the distinctions between the two dimensions and to operate a useful conflation between the nature of the object (objectivity) and its institutional domain (social).

However this conflation is not possible nor legitimate whenever appropriable things are involved. In the case of appropriable things, the *nature of the valuation process* (whether the origin of value is objective or subjective) and the *institutional domain* it is operated in (if it is used socially as means of exchange or privately as means for auto-consumption), do not overlap and cannot be conflated into a unique dimension.

The fact that a thing can be potentially appropriable, thus it can be brought under the private domain, does not mean that it should or has to be. And this is completely independent from the nature of the valuation process associated to it: objective valuation processes can still be enacted "in use", and therefore be generated in auto-consumption. After all, this was precisely the case of use value in classical economics: *objectivity* and *privacy*. This is obviously even truer in the case of subjective valuations: the fact that a good, asset or resource is subject to subjective valuations (meaning that the potential *use* value that can be derived from it if it was employed "in house" depends on the specific identity of the actor actually employing them) does not mean that it necessarily has to belong to the private sphere and be employed in-house. This is precisely the case of contemporary (ordinal) neoclassical economics: *subjectivity* in valuations coupled with institutional *monism.* In fact, the possibility to guarantee monism in value and institutions is precisely what renders the subjectivity in valuations irrelevant in neoclassical economics, as discussed above and in the following chapter.

As a result, for appropriable goods, resources or assets, what has been declared above is still valid, namely that subjectivity in value, alone, is uninformative on the nature of the valuation process if it is not also supported by exogeneity (dualism) in value. As mentioned above, the subjective dimension of (use) value makes sense only if the private institutional layer can be preliminarily clearly distinguished by (and not endogenously retrieved from) the social institutional layer. Subjective valuations become meaningless and irrelevant if occurring in a context characterised by monism in value and institutional layers.

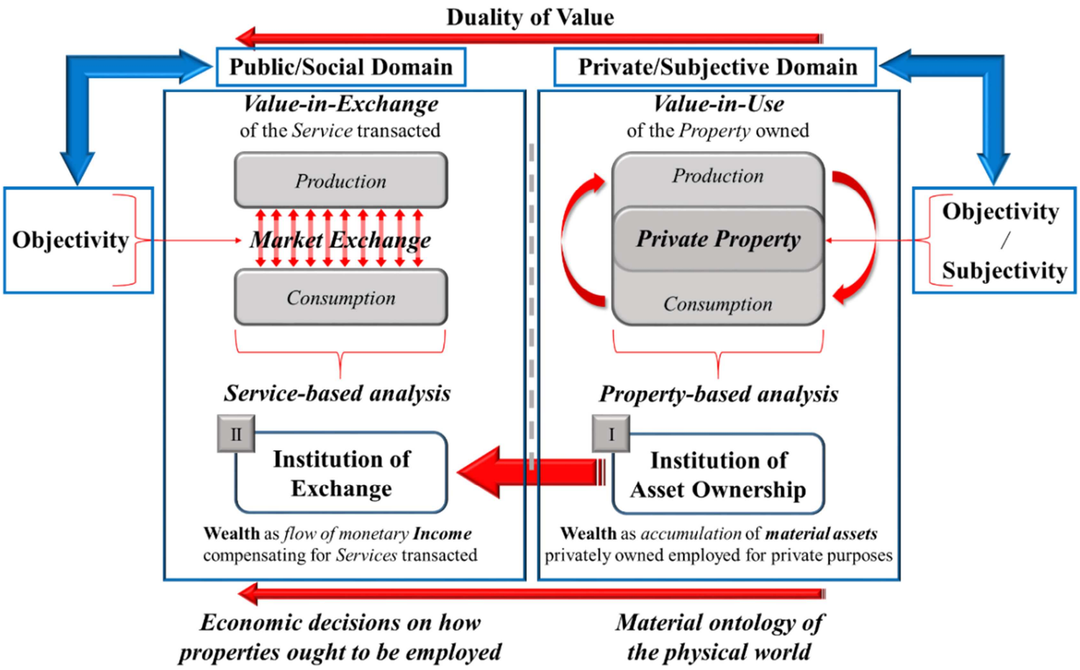

Figure 3. 6 – The two dimensions in perspective: private-social vs. objective-subjective

Figure 3.6 makes explicit the fact that with respect to value-in-exchange (supported by the transaction of money), the social dimension and objectivity are indeed the same thing and can be treated as synonyms, and this is also how they will be treated in the remainder of the work: price will be indifferently characterised as "objective" or "social".

4. Concluding remarks: a general taxonomy for the study of institutions through value theory

Figure 3.7 provides a summary of the possible situations in function of the two dimensions that constituted the two original difficulties of classical value theory already identified by economists in the early 19th century. The figure can be interpreted as a taxonomy of the possible approaches towards value, and therefore institutions.

From the figure another interesting fact can also be appreciated: while we have a theory of (i) objective and private value (classical economics), of (ii) subjective and monistic (social) value (neoclassical economics), and of (iii) objective and monistic value (monetary theory), a theory that treats value, at the same time as (iv) subjective and private, is missing.

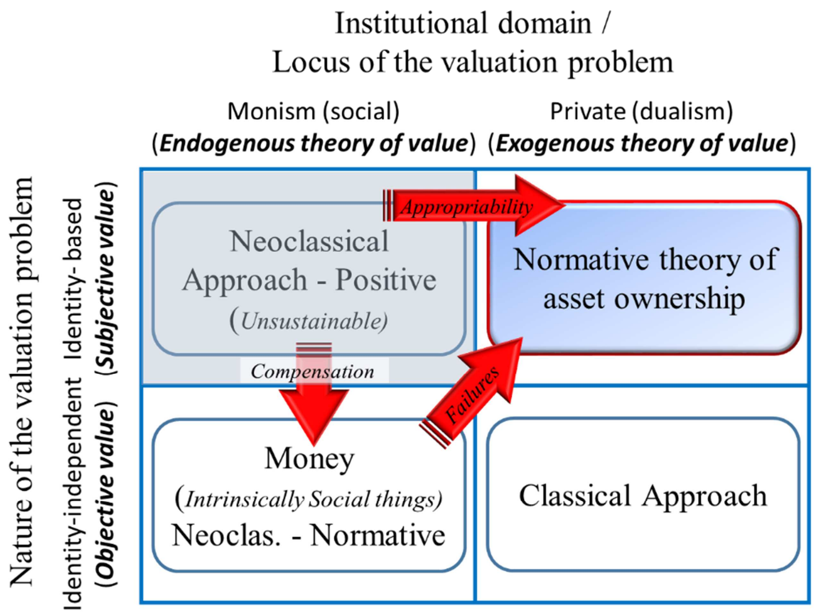

Figure 3.7 – A taxonomy of approaches to value and institutions in function of the two dimensions of value: objective-subjective and social-private

Some considerations are in order at this point. Many of these considerations summarise and anticipate the in-depth discussion that will be developed in the next chapter. In conclusion to the chapter, it is nevertheless important to provide an overview of the key themes.

First of all, from figure 3.7 an ironic fact emerges: in the unfortunate case in which the two dimensions are conflated, univocally linking subjectivity with the private domain and objectivity with the social domain, classical value theory and neoclassical value theory both become incomprehensible. Yet, while the case of classical economics is perfectly legitimate, the case of neoclassical economics is only "analytically" legitimate and deserves some further considerations (chapter 4 below).

Secondly, the figure highlights that the top-left corner is "unsustainable". The reason for this is that there can be no subjectivity in a monistic environment. As discussed in section 3 above, subjectivism in value, and monism in institutional domains is meaningless as in a monistic environment there is no distinction whatever between institutional domains. Value must necessarily always become a unique, objective, thus shared entity whenever it is developed internally to the system. For this reason, the specific top-left spot of figure 3.7 has to be abandoned, which leads us to the third point.

Thirdly, the standard neoclassical framework (Walrasian) frames the economic problem in subjectivist and endogenous terms (see chapter 4). From a mere descriptive point of view, taking the model as a mere "snapshot" of a social system, this can work. However, only a positive approach to the neoclassical treatment can lead to the unsustainable situation of subjective (identity-dependent) valuations and monism in value and institutional domain. In order to understand how valuations shape actors' decisions, and therefore in order to understand how the institutional context operates in function of actors' valuations, that specific spot has to be abandoned, and it is necessary to move either downward, or right-ward. The nature of this move depends on the nature of the object (asset, resource, or good) in question, and/or on the way in which an actor can employ the object.

On the one hand, the shift from the subjective-endogenous case to the objective-endogenous case defines the case of contemporary (normative) microeconomics, and Paretian (ordinal) welfare economics. This shift towards objectivity in valuations is ensured as long as the object in question becomes a "monetary one" or, as long as the actor can only employ the object socially, always treating it as means of exchange, and never in consumption (means for auto-consumption). After all, this is the essence of a situation where subjective valuations disappear, as they can only make sense in auto-consumption, in the private domain.

This shift in neoclassical economics is guaranteed by the normative "compensation" (or duality) assumptions defined in Paretian welfare economics. These assumptions are also usually (improperly) referred to as the "rationality" assumption. The duality, or rationality, assumptions in ordinalist neoclassical economics ensure precisely that any material things (resources or assets) are only employed in the form of *services* for third parties against monetary compensation, thus in exchange, *as if it was pure money*. This ensures the consistency of the value framework inherent in the normative neoclassical paradigm, and this is the meaning of the top-down arrow labelled "compensation".

On the other hand, the left-right arrow represents the situation of "appropriable goods". In contrast to "monetary goods", "appropriable goods" can be privately stoked in order to be also employed in auto-consumption, in order to address personal necessities independently from the social context. The mere option of self-employment of an object automatically generates a decoupling in the institutional environment between the private and the social domain. This also opens the door to subjective valuations in consumption (value-in-use) next to, and independent from, the valuations in exchange (value-in-exchange). The two aspects of value then become exogenously given to each other. And this is the essence of the shift rightwards, towards the right-column representing an "exogenous theory of value".

Fourthly, the figure shows a third arrow, from the bottom-left to the top-right spot. This is the arrow showing the effect of failures (market or behavioural). Failures in neoclassical *normative* (welfare) economics are precisely those devices introduced to explain the emergence of a decoupling in valuations and institutional domains from the (artificial and "super-imposed") condition of social monism and objectivity in valuations and institutions (bottom-left), to a condition where the private domain formulates *subjective* valuations that are *independent* from, and *exogenous* to, the valuation generated by the market (pricing) mechanism (top-right). It is for this reason that the presence of failures becomes the necessary condition in normative micro and welfare economics in order to explain and justify the presence of extra-market institutions characterised by subjective valuations formulated in the private domain (such as private ownership, firms, and so forth), all located in the top-right corner of figure 3.7.

The standard way to justify (and rationalise) this back-ward arrow bringing the analysis back to the "initial position" is twofold: a violation, or a break-down, of the rationality (duality or compensatory) assumption, or a break-down of the smooth and frictionless allocative process. In truth, chapter 4 will argue that this is an unconvincing way to explain the situation of subjective and exogenous valuations. A much more faithful explanation explains this "feedback" by looking at how (1) the ontology of the objects is treated, and how (2) the relationship between these objects and the subjects is modelled. The key distinction will be between *services* (rendered to the public in exchange for monetary wealth) and *physical assets or resources* (accumulated privately for personal purposes and necessities). It is possible to see that the two fundamentally differ for their different institutional domain: services are always defined in the social domain, assets (if privately owned and stocked) are defined in the private domain (that is, granted private ownership of things).

In conclusion, figure 3.7 provides a very clear understanding of how institutions are conceived, modelled, and explained in function of the way in which value is intended and treated. The figure shows that *there are two alternative ways in which subjective valuations in the private domain can be interpreted*.

First, the failure-based approach is a three-stage step that operates the following conceptual steps:

1. It starts from the (unsustainable and meaningless) top-left corner of subjective and endogenous valuations;

2. It moves downward to the bottom-left corner of objective (social) and endogenous (unique) valuations through the imposition of clear restrictions on the nature of the object and on the behaviour of the subjects with respect to the way in which the latter ought to behave when employing the object;
3. It retrieves institutions and valuations emerging in the private domain (top-right corner) by introducing failures either in the way in which the object can be employed (market or allocative failures), or in the way in which subjects behave (at odds with the way in which they are supposed to, or ought to, behave: behavioural or rationality failures). Both are violations of the conditions guaranteeing the normativity of the framework.

If figure 3.7 and figure 3.3 are overlapped (as in figure 3.8), it is possible to appreciate the path that leads from a situation of dualism and subjectivity in value (top-right corner) to the new situation of monism and objectivity in value theory (bottom-left corner).

The usual way to relabel these two features of value (monism and objectivity) is by stating the well-known assertion that whenever there are no allocative nor behavioural failures, then the social allocative mechanism provides a *unique* set of prices for all individuals in the *social* system (social objectivity of value-in-exchange), which are also *socially efficient* (endogeneity and monism of value-in-exchange). This point will be developed and explained in the next chapter.

From figure 3.8 t is possible to see that when value theory has the characteristics defined by the bottom-left corner, the usual dualism of value collapses, and the univocal causality (from the ontology of factors to their employment) disappears. This is the peculiarity of an endogenous framework where time, causality, and priorities collapse (section 3.3 above; figure 3.2).

The presence of failures then has two implications:

1. It re-establishes the subjectivity in valuations (upward shift)
2. It re-establishes dualism in valuations, and therefore the univocal flow of time, causality and priorities (rightward shift).

The second approach will be discussed in the following chapter, together with its comparison with the failure-based approach. As the chapter will show, in this second "new-classical approach" subjectivity and dualism in value can simply be explained by not forcing the ontology of the material world into ad hoc assumptions, also constraining the way in which objects and subjects have to interact. Subjective and exogenous valuations can be accounted for by simply allowing for (1) the private appropriation and accumulation of proper physical assets (material wealth), and for (2) their employment in auto-consumption “in use” as a means for “self-subsistence” by the part of the various actors. This is the case where asset ownership (and institutions) can be explained by simply merging the ownership and employment of assets with their (auto)-consumption, with no resort to extra-conditions or assumptions such as the ones of failures or rationality.

In conclusion, as figure 3.8 shows, the key difference between the two cases derives from the fact that (consistent with the discussion in chapter 2 above), in the failure case the ownership and (productive) employment of assets (i.e. production) is sharply severed from the case of the consumption of the services rendered by those same assets (i.e. consumption), and *this allows for a univocal retrieval of the underlying private material wealth (asset ownership) by merely looking at the way in which services are employed, transacted, and valued (priced)*. On the other hand, in the second (new-classical) case the employment of owned assets for the auto-consumption of its services is allowed by simply admitting the presence of appropriable goods, and their contextual self-employment in auto-consumption.

The fundamental implication of this second situation is that there is a critical decoupling between the services transacted and valued on the market and the actual employment of the underlying material assets and resources. With self-consumption, the monetary wealth corresponding to the reward (compensation) of services does not correspond to the underlying material wealth accumulated and possessed by individuals privately. This is what characterises a non-failure based theory of asset ownership, and this is precisely what is ruled out by all three (or four) situations guaranteeing monism described in chapter 2 above and represented in figure 3.8 (no ownership, ownership of service-specific assets, social restrictions on the private employment of assets, together with ownership of pure monetary resources).

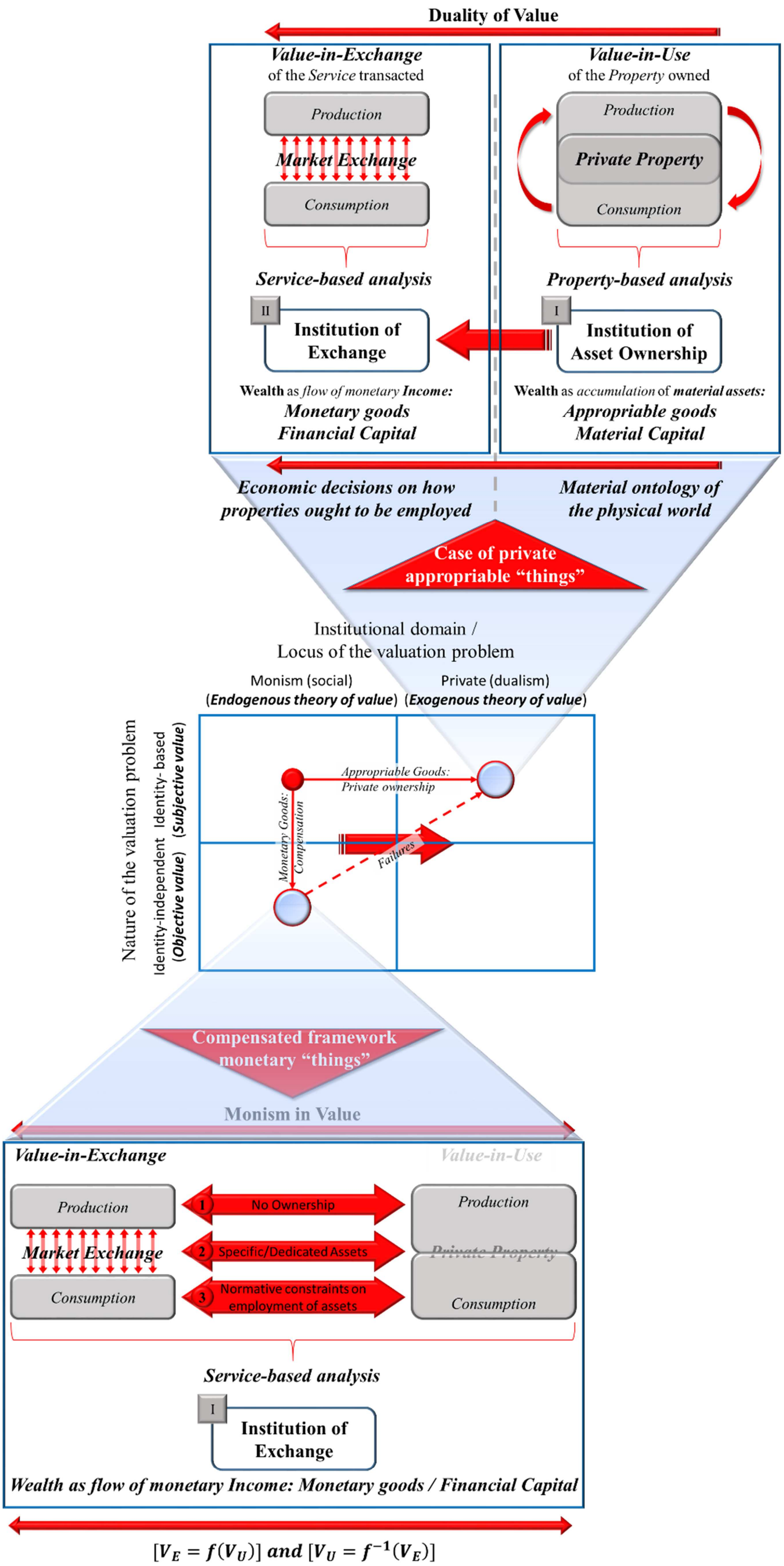

Figure 3. 8 – An overview of the conceptual path leading from the case of subjective and dual valuations to the case of objective and unique valuations and back

CHAPTER 4

A critical reconceptualization of the conditions ensuring the (neoclassical) monistic and objective theory of value: the switch from assets to services and why a "failure approach" to ownership is inconsistent

Chapter synopsis

The present chapter relies on the intuitions developed in the previous two chapters, while also providing a final synthesis and justification for many of the claims developed thus far.

The chapter concludes the historical overview of the history of (use) value theory by showing how marginalism eventually managed to achieve the "holy grail" of any normative approach to social issues: monism and objectivity of value judgments. Whenever value becomes unique (monism) and social (objectively shared by everyone) normative value judgments on how resources ought to be employed within a social system become legitimate, irrespectively of the nature of these resources, and of their relationship with the members of the society. This is the essence of the (perfectly competitive) market mechanism in the contemporary neoclassical framework: it endogenously generates a unique set of prices (monism in value), which are also defined at the social level, thus identically shared and perceived by all actors (objectivity).

In order to achieve this path-breaking result, the duality in value inherent in the classical paradox of value had to be solved once for all. Although the marginalist revolution in value theory is sometimes regarded as the fundamental event in the history of the economic discipline that achieved this goal, this is not precise, nor correct. On the one hand, neither the English, nor the Austrian schools have ever developed a theory of objective and endogenously computed value. The Walrasian framework itself did not provide a theory of value that could legitimise normative claims. The real and definitive solution to the three-century old paradox of value could only start with the marginalist revolution in production and distribution theory (mid-1890s), and was concluded by the mid of the 20th century.

Both these events profoundly modified some fundamental aspects of the original (classical) theory of value.

As the new aspects will precisely be those that will justify a failure-based approach to institutions, the aim of the chapter will be to clearly understand through an in-depth historical enquiry what these new conditions mean, and what they imply.

More precisely, the chapter provides an extensive historical overview of the various attempts that have been developed, from the early marginalist revolution to the final definition of the neoclassical framework in value theory, in order to address the outstanding issues embedded in classical value theory. The reason why it does so derives from the conviction that the contemporary neoclassical approach to institutions (based on the concept of failures) has overlooked the real essence and nature of some fundamental conceptual steps that have characterised the new, neoclassical approach to the theory of value. If certain assumptions and conditions are not placed in historical perspective, their essence might be misunderstood, or completely overlooked.

The main message of the chapter can be summarised as follows: the only reason why the (Walrasian) vNAD framework can describe an economic system founded on a monistic and objective (social) theory of value derives from the fact that some fundamental conditions are imposed with respect to (i) the ontology of the unit of analysis of the framework, and (ii) the way in which asset owners can employ their property.

More precisely, the chapter shows that the two conditions that ensure that the neoclassical framework is characterised by monism and objectivity in valuations are the following.

The first condition establishes that in order to achieve monistic and objective valuations through the exchange mechanism, only services can enter the economic system and never material “things”, which must remain strictly outside the domain of the analysis. This ontological condition represents a change of the unit of analysis of the economic system from the material welfare to the immaterial (monetary) welfare, and it ensures that the materiality of the factors is severed from the immateriality of the services provided. A clear decoupling between the physical ontology of the material support (assets and resources) and the immaterial employment of these things is introduced. The two must remain severed and can never be confused, or overlap in the analysis.

The second condition complements the former and establishes that all those actors participating to the economic system who own physical “things” (whoever is in private possession of material welfare) must respect a strict behavioural constraint norming the way in which their properties ought to be treated and employed. This behavioural assumption represent a social constraint on the way in which asset owners can employ their property and ensures that any actor will only employ his own material wealth in order to generate financial wealth by trading services with third parties (the public), and will never employ its material wealth for himself, thus in auto-consumption.

Given these two conditions, the classical paradox of value is disposed of, and value becomes an endogenously computed monism.

The first ontological condition has a descriptive role: it ensures that social value-in-exchange is indeed the only form of value, and that it can be endogenously generated by the exchange mechanism. The second behavioural condition completes the picture and has a normative role: it ensures that the (objective and monistic) valuation of the *social financial wealth* generated by the services transacted on the market univocally characterises the underlying *private material wealth* from which those same services derived. As a result, the behavioural condition ensures that it is always possible and legitimate to undertake a service-based analysis of value and wealth as the latter univocally represents a faithful reflection of the underlying material wealth. While the former is intrinsically social (it is generated through public transactions), the latter is intrinsically private (it belong to the private sphere of the various actors and characterises private ownership).

The historical inquiry of the steps characterising the development of the marginalist theory of value becomes fundamental in order to appreciate the nature of these two conditions and their respective implications.

These two conditions are usually referred to in the academic literature in two other (much less informative, and often misleading) ways. The mathematical literature defines the ontological assumption as a “convexity” assumption, and the “behavioural assumption” as the “monotonicity”, or “non-satiation” assumption. On the other hand, the organizational and institutional literature defines the ontological assumption as the “perfect market” condition and its violation a “market failure” condition, while it defines the behavioural assumption as the “rationality” assumption.

These are the conditions on which contemporary normative theories of institutions rely.

The chapter provides an alternative explanation to the essence of the neoclassical (failure-based) approach to institutions: the properties of the framework rely on precise normative prescriptions on the nature of the ontology of the objects allowed in the social system, and on the nature of the interaction that the subject is allowed to establish with the object.

A brief reconsideration of these conditions will reveal that their ultimate role is simply to rule out the real source of dualism and subjectivity in valuations: private self-employment of material property in auto-consumption.

A careful look at these conditions will reveal that they are genuinely normative, and regulatory in character. As a result, the chapter reaches a fundamental conclusion: the standard neoclassical claim that explains extra-market institutions through the presence of (behavioural and allocative) failures, asserting that legal institutions (such as asset ownership) become irrelevant whenever the market mechanism works perfectly is unfunded and illogical as it rests on circular reasoning.

The present chapter shows that the market mechanism can only guarantee monism and objectivity in valuations under specific conditions that can only be ensured and guaranteed if some legal regulatory framework is already present in the first place.

Independently form the performance of the market, some form of regulation of actors' behaviours has to be necessarily always present in order to guarantee monism and objectivity in valuations. For this reason, the chapter argues that the legal regulatory framework has always temporal, logical and causal priority over the market institution. The reverse causality is always logically unsustainable, meaning that *markets and the price mechanism can never be normative in a legal and regulatory vacuum whenever physical assets (material propriety) can be privately appropriated*.

In conclusion, some pre-existent regulatory setting must always be devised in order to reach social objectivity and monism in valuations whenever the private appropriation of material things can guarantee subjective usefulness in auto-consumption to its owner. This can be taken as the fundamental take-way of the chapter.

The most striking implication of the framework consist on the fact that the neoclassical economic model renders private ownership of assets irrelevant for the simple reason that the most fundamental property characterising ownership is forcefully ruled out: the refusal to deal. The second behavioural condition rules out self-subsistence and auto-consumption of private property, and deemed to be "irrational" behaviour.

The essence of rationality, which applies only to asset owners and never to property-less actors, is then clearly understood: any property owner acts “rationally” whenever he treats its material wealth *as if* it was pure financial wealth to be completely monetise through the sale of its useful services to third parties on the market, upon monetary compensation of those services. If this is the case, then it is possible to derive normative value judgments on how the underlying “material support” ought to be employed by solely looking at the services valued (priced) by the market. And this value will have normative force as it will be objective and monistic (endogenously calculated by the system).

For this reason the chapter concludes that, because of the two constraints on the ontology of the objects, and on the behaviour of the subjects with respect to the employment of their private property, the (Walrasian) neoclassical framework does not describe the mechanism of a perfectly working decentralised system characterised by private property of physical things and objects, but it rather describes the functioning of a social system with perfect public regulation of private property.

The chapter then anticipates the topic of the next here chapters.

On the one hand, the failure of contemporary institutional economics to recognise and take into account the difference between ownership over assets and use rights over services (Chapter 5), and the failure of contemporary antitrust and regulatory authorities to recognise and appreciate the genuine parallelism between the two public duties super-imposed on a public calling (common carriage) in common law regulation, and the two conditions super-imposed on any property owner in the vNAD framework.

The chapter is structured as follows.

Section one provides a recap of the argument developed in the previous chapter.

Section two provides a detailed historical overview of the early marginalist approaches to the topic of value, and discusses their respective solutions to the classical paradox of value. The chapter shows that, if the peculiarities of the English school, the Austrian school and the Lausanne school are studied through the two-dimensional framework of value provided in the present work, they become just alternative and consistent ways to interpret the nature of value according to the two dimensions of (exogeneity-endogeneity) and (subjectivity-objectivity).

Section three is the core of the chapter and develops the main argument. It shows that the key aspect of the new marginalist framework in order to achieve monism in value while retaining subjectivity was to change the unit of analysis from the classical materiality of assets to immaterial services. The section provides an extensive discussion of the meaning and the implications of the new conditions ensuring monism in value.

1. Resumption of the argument and preliminary overview

1.1. The two open issues in the classical paradox of value: alternative ways to address objectivism and exogeneity in classical use value

The previous chapter showed that the agenda of economists since the mid-19th century was only in appearance a consistent and homogenous one. Although the two limitations inherent in classical value theory were both deriving from the unclear and unsatisfactory role given to use value by the part of classical political economy, the respective meanings and implications of the two lines of attack were profoundly different. Certainly, consumption had to play a role in political economy in general, and value theory in particular. Similarly, the twin concepts of usefulness and utility had to perform a much more fundamental role in the overall framework. However, the way in which these resolutions had to be tackled could lead to two very different results.

The two lines of attack, corresponding to the two outstanding problems in classical value theory, can be summarised as follows.

On the one hand, it is possible to solve the paradox of value by annihilating use value, while treating it just as an aspect of exchange value. In this case, it is the "exogeneity" of classical value theory that is attacked, together with the duality of value. This approach would solve the dualism in values by interpreting the private domain as a mere by-product (a particular manifestation) of the social domain. The paradox is solved by "pushing" the hidden private domain of consumption (usefulness and use value) towards the social domain of exchange and market transactions, so to *reveal the former in function of the later*. In this case use value is explained, defined and valorised through exchange value.

On the other hand, it is possible to give relevance to the private dimension of consumption by making it the explanatory foundations of exchange value. In this case it is the subjectivity of use value that is valorised, and it is its objective interpretation that is attacked. In this case, the dualism is solved as use value is interpreted as the ultimate cause of exchange value, while the social domain is now interpreted as the by-product of the private domain. In this case it is ensured that economic outcomes have a genuine subjective, rather than materialistic, origin, and this means that they the valuations formulated in the market exclusively depend on the private identities of the actors rather than on the physical characteristics of the material objects (or of the technological production process).

The two approaches differ for their way in which they interpret the causality between the private domain of consumption, usefulness and assets' use, and the social domain of exchange, market transactions and monetary compensations. However, while the first approach directly tackles the exogeneity problem, generating a new endogenous relationship between use and exchange value, the second one does not need to. The subjectivity of use value, per se, does not need to also imply endogeneity between the two value dimensions, provided use value remains determined by exogenous factors.

A simple way to appreciate the essence of the two alternative lines of attack is by considering the well-known paradox of value as originally portrayed by Adam Smith. It is convenient to quote at length to appreciate the fact that the two dimensions are really treated as two independent (exogenous) dimensions:

> "The word value, it is to be observed, has two different meanings, and sometimes expresses the utility of some particular object, and sometimes the power of purchasing other goods which the possession of that object conveys. The one may be called 'value in use'; the other, 'value in exchange.' The things which have the greatest value in use have frequently little or no value in exchange; and, on the contrary, those which have the greatest value in exchange have frequently little or no value in use. Nothing is more useful than water: but it will purchase scarce anything; scarce anything can be had in exchange for it. A diamond, on the contrary, has scarce any value in use; but a very great quantity of other goods may frequently be had in exchange for it."

The first line of attack to the water-diamond paradox would try to dismantle the dualism in value by reducing the two to one single entity so to explain every type of valuation through exchange value. The well-known answer in this respect is represented by the concept of *social* scarcity of the "element" water, or the "element" carbon.[46] As summarised by Wicksell (1934: 18; emphasis in the original):

---

[46] So for instance this is how Hicks (1934: 338) opens his biographic article on Walras: "Leon Walras was the son of an economist. His father, Auguste Walras, was one of those excellent people (they seem to have existed since very near the dawn of history) who taught the true but unhelpful doctrine that value depends on scarcity (*rareté*)".

> "Since, as was assumed, utilities and exchange values did not always coincide, but frequently diverged, exchange value must either depend on something entirely different from utility, or upon utility and something else as well. The latter explanation was generally accepted … The result was the concept of relative *scarcity*: in order to have exchange value an object must, it was said, necessarily be useful, but in addition, it must exist in *limited quantities*."

This explanation relies on the approach already adopted by Ricardo intended value (as a combination of usefulness and something else: either scarcity or quantity of labour), with the difference that Ricardo explicitly excluded scarcity goods from the realm of economics and only kept labour goods (see chapter 2). Moreover this explanation has already the germs of the standard "two-blade scissor" where a variable in consumption (use value or utility) is intersected with a variable in production (scarcity).

This approach has a fundamental shortcoming at the eyes of subjectivists: there is an inconsistency or, better, a conflation, of institutional layers: exchange-value is made a function of a variable defined in the private sphere as a function of ownership (utility in consumption) and of a variable inherently social, which has nothing to do with private appropriation: scarcity. It is the *social* quantity of a certain *substance* available on hearth that is at stake in this approach, not the actual number of assets, goods, or proper resources owned and employed in consumption by the specific actor. Whether this is a legitimate logical jump, or an inconsistency will be addressed in the reminder of the chapter. Yet, the important point is that the introduction of the concept of "scarcity" already puts the germs of that monism whose result is to conflate the private with the social sphere.

The subjectivist line of attack would similarly find Smith's description flawed and unsustainable, but not for the same reasons put forward by those supporting monism in valuations. A subjectivist critique would just disregard the issue of scarcity and, especially, would never expand the private domain of private ownership, fundamental in the classical description of use value, to the social sphere (the institutional locus where scarcity is defined). The subjectivist critique would simply object that the assertion that "[n]othing is more useful than water [while] … diamond, on the contrary, has scarce any value in use" is unwarranted as it is not Adam Smith that decides and judges the usefulness of a thing for every single man on hearth. In other words, a critique to the paradox of value along subjectivist lines would object that the entire logical construction on which the paradox of value rests is funded on illegitimate assertions concerning the nature of (subjectivities) usefulness.

The reason why Smith was making those kind of assertions almost certainly derived from the fact that, consistent with his overall approach, usefulness “in use” was computed in function of an actor’s primary and unavoidable needs for his own bare survival (the basic level of self-subsistence): water is necessary to live, diamonds are not. Still, this does not excludes the fact that the same “element” can be regarded and employed differently (thus have a different intrinsic value-in-use) for different actors facing different necessities, operating in different (geographic) environments, and experiencing different feelings. What can be a necessity for survival in some contexts might very well be superfluous in others for climatic, geological or environmental reasons in general.

The subjectivist critique to classical value theory would therefore argue that Smith’s description of the paradox of value has to be rejected based on the fact that, differently from exchange-value, usefulness (thus value in use) is not an objective magnitude, but rather a subjective one. Moreover, it is not even quantifiable or measurable, nor comparable.

As a result, a subjectivist would have concluded that to define the distinction of use and exchange value through objective rankings of use values, and to objectively compare the use value or an element with its exchange value, across various elements and across various subjects as Adam Smith does, is simply arbitrary and illegitimate. It was the *nature* of value-in-use that had to be modified and that represents the real flaw of classical value theory, rather than its independence from exchange-value.

1.2. Conflation of the two dimensions of use value in the early classical attempts to address the paradox of value

A faithful development of the classical approach to use value should have retained the logical and analytical priorities inherent in the classical approach: the institution of property, together with the materiality of the physical world come first, and have conceptual and temporal priority over the social domain in classical political economy. From this it should follow that from an analytical and logical stand point, the subjectivist line of attack should have the precedence. Thus, the definition of an economic theory of consumption explaining how the satisfaction of wants contributes to the generation of value should represent the first logical step, followed by an analytical explanation showing how exchange and use value can be linked, and the paradox of value resolved. Although it was indeed the subjectivist critique that was eventually adopted by the two earliest marginalists (Jevons and Manger, while the case of Walras is different; see section 2.3 below), this was not really the case through the decades preceding the marginalist revolution in value theory.

Historically, if one looks at how classical contemporaries tackled the problem of use value, it is possible to see that the interest of economists since Smith's time has been first and foremost to provide a solution to the paradox of value and to find a common ground for the various manifestation of value (Schumpeter, 1954: 300-311). This research stream has been only loosely linked to the concept of utility (but see Schumpeter, 1954: 188 & 300-301), and certainly it was independent from the definition of a proper economic theory of consumption, in "subjectivist terms" (Bharadway, 1978; Sotiropoulos, 2009).

Indeed, neither J.S. Mill (see Sotiropoulos, 2009), nor the early Marshall (Bharadwaj, 1978) had developed a theory of use value grounded on the theory of subjective utility. As remarked above (section 2.2 of chapter 3), Marshall's attempts were mainly direct towards the dismissal of the dual nature of value, rather than to the definition of a theory of consumption based on subjective utility. As noted by Bharadwaj (1978: 263): "The notion of marginal utility or marginal cost did not appear in that form but was hinted at when Marshall spoke of the 'value in use to those buyers who are the last induced to buy the commodity' and the 'value in use' to the last of those who produce for sale, or 'the cost of production of that portion which is produced under greatest difficulty."

In sum, the "classical" attempts that have been made through the 19th century in order to give a new role to the private sphere of use value and consumption were not consistent. The two possible approaches were often confusingly mixed and there seems to be no clear recognition of the crucial differences distinguishing the two: exogeneity (monism) and objectivity. Even more important, the general consensus among "later classicals" seemed keener in dismantling the dual nature of value by endogenizing use value into the theory of market exchange, rather than in conferring economic meaning to actors' private identities through a subjectivist approach.

For this reason, even when a "subjectivist" approach was adopted, based on the recognition that the fundamental problem with the Smithian paradox of value was represented by the conceptual limitations that an "objectivist" approach to value necessarily entails when it comes to explain the economic role of value-in-use, the results have been inadequate and not convincing. As summarised by O'Brien (2004: 114):

"The contrast, then, is between those writers in the cost of production camp who saw the paradox of value as offering an insuperable bar to a subjective value theory, and those writers who felt that a useful theory could be constructed on such basis. But even those in the last group … failed to produce a cohesive theory in this respect, and so Adam Smith's basic reason for largely dropping the subjective elements in value theory, and concentrating on cost of production, remained effectively vindicated throughout the classical period."

2. Making sense of the subjectivist revolution in value: clarifying inconsistencies and misunderstandings among early marginalists on value and cost, and deriving preliminary implications for institutions

The key event in the history of the theory of value is represented by the marginalist revolution in value theory, occurred in the first half of the 1870s. The conjoint, but independent, works of Jevons (1965), Menger (1981), and Walras (1954) represent the first attempts to confer a new economic role to use value and usefulness, and to give to the private sphere of consumption "in use", new prominence in economic analysis.[47]

Although the revolution is sometimes regarded as a coherent monolith, this is certainly a very superficial interpretation of the marginalist revolution in value theory. Jaffé (1976) represents one of the first attempts to "de-homogenise" the contributions of the three early marginalists and, since then, many other authors followed, disentangling the heterogeneous contributions of the "triad".[48]

In principle, all three early marginalists relied on the concept of utility in order to derive implications for exchange value. Subjective utility was adopted as the ultimate explanatory factor of exchange and this renders the first marginalist revolution the epitome of the second line of attack to the classical dualism in value: the subjectivist one.[49]

[47] For a general overview of the first marginalist revolution in value theory, the reader is referred to Blaug (1996: chapters 8 and 9), Hutchison (1953), Mirowski (1989), Schumpeter (1954: 909-923).
[48] Early works that explicitly considered the contributions of the single early marginlists on which the present chapter relies are, among many others: Groenewegen (2003), Hicks (1934), Hayek (1934), Horwitz (2003), Jaffé (1976), McCulloch (1977), Moscati (2007; 2013a; 2013b; 2013c), Schumpeter (1954: 825-885 & 909-924), Stigler (1937), Streissler (1972), Walker (2003), Young (1912).
[49] As remarked by Schumpeter (1954: 911-912): "they [the early marginalists] established what A. Smith, Ricardo, and Marx had believed to be impossible, namely, that exchange value can be explained in terms of use value. Jevons, Menger, and Walras would all of them have approved of this statement. It is this which they meant when they claimed to have discovered the 'cause' of (exchange) value." In other words, "as regards the actual causation of value, the subjective theory did, as it were, stand the Classical explanation on its head" (Robbins, 1970: 175). It is worth of note

However, the way in which the subjectivist approach to use value was interpreted by the three authors was far to be univocal. Not only the three did not agree on the meaning of utility and of use value but, most importantly, they did not even agree on the way in which the two dimensions of value should have been conceptually linked to each other.[50]

2.1. Use value among early marginalists: fundamental inconsistencies on how subjectivism should be interpreted

2.1.1. Some preliminary considerations on use value among early marginalists: fundamental confusions and inconsistencies

With respect to the way in which the new subjectivist concept of use value should have been intended, the three early marginalists provided three different interpretations. Thus, Jevons (1965: 79) intends value in use as the total cumulated utility of a substance constrained by its social scarcity,[51] while Menger explicitly rejects Jevons' interpretation of use value as "cumulated utility", and intends the use value of a good simply as the relative importance of the least important need that can be directly satisfied in consumption (Menger, 1981: 228). On the other hand, Walras represents a mid-way between the other two: he redefines use value as *rareté* as the subjective counterpart of exchange-value, and then he proceeds to define the latter as the derivative of the utility curve, even though in an inconsistent manner through the *Elements*.[52]

In conclusion, although Jevons, Menger and Walras were all subjectivists with respect to the theory of use value, meaning that they all recognised that the usefulness of a good depends on the identity of the specific actor employing and (auto)-consuming it, the way in

---

the fact the fact that, still in 1888, Philip Wicksteed, the forefather of marginalism in distribution, decided to open his treatise on the nature of value (Wicksteed, 1970: 1; emphasis in the original) in the following way: "It is the object of this volume in the first place to explain the meaning and demonstrate the truth of the proposition, that *the value in use and the value in exchange of any commodity are two distinct, but connected, functions of the quantity of the commodity possessed by the persons or the community to whom it is valuable* …".

[50] Concerning the relationship between exchange value and utility in Jevons, see chapter IV of the *Theory of Political Economy* (Jevons, 1965: 75-84). See also the considerations of Allyn Young (1912: 586). Concerning Menger, it is sufficient to consider chapter VI of Menger's *Principles* (Menger, 1981). Concerning Walras, the reader is reminded to his discussion of utility and value in exchange provided in Lessons 3, 8 and 10 of his Elements (Walras, 1954: 65-72, and 143-49).

[51] "Thus Smith evidently means by value in use, *the total utility of a substance of which the degree of utility has sunk very low, because the want of such substance has been well nigh satisfied*." (Jevons, 1965: 79).

[52] Sometimes *rareté* is defined as "the derivative of the effective utility with respect to the quantity possessed" (p. 146), meaning that any point of the utility curve can have its associated *rareté*, while some other times he defines rareté in a way very similar to Menger (even though not identical), as "the intensity of the last want satisfied by any given quantity consumed of a commodity (p. 119).

which “subjectivism in use value” should have been interpreted was far to be univocally recognised.

The misunderstandings and the incongruences among the three early “subjectivists” (which led to diverging paths of the various schools in the subsequent decades) reflect the deep confusion concerning the way in which subjectivism in (use) value had to be linked with monism in value. In sum, the meaning of exchange-value, and the way in which the two dimensions of value had to be linked (exogenous or endogenous), was not univocally interpreted. In this case, it was the locus of the institutional layer that differed greatly among the three marginalists, and the way in which the private and the social institutional domains had to be linked.

2.1.2. *The relationship between value-in-use and value-in-exchange in the English marginalism*

Value in exchange was still genuinely classical in Jevons (objective and exogenous), against a subjectivist and endogenous interpretation of use value.[53] As stated by Buchanan (1969: 11), “Jevons was unique among the subjective-value theorists in his treatment of cost, and he is considerably more classical than Austrian.” The same can be said of Marshall’s partial equilibrium analysis: although it is true that the exchange value in the consumption market is endogenously determined by the interdependence of supply and demand curves, this is not the case for the market for productive services “upstream”. In this case, value-in-exchange (the various market prices of productive services) is indeed exogenously given to the actors and objective, and derives from the “real” cost of production sustained by producers.

More generally, when it comes to the problem of the valuation of productive factors and resources, the English marginalism was in perfect continuity with the English “production-cost”, where the value in exchange is determined by the real costs suffered by factors’ owners in production. All English marginalists have always shown deference to the old classical tradition when it comes to value and costs in the production camp.[54] The same concepts can be expressed by saying that in the English tradition, the value-cost sustained

[53] Jevons himself declares in the first pages of his preface to the first edition of his work (1965: 6): “There are many portions of Economical doctrine which appear to me as scientific in form as they are consonant with facts. I would especially mention the Theories of Population and Rent, the latter a theory of a distinctly mathematical character, which seems to give a clue to the correct mode of treating the whole science.” On this point, see also Robbins (1970: 177-178), Stigler (1941: 13-37), and Young (1912: 587-588), where he recognises that, apart from the theory of interest, “[o]ther features of Jevon’s theory of distribution need less consideration. He accepts the orthodox theory of rent and propounds a residual claimant theory of wages.”

[54] See for instance Buchanan (1969; chapter 1), and Schumpeter (1954: 920-24).

by factors' owners are computed as "real" production costs, rather than as opportunity costs.

In conclusion, the English tradition adopted a subjectivist theory of use value coupled with an objectivist theory of exchange value. The way in which the two interact, however, depend on the specific "side of the economy considered". With respect to the valuation formulated in the production side, the two are also exogenous, consistent with the classical tradition in the sense that the costs sustained in production are unaffected by the utility in consumption.

Conversely, value in use is endogenously linked to value-in-exchange in the consumption side. The mechanism through which this endogenous linkage occur is social scarcity. This reveals an inconsistency that will also be present in Walras: although Jevons initially computes value in exchange starting from private ownership, the way in which value-in-exchange and value-in-use are merged is by making the quantity privately possessed of any good by any actor an endogenous function of the social endowment (supply) of that specific good.[55] Use value then, becomes an endogenous function of scarcity, or supply, and therefore an endogenous function of exchange value in turn.

This is, more generally, the way in which any subjectivist and monistic approach to value (shared by both Jevons and Walras) would solve the paradox of value.

2.1.3. The relationship between value-in-use and value-in-exchange in Menger

Menger is the only early marginalist among the three to explicitly retain the classical logical and ontological approach: from the appropriable material goods (institution of private property), to their final employment (social exchange).[56] For this reason, Menger is also the sole of the early marginalists to maintain an exogenous approach to the theory of use-value, in the sense that Menger will never resolve the paradox of value and switch towards

[55] Thus, Jevons starts his chapter on exchange adopting a genuine classical approach, where exchange is subsumed to self-consumption (1965: 75): "Trade is not indeed the only method of economising: a single individual may gain in utility by a proper consumption of the stock in his possession. The best employment of labour and capital by a single person is also a question disconnected from that of exchange, and which must yet be treated in the science." Yet, by the end of the chapter he concludes by saying that "Value depends solely on the final degree of utility. How can we vary this degree of utility? –By having more or less of the commodity to consume … Labour affects supply, and supply affects the degree of utility, which governs value, or the ratio of exchange."

[56] Work specific works on the economics of Menger, the reader is referred to, among others, Alter (1982; 1990), Caldwell (1990), McCulloch (1977), Hicks and Weber 91973), Stigler (1941: 134-57), Streissler (1972).

monism in value theory. This sharp separation and exogeneity of the two spheres, typical of the classical doctrine, is clearly summarised by Alter (1990: 336):

> "it does not suffice to treat prices as magnitudes which are simply calculated from values by setting price equal to value. Here we are dealing with fundamentally different spheres of discourse which obey different epistemological and methodological laws because these spheres are constituted of completely different elements."

The value dualism of Menger is a direct consequence of the fact that Menger does never use a concept defined at the social layer of analysis (such as scarcity), to explain the entity of usefulness in the consumption of a given asset, resource or good.

In the same way in which classical economists intended usefulness as a concept only meaningful if associated to the appropriation of "material wealth", and subsumed the concept of usefulness to the preliminary possession and appropriation of a good (see chapters 2 and 3 above),[57] Menger links the concept of use value to the preliminary "command" of the goods, in the form of private property. Most importantly, Manger decouples the act of the consumption of owned goods from the act of exchange as the former has causal and temporal priority over the latter (1981: 77-78):

> "In practice, the concern of men for the satisfaction of their needs is expressed as an attempt to attain command of all the things on which the satisfaction of their needs depends. If a person has command of all the consumption goods necessary to satisfy his needs, their actual satisfaction depends only on his will. We may thus consider his objective as having been attained when he is in possession of these goods, since his life and well-being are then in his own hands."

---

[57] In classical economics resources' appropriation is preliminary to their employment, which is in turn preliminary to the computation of economic magnitudes (in exchange). As summarised by Ricardo, 1911: 39; emphasis added): "If air, water, the elasticity of steam, and the pressure of the atmosphere were of various qualities; *if they could be appropriated*, and each quality existed only in moderate abundance, they, as well as the land, would afford a rent, as the successive qualities were brought into use." And again (p. 34; emphasis added): "for no one would pay for the use of land when there was an abundant quantity *not yet appropriated*, and, therefore, at the disposal of whosoever might choose to cultivate it."
The concepts of wealth, scarcity and value all depend on the preliminary assumption of private appropriation of the resources in classical economics. As already remarked in chapter 1, with no ownership, there can be no accumulation of wealth in classical economics, and therefore no use value.

Menger’s assertion is a perfect description of property and ownership in classical terms: as the faculty to employ the goods “at will” with no dependence on the external environment. As discussed in chapter 2 above, this is also the way in which property law defines ownership: as the right to exclude others, and therefore to employ and benefit from the property “*at will*.”

In sum, it is the given quantity of material wealth possessed by, and at the disposal of, an individual that, depending on one’s set of subjective wants, determine use value in Menger and *not* social scarcity, as in Jevons and Walras. In other words, consumption depends on the *private endowment* of goods (defining private wealth), not on the *public endowment* of potentially purchasable *services* (defining social scarcity). As scarcity is irrelevant in determining use value, the Mengerian framework has been interpreted as a framework where only one primitive does the job (utility in consumption), while supply is rigid, or inelastic, and therefore could be ignored. As stated by Viner (1953: 200):

> “The Austrian School starts with the assumption, usually tacit, never emphasized, that the supplies of all the elementary factors of production are given and independent of their rates of remuneration.”

It is true that the independence of use-value derives from the fact that “the supply of factors is given”. Yet, this precisely derives from the fact that the private sphere of factors’ ownership in Menger’s framework has priority over the social sphere, meaning that only goods under the direct command of an actor can be employed by the latter. Goods outside its control must first be purchased in order to be consumed (Menger, 1981: 119):

> “The circumstance that a good has value to us is attributable, as we have seen, to the fact that command of it has for us the significance of satisfying a need that would not be provided for if we did not have command of the good.”

As ownership and appropriation are always a precondition for utility and usefulness (as in classical economics), meaning that a certain good has always to be appropriated in the first place, then the given amount of goods that can be employed by its owner is always “rigid” in a specific instant of time. This is why social scarcity (characterising supply) also becomes irrelevant in order to compute subjective use value and the Mengerian framework remains dualistic, with exchange value following from use-value, exactly as in classical economics.

For this reason, Mengerian framework can be considered a genuine example of a subjectivist, yet exogenous, theory of use value.

2.2. English and Austrian value and cost theory: making sense of a long-lasting dispute and some implications for asset ownership

2.2.1. *Some preliminary considerations on the differences between Austrian and English approaches to the classical paradox of value: the unit of analysis*

In sum, while the English marginalist school retained the objectivity and exogeneity of exchange-value from classical economists (with respect to cost-values in production), the Mengerian school (Menger and his direct disciples: Wieser and Böhm Bawerk) retained the exogeneity of use-value from classical economics.

The latter derived from a profound affinity between classical and Austrians in the way in which the private and the social sphere had to be linked: not as two inter-dependent sides of the same coin, but as two severed institutional locus that are intermediated by the institution of ownership, whose role is precisely to decouple the available quantity of a certain thing *in the social sphere* (scarcity or supply), from the quantity of a certain thing readily available (at will) *in the private sphere*.

The difference between the two cases are profound, and they cannot be treated as two indistinguishable aspects of the same phenomena. Two important have to be highlighted here.

One, the classical-Austrian approach to use value relies on a definition of private ownership (as the faculty to address private needs and satisfy private requirements independently from the external environment) that becomes impossible to justify in a "supply-and-demand" Marshallian paradigm. More generally, every time that use-value becomes an endogenous function of social scarcity and social supply (thus, of value-in-exchange), the satisfaction of personal needs is always dependent on the purchase of inputs from the outside (form which utility becomes a function of the *purchased* quantities of inputs $q: U(q)$).

Two, this means that in this second case, the satisfaction of one personal needs exclusively depends on his monetary wealth (income), and his purchasing power, not on his actual cumulated material wealth to be employed in self-consumption.

These two points are themselves the outcome of a most fundamental distinction between the two schools: the difference in the unit of analysis of the economic framework. While in the English case, and *in any other* monistic framework, the unit of analysis is the *quantity of a given service purchased and consumed*, in the Austrian (and classical) case the unit of analysis is the *given material "thing"* that has to be first purchased and appropriated before it can actually be used and consumed later on. More generally, the solution of the paradox of value through the concept of social supply (scarcity) always relies on a preliminary change in the unit of analysis from the material wealth to flowing services, as we shall see in a moment.

Although the change in the unit of analysis will represent the topic of section 3 below, it is interesting to briefly discuss one of its fundamental implications that has spurred harsh debates between the English and the Austrian school: the meaning of cost.

2.2.2. *The change in the unit of analysis from properties to services implies a radical change in the way in which cost is defined: opportunity costs vs. production costs*

The English approach shared with classical economists the notion of cost intended as a "*production cost*". The objectivity and exogeneity of production costs not only is what confers objectivity and exogeneity to value-in-exchange, but it is also what ensures that value is the outcome of two distinct primitives among English marginalist: utility and, precisely, production costs. The Austrian subjectivist approach did not rely on the concept of production cost. Austrians were the first to introduce the concept of "*opportunity cost*" or, "alternative cost" as Wieser (1930: book v) first defined it.[58]

Soon after the translation of the Austrians texts in the early 1890s, a fierce debate developed between English marginalists (mainly Edgeworth, 1892; 1894) and Austrians (mainly Böhm Bawerk, 1894a & 1894b). The debate developed along two lines. The ultimate standard of value, and the meaning of cost.

The two aspects of the debate were in reality strictly linked: the presence of (real) production costs was the reason why English marginalist relied on two distinct sources of value: production and consumption.[59] On the other hand, the Austrian's subjectivist approach gave prominence to use-value and relegated everything else to a subordinate position. The concept of "opportunity cost" is the cornerstone on which the Austrian

[58] The concept of opportunity cost was introduced by Wieser in his Natural Value, originally published in German language in 1889, and defined as "alternative cost doctrine". In English speaking countries the book became available only in 1893, rapidly fostering the American legacy of the opportunity cost doctrine, that took off soon after the translation of Wieser's book with the articles of Green (1894) on Pain Cost and Opportunity Cost and Davenport (1894). The history of the concept is very clearly summarised by Buchanan (1969), Buchanan and Thirlby (1973), and Schumpeter (1954: 917).

[59] Thus Edgeworth (1894: 519) starts his critique by saying: "I speak on behalf of those who hold, in opposition to the Austrian School, that there is not one ultimate standard, but two ultimate standards: utility and disutility." See also the account provided by Stigler (1941: 182-187), and Edgeworth's review of Böhm-Bawerk Positive Theory of Capital (Edgeworth, 1892) for an earlier critique along the same lines.

subjectivist approach to use value relied. The concept of opportunity cost is describe by Wieser (1930: 175) as follows:[60]

> "In this way we reach the point of view from which production goods are conceived as costs … To say that any kind of production involves cost, simply implies that the economic means of production, which could doubtless have been usefully employed in other directions, are either used up in it, or are suspended during it. Costs are production goods when these are devoted to one individual employment, and on account of their capacity of being otherwise employed, take the shape of outlay, expenditure."

It is the concept of opportunity cost that give operational meaning to value in use. Moreover, as the various employment of a good are identity-specific, the concept of opportunity cost became the cornerstone of the early (not exclusive Austrian) subjectivist approach to the theory of use-value, and in sharp contrast with the old objectivist tradition, still hinged in the classical notion of production cost.

As a result, the original debate of the late 19th century on the *origin* of value soon conflated into the debate on the meaning and the *nature* of costs that ensued for various decades, until the mid-20th century at least (see Robbins, 1930, 1933 and Knight, 1928, 1967a & 1967b).[61] This debate on the meaning and the nature of opportunity costs vis-à-vis production cost was the incubator, among others, of Frank Knight (1924) "*Some Fallacies in the Interpretation of Social Cost*" that can be regarded as the ancestor of the Coasean legacy in Law and Economics (Coase, 1960), and of modern (post-Classical) trade theory (Haberler, 1929, 1936, 1950; Leontief, 1933; Lerner, 1932 & 1933; Mason, 1926; see Vanek, 1959 for a review article).[62]

Here it is interesting to highlight a possible source of the misunderstandings on the nature of cost. A careful look at this debate on the meaning of cost (and the nature of value) reveals that the crux of the argument derived from the fact that the two definitions of cost apply to *two different units of analysis*: the given service intended as the output, in the case of production cost, and the given good, intended as the input of the consumption process in the case of opportunity costs.

[60] For contemporary definitions of the term see Alchian (1968) and Buchanan (2008).
[61] See Buchanan (1969) for a review of the debate.
[62] In fact, Knight's famous article has to be intended as a response to the debate on the meaning of costs in international trade.

The distinction is between the production cost sustained in order to render a given service, and the opportunity cost sustained in order to operate a given resource. Not surprisingly the concept of opportunity cost is defined by looking at how a "good" is employed in Wieser's definition, while production costs, which only make sense in partial equilibrium analysis, are always defined in function of a given outcome-service to provide (produce). It should also be noted that the distinction between the two types of costs also reflects a distinction between the two types of values: production costs is always a primitive for exchange value, while opportunity cost, consistent with the unit of analysis, is a primitive for use value. Furthermore, there is a further distinction characterising the two natures of costs: production cost is a genuine concept rooted in distribution theory, while the concept of opportunity cost developed in the camp of (use) value theory.

2.2.3. Making sense of the debate over the meaning of cost: Austrian opportunity costs and English production costs are the two sides of the dualism of value

Nowhere this change in the unit of analysis is more evident than in the definition of the economic discipline itself. "Opportunity costs" are those costs generated by the allocation of *a given thing* among alternative uses, or employments. Once again, the unit of analysis is *the underlying thing*. There is no possible allocation problem without an opportunity cost, and certainly there is no allocation problem inherent in a production problem.

It is for this reason that, since Wieser's introduction of the concept, the *allocation problem* became a synonym of the *valuation problem* (in the use or employment of a resource) from which Robbins' (1935: 16) well-known definition of economics: "the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses."

Robbins new definition of economics is simply a definition of subjective value-in-use, simply be defined as the value judgment informing how to employ and consume given resources. The interpretation of economics as a discipline studying the allocation of resources is a therefore a heritage of the early subjectivist wave in value theory, and this made the concept of economics, use value, and allocation interchangeable synonyms. Robbins' book is a clear example of a subjectivist and exogenous approach to use value (i.e. the allocation problem).[63]

[63] On the "subjectivity of Robbins' value theory there is hardly any doubt (O'Brien, 1990). The reason why Robbins' is defined as a proponent of the "exogenous" theory of value derives from his refusal to automatically express the concept of use value as an endogenous outcome of exchange-value and, vice versa, to univocally derive a market price given any valuation in use (on this see also

The previous considerations show that the debate on the nature of cost was relying on an underlying duality that runs through the discipline of economics and that touches upon the various concepts.

In this respect, three important considerations are in order.

First, if value-in-use is considered the essence of the valuation process, then only opportunity costs can be legitimate economic costs as they are the only one clearly modelling the problem generated by the employment and use of a given assets or resources among alternative uses. This was the original way in which usefulness in self-consumption was defined by classical economists, and it will remain the way in which value-in-use is defined.

From the first point it also follows that (second), the notion of production cost is not an economic concept, but a mere technological concept: it does not involve any economic problem, meaning that it does not describe, nor model, any valuation (allocation) problem, provided that economics is intended as the discipline relying on use-value, describing how resources ought to be employed. The description of how a given service ought to be best produced (or a given market ought to be best served) is a scientific and technological question, thus a *social* and *objective* one, and not a (socially relative) and subjective question. This was, in my view, Hayek's long-standing critique against objectivism and "scientism" (which we can reinterpret as "technologicism") in economic analysis (Hayek, 1937, 1942, 1945, 1952).

Note that the subjectivism and the pluralism of valuations and value judgments that characterise the allocation problem inherent in the opportunity cost approach cannot be reproduced when the question entails the definition of the optimal solution for the production of a given service. The latter is an objective and social problem, in the same way in which English classical economics, and Marshallian production economics are.

For these reasons, it is also possible to say that any allocative problem in general equilibrium has to employ the concept of opportunity cost, while the concept of production cost can only be legitimate in a partial equilibrium framework (which models and studies technological properties of production functions, not allocative problems of resources).

Robbins, 1933, 1934, 1981; see chapter 1 of the *Nature and Significance*). Buchanan (1969, 1973) is another example of late subjectivism in use value coupled with a mild exogeneity. The situation of late Austrians is much more complicated, and often inconsistent, as we shall see.

Third, a comparison of the specular features of the English and Austrian marginalism shows that the two are the two sides of the duality inherent in any non-endogenously determined theory of value.

As a result, it is possible to wrap-up the discussion, and summarise the inherent dichotomies between the English and the Austrian early marginalism as in figures 4.1 and 4.2 below.

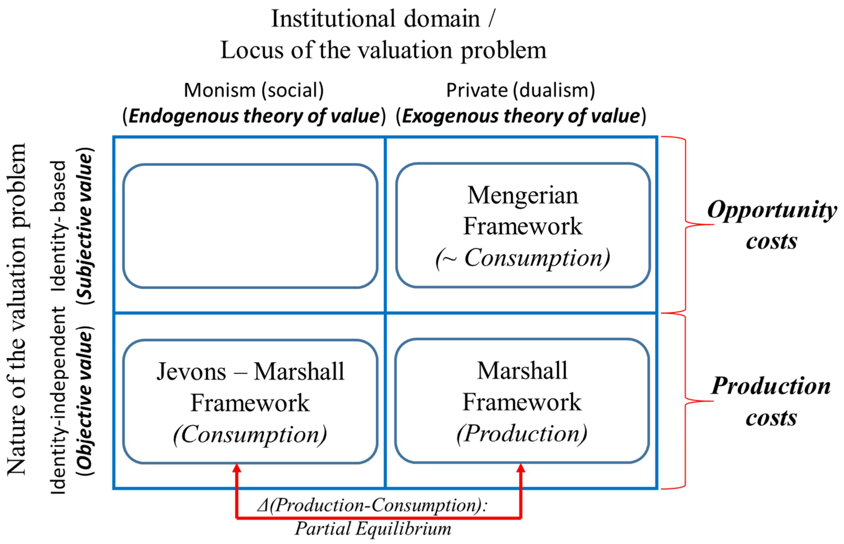

Figure 4. 1– English and Austrian early marginalism in perspective

2.2.4. *Wrapping up English and Austrian value theory: making sense of the differences by interpreting them as two alternative but compatible approaches to value theory*

Figure 4.1 shows three things.

One the English framework did not have a unique economic model of value (use-value) and distribution (exchange value) that could explain both in a monistic way.

The taxonomy of value introduced in chapter 3 above can be used in order to easily understand what partial equilibrium and what general equilibrium mean: *in case of social (objective) valuations* (prices),[64] a partial equilibrium analysis is characterised by an exogenously given valuation: actors receive prices from the outside. Value-in-exchange in a partial equilibrium analysis is exogenous as it cannot be (endogenously) modified by the economic activity of the actors. Vice versa, value-in-exchange is always endogenous in a

[64] Remember in the specific case of monetary valuations (prices), the adjective objective and social are identical, and can be used as synonyms (see chapter 3: section 3.4).

general equilibrium analysis (Walrasian and Marshallian in the case of consumption markets).[65]

Two, the figure shows that (because of the fundamental rupture inherent in English partial equilibrium analysis), no one of the two earliest schools dared to switch to monism in valuations. When it comes to the employment of assets and resources both early schools are located in the right-hand side of the table.

Three, the Austrian *subjectivist and dualist* approach and the English *objectivist and dualist* approach can be intended as the two antithetical sides of the same dualism, as represented in figure 4.2.

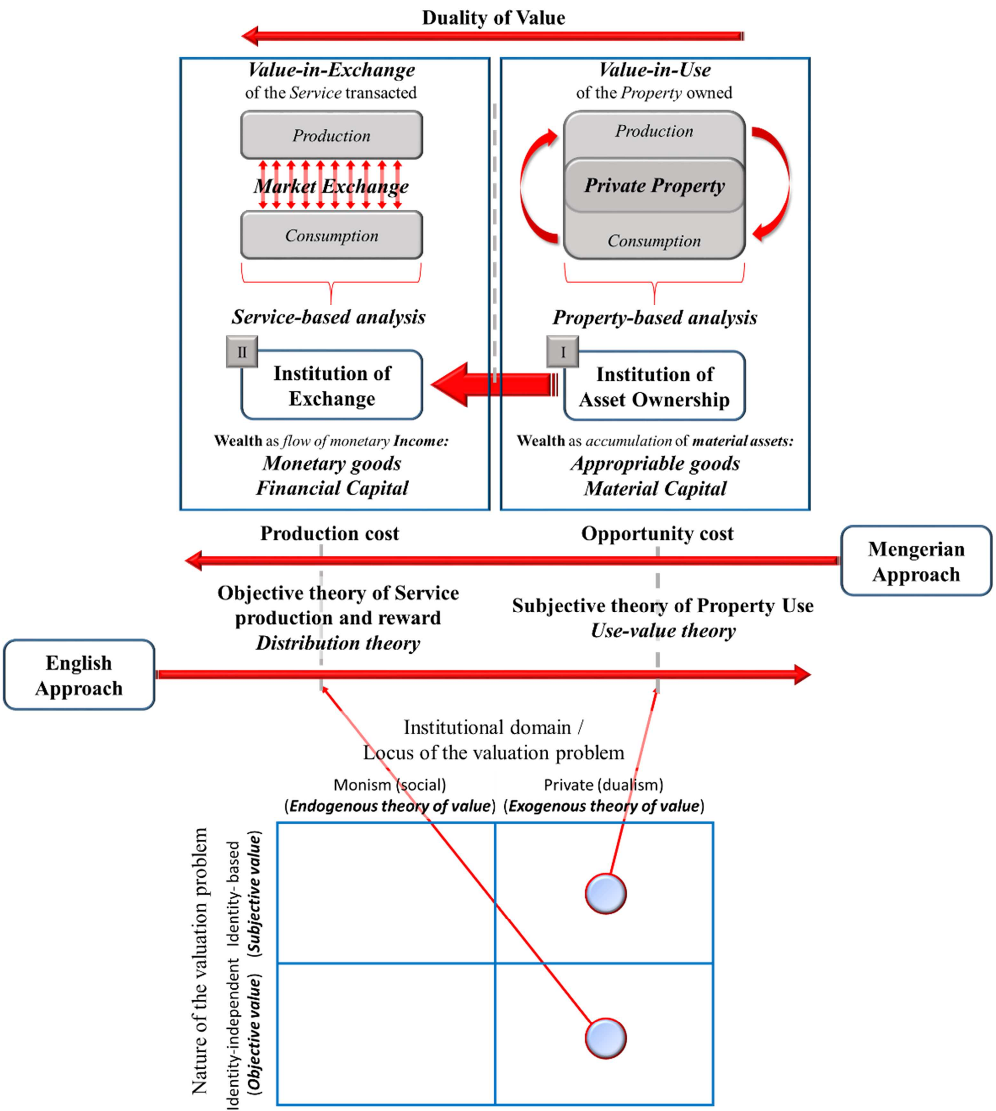

Figure 4. 2 – Early English and Austrian schools in perspective

[65] Note that this is perfectly consistent with the definitions provided above (chapter 3: section 4.2).

As both approaches are dualist, they retain the separation between the two sides of the economic problem: the property side (right) and the exchange side (left).

However, figure 4.2 also highlights a fundamental difference between the two: a service-based analysis based on the concept of production cost, defines every magnitude in function of the outcome (service) to be produced, or rendered, and retrieves everything "backward" from there. As figure 2.2 shows, as the service is the unit of analysis, the logical and analytical approach is reversed; form the service-outcome to the properties to appropriate in the input side.

Three points have to be made in this respect.

One, this is in antithesis with the original classical approach, which was adopted by Austrians, where the initial property is given and the allocation (value-in-use) problem ensues.

Two, this is the essence of any partial-equilibrium analysis, also adopted by antitrust authorities, and by industrial organization in general, where assets and resources are always purposefully acquired in order to enter a given service market (in output). For this reason, property ownership in industrial organization is not regarded as a source of value-in-use across outcomes (opportunities), but merely as a rigidity (indivisibility), generating *production* economies *given a certain service*.

The concept of scale and scope economies are consistent with a production-cost analysis (in fact they are computed through the standard Marshalllian production-cost curves), certainly not with an opportunity-cost analysis, as they can only be defined *given a service to be produced* (and a market to be served), and not given a property to be valued or, in other words, allocated (among a bundle of alternative opportunities). Chapter 2 above already developed this point with respect to the topic of the division of labour and the meaning of use-value.

Three, it should be noted that a production-cost framework, which is also a service-based approach, is consistent with the presence of service-specific (dedicated) assets, *à la* Williamson. In chapter 2 it was mentioned that dedicated assets make the valuation problem irrelevant and for this reason it is one of those four conditions making ownership, and institutions in general, irrelevant from an economic point of view.

The present analysis is already foreshadowing the most fundamental reason why in a production-cost, or service-based, analysis (subjective) use value and institutions lose meaning: the possibility to generate a reversed-causality from the service rendered as output to the material underlying property employed as input. This violation of the classical (and Austrian) univocal causality flow, which is implied in the change in the unit of analysis, will be one of the two fundamental conditions on which the neoclassical “failure-approach” to institutions relies on.

In conclusion, the quarrels that developed among the English and the Austrian school of thoughts can be much better understood, and can be made sense of, if there are studied in function of the taxonomy here suggested. The misunderstandings can be reinterpreted as the most evident example of the incomprehension concerning the nature of use value (subjectivism and objectivism) with respect to the nature of the two dimensions characterising the theory of value: exogeneity and endogeneity.

Figure 4.3 completes the picture by showing that while the right-had side of the table (exogeneity and dualism) necessitates some form of temporal and logical causality between the two exogenous elements of the framework (at least in order to understand the temporal order and priorities between the two), as soon as we move to the case of monism in valuations time loses meaning, and the framework collapses into a singularity where time disappears and the causality can go either way indifferently. This is the case of the supply-and-demand endogenous adjustment in the case of Marshallian theory of consumption. As we shall see in the next section, and consistent with what has been already foreshadowed, it is not only time to collapse, it is also space, in the sense of the ontological materiality of the physical world.

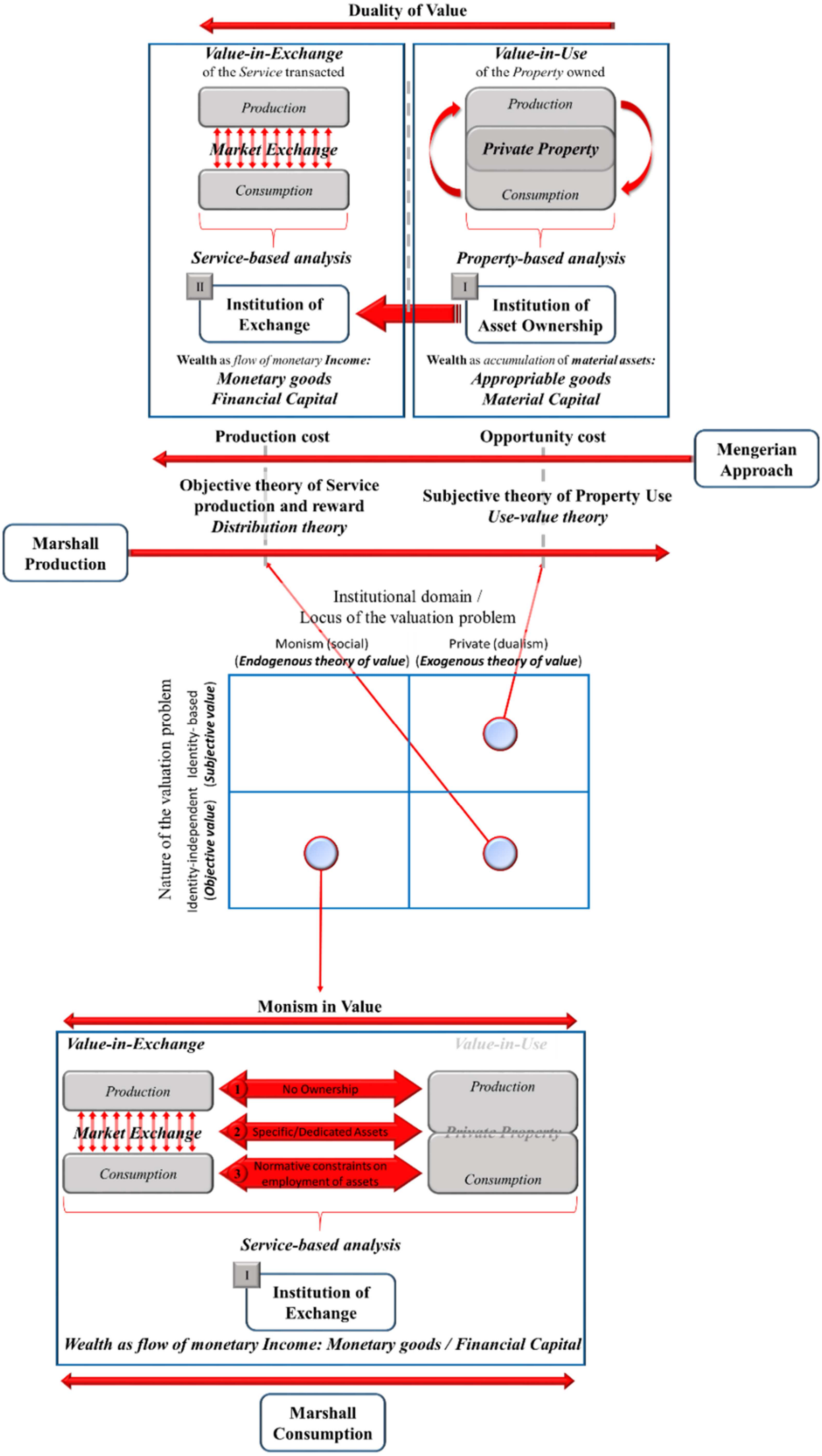

Figure 4. 3– Early Austrians and Marshallian framework in general

2.3. Walras value theory as a final synthesis of English and Austrian value theory: the first subjectivist and endogenously computed theory of value and the final (only apparent) solution of the classical paradox of value

Last, among the early marginalists, came Walras. Both Menger and Walras adopted a subjectivist approach to both use and exchange value, in sharp contrast with the English (classical and "post-classical") approach. As a result, Walras, in line with the Austrian subjectivism, frames his economic problem by adopting "opportunity cost schedules", which are, at least in principle, always a symptom of subjectivism in value, and therefore cost.

Walras' opportunity cost schedules are represented by the well-known capacity-constraint schedule, and are defined in lesson 20 of his *Elements* (Walras, 1954: 240; relations 3).

$$Q_R = a_R q_1 + a_R q_2 + a_R q_3 + \cdots + a_R q_j + \cdots + a_R q_n = \sum_{j=1}^{n} a_R q_j$$

Where Q_R defines the (constrained) productive capacity of the resource or asset, while the various n employments (opportunities) are addressed with coefficients a_R and in quantities q_j.

For this reason, the Walrasian framework also retains the key characteristic implied by the presence of an opportunity cost schedule: economics and value theory intended as an allocation problem. From this point of view, Walras is much closer to the Mengerian approach than to the English (Marshallian) one.

Although Austrians and Walras share the subjectivist "opportunity-cost approach" to value, and the same approach to the economic (thus valuation) problem, Walras' framework has a substantial difference: while the former interpreted value-in-use consistent with the classical way, thus as exogenously given to the social (exchange) domain, Walras coupled subjectivity with endogeneity, substantially solving the paradox of value, and putting the germs of contemporary neoclassical economics (in general equilibrium, thus consumption *and* productive services' market).

Differently from English marginalists that locate the origin of (production) costs and value-in-exchange outside the economic framework (they are exogenously given by the technological and social context, in case of "real" costs), and from Austrians that also based they value theory on an exogenous source of use value ("natural value" to use Wieser expression), the Walrasian framework computes everything endogenously: the opportunity-cost schedule denoting subjective value (in use), together with the production-cost schedules for production and distribution.

The two elements of the Walrasian framework are, precisely, the Austrian opportunity cost schedule, and the English (classical) production cost schedule, interwoven and endogenously linked together as follows:

$$Q_R = a_R q_1 + a_R q_2 + a_R q_3 + \cdots + a_R q_n = \sum_{j=1}^{n} a_R q_j \quad \text{Austrian opportunity cost schedule}$$

$$a_R p_R + a_S p_S + a_T p_T + \cdots + a_Z p_Z = \sum_{x=R}^{Z} a_x p_x = p_a \quad \text{English production cost schedule}$$

The second schedule (Walras, 1954: lesson 20; equation 4) defines the sum of inputs' prices (p_x) multiplied per the respective coefficient of production (a_x), giving the total production cost ($\sum_x a_x p_x$) to be equated to the price p_a of the final output a to be sold in the consumer's market. (It should be noted that both are sets of equations.)

As summarised by Schumpeter (1954: 913):

> "This means, on the one hand, that the marginal utility principle now covers the cost phenomenon and in consequence also the logic of the allocation of resources (structure of production), hence the 'supply side' of the economic problem *so far as all this is determined by economic considerations*."

The importance of the achievement can be better appreciated by saying that Walras, first, put the foundations for a theory where the *subjectivist theory* of *use value* (founded on usefulness in consumption) conflated with the *social theory of income distribution*. This laid the foundations of a monistic framework in value and distribution for the very first time in the history of the economic discipline.

The reason why Walras is (rightfully) regarded as the forefather of neoclassical economics derives precisely from the fact that he could show first, and after nearly a century of struggles, how to solve the classical paradox of value (*endogeneity* or monism), *while also modelling values in "subjectivist" terms.* And this precisely corresponded to the original aim of the marginalist revolution.

It finally represented the long-awaited marginalist synthesis that brought together Austrian subjectivism with English monism at a general scale. As summarised by Schumpeter (1954: 913):

> "Most of the problems that arise from this [Austrian] set-up can be discussed only on a level on which Walras rules supreme. But, though I believe that Jevons should be credited with a vision of the facts above and if so holds priority, the credit for having worked out that theory systematically, on the plane on which we are moving now, should go to the Austrians and particularly to Menger, whose *Grundsätze* contain all the essentials."

As the subjectivity of use value has been usually reframed as *value theory* proper since the marginalist revolution, while the social objectivity of exchange value (in production) has been usually referred to as *distribution theory* (computing the income shares of the resources, assets and factors participating to the production process), the final Walrasian solution can also be intended as framework conflating value with distribution theory. After the rough Menger-Wieser attempt (theory of imputation through the opportunity cost schedule), this was also a primer.

Thus, the resulting *monism* is not only in institutional layer and in temporal causalities, but it also involves the two classical departments of the economic discipline. As value theory since the subjectivist revolution of the Austrians has become a theory of allocation of factors and assets, the Walrasian framework can also be reinterpreted as a comprehensive *theory of allocation and distribution*. The former is inherently *subjective*, the latter is inherently *social and monistic*.[66]

As a result, we can finally complete the picture placing Walras in the adopted taxonomy (figure 4.4).

[66] It is for this reason that Blaug (1996: 408) concludes that "marginal productivity theory is a theory of factor pricing, not a theory of the distribution of relative shares".

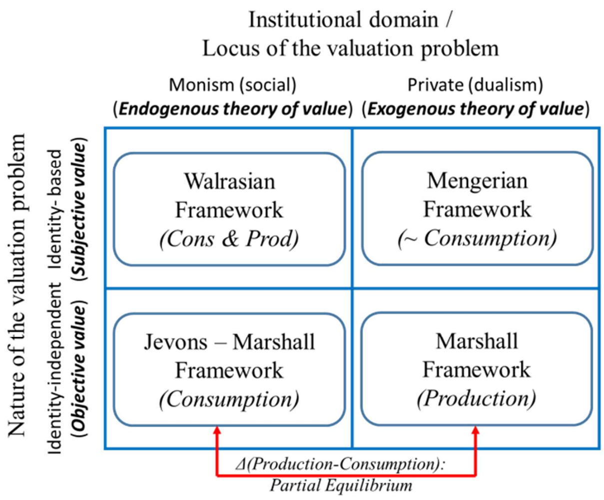

Figure 4. 4 - The 1870s marginalist revolution in value theory: A taxonomy

In truth, the conflation of subjectivism in value and monism is not acceptable, nor meaningful, as remarked above. This means that some further assumptions have to be made. The introduction of these new assumptions will fundamentally modify the way in which private property and asset ownership are treated in economic analysis. Their implications for the role of institutions and for the distinction between public and private domains will be profound.

3. Foundations of a monistic and social theory of value: the two conditions legitimising the shift from the material nature of assets to the monetary nature of services

3.1. Preliminary remarks: uneasiness of English and Austrians with a subjectivist and monistic theory of value

The advantage of studying the different economic approaches in function of value theory derives from the fact that it allows to immediately understand and perceive a fundamental inconsistency of the Walrasian framework: in principle, subjectivity cannot exist in a monistic framework, nor subjectivity in valuations can make any sense when the private and the social sphere conflate and collapse into a time-less and identity-less “singularity”. The combination of subjectivity in (use) value and mutual endogeneity in valuations is just not sustainable, nor logically defensible. This point was the essence of chapter 3 above.

The incompatibility of subjectivism in valuations and mutual interdependence of the institutional layers certainly represents the underlying reason why all economists willing to overcome the dualism and the objectivity of classical use-value theory has struggled with the issue for so many decades. Moreover, it certainly also represents the reason why neither the English (objectivist and exogenous) school, nor the Austrian (subjectivist and exogenous) school were ever completely convinced by the Walrasian solution, and rightfully so.

The treatment of value-cost (in production) in subjective and opportunity-cost terms looked like an unacceptable refusal of the "real" nature of production at the eyes of English economists. Conversely, the presence of socially-dependent private valuations (endogeneity) was clearly regarded as a betrayal of the essence of the subjectivist revolution by early Austrians.

Although neither of the sides involved spelled their dissatisfaction along the two dimensions adopted here, it is clear that while the uneasiness of the English tradition derived from the Walrasian's treatment of value-cost as the outcome of *subjective* value judgments of private actors (across alternative options), the dissatisfaction of the Austrian school was dictated by the presence of subjective valuations that were however at the same time *endogenously* dependent on (determined by) the social system.

3.1.1. Two fundamental problems undermining Walras' subjectivist and endogenous theory of (use) value

The complaints of Walras' contemporaries were not misplaced. The framework proposed by Walras was indeed problematic and had many conceptual and analytical mistakes at the point that it could not even be solved. There were two main problems embedded in Walras' theory of value. One, equalities in the opportunity cost schedules (as written by Walras above) cannot guarantee a solution in linear programming.[67] Two, the original version of Walras' equations relied on fixed coefficients of production a_x and this made the production schedules inelastic and (non-continuously derivable) broken-curves that, again, could not guarantee the existence of an optimal social solution to the social valuation problem. This "Leontief-type" curves (Leontief, 1946)[68] would have simply bring the framework back to the old classical theory of distribution where the retribution of factors would have depended on some *exogenous* source of value, thus independent from the pricing (market) mechanism.

In other words, the original *subjectivist* and *monistic* Walrasian framework was not sustainable, and had to be modified, as our understanding of the taxonomy of value theory

[67] Not all constraints in a linear programming model can be binding at the same time, meaning that there might be no solution at all if constraints are conceived as strict equalities. On this see the basic treatment in Dorfman et al. (1958: chapter 2).
[68] See also Blaug (1996: 411-417), and Dorfman et al. (1958: chapter 10).

should have already told us. Some further *ontological* and *behavioural* assumptions had to be spelled out.

The problem of fixed coefficients represented an *ontological limitation* of the Walrasian framework and was solved first. The problem of fixed coefficients can be considered an "ontological flaw" as it required the switch from the classical notion of *physical assets* and goods to the (neoclassical) of flowing and divisible *services* or, in other words, the switch from *material wealth* to *money income* (financial wealth). It was the marginalist revolution in production and distribution theory, occurred nearly twenty years after Walras' first editions of the *Elements*, which provided the answer: Barone and Wicksteed pointed out the flaw in 1896, and Walras fixed it in his third editions of the *Elements*.

On the other hand, the problems deriving from the unique and optimal solution of a constrained linear program required new *behavioural assumptions*. The overcoming of this "behavioural flaw" required a much more advanced knowledge in linear programming and convex analysis that could only arrive much later, when Walras was already dead. In sum, Walras needed some extra-assumptions on the behaviour of the subjects employing the physical assets and resources according to their respective (subjective) *opportunity-cost schedule* in order to ensure that his subjective and endogenous theory of value could stand.

In this case, it was von Neumann that corrected Walras and, first, demonstrated the existence of *unique social valuations*, *endogenously* computed by the exchange mechanism, while starting from *private wealth*. Von Neumann did it in 1933 (translated and published in English language in 1945), nearly sixty years after the first edition of Walras' *Elements*, while Arrow and Debreu developed it in 1954.

3.1.2. A preliminary overview of the new conditions guaranteeing Walrasian value theory: changes in the unit of analysis, in the ontology of material objects, and in the relationship between the object and the subject

What the contributions of Barone, Wicksteed and von Neumann were indirectly saying can be simply summarised as follows: unless some extra-assumptions are made, it is just not possible to address both dimensions of classical value theory at once, while still providing a meaningful economic solution, if the classical premises in terms of actors' behaviours (time and causality), and world's ontology (space) are retained. Given the classical premises of *appropriable material wealth*, value will never be, at the same time, *subjective* and *endogenously* computed at the *social* level (a monism). This is just not logically sustainable for the simple reason that the private ownership of appropriable material goods, and social monism in valuations are incompatible.

Either the assumptions on the ontology of assets is modified, or the assumptions on the way in which these assets can be employed if privately owned are modified. Otherwise the Walrasian "dream" of private ownership of assets coupled with monism and objectivity in valuations is simply not sustainable.

While the 1890s marginalist revolution in production and distribution theory changed (for ever) the way in which the *material ontology* of the world is modelled by economics by changing the unit of analysis from the material, physical assets, factors and resources to the immaterial flowing services, the duality revolution of axiomatic set theory and convex analysis of the 1940s-1950s changed (for ever) the way in which private actors can use and employ their own material property by introducing the "much cited but little used" behavioural concept of rationality. What will be discarded in this case will not be the material spatiality of the physical world (space), but the causality relationship between the private and the public domain (time).

Both will have to be implemented in order to have an economic theory of value that is not just correctly spelled from a descriptive point of view, but that is also useful in normative terms. The final outcome, will be the space-less and time-less singularity in which neoclassical economics has been precipitated since.

It is in this specific point in time that, in order to address these two shortcomings of the Walrasian framework that prevented the development of meaningful economic conclusions based of a theory of value that could be *subjective and monistic* (endogenously computed) at the same time, the role of institutions in economic analysis (and asset ownership or private property in particular) will be completely erased, and made completely dependent on, and subsumed to, the performance of the market mechanism. In a sense, it is possible to say that in a singularity where time and space do not matter, institutions do not matter either.

3.1.3. Overview of the argument: conceptual circularity and logical paradox inherent in any failure approach to institutions and private ownership relying on a monistic and objectivist (social) theory of value

As the section will show, to assert the irrelevance of institutions by imposing monism and objectivity in valuations is only an illusion: the ontological and behavioural assumptions introduced into Walrasian theory of (use) value between 1890s and the 1950s in order to elevate it to an instrument for normative value judgment on how privately owned resources ought to be socially employed are themselves relying on very precise normative constraints on actors' behaviours with respect to how the latter ought to employ their private property in the first place.

The misunderstanding of the real nature of the assumptions super-imposed on Walrasian theory of value through the decades has generated the strange paradoxical assertion (retained in both institutional economics and law and economics in general) that sounds as follows: institutions (and asset ownership) become irrelevant in order to normatively judge how resources ought to be socially employed whenever specific normative restrictions are super-imposed on private actors by the part of the institutional context.

In other words, the aim of the section is to show that any neoclassical (i.e. failure-based) approach to institutions concludes that extra-market (legal) institutions become irrelevant whenever certain conditions that can only be guaranteed if some legal institutions regulate actors' behaviour hold in the first place. The paradox derives from the dismissal of the role of the legal environment by relying on preliminary conditions and premises that can only be guaranteed if a legal environment is already in place.

The takeaway of the section will therefore be the following: *no matter whether the valuation problem is affected by "failures" (in the form of allocative or behavioural) or not, in either case the legal and regulatory context will always have temporal, logical and analytical priority over the competitive market mechanism whenever material (physical) assets are privately own*. As private ownership of physical assets always matters, the legal environment regulating private ownership always matters in turn (and no matter what).

The reminder of the section will be spent to briefly illustrate why this is the case, to justify this claim, to show the implications and to provide alternative solutions consistent with the arguments developed in chapters 2 and 3 above. The aim is to put the study of institutions in economics on firmer grounds, so to better understand the legitimacy of certain normative conditions justifying certain institutional arrangements, and to better understand the role of regulation, which will be much better developed in chapters 7 below.

3.2. The ontological condition derives from the marginalist response to Ricardian production and distribution theory: changing the unit of analysis of value theory from material assets to flowing services

The marginalist revolution in production and distribution theory that occurred in 1890s had the precise aim to dismantle classical (Ricardian) distribution theory, which also represented the cornerstone of the entire Ricardian edifice.[69] The revolution can be decomposed into two contributions: one reforming the way in which distribution had to be intended, the other reforming the theory of production.[70]

3.2.1. Ricardian distribution and production theory: key features, outstanding problems and marginalists challenges

First, we address the issues emerging in distribution theory. There were many aspects that resulted problematic (at the eyes of neoclassicals) in Ricardian theory of distribution, but one was perceived as particularly serious: the fact that factors' owners could command surplus, or rents, that were generated by the social system independently from, and outside, the economic mechanism.[71]

Surplus in the Ricardian theory of distribution have two peculiarities. One, they reward the private ownership of physical factors, such as land and capital assets and, two, they are price-determined, meaning that the remuneration (or monetary compensation) of these factors does not depend on the pricing system, but it is a left-over or, a residual. From which, the term surplus. Moreover, as in the classical production-cost approach (which took over soon after Ricardo's death), the sole factor enjoying surplus remained land, since then the two terms (surplus and rent) became synonyms.[72]

This peculiarity of classical distribution theory derived directly from its value theory: factors enjoying rents were those same factors that did not participated to the value process in the

[69] As remarked by Blaug (1958: 3): "The heart of the Ricardian doctrine consists of the proposition that the yield of wheat per acre of land governs the general rate of return on invested capital as well as the secular changes in the distributive shares … It is the presence of this element, rather than the Law of Markets of the special theory of value which Ricardo employed to obtain his results, which identifies a 'Ricardian Economist'."

[70] The main authors leading the revolution are Barone (1896), Clark (1891, 1899), Edgeworth (1904), and Wicksteed (1991). For a historical overview of the marginalist revolution in production and distribution theory the reader is referred to, among many others: Blaug (1996: chapter 11), Hicks (1932), Kaldor (1956), Robinson (1934), Scitovsky (1964) Stigler (1941).

[71] Overviews of Ricardian theory of distribution can be found in Blaug (1958: chapter 1; 1996: chapter 3 and 4), Buchanan (1929), Cannan (1903), Edelberg (1933), Pasinetti (1960), Peach (1993: chapter 3), O'Brien (2004: chapter 5), Schumpeter (1954: 645-679).

[72] And yet, interest on capital has long been considered a surplus. It is not clear whether rent is the only factor left enjoying a rent because it was immobile, or because it did not suffer any "pain" in participating to production, differently from labourers and capital lenders.

production side, where value has to be intended in production-cost terms. In other words, “costless” factors were those factors that could be rewarded above the costs that they experienced. From which, a surplus was created in the form of residual.

The introduction of the Austrian concept of opportunity cost provided the first, and most important, key to solve the puzzle: the only reason why factors such as land “appear” to face no cost and can be ignored derives from the fact that in the classical production-coast approach these factors are treated as rigid, indivisible, and unique entities, which can only be fully dedicated to one single specific employment (Samuelson, 1959): as soon as it is understood that any factor can be simultaneously employed in alternative employments, the cost sustained in their employment will be apparent as their opportunity cost.

In other words, whenever the classical production-cost approach is replaced by the new marginalist approach based on the concept of production cost, any factor that can be alternatively employed in more than one employments at the time will never be perceived as “rigid” and inelastic, and therefore costless, in that specific employment, unless there is no opportunity lost. This means that rents in an opportunity-based analysis only appear for dedicated assets, or factors, and this makes classical distribution theory (with respect to land) a theory of specific assets or, as Robbins (1970: 14) puts it:

> “If we are prepared to assume a more or less homogeneous labour supply, tending to multiply to some conventional subsistence level; if we regard land as having almost complete specificity in one agricultural use, … then it is not difficult to make out a case for many of the characteristics of the Ricardian propositions … at least in terms of logical consistency.”

On the other hand, with respect to the production side, the shortcoming of classical economics was even more basic: there was no theory of production whatever. At least there was no theory of production as we now intend it: as the combination of various inputs that can be combined in varying proportions among each other’s. Production in classical economics occurred through rigid factors employed in fixed proportions.

As a result, with respect to production theory it was just sufficient to allow for variable coefficients of production that could allow for continuous substitution among the various inputs as described by Cobb-Douglas (1928).

3.2.2. *Change in the unit of analysis in the switch from Ricardian to neoclassical production and distribution theory: from physical "things" to flowing services*

The common solution to both these problems was represented by a change in the unit of analysis of the economic problem: not factors, assets or physical resources anymore, but services. The adoption of variable and flowing services as the new unit of analysis all economic variables and concepts should refer to changed completely the picture and could offer two solutions at once. For these reasons, the adoption of services has been regarded by all marginalists as the *pons asinorum* of the marginalist revolution in production and distribution.

This point is so crucial, and its implications have been so overlooked in recent times, that is necessary to cite at length.

Wicksteed (1991: 54-55; emphasis in the original), in the pamphlet that will represent the cornerstone of the marginalist revolution writes:

> "In investigating the laws of distribution it has been usual to take each of the great factors of production such as Land, Capital and Labour, severally, to enquire into the special circumstances under which that factor co-operates in production, the special considerations which act upon the persons that have control of it, and the special nature of the service that it renders. And from all these considerations to deduce the special law regulating the share of the product that will fall in distribution to that particular factor … [However] it is difficult even to conceive any calculus by which the share of land and the share of capital could be added together and an investigation then instituted as to whether the residual share will coincide with what the theory assigns as the share of wages … The basis of those laws is being sought not in the special nature of the services rendered by the several factors but in the common fact of *services rendered*."

In the same way, this is how Walras introduces his chapter on the general economic equilibrium with production in his *Elements* (1954: 212: lesson 17; emphasis in the original):

> "The elementary factors of production are three in number. In listing these factors, most authors employ the terms: *land*, *labour*, and *capital*. But these terms are not sufficiently rigorous to serve as a foundation for rational deduction. *Labour* is the service of human faculties or of persons. We must rank labour, therefore, not with

land and capital, but with *land-services* [‘*rente*’] rendered by *land*, and with *capital-services* [‘*profit*’] rendered by *capital goods*.”

The switch in the unit of analysis is so important for the marginalists purposes that Knight considers the employment of material factors one of the main “aberrations” of the classical system (Knight, 1935: 4-5):

> “The second main aberration [of the Ricardian theory of production and distribution] is a false conception of production, another aspect of which is a fallacious view of the nature of wealth or capital. Production was defined as production of wealth. But in fact, primary production consists in the rendering of services. Wealth is an agency by which services are rendered, not a product in the primary sense. (As a quantity it is abstract capacity to render service, not a collection of concrete things.)”

And again (Knight, 1935: 8):

> “What is called ‘distribution theory’ has to do with the pricing of productive services … What is in fact consumed in economic life is exclusively services, and accordingly, the primary meaning of production is the rendering of service.”

Similarly, Stigler (1941: 2-3) highlights the fact that the most fundamental shortcoming of production and distribution theory before the marginalist revolution was the inability to operate a change in the unit of analysis, towards services:

> “The branch of economics which was in most urgent need of reformulation was, in fact, distribution. In 1870 there was no theory of distribution … Extended criticism is unnecessary at this point; the fundamental defect was clearly the failure to develop a theory of prices of productive services.”

This list is probably enough for the present purposes. Yet, nearly any economist wiring in the first three decades of the 20th century made this concept very explicit in his writing in order to remark the essence of the departure from the old Ricardian framework.

The switch to services as the new unit of analysis of value theory has several advantages. Three are worth of mention.

One, services are flows, not stocks. In other words, they are derivative or, intensities, and cannot be frozen or accumulated, nor appropriated. Services disappear and are “consumed” in the same exact moment in which they are used or rendered. This implies that any transfer or exchange or services does not entail any actual transfer of ownership or property. As we will discuss in the next chapters, only use rights are involved, never actual property.

Two, services are defined only as intensities changing over time as they are derivative of certain quantities, but they are not defined in space. In other words, services can be infinitesimal. This implies two things: that the underling "support" the actual material factor can be used at infinitesimally varying intensities, and that production can recombine any service, at any chosen intensity. In other words, the supply of services is always infinitely elastic, due to the inherent ontological characteristics of a service.

This second point not only means that the problem of factors' dedication that characterised Ricardian economics will never appear, but it also means that inputs in production can be recombined in infinitesimally varying proportions, as any "smooth" curve requires.

With respect to this point, it is interesting to re-examine the well-known marginal equivalence characterising welfare (Pareto) optimality in welfare economics:

$$\frac{\partial F_i/\partial q_c}{\partial F_i/\partial q_j} = \frac{dq_j}{dq_c}$$

Above, the marginal contribution to production of factors c and j, defined respectively by the two marginal rate of transformations $\partial F_i/\partial q_c$ and $\partial F_i/\partial q_j$, depends on the two factors' marginal rate of substitution in the production process (dq_j/dq_c).

It is important to note that only derivatives with respect to quantity enter relation. Derivatives are intensities or, infinitesimally varying flows. In other words, factors' services enter the welfare optimality condition, never the actual factors. These physical analogies were particularly stressed by Knight (1935c: chapter 6).[73]

These first two peculiar characteristics of services were already enough to address the two main problems embedded in Ricardian production and distribution theory, and to dispose it once for all.

There is also a third characteristic of service that shall interest us here: services correspond to the way in which the material factor is employed. In other words, they are the "visible"

[73] And for a treaty on the commonalities between physical concepts and early marginalism, see Mirovsky (1989).

outcome of an economic decision, but they never characterise the underlying input that remains the underlying material support. As a result, the switch to services is also a switch to the dual of the material assets, intended as inputs: marginalist productivity theory is an output-based theory, is not in input-based theory.

The switch from material factors (which can be appropriated) to immaterial services (which can never be appropriated) is the first main change to the original Walrasian framework. Here we define it as a change in the *ontology* of the old frameworks.

3.2.3. *Material assets and not services: the fundamental reason behind the distinction between Walras endogenous value theory and Menger exogenous theory of (subjective) use value*

It should be noted that the change in the unit of analysis, and the inevitable replacement of appropriable material factors or assets with the concept of services was already implied by the framework developed above.

The ontological "flaws" already perceived by the second generation of marginalists can be interpreted as follows: given Walras' initial position in the "top-left corner" of figure 4.4, as long as physical, appropriable, assets were considered, self-consumption through the accumulation of private (material) wealth could not be ruled out, and the framework would have inexorably slid right, towards dualism.

The presence of material "things" defined in space and over time always introduces a decoupling in the valuation framework between use and exchange value, and therefore between the private and the social layer. Any material thing that can be accumulated, stored, and appropriated will always provide to its owner the opportunity to be employed "in use", for self-subsistence. As already recalled, this will, alone, generate a decoupling between the private domain defined by auto-consumption, and the social domain defined by the "public consumption" of the same "thing". For this reason, private ownership of "appropriable" goods necessarily leads to the "weakness" of value theory that is its dualism.

Given the unsustainability of Walras' original position, with no assumptions on the ontology of the object considered, the original Walrasian framework would have necessarily slid right-ward as figure 4.5 shows, and consistent with what was illustrated by figures 3.7 and 3.8.

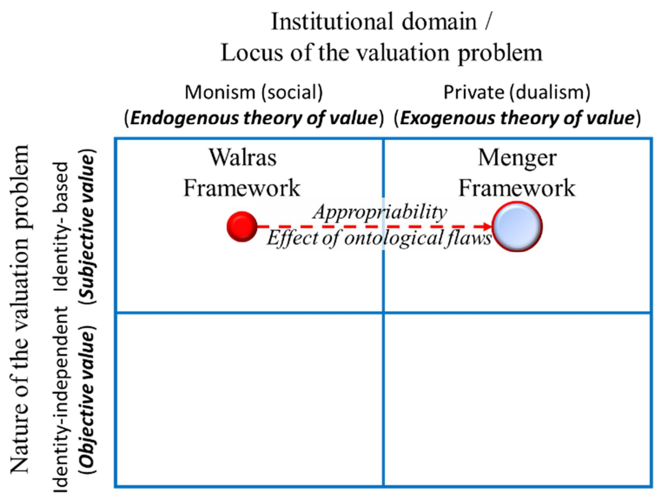

Figure 4. 5 – The presence of appropriable goods undermines the original Walrasian framework

Figure 4.5 shows that a Walrasian framework where actual material factors are involved in the economic process slid rightward, towards the case of exogenous value (valuations not necessarily determined by the exchange, or price, mechanism). This was precisely were the Mengerian framework was located that, not surprisingly, always retained the classical notion of material good or factor, and never switched to an ontology of immaterial services.

This insistence of material factors was, however, also Menger's major flaw at the eyes of later neoclassicals so that Stigler (1941: 137; emphasis in the original), among many others, boldly asserts:

> "Menger follows the classicists, however, in failing to distinguish between *goods* and *services from goods* on the basis of the time dimension involved …"

So that (p. 154):

> "One general weakness in Menger's exposition which clouds his value theory is the failure, previously mentioned, to differentiate between durable goods of any order and their services".

The comparison between the Mengerian and the new Walrasian framework highlights an important point: any theory of value founded on the opportunity cost concept (as it is the case for both Walras and Menger) necessarily entails dualism and subjectivism in value theory unless some specific ontological restrictions are imposed. These restrictions require that *the underlying object (the opportunity-cost schedule applies to) is never directly involved in the economic process, but it is only indirectly involved through the provision of services*.

Differently from the case Walras, factors and assets were always *directly* involved in the economic process in Menger (and early Austrians in general).

3.3. From material assets to flowing services: discussion of some fundamental implications and shortcomings

The switch in unit of analysis from the materiality (spatiality and temporality) of the physical things to the immateriality of the services represented the key factor in order to complement the marginalist revolution in value theory with the marginalist revolution in production and distribution theory.

This change has some key implications, and shed light on some fundamental issues, especially with respect to the role of rights and ownership in economics.

Two remarks are particularly relevant in this respect.

3.3.1. There is no need to switch to services in order to justify monism in value if assets are specific: legitimacy of the English approach in production and distribution

One, the assertions above also help to understand why English marginalism could develop its economic analysis without worrying about the problems posed by the ownership and direct employment of material factors.

Consistent with their partial-equilibrium approach, and consistent with the production-cost approach, English (Marshallian) analysis employed no opportunity-cost analysis. This makes English marginalism perfectly in line with the Ricardian approach to production and distribution, where assets are specific. Assets are specific because once that they are owned (appropriated), they can only be employed in one employment, in order to render one given service in one single market.[74]

For this reason the sole inconvenience arising from the institution of asset ownership in Marshall's production cost analysis is the emergence of rigidities and indivisibilities in supply, leading to increasing returns in production (non-convexities in production), and therefore, ultimately, monopoly rents. Marshallian monopoly rents are not the same as Ricardian rents, but the principle they rely on is substantially identical: inelasticity in the supply of a factor's services deriving from the absence of an opportunity-cost analysis. This is the unavoidable implication of ownership for any production-cost framework in partial equilibrium: the presence of failures in the form of non-convexities.

[74] These can also be set of services sold in a set of markets, without that this changes the sense of the statement. The key point is that the Marshallian framework does not frame any allocation problem leading to possible substitution across alternative services.

Conversely, the absence of an opportunity cost schedule framing the way in which properties are employed has two main advantages: one, it establishes a one-to-one mapping between the way in which an asset is employed (service rendered) and the asset itself and two, it rules out any possible subjective redeployment of the asset "in-house" for auto-consumption leading to subjectivity and dualism in value.

This is consistent with what has been stated in chapter 2 above as "asset specificity" (or partial equilibrium analysis) has been included among those peculiar conditions that render asset ownership irrelevant, even when possible and legitimate.

3.3.2. Services vs. assets and decoupling between use and ownership rights

Two, the fact that factors, resources, assets and material things in general (any discrete object defined in the space-time) are never "directly involved in the economic process" means that they are never the object of a transaction, the object of an exchange, nor the object of a valuation (pricing) problem. Conversely, services are transacted in the market, are exchanged among actors, are priced by the market, and services are purchased, organised and employed in the production process, as we saw above. Physical things are never directly involved in any economic activity, they simply provide the underlying "material support" for the economic activity, and they operate "remotely".

This condition has a fundamental implication: there is no transfer of actual material property in the Walrasian framework, nor of ownership rights. Property is "static" and the ownership rights of the material appropriable things are never at stake in the economic system. And this means that the ownership over the various physical things remains unchanged between the beginning and the end of the economic process.

Conversely, what does change through the economic process is the distribution of rights over the *employment* of these things. *Use rights* over the things are transacted, exchanged and priced, but never the rights over the ownership of the things themselves.

The distinction between ownership and use rights will be the topic of the next chapter.

The fact that appropriable goods cannot be transacted does not prevent for monetary goods to be transacted. Similarly, money is exchanged and transacted in the same way in which any "intrinsically social" good can be transacted and exchanged.

3.3.3. Services vs. assets and the switch from material to financial wealth

This highlights another important aspect of the new Walrasian framework: the switch towards monetary, or financial, wealth at the expenses of material wealth: the provision of services (or use rights) generates income for those providing the services, while the purchase of services (or use rights) generates current (operational) expenses. In sum, circulating capital, or income is involved, but never material (physical) and stackable capital.

This also means that the concepts of distribution (and redistribution) have to be intended as applied to money, not to material things.

The previous characteristics of the Walrasian framework are summarised by figure 4.6.

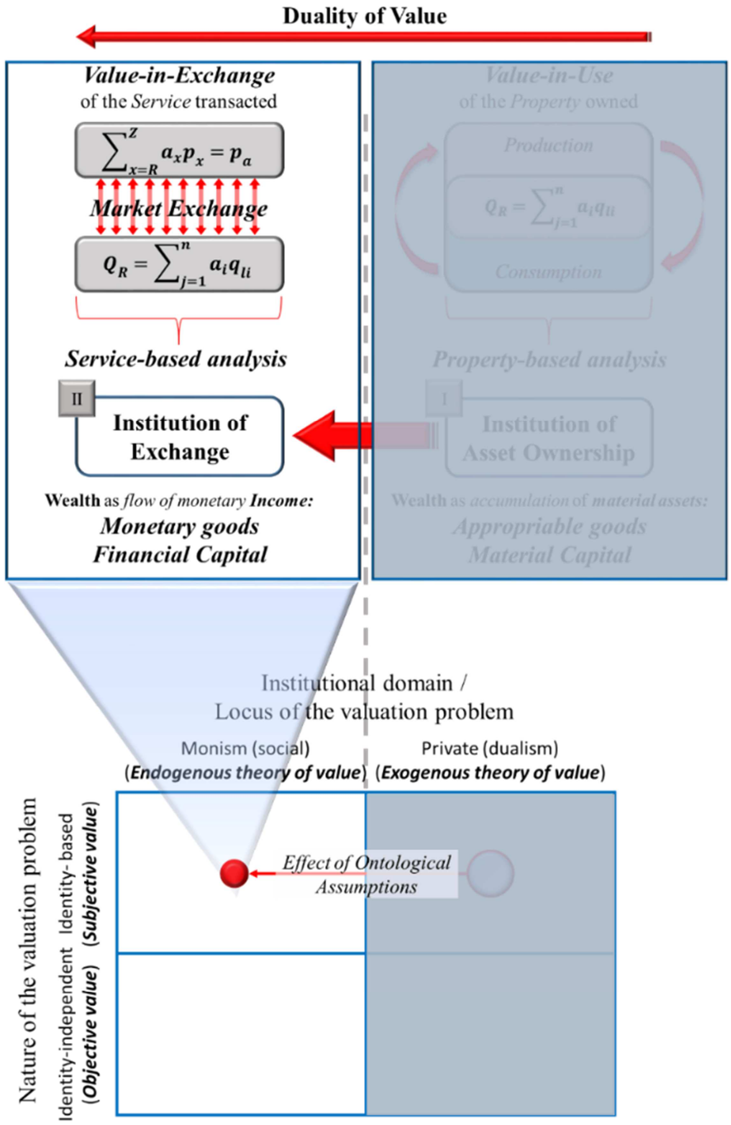

Figure 4. 6 – The Walrasian framework after the imposition of the ontological assumptions in the form of the shift in the unit of analysis from the material things, to the flowing services

Figure 4.6 provides a synthesis of all the key characteristics of the Walrasian framework characterised by the new ontological assumption: the shift in the unit of analysis from material things to services.

3.3.4. Walrasian subjectivist and endogenous value theory in the shift from assets to services: limitations and open problems

Two further points that emerge from the figure should be pointed out.

First, the figure shows what the characteristic of the Walrasian framework that makes the "ontological restriction" operational is: the clear-cut separation between the locus of the opportunity cost schedule and the locus of production. The latter in the Walrasian framework is still represented in production-cost terms, as in Marshall.

The fact that those actors that have to produce are only characterised by the production schedule ($\sum_{x=R}^{Z} a_x p_x = p_a$), and never internalises the opportunity cost schedule ($Q_R = \sum_{j=1}^{n} a_R q_j$) entails that the resource characterised by the productive capacity Q_R is never appropriated, and can only interact with the producer through the provision of services.

One the one hand producers are univocally characterised by the output that they produce and sell (p_a) in line with a Marshallian partial-equilibrium framework, where a producer is identified by what he produces, never by the opportunities that that opportunity-cost schedule presents him. On the other hand, producers are forced to purchase services in order to undertake production (a_x), so that the exchange mechanism only exchanges and prices services (p_x).

In sum, the clear decoupling between the locus of the production cost schedule (exclusive domain of the producer) and the locus of the opportunity cost schedule (exclusive domain of the assets' owners), and the impossibility to generate any overlap between the two is the essence of this first ontological assumption. The result is that, consistent with what has been discussed above, the only entities involved in the economic process are the services (prices p_x and quantities q_a).

This has a major implication: *in the Walrasian framework producers do not own any physical property*.

From which, a slightly different assertion follows: *in the Walrasian framework, producers that purchase use rights over certain material assets can never also exercise ownership rights over those same assets*.

From which it also follows that in the Walrasian framework producers will never be identified and characterised by the assets that they own, or by the opportunity cost schedules that they face, but only by the use that they make of certain assets (how they employ certain given assets).

As a result, in Walras the institutional domain of private property (private domain) does never overlap with the social domain of production and exchange. The separation of the locus of propriety from the locus of production is the necessary precondition in order to rule out dualism in value theory *from the producer point of view*.

The separation of ownership rights from use rights in production does not exhaust all the problems. It addresses and solves only one half of the problems inherent in the Walrasian framework. There is a second half of the coin that still remains a potential source of troubles. This represents the second important remark emerging from the figure.

There is a problematic omission in Walras' framework: he does not characterises the private domain or, the opportunity cost schedule. This is a serious problem as the opportunity cost schedule is precisely that device that characterise the presence of subjective valuations associated to the ownership of a property.

Walras does not even say who own the factors, what their identities are, nor how these owners behave. In other words, the fact that producers are characterised by the Marshallian production-cost schedules is not informative of how the respective assets' owners are actually behaving and employing those assets.

This problem can be reframed in a significant way as follows: the fact that producers cannot be the suppliers of their own services (they cannot auto-consume the services of their own property), does not imply that the property owners (whoever they are) cannot also consume the services of their own property in-house.

This means that the Walrasian framework is incomplete. It still necessitates some other assumption that ensures that auto-consumption of private properties for self-subsistence is ruled out *tout court*.

In a sense, the mere imposition that producers only purchase services does not completely solve the problems embedded in the institution of ownership, but it simply avoids them by removing them from the picture. Appropriable assets were, of course, not ruled out. However, their private employment was just placed outside the realm of the new framework, which only focussed on the valuation of transacted and exchanged services, and focussed on their purchased and employment by the part of producers.

This way to solve the duality of value reproduced exactly the way in which classical economists ruled out use value, and disposed of the paradox of value: they recognised and accepted its existence, yet use value was brought outside the realm of economics. In a sense, this is exactly what the decoupling between the production cost schedule and the opportunity cost schedule represented: a way to leave material wealth at the margin, and to focus the analysis on the former. Monism in value was not really guaranteed.

It is for this reason that in the figure does not provide the standard monistic and bidirectional framework. This is consistent with what was discussed in chapter 3 above: logically, under monism in value theory it does not make sense to define “subjective valuations” as there is no subjectivity in the first place. Monistic valuations are always identity-independent, as it is well-known from the standard microeconomic framework. Moreover, it is for this reason that in the Walrasian framework with no additional assumptions it is not possible to establish a directionality from the material property to the services and vice versa. As there are no assumptions guiding the behaviour of the ultimate resources’ owners, it is not really possible to know with certainty how properties are used by their respective owners.

For these reasons, as subjectivity in valuations by the part of the property owners cannot be ruled out, a univocal, and causality-free, mapping form the transacted services to the underlying things cannot be established. And for this reason the original “marginalist” Walrasian framework is located in the top-left spot, not on the bottom-left, where it should, in case it can ensure that no subjectivity in valuations is allowed in the picture.

The solution will be the imposition of another set of assumptions, this time behavioural, in the form of “rational behaviour” of asset owners.

Figure 4.7 anticipates this last logical jump, to which we now turn.

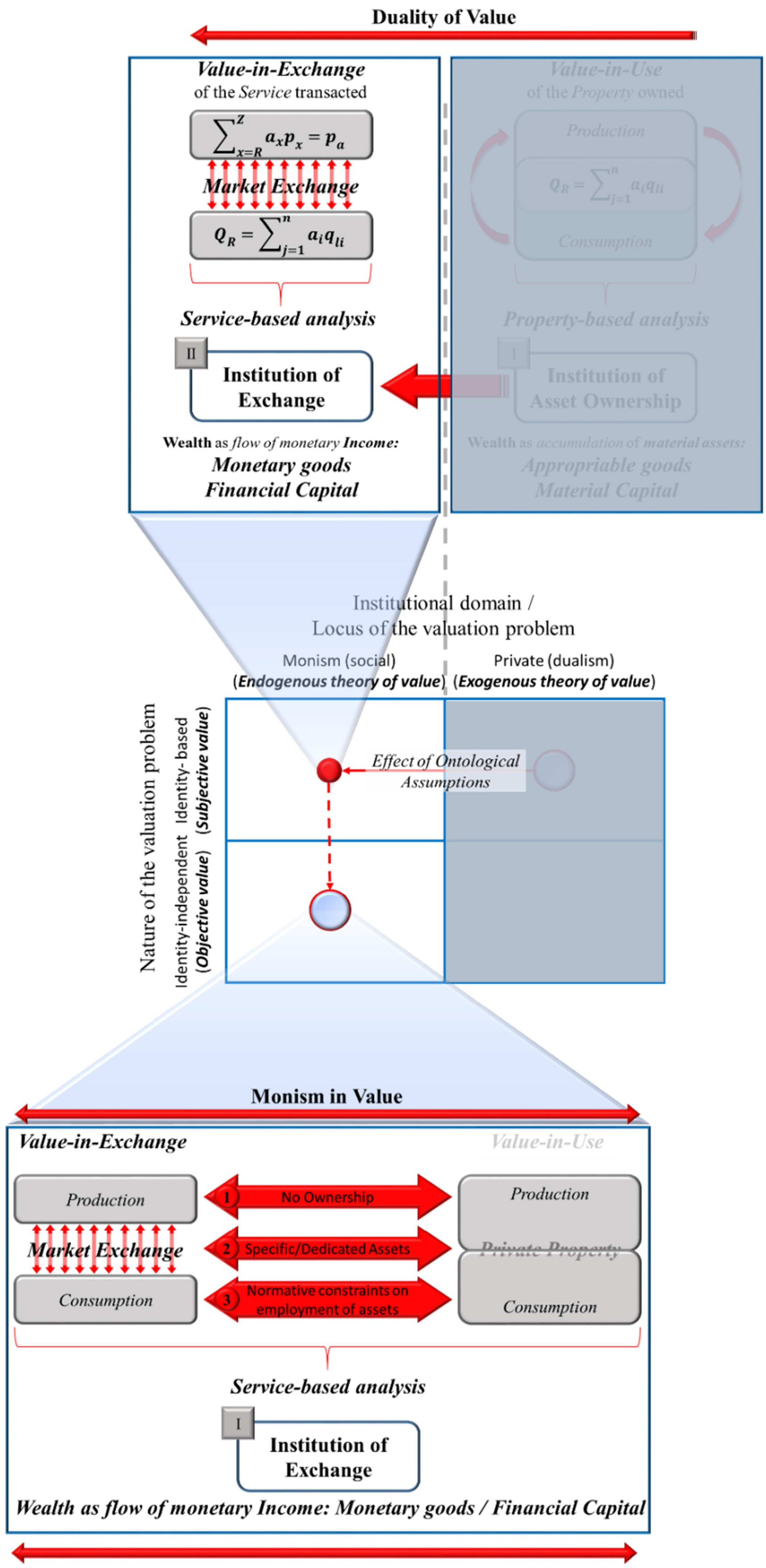

Figure 4. 7 – General competitive analysis: form the ontological assumption to the behavioural assumption in order to guarantee objectivity and unicity (monism) in value judgements

3.4. The behavioural condition derives from the need for a normative theory of value: prescribing a univocal mapping from the monetary wealth generated by the services to the material wealth generated by private appropriation of physical assets

The original Walrasian framework with services as unit of analysis was a *descriptive* model of how flowing services are priced in the market, but it could not have any normative, or prescriptive content. With no extra-assumptions the model could not even be solved, nor could it guarantee the existence of any solution, let alone its uniqueness.

What was still required was that the underlying material things from which these services derive, and that were artificially kept outside the picture, could be univocally identified with the services that they ultimately provide, and that are valued and transacted in the market.

In other words, it was not sufficient to have only services transacted, exchanged and employed by the actors, but *it was also necessary to express all forms of wealth in service-terms*, so that it is always possible to assert with absolute certainty that all services priced (valued) and transacted by the social mechanism can univocally retrieve (backward) all the underlying material (stocked) wealth from which they originally derive. These new assumptions are necessary as once that material things are brought outside the realm of the economic analysis, it becomes necessary to make sure that they will effectively be employed "as means of exchange" in the form of services, so that the valued services can really be considered the correct measure of all underlying wealth possessed and cumulated by the various private actors.

This is the essence of the normative content of the Walrasian framework, and this is also the essence of the behavioural assumptions that, starting with von Neumann, were introduced in the original Walrasian framework.

The role of these assumptions is to impose restrictions on the behaviour of any private owner on the way in which the latter can employ and conceive his private wealth: not as a stock of *appropriable* goods but only and exclusively as monetary goods. If all the private wealth possessed and cumulated by any individual cannot be privately employed and auto-consumed "in-house", in the same way in which money cannot be, then the observed services transacted on the market, and valued by the market, necessarily exhaust the underlying stock of wealth privately possessed by the various individuals.

In sum, the switch to monetary goods allowed the Walrasian framework to eventually move downward, in the endogenous-objective spot (figure 4.3). By making all privately cumulated wealth "intrinsically social", the new Walrasian value theory achieved the astonishing result of being at the same time *objective* (social) and *endogenously* computed internally (unique: monism), and therefore also *normative*.

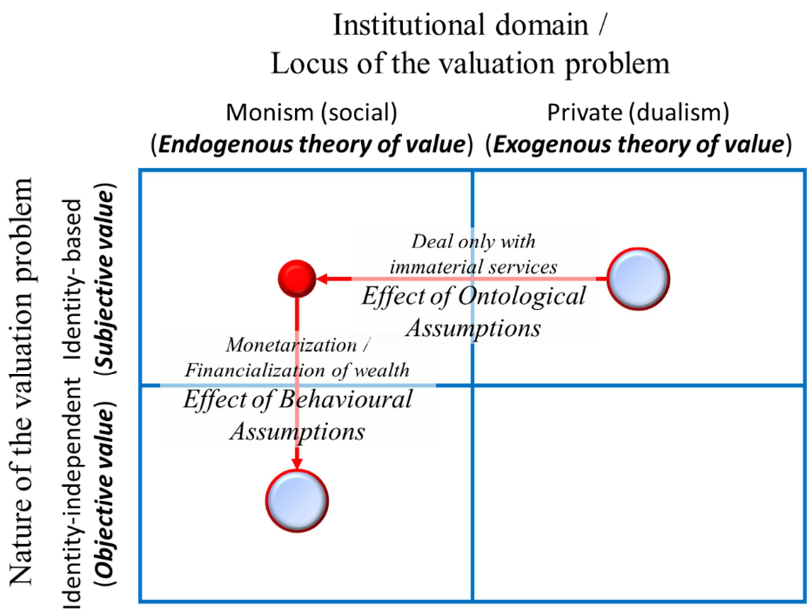

Figure 4. 8 - Overview of the logical steps that ensure monism (uniqueness) and objectivity (social nature) of value in the final vNAD Walrasian framework

The vNAD framework is a normative Walrasian framework that contemporaneously ensures *endogeneity* (uniqueness) and *objectivity* (identity-independence) of value. It is not a case that the well-known fundamental proprieties of the general-equilibrium framework are its uniqueness (rather than dual) and its social (rather than private) nature: these were the two open issues of classical value theory, and these were for the first time contemporaneously ensured by vNAD value theory.

These achievements came at a cost; ontological and behavioural costs. The old classical framework where appropriable material and physical things (assets, resources, factors) could be accumulated, stoked and used for private consumption left the picture. At their place were introduced "monetary" goods that:

1. could only be used in order to provide services (had to be completely transformed into services: ontological assumption);
2. could only provide services by means of market exchanges to third parties upon an equivalent monetary compensation (had to be completely employed on the market as means of exchange: behavioural assumption)

The first assumption ensures monism, the last ensures objectivity or, independence from actors' private subjectivities.

It is fundamental to appreciate the essence and the implications of these new assumptions whose effect is to act and modify (1) the *ontology* of the *objects* possessed, employed and transacted by the *subjects*, and (2) the *nature of the relationship* allowed between the object and the subject (assumption 2).

The two assumptions are usually referred to as, respectively, the (1) convexity assumption and the (2) monotonicity (non-satiation) assumption. While the former ensures that there are no physical "indivisibilities" creating non-convexities, thus increasing returns to scale, the latter ensures that value-in-exchange computed by the market (market prices) is indeed existent, unique, and optimal from a *normative* point of view Pareto-optimal. The two also correspond to the assumptions that have to hold in order to ensure, respectively, the second and the first welfare theorem.

As the above discussion illustrated, the assumptions ensuring social and endogenous values are assumptions on the *ontological nature* of the objects employed and transacted, coupled with an assumption on the way in which these objects ought to be regarded and employed by the subjects.

4. Concluding remarks and final discussion of some fundamental points

Three key points emerge from the chapter.

The first relevant point of the chapter deals with the definition of ownership emerging from the discussion above. Asset ownership is always defined by an opportunity-cost schedule. These opportunities must define either options "in exchange, or options "in use". It is important to note two things about the parallelism between the two options.

One, the two options define the two ways in which a property can be used: as means for self-consumption or as a means for exchange. Two, the two correspond to two different definition of asset ownership in law and economics.

On the one hand the internalization of an opportunity schedule represents the faculty to exercise control over the resource, and to decide how the resource ought to be employed among alternative potential employments. This is a definition of ownership that relies on the concept of control. In case of a dynamic framework, this also correspond to the definition of ownership provided by Transaction-cost Economics: as the faculty to decide how a resource ought to adjust among alternative options ex-post.

On the other hand, the internalization of the opportunity schedule represents the faculty to employ the resource for private purposes, and therefore to exclude others from establishing transactions and dealing with the resource. This is the definition of ownership intended as the “right to exclude”, and “refusal to deal”, mentioned in chapter 2 above. In this case the relevant option would be “internal”, thus in self-consumption.

The third possible definition of ownership, intended as the rights to appropriate residual rents from the employment of the resource cannot be merely appreciated by looking at the opportunity-cost schedule, but has to be coupled with a production program.

The second relevant point discussed in the chapter deals with the normative foundations of the economic framework, and the nature and the implications of the ontological and behavioural assumptions that shape the institutional framework in the neoclassical framework.

A world where material assets are employed, and where these material assets can be appropriated and accumulated by the part of private actors, the legitimacy of the normative conclusions rest on the legitimacy in the switch in the unit of analysis of the economic framework: the underlying material support has to be abandoned, and it has to be replaced by an analysis of the services operated “over the top”.

The following argument is made: *it is legitimate to derive normative conclusions with respect to the way in which privately owned physical things ought to the employed if and only if it is possible to univocally retrieve the nature of any underlying material “wealth” by merely looking at the way in which actors transact its services in the market*.

In other words, normative conclusions on the way in which the cumulated material wealth of a society (assets, resources, and factors) ought to be employed can only be based on the observation of services, and are only legitimate if it is possible to approach these material things “as if” they were pure services or, only if a service-based analysis is completely informative on the nature of the underlying material supports (the assets). In this case, the private ownership of private properties becomes irrelevant.

The reason why a switch to services is necessary in order to derive normative conclusions derives from the fact that it rules out the subjectivity in value that is necessarily introduced whenever self-consumption becomes possible.

This condition that allows to retrieve materiality from exchange services is guaranteed by the imposition of conditions affecting the ontology of the underlying assets, and the behaviour of the subject dealing with the object.

On the one hand, a service-based analysis relies on a sharp distinction between services and assets, which is also a distinction between use rights and ownership rights. The distinction of the private ownership of the material assets from its employment by the part of third parties guarantees that auto-consumption is ruled out, and that the private domain loses economic meaning. This also guarantees that the identity of the actor providing the underlying material support does never correspond to the identity of the actor employing and exercising use rights over the support itself.

As a result, value becomes a monism, and the private sphere ceases to be economically relevant whenever the underlying material support is only a passive support for the provision of services to third parties subject to monetary compensation.

The failure to sharply separate ownership rights from use rights in employment is usually interpreted as a violation of the "compensatory assumption" or, as the presence of allocative failures. As a result, it is important to understand what the concept of "market failure" really entails: the social system is considered to fail whenever an analysis of the services deriving from a given asset does not allow to univocally retrieve (backward) the nature of the underlying property employed in order to render those services. The expression of allocative failures denote a situation where a service-base analysis cannot satisfactory replace an asset-based analysis.

On the other hand, a service-based analysis also relies on an ontological assumption, which is the outcome of an underlying behavioural assumptions: actors can only treat their own properties as if these were "intrinsically social" goods or, stated differently, "monetary goods". This second way to interpret the legitimacy of a service-based analysis is usually relabelled the "rationality assumption", which is the precondition to "behavioural failures" in the standard analysis of institutions.

It is important to understand what the rationality assumption really entails when it comes to the institution of asset ownership and the way in which the latter is treated in the standard framework.

Three strictly linked interpretations have to be provided.

One, the behavioural constraints imposing rationality on asset owners correspond to socially super-imposed constraints preventing actors from using their own properties as means for self-subsistence. Actors are considered rational whenever they can only use the properties for the satisfaction of third parties' needs and, vice versa, whenever they can only satisfy their own needs by purchasing services from third parties' properties. In other words, rationality imposes the absence of self-consumption.

Two, the behavioural constraints imposing rationality on asset owners correspond to socially super-imposed constraints making sure that actors treat their cumulated material wealth as if it was mere financial capital. This aspect of the rationality condition has a specific effect also on the ontology, or the nature of the object employed: rationality requires that any appropriable good is treated as if it was a monetary good. Or, stated differently, rationality requires that any form of material wealth is treated by its respective owner as if it was monetary (financial) wealth to be sold in the market.

Three, there is also another subtle implication of the rationality constraint: property owners cannot refuse to deal, nor exclude third parties from exercising use rights over their own property.

Consider for instance Barone (1935: 248) statement:

> "Let us begin with individual budgets. It is convenient to suppose – it is a simple book-keeping artifice, so to speak – that each individual sells the services of all his capital and re-purchase afterwards the part he consumes directly."

The same "artifice" is employed by all other works in welfare economics in general. The essence of Barone's argument is clear: although it is true that in the Walrasian framework it is only possible to transact services (as just recognised by Barone himself), this does not rule out private self-consumption as it is possible to treat self-consumption as that part of wealth that is simply re-purchased.

This approach suffers from two main shortcomings. One, any transfer actual ownership must necessarily be associated also to a transfer of the opportunity-cost schedule. It cannot just be treated as a flow quantitative wealth from one owner to another one. Even most importantly, it is clear that this assertion relies on the fact that actors, once that become legitimate owners of a given property, cannot refuse to deal in turn. In principle, nothing guarantees to an actor relinquishing his property to third parties to have it back as the first fundamental right always coming with private ownership is the right to refuse to deal.

As a result, the real effect of behavioural constraints is to force all property owners to treat their material property as if it was mere financial capital and to always and exclusively "hold out" to third parties requiring to use the property.

These are the implicit conditions brought into the picture by a pure service-based analysis.

Even more interesting, these are also the conditions that characterise public regulation and the institution of common carriage: mandatory dealing, and delegation of the use rights over the property to third parties under public constraint (*juris publici* over *juris privati*).

This makes the vNAD a framework modelling and characterising optimal public regulation of privately owned assets, and not the optimal decentralised operations of privately owned assets.

Yet, far to be a discouraging conclusion, the previous assertion does not have to be intended as an advocacy for mere positivism. Conversely, the opposite is true: the implicit conditions established by the vNAD framework testify that "the law", in the form of the institutional and legal context always logically, analytically and temporally precedes, and has priority over, the market mechanism.

Any normative judgment, in order to rule out the inherent subjectivity and dualism generated by the materiality and temporality of physical goods, has to first define a legal and regulatory framework, and only afterwards, it can define value judgments on the operations of the markets within this legal framework.

In conclusion, the following recap can be provided:

1. Any framework willing to derive normative conclusions on how private properties ought to be efficiently and optimally employed must necessarily guarantee monism in valuations.
2. In order to do so, it is necessary to impose some ontological and behavioural conditions that make the private sphere irrelevant and indistinguishable from the public sphere so to rule out subjectivity and dualism in valuations. The latter are necessarily introduced into the picture whenever material properties can be appropriated, stocked, and cumulated in the private domain (because of auto-consumption).
3. In order to rule out subjective and dual valuations, it is necessary to switch to a service-based analysis and to make the material world irrelevant, leading to an irrelevance of the private appropriation of physical assets.

4. In order to make the private ownership of physical assets irrelevant it is necessary to make sure that the respective asset ownership behave rationally, from a social point of view, which simply means that they cannot treat their private material wealth independently from monetary and finical considerations.
5. In order to do so, social constraints are needed in order to oblige private asset owner to (i) "hold out" their properties always and only to third parties and (ii) relinquish any control over the employment of the property (use rights) in faculty of their detention of ownership rights over the property. In this way, the framework ensures that private asset owners treat their physical endowments as if it was pure financial capital

Granted all these five steps, two fundamental implications follow:

One, normative conclusions on how privately owned resources ought to be employed can be legitimately reached.

Two, it is possible to legitimately reach conclusions on the optimality and efficiency of institutional solutions (such as the institution of asset ownership) merely based on the performance of the allocative mechanism (allocative failures) and/or on the presence of behavioural failures. Yet, some form of regulation will always precedes the optimality and the efficiency guaranteed by the market.

The next chapters will develop the two main topics raised by the present one, and derive implications with respect to the current approach of institutional economics, and with respect to the meaning of asset regulation and public control of infrastructures.

Chapter 5 will show that contemporary literature in institutional economics has completely overlooked and misunderstood the essence of the marginalist revolution: the shift in the unit of analysis from physical assets and ownership rights, to services and use rights.

Chapters 6 and 7 will show that regulatory law deriving from the common law tradition is relying precisely on the same two assumptions defined by the vNAD framework in order to justify public intervention. This will show, once more, that the standard neoclassical framework vNAD is nothing but an ideal-type of public regulation (in common law), certainly not a description of decentralised asset ownership.

CHAPTER 5

Some inconsistencies in the meaning of transaction cost: a reconsiderations of the role of value, information, asset ownership and rights

Chapter synopsis

This chapter represents a first discussion of the framework developed in the previous chapters.

The focus of the chapter is new-institutional economics, and the Coasean legacy. The very well-known key concept of the Coasean framework is the one of transaction cost. The concept has been adopted and applied by a vast scope of fields and disciplines, and in various ways. However, the fame of the concept is due to the fact that it officially started the study of institutions into economic analysis.

The main problem with the concept is that its definition is not clear, and this also implies that the way in which it should be intended, and the way in which it should be applied have not been unanimously agreed upon by scholars. There is a certain degree of inconsistency in the way in which the concept should be regarded.

The present chapter starts from there. The essence of the chapter is to show that the same inconsistencies and misunderstandings that have plagues the discipline of economics carry over in the fields of institutional economics, and transaction cost economics. The chapter focusses on the two key sources of confusions that have been highlighted in the previous chapters, and shows that these are the same type of confusions that preoccupy scholars in the field of law and economics, and institutional economics.

First of all, the chapter distinguishes between two distinct legacies of the Coasean production: a regulatory legacy, and a public economic legacy. The fundamental reason why this distinction is relevant is that the very concept of transaction cost, and its nature change between the public economics and the regulatory approach. The distinction between the two is not clear-cut in the literature, neither in Coase's writings, and scholars do not completely agree on the specific peculiarities of the two interpretations of the Coasean thought.

The aim of the chapter is to show that this confusion trace its roots back to the original confusions on the two key concepts of value and asset ownership that the work has developed so far. For this reason, the aim of the chapter is to show that a better understanding of the inconsistencies inherent in the concepts of value and asset ownership can help in throwing some light on the underlying reasons that make the concept of transaction cost so confusing and elusive.

The chapter is structured around two main sections. Each section focusses on one source of misunderstanding.

Section one focusses on the first misleading concept highlighted in the previous chapters: the one of value. The section shows that the concept of transaction cost is not well defined as the correct causality between subjective valuations and transaction costs is not spelled in Coase's writings. The mediating factor identified in the literature is the one of information (asymmetry).

The public economics legacy of Coase interprets subjective valuations as the outcome of some failures in the allocative process, typically uncompensated transactions, or missing markets. In this case it seems that Coasean transaction costs remain the necessary condition in order to explain the emergence of subjective valuations as, in their absence, markets would perfectly compensate.

Things are different in the regulatory legacy, where subjective valuations pre-exist the presence of transaction costs, intended as missing markets. When valuations are inherently subjective independently from the allocation mechanism due to information asymmetries, it is not clear how transaction costs should be treated.

The essence of the presence section is therefore the one of the causality between subjective valuations and transaction costs: the role and the nature of transaction costs can be better understood if its relationship with the emergence of subjective valuations is clarified.

Section two focusses on the second misleading concept highlighted in the previous chapter: the unit of analysis and the difference between services and assets. The section shows that there is a great deal of confusion in the literature when it comes to define the role of transaction costs due a constant confusion in the unit of analysis. Some traditions in institutional economics such as transaction-cost economics explain the emergence of institutions while taking the (physical) asset as unit of analysis. Transaction-cost economics, and new-property right theory are among the most famous ones. On the other hand, some other traditions consider the service as their unit of analysis.

The difference between the two is profound not only because transaction costs are defined in different ways in the two approaches, but especially because it highlights a profound confusion in the literature with respect to the concept of "right".

In the same way in which the concept of transaction cost is unclear in Coase depending on the unit of analysis adopted, the concept of right becomes unclear in the Coasean treatment depending on the unit of analysis adopted: The well-known Coase theorem that generated the discipline of law & Economics does not clarify whether with the term "rights" we should intended the ownership rights over physical assets or resources, or the use rights in the allocation and employment of these resources.

This has also implications for the economics of contracts as asset ownership necessarily introduces some temporality in the context. As a result, while transaction costs associated to asset ownership are usually deriving from the incompleteness of inter-temporal decisions, transaction costs associated to services usually derive from the static failures of the allocative failures.

Section three concludes the chapter.

1. Alternative legacies to the Coasean paradigm: confusing causation between value, information, and transaction cost

The section highlights two alternative approaches to the Coasean work that are here renamed the public economics and the regulatory legacy.

The section shows that the two differ in their interpretation of the causal relation between the origin of value and the presence of failures, or transaction costs. Coase himself is not consistent in his writings as sometimes information asymmetries are considered preliminary to the presence of transaction costs, while some other times they are not.

The section argues that these problems can be better understood if the unit of analysis, and the concept of value are used in order to shed some clarity on the matter.

1.1. Regulation, public economics and utility pricing in Coase

The concept of market failure has been the cornerstone on which contemporary economic theories of institutions have developed in the last half-century. The concept of transaction costs has permeated an incredible scope of fields, including the theory of social organizations, and the theory of public control and regulation.

Usually, when dealing with the interplay between the economic theories of institutions and regulatory economics, the standard approach consists in focussing on the influences that, starting from 1960s, the former had on the latter. This way to put the issue is, however, not correct from an epistemological and methodological point of view as it conveys two misleading ideas. On the one hand, it might induce regulators to think that the new theories and disciplines introducing institutions into economic theory have been imported into regulatory economics from somewhere else. On the other hand (and more importantly), it might induce strategy and organizational scholars to think that the set of theories developed to analyse institutions and organizational arrangements in economics have been developed with the primary intent to solve *positive* problems dealing with the behaviour of economic actors.

In truth, the causal connection between the two is likely to be the other way around: the theories explaining and justifying the role of institutions into economic analysis have been developed starting from regulatory and public economics, as answers to very specific *normative* problems faced by those two disciplines. In what follows, I employ the terms *positive* and *normative* as synonymous of the terms *descriptive* and *prescriptive*.

Thus, it is important to start from the recognition that the problems and the challenges arising in the fields of regulatory and public economics have historically always represented the original incubators for the subsequent developments not only of a multiplicity of new theories dealing with the role of institutions in economics, such as the theory of the firm and the theory of property rights, but also of proper new disciplines, such as neo-institutional economics (NIE), and Transaction Cost Economics (TCE).

The first straightforward example is represented by the immense influence that the studies in the regulation of infrastructures and public utilities, and the field of regulatory economics in general, had on Ronald Coase for the developments of his thoughts concerning the fundamental role played by transaction costs in generating institutions and ownership structures.

1.1.1. *The Regulatory influence*

On the one hand, Coase has always acknowledged the influence that the debate on social planning in 1930s and 1940s had on the developments of his ideas (Coase [1988: 8]; [1992: 715]; [1993: 248]). Yet, "contemporary" regulatory economics[75] is itself probably the most relevant sub-field of the economics of planning. One of the first and harsher critiques to the "planner's marginal-cost rule" have been developed mainly by London School of Economics (LSE) scholars during the 1940s, with the specific purpose to criticise the theory of *public utility pricing* and marginal-cost pricing proposed by Lerner (1944), Meade (1944) and Hotelling (1938).[76] The debate that developed from these critiques is usually defined as the "Marginal Cost Controversy"[77] and Coase's (1946; 1947) contribution to the debate is usually regarded as one of the fundamental cornerstones for the advancement of a theory of regulation (and institutions) based on incentives and information asymmetries, as we shall see later in this chapter (see Laffont and Tirole [1993: 24-29] for a specific discussion of the relevance of Coase's original criticisms on the contemporary incentives-based theory of regulation and Laffont [2000: chapter 6] for a comprehensive reassessment of the marginal cost controversy).

More generally, the studies on the regulation of natural monopolies and on the theory of public utility pricing have been two fundamental pieces of the Coasean legacy (Coase, 1964; 1970) as also recalled by Coase himself in his numerous memories (1988: 21-22; 1993: 248). His studies in cost accounting (Coase, 1973; 1990) have to be intended in this perspective, the theory of accounting being a fundamental building block for anyone interested in the theory of public utility pricing.

For this reason, Coase's studies in cost accounting and his critical arguments towards the marginal-pricing rule cannot really be treated independently. They both involve the

[75] For "contemporary" regulatory here we intend the Post-Pigouvian developments in the theory regulation and of monopoly that took off starting from 1930s: after the Chamberlin-Robinson-Lerner theory of monopolistic competition and of monopoly power (Chamberlin, 1962; Robinson, 1969; Lerner, 1944), and after the new formalization of the (Paretian) theory of consumer behaviour, first provided by Allen (1932) and Hotelling (1932). With respect to this second temporal classification, we follow Samuelson (1950: 356): "We may now move into the modern era [of the theory of consumption], which I arbitrarily date from R. G. D. Allen's foundation article of 1932". Basically, the new era in regulatory economics corresponds to the period covered by the "new welfare economics" (starting with Kaldor [1939] and Hicks [1939] articles).

[76] The usual reference is to Coase's (1946; 1947) articles, although those should be regarded as two (important) pieces of a broader critique that also includes other LSE professors, such as Thirlby (1946; 1947; 1960) end, especially, Wiseman (1953; 1957; 1959).

[77] For a review of the controversy see Ruggles (1949a; 1949b) and Frischmann and Hogendorn (2015). For a technical, other than historical, reinterpretation of the controversy in a context of asymmetric information and incomplete contracts, see Laffont (2000: chapter 6).

problematic issue of the centralised cognoscibility of the "correct" subjective marginal cost on which the marginal rule should be benchmarked.[78] *Transaction costs*, in the theory of accounting and public utility regulation should be intended as the impossibility to retrieve the correct ("natural") marginal cost by the part of the central regulatory body. This impossibility lies at the hearth of contemporary incentive-based regulatory economics.

As a result, Coase's attitudes towards centralised planning, together with his theory of the firm based on the transaction-cost concept, really cannot be fully understood if the fundamental influences of his studies in cost accounting applied to the theory of public utility pricing are not taken into account.[79]

1.1.2. *The public economics influence*

On the other hand, it is probably superfluous to recall the influence that the issues raised by the regulation, and the allocation, of the radio frequency spectrum had on Coase (1959) for the developments of the "*Problem of Social Cost*" (Coase, 1960).[80] The subsequent new discipline of "law and economics" (L&E)[81] and the "early" Property Right Theory (PRT)[82] both derive from Coase's 1960 famous article. Thus, there would be no theory of property-rights without Coase's remarks on the impact of positive transaction costs on the ability of the central planner, or the regulator, to undertake a centralised allocation of the shared resource, together with a centralised mutual compensation of the affected parties in case

---

[78] These considerations have been very clearly developed by Medema (1994a: 41): "Perhaps the most traditionally micro-theoretic elements of [Coase] work are his forays into the areas of pricing and cost. His work on pricing is confined primarily to situations of monopoly, and his most enduring impact in this area is his analysis of price regulation under conditions of natural monopoly." See also Coase (1988b: 16-22) for a discussion of the essence of the "Marginal Cost Controversy."

[79] It is important here to quote extensively from Coase (1990: 12): "As a result, the activities which we find firms undertaking must be influenced to some degree by their effect on the efficiency with which the accounting system operates … In understanding how in a competitive society the choice is made between these alternative but interrelated means of organization, we must take into account the role of the accounting system. The theory of the accounting system is part of the theory of the firm. It is not my belief that the
secret to the determination of the institutional structure of production will alone be found in the accounting system, but it certainly contains part of the secret."

[80] See also Coase's own memories (Coase, 1993: 247-249).

[81] For a general review of Law and Economics, see Mercuro and Medema (1997). Otherwise, the reader is referred to the magnus opus *Encyclopaedia of Law and Economics*.

[82] It is important to distinguish between "early" and "late" property right theory, as we shall discuss later in the work. Among the very few works that have seriously discussed the differences between the two strands of the much broader PRT, the paper of Foss and Foss (2001) deserves credits. The distinction between "early" and "old" is defined by Foss and Foss (2001) as a distinction between "old" and "new" PRT (OPRT and NPRT). The distinction between the two will be developed in more detailed below. For the moment, it is sufficient to anticipate that while EPRT deals with the allocation of use rights, NPRT deals with the implications of incomplete contracts on assets' ownership.

of externalities.[83] Similarly, there would be no economic theory of liability rules and of property rules (basically, no need for any economic theory justifying the presence of a legal institutional framework enforcing property rights and bilateral transactions) without Coase's remarks on the role of transaction costs and "market failures" (Calabresi, 1968; Calabresi and Melamed, 1972).

Thus, even in this second case, the problems raised in the public economics literature with respect to the definition of a centralised optimal taxation plan, together with the definition of a system of rights, turned out to be a very fertile field for the development of a theory of rights and of property. In this respect, Coase's 1959 and 1960 works really represent the "transmission belt" between the old (Pigouvian) approaches to externalities as compensated allocation by central taxation, and the new approach, subsequently developed by Arrow (1969), where externalities are interpreted as "missing markets". Where the expression "missing markets" simply reflects a lack of a complete set of rights.

In sum, Coase contribution is usually regarded as a unique and coherent body of thoughts. This is the way in which the author himself interprets his overall contribution through the years (Coase, 1988b; 1992),[84] and this is also the way in which his contribution is usually portrait by others (Medema, 1994a; 1994b; Barzel and Kochin, 1992). Yet, things are slightly more complicated than it may appear from the previous discussion. The reason why it is important to distinguish between the two distinct sources of Coase's contributions (the "public economics" influence leading to the property-rights legacy and the "regulatory" influence leading to the legacies in the theory of centralised planning and hierarchies) is that the two approaches deal, in reality, with two different types of market failures, which might look identical only because of the nebulous definition of the transaction cost concept.[85] This ambiguity is very important as it carries over once that an even more ambiguous concept will have to be defined, the one of ownership (which in fact, has a different meaning according to the specific interpretative paradigm adopted, as we shall see).

---

[83] As summarised by Lueck and Miceli (2007: 187) in their exhaustive overview: "The economics of property law begins with Coase (1960), who provides a property rights perspective on the problem of externalities, or 'social costs'."

[84] This is how Coase introduces his collection of essays (Coase, 1988b: 1): "The core of this book consists of three papers … Other papers which extended, illustrate, or explain the arguments in these three papers are also included. As will become apparent, these essays all embody essentially the same point of view."

[85] Using the expression of Milgrom and Roberts (1992: 29), transaction costs "is a simple, but profound idea. However, Coase was not very explicit about the origin and nature of these transaction costs, and without a systematic understanding of the issues, the idea is not very useful."

1.2. The "Public economics" and the "regulatory" strands: bilateral bargaining and the role of information

First of all, it should be recalled that both legacies roughly start from the same verification: in the presence of transaction costs, the marginal equivalence between the price (value) defined in a market transaction and the cost (value) faced by the actors cannot be reached. This discrepancy creates the market failure. It should be noted that this is just another way to reframe the so-called "Coase theorem" (Stigler, 1966: 113): "The Coase theorem thus asserts that under perfect competition private and social costs will be equal".

Thus, the common nature linking all Coase's works is the rejection of the Paretian (Lerner-Hotelling) marginal-cost rule; as boldly concluded by Coase (1988b: 19): "Marginal cost pricing as a policy is largely without merit". Yet, the key question at this point becomes: why does the break-down of the marginal equivalence occur? A more accurate look will reveal that the underlying reasons why this discrepancy is occurring are different based on the specific approach adopted. The underlying problem can be clearly understood if one looks at the way in which the two types of failures are framed and solved.

In the "public economics" framework (mainly associated to Coase's 1959 and 1960 papers), the reason why the private marginal costs of an actor are not aligned with the social (Paretian) optimum (mainly) derives from the fact that markets are incomplete. In other words, not all transactions are correctly compensated by an associated market exchange for the purchase of use rights. This is Arrow's missing market approach. Thus, in principle, this specific interpretation of market failures would guarantee the Paretian efficiency is reached as soon as all interactions across actors are covered by a respective endowment of *purchasable* rights.

On the other hand, in the "regulatory" framework (stimulating Coase's writings on central planning and the nature of the firm, public utility pricing and accounting) the real nature of transaction costs is linked to the missing information about one actors' private (subjective) valuations.[86] In this second case, there is still a discrepancy between the private marginal costs of an actor (her subjective valuations) and the objective (social) valuation, but this failure is not necessarily deriving from an incompleteness in the set of exchanged (and priced) rights. Conversely, the failure lies in the dispersion of the information.

[86] This aspect of missing information concerning one another subjective valuations is evident in Coase's MCC, but it is also the essence of the Nature of the Firm, where transaction costs derive from the information asymmetry between the actors, and not from incompleteness in the transacted property rights among the different parties.

It is difficult to appreciate the distinction between the two cases as both substantially describe the same phenomenon (marginal valuations are context dependent) and both can be solved in the same way: internalizing the failure through integration and alignment of incentives and information: Frank Knight (1924) solution to the problem of the overcrowded highway is an example of centralization and internalization in the case of non-compensated externalities among actors, while Williamson's hierarchy is an example of centralization and internalization in case of non-compensated idiosyncratic information among actors (integration as an efficient governance solution to reduce "haggling" and hold up).

Nevertheless, the two cases differ in two fundamental respects. Both involve the way in which the unit of analysis is considered, and the causality between different elements of the system. On the one hand, there is a decisive difference in the causal relationship between actors' subjective valuations, and market failures deriving from the presence of transaction costs. On the other hand, and strictly linked to the first one, there is a decisive difference in the unit of analysis considered, and in the way in which "cost" is defined.

We now examine these two cases in detail.

1.2.1. The causal relationship between transaction costs and subjective valuations: the role of information asymmetries

With respect to the role of information, it is important to note that subjective costs can be the outcome or the cause of the presence of transaction costs.

On the one hand, with mere market incompleteness, the establishment of complete and perfectly competitive markets can still be sufficient to guarantee optimality through decentralised (bilateral) bargaining.[87] This is not the case in the case of information asymmetry any more: with information asymmetry among actors, decentralised bargaining will still remain inefficient also in the presence of complete and perfectly competitive markets for rights. In this single case the single actor will always enjoy some rents, no

[87] Sufficient but still not necessary, as the introduction of complete markets for rights can still create non-convexities and therefore inefficiencies as demonstrated by Starret (1972) (See the discussion in Laffont (1988: 17). Yet, another problem with the pure version of the Coase theorem is that it does not take the "shape" of actors' primitives into account in order to understand their welfare implications. In Farrell words (1987: 114): "Like the welfare theorem, [the Coase theorem] says that if everything is tradeable then Pareto-efficient outcomes result. Unlike the welfare theorem, it makes no strong assumptions about convexity, price-taking, and complete markets." A similar critique is also provided by Samuelson (1995). See also the discussion in Laffont (2002).

matter whether the counterpart (decision maker) is a central (regulatory) authority, or another bargaining actor.[88]

It is this last aspect that generated some harsh critiques by the part of some economists (namely, Paul Samuelson)[89] on the correctness of the Coase theorem.[90] In effect, if one just considers Coase's 1959 and 1960 papers, it is not really clear whether transaction costs also involve the intrinsically subjective nature of private valuations, or if the underlying argument is the following, *specular* one: optimality will be reached through bilateral decentralised transactions as long as markets for use rights are complete, exclusive, work smoothly and can be perfectly enforced. The literature reveals that the position is not unanimous. Medema and Zerbe (2000) in their extensive review of the Coase theorem identify two distinct interpretative frameworks of the Coase theorem: the quasi-competitive framework and the game-theoretic framework. This categorization reflects the distinction between the "public economics" strand and the "regulatory" strand only in a partial and indirect way, as discussed below.

The fundamental distinction between Medema and Zerbe's dichotomy lies in the role of information in shaping individual valuations and in its relationship with the transaction cost concept. For this reason it is possible to also reframe the two strands as the "Paretian legacy" and the "Nash legacy".[91] More precisely, while in the Paretian (quasi-competitive) approach interprets informational failures as a result of the presence of transaction costs, the Nash (game-theoretic) approach does not considers private information as the result of positive transaction costs. As already anticipated above, there is no univocal interpretative

[88] This fact reflects the so-called "rent extraction-efficiency trade-off", which constitutes the fundamental intuition of contemporary principal-agent theory, as we shall discuss later in this chapter; see Laffont and Martimort (2001: 41-46) for a technical overview of the trade-off.

[89] The reader is referred to Samuelson's unusually harsh critique of the Coase theorem "*Some Uneasiness with the Coase theorem*" (Samuelson, 1995). The misunderstandings clearly derives from the fact that Samuelson does not interpret the Coase-Stigler "perfect competitive market" solution as one that also implies perfectly centralised collection and processing of information. His reference to Hayek is quite revealing in this respect (Samuelson, 1995: 5): "The Coase-Samuelson generation were brought up witnessing the great debate between von Mises and Lerner-Lange concerning the feasibility of socialist rational pricing to produce Utopia. (That was a reprise of earlier Pareto-Barone-Wieser-Taylor debates.) Many contemporaries believed Lerner-Lange triumphed in the debate. I came to believe that Friedrich Hayek was the true victor."

[90] For a broad overview of the critiques to the Coase theorem, see Medema and Zerbe (2000).

[91] The Paretian approach assumes given actors' primitives (technologies in production and consumption), and does not admit strategic (endogenous) responses among actors. For this reason, the Paretian efficiency frontier can always be reached as long as markets work perfectly as portrait in the Arrow-Debreu framework. The Nash approach allows for actors' multiple types, admitting endogenous arbitrage *across primitives* according to the external situation. In this case, the Paretian optimum will never be reached even though the markets for rights are competitive and complete, as previously remarked.

answer among the various authors, although the way in which Medema and Zerbe introduce and define the "theorem" seems to be much more in line with the former approach (Medema and Zerbe, 2000).[92] The same holds for Barzel (1997: 11) who seems to draw a univocal *causal mapping* from rights to transaction costs:[93]

> "The presence of positive transaction costs is what makes the study of property rights significant. On the other hand, in the Walrasian, perfectly competitive model, rights are perfectly delineated and transaction costs are zero."

1.2.2. *The causal relationship between transaction costs and subjective valuations: the role of the unit of analysis and the definition of cost*

Beside the mutual interaction between markets (the allocation process) and information, there is another source of confusion between the two strands of Coase's works: the unit of analysis considered, and the meaning of cost.

The public economics framework always take the given resource as the unit of analysis. As a result, costs are defined in opportunity- cost terms. It is sufficient to look at the *Problem of Social Cost* (Coase, 1960), to realise the two key characteristics of the article. One, resources are the unit of analysis, and the valuation process is considered as an allocation process of the resource. Two, transaction costs are computed in opportunity-cost terms (p. 43):

> "Economists who study problems of the firm habitually use an opportunity cost approach and compare the receipts obtained from a given combination of factors with alternative business arrangements. It would seem desirable to use a similar approach when dealing with questions of economic policy and to compare the total product yielded by alternative social arrangements."

The opportunity cost concept is indeed the way in which cost and value are computed when the given resource is the unit of analysis.

---

[92] For instance, Medema and Zerbe (2000: 837; emphasis added) define the theorem as follows: "Negotiations among affected parties would result in an efficient and invariant outcome under the standard assumptions of competitive markets (especially if the costs of transacting are zero), *as long as rights are well-defined*." This definition gives a clear causal prominence to the "market completeness" position.

[93] For instance, Barzel (1997) does recognize the role of information. However, from his treatment of transaction costs it clearly appears that the fundamental failure of the system derives from the incompleteness of markets and the break-down of the price (allocative) mechanism, consistent with the Paretian tradition of PRT.

Things are very different in the regulatory approach. In this case the analysis does not deal with resources, nor does it deal with the associated opportunity costs deriving from the allocation problem. Subjective valuations in this case derive from a certain technological process for the production of a certain outcome (a service), and the associated (hidden) costs are in production.

In the two cases the unit of analysis differs. Moreover, the causality differs: while subjectivity in value in the first case arises given a misallocation of the given resource across services, subjectivity in value in the second case arises given a misuse of the various resources in production in the production of a given service. The causal relationship between the two elements (resources and services) is inverted.

It should be noted that the confusing interaction between market incompleteness and information asymmetry is just another aspect of the more fundamental confusing relationship between units of analysis in the two cases. Thus, if resources are the unit of analysis and the opportunity cost schedule is socially defined by the market mechanism, transaction costs will be only defined by incomplete markets and uncompensated transactions. Conversely, if the service outcome of a production process is the units of analysis, then the incompleteness of the market is not informative of the various subjective capabilities in production (or technological capabilities) of an actor. In this second case, information asymmetries leading to subjective valuations precedes and disregards missing markets.

In sum, the issue pins down to the definition of the "transaction cost" concept and, ultimately, to the definition of the "market failure" category. Certainly, in both cases we do have some waste of resources employed in order to "run the economic system".[94] Yet, this

[94] This is the standard definition provided by Arrow (1969: 48). This general definition is usually shared by a broad set of authors (see Furubotn and Richter, 2005: 48-50). In general, transaction costs are defined as the waste of resources in the Paretian tradition adopted by old PRT (see Allen (2000: 898) who defines transaction costs as "the costs establishing and maintaining property rights"). However, the same approach is also adopted by one of the forefathers of the Nash tradition, and forefather of mechanism design, Leonid Hurwicz, who describes the issue in the following terms (Hurwicz, 1972: 299): "To calculate the 'net' welfare generated … one must take into account the drain on resources by the informational activities (the cost of information processing) and the incentive-inducing activities (the cost of enforcement). Thus there are two types of 'givens' in our problem … these two types of givens are interrelated, of only because resources used to operate the organization cannot be used for 'substantive' purposes." For instance, this is also the way in which Milgrom and Roberts (1992: 29) define transaction costs: "transaction costs are the costs of running the system: the costs of coordinating and motivating."

cannot be a satisfactory way to define transaction costs as it still does not tell us whether decentralised bargaining will still be Pareto-efficient, once that these extra-resources are spent, or if some more fundamental (non-solvable) flaws will still be present, independently from the efficiency and the completeness of the market system. As concluded by Medema and Zerbe (2000: 857):

> "In the end, then, whether the game-theoretic and other challenges to the Coase theorem go to its correctness or its relevance comes down to how one interprets the almost mystical world of zero transaction costs."

1.3. Transaction cost and the origin of value in comparative institutional analysis

1.3.1. *Confusions on the original institutional domain from which value derive in the case of zero transaction cost: private vs. public*

The distinction between the Paretian and the Nash approach is an important one as it reflects two different ways to interpret one of the most important concepts in economics: the one of value. The important point to remark here is that in the Paretian approach also adopted by early writers in PRT and L&E, the subjectivity of actors' valuations can be solved *ex-machina*, by simply modifying the way in which the central market mechanism operates. Conversely, in the (Nash) information-asymmetry approach, subjective valuations will remain endemic to the system, regardless of how the centralised (social) allocation mechanism performs (in other words, regardless of the completeness of markets).

It is only in this second case, when subjectivity in actors' valuations is admitted, that a genuine assessment of the relative performance of alternative institutional (governance) arrangements makes sense. It is from these premises that the later contract-theories of institutions (such as TCE) could develop.

For instance, the subjectivity in private valuations in TCE is given by the behavioural assumption of "opportunistic behaviour", which reflects the fact that actors can act according to their own subjective valuations (interests), independently from the external (objective) constraints. Not surprisingly, the assumption of opportunism corresponds to the assumption of *hidden action* (or moral hazard) usually developed by the literature in principal-agent theory.

More generally, the entire field of "incentive theory" starts from the recognition that private information is a (non-solvable) problem and that "these informational problems prevent society from achieving the first-best *allocation of resources* that could be possible in a world where all information would be common knowledge" (Laffont and Martimort, 2001: 3).

Farrell (1987) provides a clear exposition of the problem, underlying the following fact: efficient (ideal) bilateral bargaining can always be an efficient solution to the underlying market failure (thus reproducing the Paretian optimality) if and only if it is assumed that no private (subjective) valuations is involved in the decentralised bargaining process.[95] In this case, where there is no room for subjectivity in actors' private valuation, the Coase theorem perfectly holds and objective marginal costs (values) are always identical to subjective marginal costs (values). Yet, the paradox is that under these assumptions, there would be no need for bilateral decentralised bargaining any more, as centralised planning would work just as fine.[96]

If positive transaction costs also include the presence of private information,[97] then it is possible to safely assert that, in the absence of transaction costs, decentralised bargaining will always lead to the efficient Paretian optimum where subjective and objective valuations always correspond. The important point is that in this case, objective (or "natural") value and efficiency always pre-exist the presence of failures (consistent with Paretian welfare economics), while subjectivity in valuations can only arise as consequence of the presence of these market failures. Similarly, under this broad view of market failures, institutions have their economic legitimacy only as a sub-optimal result of the failures in the system. The same holds true for the institution of ownership.

---

[95] Technically speaking the property-right and L&E developments of Coase's argument make statement about the efficiency of the Pareto equilibrium, but still remaining in a Paretian world, while ruling out the possibility of sub-optimal Nash equilibria (that emerge as soon as actors can act strategically based on their *private* valuations).

[96] It is worth to quote some very clear passages from Farrell (1987: 115-116), to give a more precise, but concise, restatement of the argument: "So we cannot assume that all mutually beneficial contracts are signed, unless we assume that everyone knows everything about everyone, which they do not. The strong form of the Coase theorem -the claim that voluntary negotiation will lead to fully efficient outcomes- is implausible unless people know one another exceptionally well. As I shall argue, that case is not only unlikely, but is also the case where decentralization is least useful … True decentralization consists in delegating decisions to those who know more about them. If there were no private information, to take an extreme version of Hayek's view, decentralization would have no point: a central authority would be perfectly able to take fully efficient decisions. Coase points out that decentralized negotiation would also work well then, but that is not a very exciting observation. It is much more instructive to compare how different systems do when private information does exist." It should be noted the interesting point that a theory based on pre-existing subjective valuations necessarily lead to an economic comparison of alternative institutional arrangements in the vein of TCE.

[97] This is usually the Paretian case adopted by early-PRT and L&E as remarked by Medema and Zerbe (2000: 856-857).

As already remarked above, the problem with this type of all-encompassing approach is that it can explain why extra-market institutions arise, but it cannot really distinguish between alternative types of institutions when transaction costs are absent. More precisely, if this interpretation is adopted, the world of zero transaction cost is the Walrasian world where allocation occurs centrally, and not by means of bilateral decentralised market transactions. If it is possible to assume that no subjective valuations (different from the objective valuation) can be present, then it would certainly be more convenient to collect and process all the information centrally, and to allocate accordingly. A world of zero transaction costs in the Paretian approach is a world of complete centralization more than a world of decentralised independent bargaining.[98] After all, this is the sense of the Hayekian message on information (1945), also backed-up by Sameulson (1995) as already illustrated above, and also developed by Williamson in his theory of incomplete contracts based on the concept of adaptation (see for instance, Williamson, 2010: 679), as we shall see later in the work.

It should be noted that this conflation between centralised planning and decentralised bilateral bargaining in the Lausanne (Pareto-Walras) system should not be surprising. The Lausanne (general equilibrium) system has been originally developed, and has always been interpreted, as a framework whose primary aim was to overcome actors' subjectivity in marginal valuations so to justify the central allocation of resources. This was already clearly recognised by Schumpter (1954: 987; emphasis added) that is convenient to quote extensively:

> "If we consider a socialist economy, it is still more obvious that, for instance, maximization of satisfaction requires that the ratio of marginal utilities for each pair of consumers' goods must be identical for all comrades; that in every line of production must be so organized as to make the technologically optimum use of all means of production; and that the marginal value productivity of all scarce means must be the same in all their uses or, at all events, must in every use be at least as great as it would be in any other. But all this amounts to saying that *any attempt to develop a general logic of economic behaviour will automatically yield a theory of the socialist economy as a by-product*."

[98] This is after all the way in which mathematical economists dealing with linear programming in the early 1950s responded and solved the central-planning debate.

In a nutshell, in a world where actors' subjective valuations are subordinated to the efficiency of the market mechanism (and therefore to the presence of transaction costs intended in the "Paretian" way), then it is because centralised allocation is possible, that bilateral (decentralised) bargaining becomes (Pareto) efficient. It is important to note that contemporary theory of incentives has retained the standard Walrasian approach, where the standard benchmark for the computation of efficiency in allocation is given by the "central planning" solution, and not by the bilateral bargaining one).[99]

1.3.2. Reinterpreting the two Coasean approaches in value terms

The essence of the "controversy" on the correctness and relevance of the Coase theorem thus derives from the different understandings on the origins of actors' subjective (marginal) valuations. Not surprisingly, the Coase theorem can be reframed in this equivalent, but more explicit, form: "***value is an objective and endogenous parameter of the system, unless transaction costs are present***". Or, alternatively, as: "*value and distribution conflate in a world of zero transaction cots*".[100] Not surprisingly, this strong (Paretian) version of the Coase theorem corresponds to the way in which Walras was enthusiastically defines the fundamental characteristic of his system in the *Elements*.[101]

The reason why this point is a very important one is that it highlights different ways to understand and interpret the origins value: whether subjective valuations are a result of a market failure or, conversely, whether it is the presence of private (thus subjective) valuations that inherently generates market failures. The same logic extends to the theories explaining institutions. As a result, the same dichotomy can then be reframed as a divergence concerning the logical relationship between the valuation process (the origins of value) and the origins of institutions. This point becomes very important once that the

---

[99] For example, this is how Laffont and Maskin (1982: 31; emphasis in the original) introduce the incentive problem: "For an incentive problem to arise, noncoincidence of goals is not enough; the planner must care about either what agents know or what they do. That is, his objective function must depend either on agents' information or on their behaviour. An example of pure informational dependence is provided by the literature on resource allocation mechanism. There, the planner's objective – social welfare- is a function of consumers' (agents') preferences and endowments. The incentive problem is, typically, that of eliciting this information."

[100] Note that this way to put the issue retraces the underlying logic and "*raison d'être*" of L&E, which can be summarised as follows (Mercuro and Medema (1997: 22): "Altering the Law, that is changing the legal relations governing society … or changing the working rules, will ultimately alter economic performance."

[101] This is how Walras (1954: Lesson 14; § 143) succinctly describes the essence of his general equilibrium framework: "Given several commodities in a market in a state of general equilibrium, the current prices of these commodities will remain unchanged no matter in what way the ownership of the respective quantities of these commodities are redistributed among the parties to the exchange, provided, however, that the value of the sum of the quantities possessed by each of these parties remains the same." This is simply the Coase theorem one century before Coase-Stigler.

nature and the role of ownership are enquired; not surprisingly, the concept of ownership has different meanings in the two branches.[102]

However, regardless of whether the adopted interpretative approach is Paretian or Nash and regardless of the specific interpretation of the transaction cost concept, one point is clear: in contemporary normative (welfare) economics subjective valuations arise as the result of some failures of an "ideal" system that include: (a) complete and competitive markets (Paretian dimension), and of (b) actors' "omniscience" (Nash dimension). A "market failure" approach towards value and institutions usually includes both dimensions. After all, both can be regarded as different aspects of a failure of a primitive allocative system where (a) all interactions among individuals are covered by exchangeable rights and (b) everyone knows everything about everyone else.

Under this market-failure interpretation of subjective value, and therefore of institutions, the Coase theorem holds and value necessarily always remains objective and endogenously calculated by the market mechanism. In this case it becomes possible to consider the multiple reinterpretations of Coase's overall production as a coherent body of literature. This is consistent with Coase's own interpretation of his work which was summarised by the famous statement: "let us study the world with positive transaction costs" (Coase, 1992: 717). Thus, in this case the Coasean legacy can be summarised as follows: *under transaction costs, subjective (marginal) valuations are not completely compensated by the respective objective (marginal) valuations transacted in the market*. This is, once again, just a different way to restate the Coase theorem.

Once again, the reason why actors' subjective valuations are not completely compensated by the relative monetary exchanges on the market can be because markets are not competitive or complete (Paretian explanation), or because actors can act strategically "hide" their subjective valuations to the transacting parties (Nash explanation). However, this distinction becomes irrelevant at this point, in case it is admitted that both are manifestations of some failure in the system.

It is clear that an alternative theory of institutions (and of ownership) that does not rely on the concept of failures should be able to go beyond both types of failures. However, the

[102] Note that the same statement previously made about the nature and the role of value, can be extended to ownership: is ownership the *result* of some market failure, or it is the inherent *precondition* in order to generate subjective valuations and private information, and therefore market failures? The reminder of the work will develop the point much further.

reinterpretation of the Coase theorem in value terms has the advantage of clarifying the task.

The definition of an alternative theory of institutions (and ownership) reduces to the (not-so easy task) to define a causal connection between value and market failures or, alternatively, between value and institutions, that goes beyond the concepts of market failure and transaction costs *tout-court*. In order to do so, the classification here adopted between the "public economics" and the "regulatory" reinterpretation of Coase's influences reveals to be quite useful. As a result, an explicit connection between the two classifications (Pareto vs. Nash on the one hand, and public economics vs. regulatory on the other) has to be provided. To this we now turn. This becomes important in order to generate a comprehensive taxonomy of the transaction cost concept, so to better understand the relationship between the concept of transaction cost, and the one of ownership.

2. Property rights vs. use rights: implications for transaction costs and institutional analysis

2.1. 'Public Economics' vs. 'Regulatory Economics': possession of resources and use rights after Coase

2.1.1. A recap on the role of information in shaping transaction costs

The distinction between the Paretian and the Nash reinterpretations of Coase's writings is a fundamental one in order to shed more light on the relationship between information, value and institutional arrangements. This dichotomy was developed based on Medema and Zerbe's (2000) interesting discussion of the different reinterpretations of the Coase theorem provided through the years starting from 1960. This means that the previous dichotomy reflects a classification of what was previously called the "public economics" strand of Coase, mainly developing from his 1959 and 1960 articles. However, the dichotomy here suggested between the public economics and the regulatory influence of Coase's contributions is not limited to Coase's two papers discussing the problems deriving from "the optimal allocation of a scarce resource among alternative uses".[103] In fact, the

[103] The two "public economics" papers of Coase really are on the basic (Robbinsian) economic issue of the allocation of scarce resources. For instance, with respect to the original work on the radio frequency spectrum, this is how the original aim is recalled by Coase (1993: 248; emphasis added): "I used my time [in Stanford] to investigate the work of the Federal Communications Commission (FCC), particularly its policies in *allocating the use* of the radio frequency spectrum." The extension of the 1959 paper to the more famous 1960 paper is then commented in the following way (Coase, 1988b: 11; emphasis added): "There is no difficulty in employing the same approach which I found

constrained allocative problem represents the essence of the sole "public economics strand".

The regulatory strand (developing from Coase's works in public utility pricing, accounting and regulation of natural monopolies, and from his 1937 article on the vertical and horizontal integration of firms) does *not* deal with the standard allocative problem. Conversely, it mainly deals with the implications of *assets and resources' ownership* for the overall economic efficiency of the system (see also Demsetz [1988] on that).[104] The switch in focus, from the allocative problem of the public economics strand, to the integration and ownership problem of the regulatory strand requires some careful analysis as it has further implications for the way in which failures and transaction costs should be defined.

First of all, the distinction between the "public economics" interpretation of transaction costs and market failures and the "regulatory" interpretation constitutes the preliminary foundations for the subsequent distinction between (early) property-rights theory and L&E on the one hand, and the incomplete-contract approaches proposed by later theories on the other. Among the second set of theories, we place TCE and the Grossman-Hart-Moore (GHM)[105] framework. The latter is also usually defined as "property-right theory". Following the nice classification provided by Foss and Foss (2001), in the reminder of the work we will call the first developments in PRT as Old PRT (OPRT), while the later GHM developments based on incomplete contracts and renegotiation as New PRT (NPRT).

Second, the mere distinction between the "public economics" and the "regulatory economics" strands is not informative concerning the role of information and value in

---

useful in discussing *the allocation* of the radio frequency spectrum for the analysis of problems which economists are more accustomed to handle." These are just illustrative quotes to remark a very simple point: the public economics legacy of Coase really is on the standard allocative problem in general equilibrium. Note that, however, this is at odds with the Coase's critical attitude towards Robbins' methodological approach that he has manifested elsewhere; see for instance Coase (1988a: 24-25; 1988b: 1-5; 1992: 714). See also Coase (1982) reprinted in Coase (1994: 208-214).

[104] Note that very similar remarks are made by Demsetz (1988: 12), where the author also adopts a very similar classification between Coasean theory of externalities (public economics) and regulatory economics: "A general peculiarity of recent literature is that its focus often is not on ownership itself. Nowhere is this more true than in Coase's famous paper on social cost. His topic is externalities not ownership … The study of regulation is another example … Without recognising the fact, the authors of this vast literature have been writing about the economics of ownership."

[105] The standard references for the GHM framework are: Grossman and Hart (1986), Hart and Moore (1990), Hart (1995). For a comparison between TCE and GHM see the two respective chapters in the *Handbook of Organizational Economics* by, respectively, Tadelis and Williamson (2013) and Segal and Whinston (2013). Gibbons (2005) offers a further general comparison of the two theories. A review of the economic theory of incomplete contracts is presented by Tirole (1999), while a more technical collection of papers on later developments of the incomplete contract literature is provided by the special issue in the Review of Economic Studies (1999, vol. 66: issue 1).

shaping allocative outcomes. For instance, the public economics strand can be interpreted either as Paretian or as Nash, depending on the peculiar approach adopted. This point was already highlighted above. The case of the regulatory strand is not less complex. The theory of incomplete contracts is not univocal with respect to the role of information and subjective valuation. Some developments of the literature explicitly assume the presence of private information either in the form of moral hazard and hidden action (agency and principal-agent theories) or in the form of opportunism (TCE).[106]

Williamson has long stressed the fact that the way in which information enters TCE is different from the way in which information enters both the Nash developments of Coase's "public economics", (economic theories of incentives in general based on the principal-agent framework), and GHM theory of incomplete contracts. With respect to the former he highlights the difference between the *governance approach* adopted by TCE (Williamson, 1985: 26-29) and the *incentive approach* adopted by the other Nash-based theories. With respect to the latter, he highlights the fact that TCE deals with *ex-post adaptation* in the use of an asset, and not with *ex-ante investment incentives* in the appropriation of an asset (Williamson, 2000: 605-606).[107]

However, other developments of the theory of incomplete contracts still assume symmetry of information among the parties. Thus, while private information about subjective valuations was a necessary requirements in the Nash approach of the public economics strand, hidden information among the contracting actors is not necessary in incomplete contract theories. [108] Tirole (1999) offers an exhaustive review of the different approaches

[106] In a principal-agent setting, there would be no need to establish a contract ex-ante in order to "curb" actors' incentives if rewards can be optimally redistributed ex-post by merely observing the "effort" of the agent. See Tirole (1999) for a review.

[107] Gibbons (2005) reformulates the same distinction as a dichotomy between a theory of "rent seeking" (governance) and a theory of property-rights (which he distinguishes form the theory of incentive properly said): "where the rent-seeking theory envisions socially destructive haggling ex post, the property-rights theory assumes efficient bargaining, and where the rent-seeking theory is consistent with contractible specific investments ex ante, the property-rights theory requires non-contractible specific investments" (Gibbons, 2005: 205).

[108] In contractual terms, this difference can also be reinterpreted as a difference between classical contracts and neoclassical contracts (Williamson, 1985: chapter 3). It has to be said that there seems to be a difference in the way in which TCE intends the two concepts and the way in which the "incomplete contract literature" defines them. While in Williamson, neoclassical contracts still ensure court enforcement (while "relational bilateral governance" is the situation when courts cannot enforce contracts), the literature in incomplete contracts do not seem to make any distinction, and always assumes the non-enforceability of contracts by the part of third parties (usually called non-verifiability).

taken by the incomplete contract theory.[109] The general remark is that incomplete contract models "usually assume that information is symmetric among the parties … although we should point out that informational symmetry is not central to the debate between complete and incomplete contracting. The *symmetric information implementation* literature studies one prominent class of complete contracting models in which agents have no private information" (Tirole, 1999: 754; emphasis in the original). In other words, incomplete contract theories usually assume *observable but non-verifiable* information.[110]

2.1.2. Asset ownership vs. allocation: the role of time

What all incomplete contract theories do have in common is their focus on inter-temporal decisions, where decisions taken ex-ante affect the way in which actors behave ex-post following some unforeseen circumstances (contingencies). It is precisely this inter-temporal dimension that distinguishes this second set of theories from the other theories belonging to the "public economics legacy". And the reason why inter-temporal decisions become relevant in the regulatory strand of Coase is that *appropriation* and *asset ownership* are brought into the picture. This renders the regulatory legacy not a *theory of resources' allocation*, but a *theory of resources' ownership* (a theory of appropriation and adaptation, as we shall discuss later).

The difference between the public economics and the regulatory strands is usually interpreted as a difference between *theories of resources allocation*, and *theories of the firm* (Barzel, 1997: 75-76; Furubotn and Richter, 2005: chapter 8).[111] In a sense, this is certainly correct as the aim of all incomplete-contracts theories is precisely to explain the ownership patterns over assets, and therefore the boundaries and the operations of productive organizations.[112] Yet, the simple reference to the theory of the firm is too broad and not very informative as the firm itself can be defined in multiple ways in the economic

[109] For instance, both Grossman and Hart (1986: 696) and Hart and Moore (1990: 1126) assume complete and symmetric information (observability). See Tirole (1999: 749-751) for a complete classification of the different approaches adopted in the incomplete contracting literature.

[110] Both parties can observe their reciprocal information (there is no subjectivity in the information). However, no third party (a court or a central planner) can actually enforce the contract (this is the meaning of non-verifiability); see Tirole (1999: 753-755).

[111] This conclusion is also shared by Tirole (1999: 742-743): "The recent upsurge in incomplete contract modelling was primarily motivated by organizational issues: what determines the size of the firm, how authority is distributed within the firm, and how the corporate charter and the financial structure (voting rights, powers of the board of directors, feasibility of takeovers, debt-equity structure) organize the control of insiders by outsiders."

[112] The reference to Coase's 1937 is explicit in all incomplete contracts articles; after all, it is sufficient to check the titles of the articles of Grossman and Hart (1986), Hart (1989) and Hart and Moore (1990) to clearly see this.

literature (see the taxonomy presented in Hart, 1989).[113] More precisely, as previously stated, the key difference between the "public economics" and the "regulatory economics" strands derives from the object of study and the *unit of analysis* adopted: while the former deals with the *allocation* of mere *use rights*, the latter deals with the implications of proper *property rights* and the *ownership (control) of physical assets*.

Differently from the proper ownership of an asset (property rights) that focus on the physical control and possession of the latter, use rights merely look at the rights that actors have in using the resources. OPRT is only involved with use rights, not with property titles over physical assets. For example, Alchian (1977: 132-133) defines use rights in the following manner:[114]

> "By this I refer to the fact that at the same time several people may each possess some portion of the rights to use the land … In sum, private property rights to various partitioned uses of land are 'owned' by different persons."

Where the important point to note in the above definition is that the concept of "ownership" above does not refer to the ownership of the land, but to the ownership of the single use right.

This distinction is by any means a new one, even though the classification between property and uses has never been consistent or homogeneous across the various authors.

Thus, for instance, Barzel (1997: 3) distinguishes between "economic property rights" and "legal property rights" where the former corresponds to the ability to enjoy a certain piece of property, while the latter corresponds to the effective endowment of a property to a person. On the other hand, Furubotn and Richter (2005) in their *magnum opus* distinguish between absolute property rights and relative property rights. While absolute property rights refer to the actual property (appropriation) of a physical object,[115] relative property rights are mutual (contractual) obligations concerning the way in which actors are supposed to behave. Thus, instead of answering to the question "who should own what",

[113] Thus, Holmstrom and Tirole (1989) well-known survey conceives firms "as a contract between a multitude of parties". However, the firm intended as a "bundle of contracts" is not in line with the vision of incomplete contract theory (Hart, 1989).

[114] Armen Alchian is, together with Coase, Demsetz (1967) and Cheung (1969) the founder of OPRT. Calabresi (1961) is the other fundamental reference in L&E as recalled above. For a treatment on OPRT see, among others, Eggerston (1990: chapter 2).

[115] As stated by Furubotn and Richter (2005: 79): "The existence of physical objects and things in general requires that there be regulation among men regarding the appropriation and use of things."

relative property rights answer to the question: “what people are expected to do with what they have vis-a-vis the others?” Similarly, Hodgson (2015) in a recent article distinguishes between property (the physical control of a resource) and possession (legally sanctioned rights). On the other hand, Foss and Foss (2001) distinction between OPRT and NPRT relies on the same distinction here adopted, between use rights and ownership rights on physical assets or resources.

The distinction between use rights and ownership rights (possession or control) is quite straightforward in OPRT. The same cannot be said of the incomplete contract approach in NPRT. This is because asset ownership (property rights) among incomplete contract theorists is still defined *in function of use rights*. Under GHM ownership (asset possession) is defined as the bundle of residual rights of control over an asset concerning the way in which the latter can be used by the owner. In other words, ownership is defined in function of the set of ex-post (or residual) use rights enjoyed by the owner herself.

This conflation between property rights and use rights may generate some confusion. For this reason, it is better to distinguish the two based on their relevant unit of analysis and economic mechanism: while OPRT deals with the efficient allocation of a resource by means of efficient market exchange of use rights among actors, the incomplete-contract strand (NPRT) deals with the efficient appropriation of (or investment in) a resource, by means of efficient ownership patterns. Once again, the key distinction is between use rights exchanged in the allocation process, and property rights acquired in the investment (appropriation) process.

2.2. Use rights and ownership rights: implications for comparative institutional arrangement

The distinction between use rights and ownership rights is fairly well understood, at least in principle. What has not been sufficiently highlighted and discussed in the literature are, on the one hand, the implications of this dichotomy and, on the other hand, the ultimate drivers behind the distinction between property and rights.

The first important implication that has to be remarked concerns the difference in the respective institutional trade-offs proposed by the two approaches (public economics and regulatory approach). It will be recalled that in the “public economics” strand the institutional trade-off is between centralised planning and bilateral (decentralised) transactions. The way in which the two are linked depends on the role of efficient markets in solving the information and the subjective value problems. In the Paretian framework the optimality and the feasibility of the former is the necessary precondition for the efficiency of the latter. On the other hand, in the Nash framework, centralised planning is implemented as a last-resort remedy in order to solve the inherent inefficiencies arising from decentralised bilateral bargaining.

Conversely, the institutional trade-off in the "regulatory" strand is between two alternative ownership patterns of the same asset. Different ownership arrangements modify the relative structure of rights and control over the asset, thus generating different incentives to invest in the asset itself. Yet, this means that the trade-off is not between centralised planning and decentralised bargaining for the *allocation* of resources. Differently, the trade-off is between two different types of ownership structures in a bilateral exchange. This is how hierarchies enter the economic analysis in the theories developing from the "regulatory strand".

In sum, while in the "public economics" strand where the allocative process is considered, the comparative institutional analysis is between centralised allocation and decentralised allocation, in the "regulatory" strand where the investment/acquisition process is considered, the comparative institutional analysis is between unified asset ownership (integration through hierarchy) and dispersed asset ownership.

As remarked by Tirole (1999: 764-765; emphasis added), the various Nash approaches of the "public economics" strand based on "discretionary behaviour under the headings of 'hidden knowledge' and 'moral hazard' … do not quite fit the spirit of Simon's or Coase's contributions. In particular, *the notion of hierarchy is absent*."

The standard explanation for this shortcoming of the public economics strand is that complete contracts are assumed in both the Nash and Paretian approach (in other words, no bounded rationality or unforeseen contingencies appears).[116] This is the standard position of Williamson (1985: 27-29, 69) (but see also Tirole [1999: 754, 764-66]).

However, this is not completely correct. Here we argue that the fundamental reason why hierarchies are absent in the public economics strand has less to do with contracts, and much more to do with the nature or rights considered. Consistent with the previous discussion, it seems that the reason why the pubic economics strand is not able to explain incomplete contracts, and therefore hierarchies, simply derives from the fact that no asset ownership (possession or appropriation) is involved in the general allocative framework.

[116] Complete contracts are also defined as "classical" contracts.

This point will be the subject matter of the next chapter. However, it is sufficient to go back to Alchian's description of rights to see that actors do not possess or own the tract of land, they merely exercise use rights on it. This should not be surprising as it simply derives from the way in which the Walrasian framework is modelled.

Conversely, it is the presence of asset ownership and the possession of resources that introduces the two-stage game in the regulatory legacy, where incomplete contracts, unforeseen contingencies and contract renegotiation suddenly become relevant concepts of the economic analysis. Any incomplete contract model has an investment stage followed by a production/trade stage. The appropriation stage (where possession or ownership rights are exercised) precedes the production/allocation stage when trade is undertaken. It is therefore the introduction of asset ownership (in the form of investment) that generates the temporal decoupling between ex-ante contract formation and ex-post renegotiation. In turn, this justifies the following economic question concerning the comparative institutional analysis: "should a certain asset be owned within a hierarchy, or should it be owned on a stand-alone basis?" this is the contemporary way in which contemporary incomplete-contract theory reinterpreted Coase's original 1937 paper.

On the other hand, in a mere Walrasian world assets or resources are *allocated* according to actors' idiosyncratic use rights over the resource. However, as there is no exchange of resources, assets, or goods (in other words, there is no transfer of ownership rights among actors), then transactions and bilateral bargaining can occur in an immediate and a-temporal fashion. In other words, it is the absence of the institution of ownership that makes so-called "neo-classical" (or incomplete) contracts irrelevant, and *not* the other way around.

The previous discussion highlighted, on the one hand, why the public economics strand cannot deal with decentralised hierarchies but it can explain socialist centralised planning and on the other hand, why the regulatory legacy mainly based on incomplete contracts can deal and explain hierarchies. What has not been explained is whether the regulatory legacy can deal with the issue of centralized panning in the use of resources. The answer is no it can't for several reasons. One of them, probably the most straightforward one, is that the verifiability of actors' private information by the part of third-parties (such as courts or planners) is ruled out by assumption. This forces the analysis to remain on the ground of bilateral (decentralised) interactions.

In conclusion, the distinction between use rights and ownership rights is important as it provides a firmer ground to understand the interplay between the two Coasean legacies (public economics and regulatory economics) on the one hand, and the different ways in which comparative institutional analysis is intended on the other. Thus, as we shall alter discuss, once that the ownership rights-use rights dichotomy is adopted, it will become easier to see that the way in which the new incomplete contract literature has interpreted Coase's 1937 paper is not in line with Coase's original article. Here, it is sufficient to anticipate that a closer look to Coase's *Nature of the Firm* will reveal that his 1937 paper deals with the (static) resource allocation problem and *not* with the (dynamic) resource appropriation/possession problem as it is usually portrait both TCE and GHM (see Tirole's quotation above).[117] This point will be developed more in depth in the following chapter.

2.3. Use rights and ownership rights: implications for the analysis of transactions

The second relevant implication deriving from the distinction between use rights and ownership rights (embedded in the public economics-regulatory economics dichotomy) concerns the definition of the concept of transaction, and the relative concept of transaction cost, in the two frameworks.

2.3.1. The transacted entity: services vs. goods

On the on hand, a potential distinction reflecting the previous dichotomy between property and use rights divides transaction costs into *market transaction costs* (costs of using the market) and managerial transaction costs (cost of exercising the rights). This distinction is provided by Furubotn and Richter (2005: 51). This distinction might be consistent with the public economics-regulatory economics dichotomy as it derives for the verification that the first one is a theory of market allocation, while the second one is a theory of the firm. This is also consistent with Williamson's distinction between transaction costs properly said, and governance costs (1985: 90-94).

[117] Consider the following passages, among many others (Coase, 1937: 387; emphasis added): "in economic theory we find *that allocation of factors of production between different uses* is determined by the price mechanism … If a workman moves from department Y to department X, he does not go because of a change in relative prices, but because he is ordered to do so." And again (p. 389; emphasis added): This distinction between *the allocation of resources* in a firm and the allocation in the economic system has been very vividly described by Mr. Maurice Dobb when discussing Adam Smith's conception of the capitalist". These are just few of many more examples in which it becomes clear that even Coase's 1937 article should be operationalised following the "public economics" strand, rather than the regulatory strand.

The problem with the previous categorization is that it does not establish a precise link between the transaction cost concept and the unit of analysis. This is a problem for a framework that explicitly wants to take into account the role of asset ownership in influencing actors' opportunities in the use of a resource. For a proper institutional theory of asset ownership has to be developed, it is important that the thing transacted on the market is defined with precision. The problem of the definition of the appropriate unit of analysis is much more fundamental than it might appear at first, as it is the preliminary step in order to properly develop a theory of ownership.

Take, for instance, TCE. Williamson (2010: 680) clarifies in a very explicit terms that the transaction is the unit of analysis of TCE:

> "Various units of analysis have been proposed for studying organizations, yet efforts to name the defining attributes of proposed units are rare. The unit of analysis in the transaction cost economics set-up is the transaction".

Yet, when defining a transaction, Williamson (1985: 1; emphasis added) employs the following ambiguous expression:

> "A transaction occurs when a *good or a service* is transferred across technologically separable interfaces."

Thus, a transaction might involve a good *or* a service. However, the distinction between good and service is not trivial. In fact, it is the real essence of the matter. The reason for that is a simple one: while the transaction of a service leaves the ownership patterns unaltered (there is no exchange of physical property rights whatsoever), the transaction of a physical good does involve changes in the relative property rights among the actors. In other words, if the distinction between use rights and ownership rights has any relevance at all, the transaction of a service should be regarded as the outcome of the standard *allocative problem*, while the transaction of a good should already be regarded as the outcome of a theory of vertical and lateral integration or, in other words, as the outcome of a theory of *asset ownership*. This is true unless the physical good itself is treated as if it has the same nature of a service.

This confusion concerning the nature of the entity transacted is certainly a heritage of the vocabulary of the standard microeconomic textbook where, in fact, no clear distinction between goods and services are made. However, this cannot constitute an excuse as the conflation of physical goods (subject to ownership rights) and pure services (subject to mere use rights) is really the essence of the problem of the standard Walrasian framework that any economic theory of institutions is trying to address.[118]

Two main reasons can be advanced in order to explain why this distinction between goods and services is the hearth of the matter.

2.3.2. Services, goods and the meaning of transaction costs

The first reason (more immediately relevant for the present purposes) why the constant confusion between goods and services is a relevant one is that it then carries over to the definition of transaction costs. For example, Barzel (1997: 4) defines transaction costs as "the costs associated with the transfer, capture, and protection of rights." Yet, once that the definition of rights is examined, the following ambiguous statement is provided by the author (1997: 3; emphasis in the original):

> "I define the economic property rights an individual has over a commodity (or an asset) to be *the individual's ability, in expected terms, to consume the good (or the services of the asset)* directly or to consume it indirectly through exchange."

Once again, there is no clarity on whether the individual consumes or transacts assets, or if she consumes and transacts services. As previously remarked, the point is relevant as it has implications for the relevant ownership rights over the asset (or the good). It is evident that no clear theory of hierarchies can be developed from this confusion in vocabulary.

It should not be surprising to see that in NPRT (theories based on incomplete contracts in general) the stage involving the transaction of the asset is clearly distinguished from the production stage, involving the provision of the specific output. Consistently, transaction costs in incomplete contract theory are defined over time, in a dynamic setting (this is the uncertainty/bounded rationality dimension in TCE).[119]

In general, the entire literature in economics intended as the science studying the allocation of resources (no matter whether it is in standard microeconomics or in institutional economics developing from the Coasean "public economics strand") tend to

---

[118] As recalled by Furubotn and Richter (2005: 80): "For the zero-transaction-cost economy of general equilibrium theory, it does not matter whether the factors of production are owned or rented by their users." More precisely, in ordinal neoclassical economics it is not even true that services (for rent) and assets (to own) are indifferent. In order to guarantee the assumptions behind the ordinal framework, firms can only transact and purchase pure services among them. This is the heritage of the marginalist revolution in distribution theory as we shall see.

[119] See Williamson (1985: chapter 3). Alternatively, Tirole (1999: 743-744) distinguishes three types of transaction costs: unforeseen contingencies, costs of ex-ante writing contracts; cost of ex-post enforcing contracts.

treat goods, factors, or even assets as interchangeable synonyms of services when dealing with the entity transacted on the market. This cannot be correct, nor it can help in trying to provide a proper theory of ownership based on actors' perceived opportunities in the use (and consumption) of the goods. As already remarked, the problem derives from the fact that, while the transaction of assets and goods between actors entail a transfer of property over the physical "things", services are the genuine manifestation of the use of an asset (either for private and personal reasons or for indirect usage through market transaction).

Thus, a proper classification of transaction costs must retain the useful distinction between use rights and property (ownership) rights, which is also a distinction between the allocative problem of the Coasean "public economics" strand, and the integration/appropriation problem of the "regulatory" strand. As previously illustrated, the same dichotomy can also be reinterpreted as a distinction between market failures in value theory (the allocation of the resources to alternative uses) and market failures in distribution theory (the remuneration of the resources' owners for their

The classification provided by Allen (2000) is probably the most appropriate one, as it starts from the clear distinction between use rights on the one hand, and assets or goods' property on the other. While the former are associated to "property-right" transaction costs,[120] the latter are associated to neoclassical frictions,[121] depending on whether rights concerning the use of a resource, or proper property entitlements over assets are transacted. While the "public economics legacy" of OPRT and L&E deals with the former type of transaction costs, the "regulatory legacy" *mainly* deals with the latter. However, in this second case the distinction between the two is much more blurred.

2.3.3. From regulation to antitrust law in the ICT industry: mandatory unbundling and service-based competition

The second reason why the distinction between goods (resources or assets) and services is a relevant one for the present purposes has to do with the specific case study here examined: the deregulatory process in the ICT industry.

---

[120] As explained by Allen (2000: 898): "Although economic property rights are enhanced by the law, they are ultimately use rights and the greater extent one can exercise these uses and bear the consequences the greater are the property rights, regardless of the law."

[121] Allen (2000) provides two definitions of transaction costs. What he calls the "property right definition" defines transaction costs as "the costs establishing and maintaining property rights" (p. 898). The "neo-classical approach" deals with ownership titles, and is defined as "the cost resulting from the transfer of property rights" (p. 901).

More precisely, one of the most prominent characteristics of the new ICT sector is the digitalization of the signal transmitted. The digitalization process has two main consequences:

1. It allows the convergence of alternative platforms towards one single final market, thus fostering inter-platform competition
2. It generates a decoupling between alternative logical layers of the transmission facility according to the TCP-IP protocol.

The second point is the one that particularly interest us here. The two logical layers of the TCP-IP protocol under consideration are the physical layer of the infrastructure, and the layers in charge for the management, processing and routing of the services composing the actual communication between two end points. The layers in charge for the routing of data and the establishment of the connection usually correspond to the so called "network" and "transport" layers (see Tannenbaum and Wetherall, 2011).

This peculiarity of the new digital environment is particular attractive to regulators as it can shift the boundaries of competition from the realm of the physical facility to the sole sphere of services, leaving the physical infrastructure (thus the asset or resource) untouched. In other words, the strategy for the introduction of competition in the formerly regulated industries is to replace infrastructure-based competition with a new service-based competition, while retaining common infrastructure. Yet, this means that the regulatory environment in the new digital (and converged) ICT sector should be able to distinguish between the regulatory issues deriving from the *ownership of a facility* (vertical or lateral integration) and the issues deriving from the *allocation of the services* running on it.

This distinction reflects the one between value and distribution theory previously highlighted and, in turn, the one between use rights and ownership rights. Of course, the latter represents the foundations of the dichotomy between the public economics and the regulatory strands of the Coasean legacy.

This point will be the focus of chapter 7 below.

2.4. Use rights and ownership rights: going beyond the public economics – regulatory dichotomy

Although many (but not all) authors have recognised the distinction between use rights and property rights, the two economic approaches have proceeded quite independently, while sometimes confusingly conflating into one another. The closest point of contact is represented, as previously discussed, by the acknowledgement of theories dealing with the "theory of the firm", beside those theories dealing with the allocative process. However, the simple identification of the regulatory strand with the simply "theory of the firm" is at the same time reductive and misleading. It is reductive as there are many other ways to define a firm. It is misleading as, consistent with what was already anticipated, Coase's 1937 article is not really about asset ownership, but it is an article about alternative ways in which services are valued and allocated.

What is missing in the literature is therefore a more precise recognition of what the fundamental difference between the "public economics" theory of institutions and the "regulatory" theories of institutions is. At the same time, there is the need for a general framework that can shed some light on how one approach inform, and is influenced by, the other and vice versa. In other words, it is important to understand how the ownership of a resource and its usage interact so that the two strands can be reconciled. This not only constitutes a fundamental starting point for a theory of ownership exclusively based on capabilities and opportunities, but it also represents a preliminary starting point for any theory of regulation.

In effect, a reconsideration of the essence of regulatory economics can provide a preliminary clue on the way in which the two strands interact and can be merged. It is interesting to note at this point that, although Coase's 1960 article explicitly deals with the "public economics" problem of externalities, the other forefather of PRT (Harold Demsetz) has developed the his theory of property rights in close connection with regulatory economics. The way in which Demsetz explicitly links regulatory economics with PRT (1988: 12) closely recalls the way in which regulatory practices are usually defined; as a restriction on the options and opportunities of economic actors (see Spulber, 1989: 32-40). The fundamental difference is, of course, that while regulatory economics focusses on the options and opportunities *in purchase*, Demsetz and the other founders of PRT look at how the external context shapes the opportunities *in the use* of a resource. The distinction between purchase and use here is fundamental.

More generally, a better understanding of the way in which asset ownership and actors' perceived opportunities in use can be merged becomes important not only in order to provide a new interpretative framework on the way in which asset ownership should be approached in regulatory economics, but also for other more or less connected, reasons:

- In order to provide a theoretical framework in order to analyse the phenomena of convergence and deregulation in the ICT industry;
- In order to reinterpret the make-or-buy problem and the problem of the boundaries of the firms;
- In order to define a new (opportunity-based) theory of subjective value;
- In order to provide a new understanding of the historic change from the Ricardian (surplus-based) theory of distribution, to the marginalist theory of distribution.
- In order to better understand the microeconomics of resource-based theories in the managerial and organizational literature.[122]

The distinction between the economics of resources, or assets, and the economics of uses, or services, is the cornerstone of the present work. It is then very important that the two dimensions of economics are first properly identified and, second, correctly linked to each other.

Here the following interpretative framework is proposed: in order to understand how assets and opportunities in use interact or, alternatively, in order to understand the relationship between the "public economics" approach to institutions and the "regulatory economics" approach it is important to reconsider the following dichotomies:

- Partial vs. general competitive analysis
- Distribution theory (generating a theory of resources' ownership) vs. value theory (linked to a theory of services allocation, and therefore of use rights)
- Adaptation towards "shadow" opportunities vs. allocation across given opportunities.

The previous analysis has already highlighted that, while the public economics legacy tend to focus on the second of the three dichotomies, the regulatory legacy tend to focus on the first terms.

Thus, while the "public economics" strand is a genuine analysis in general competitive analysis, the "regulatory" strand developed from partial equilibrium considerations. An exclusive focus on asset ownership (property rights) denotes a partial equilibrium approach (where Marshallian indivisibilities are the prominent focus of the analysis). On the other hand, an exclusive focus on use rights and on the behaviour of actors in using a given

[122] Concerning the literature dealing with the micro-foundations of resources, capabilities see Abell et al. (2008), Felin and Foss (2011; 2012), Felin et al. (2012), Foss and Lindenberg (2013); Helfat and Peteraf (2015).

resource denotes a general equilibrium approach (where the Walrasian allocative process is the prominent focus of the analysis).

The following chapter will shed more light on that, and will try to expand the previous considerations concerning the role of subjective valuations, and to link them with the concepts of assets, services and, most importantly, adaptation and allocation.

3. Conclusions: General and Partial equilibrium analyses in a theory of value and market failures

In conclusion, it is really not possible to appreciate the origins and the meaning of Coase's contribution, and therefore the origins of contemporary NIE, if the original long-standing issues and debates in: (a) the theory of optimal public utility pricing (and cost accounting) in regulatory economics, together with (b) the problems arising from the shortcomings associated to the theory of centralised allocation and centralised taxation in public economics are not taken into account. Both fields really represent the two original forefathers of the many theories of institutions that developed shortly after Coase's contributions: the theories on the role of property rights and property rules, the theories on the origins (and the boundaries) of the firm, the theories on the origins of centralised institutions (such as the regulatory agency), and on the origins of legal institutions such as, among others, the institution of ownership.

The fundamental influences of the long-standing debates in public economics (such as the problems of taxation, public utility pricing, and externalities) on Coase's original works on property and liability law, and on the implication of the optimal centralised allocation of the scarce resources (i.e. the radio spectrum) have been acknowledged and recognised shortly after Coase's contribution.[123] The same cannot be said for the other, equally important, types of influences on Coase's intuitions: cost accounting and the regulation of natural monopolies, especially with respect to its implications for antitrust economics and the theory of the firm. In this respect, it is famous Coase's bold assertion concerning his "much cited and little used" article on the theory of the firm (Coase, 1972: 63) that, not surprisingly, appears in the only work that Coase has ever explicitly dedicated to industrial organization and therefore to antitrust economics.

[123] As it is well-known, already in 1966, Stigler reinterpreted the message of Coase's *The Problem of Social Costs* coining the expression "the Coase theorem" (see Medema, 2011). Likewise, Buchanan and Stubblebine (1962) fundamental paper on externalities will come only two years after Coase's paper. With respect to property right theory, by the beginning of the 1970s there was already a flourishing literature on the topic, as discussed by Furubotn and Pejovich (1972).

While Buchanan (1969) and Buchanan and Thirlby (1973) represented two very rare exceptions concerning the neglected influence of the role of cost accounting on Coase's writings,[124] Oliver Williamson is, of course, the other big exception, being the first one that, starting from 1970s, clearly took seriously the influences that natural monopoly regulation and competition law had on Coase's argument concerning the role of institutions in economic analysis.[125] The influence of the regulatory and antitrust literature on Williamson's TCE is enormous.

A standard way to proceed would therefore now switch to a discussion of Williamson's contribution to the Coasean transaction cost legacy. However, this will not be the path taken here for one simple reason: differently from all other developments in welfare (normative) economics, the contractual approach adopted by TCE introduces something completely new into the picture: the concept of *asset ownership*. Assets, goods or resources in general are fundamentally absent in all developments of the Coasean "market failure" approach up until 1970s.[126] More generally, the concept of asset or resource has been ostracized from the realm of economics long before Coase and more precisely since 1894, when Ricardian distribution theory (based on the *ownership* of resource, namely land) was definitively replaced by the new marginalism in distribution theory, as illustrated in chapter 4.[127] The standard reference usually regarded as the beginner of the new era is Philip Wicksteed's *Coordination of the Laws of Distribution*, published, precisely in 1894.

As already remarked above, it is only after 1970, when the new "contractual" approaches *in partial equilibrium* based on incomplete contracts and unforeseen adaptation were developed (first TCE and later GHM), that assets, and the institution of asset ownership, make their return back into the discipline of economics in informing and explaining economic outcomes. Not surprisingly both these two fields are genuine developments of what was called the "regulatory legacy" of Coase's writings and, more precisely, of his 1937 paper. In truth, Coase'1937 paper was much closer to the spirit of the public economics

---

[124] Buchanan links Coase with what he calls the "LSE tradition" in cost theory. Yet, it is not clear how much weight Coase puts on this categorization (Coase, 1982: 33; 1990: 8-9).

[125] For instance, it is interesting to note that, beside Coase's seminal articles, Williamson (1994: 203) suggests that for "[t]hose who wish to continue should read 'The Regulated Industries: Discussion' [Coase, 1964], 'Durability and Monopoly" [Coase, 1972] …" These are specific articles of Coase on the regulation of natural monopolies that are not very well known.

[126] This is not a striking statement; See for example, among others, Allen (2000) and Foss and Foss (2001).

[127] It is interesting to note that early writers in PRT took land as the standard case for their illustrations (see Coase's farmer and cattle raiser example in the *Problem of Social Cost*, or Alchian's definition of rights provided above).

branch rather than to the interpretation provided by Williamson. Nevertheless, it is a fact that the *Nature of the Firm* from a writing dealing with the *allocation* of resources (and thus a work in *value theory*) became, under Williamson, a work on the *ownership* of the resources (and therefore a work in *distribution theory*).

Yet, it should be highlighted since now that this neo-Ricardian wave in organizational and institutional theory (whose main exponent and precursor is Armen Alchian as far as I can see)[128] was not without problems.[129] The single most important deficiency of this new institutional and organizational economics is that assets, resources or goods, were all introduced into the picture at the expenses of another concept that, until the 1970s, really represented the cornerstone of the economic discipline: the allocative process or, using Robbins' (1935: 16) definition, "the relationship between ends and scarce means which have alternative uses" which has long been the official definition of economics and that I think should still be regarded as the correct way to approach an economic problem. This was certainly not the position of Coase (1988b: 22-25; 1992) or Williamson (2003: 920): as we shall see, there is no proper recognition of the *constrained allocative problem* in the new theories of asset ownership based on incomplete contracts. The sharp dichotomy between the early theories in property-right theory (and L&E) and the latest developments in incomplete contracts derive precisely from this shift of the analysis focus from the

[128] The figure of Armen Alchian is quite eclectic as he not only represents one of the forefathers of early PRT, but he is also one of the forefathers of, what Gibbons (2005) call, the "rent-seeking" theory of institutions. In fact, not only Alchian decisively reintroduced the concept of rent in economic analysis, together with Klein and Crawford, in their path-breaking 1978 paper (Klein et al., 1978), but he also explicitly linked ownership with the rights to receive the residual of production, or, more simply, the *production surplus*, in his very well-known paper on production costs and organization, written together with Harold Demsetz (Alchian and Demsetz, 1972). This paper will then become the precursor of contemporary property-right-theory (NPRT) approach, where the key intuition, as summarised by Grossman and Hart (1986: 696; emphasis in the original) is that the "distribution of ex-post surplus … will be sensitive to ownership rights". It is probably not possible to give a better synthesis of a surplus-based theory of distribution, the distinctive trait of Ricardian economics. As a result, in Alchian's works the Ricaridan echoes are undoubtedly powerful so it is not surprising to find that Alchian was also the author of the entry "Rent" in the *New Palgrave Dictionary of Economics* (Alchian, 2008).

[129] More precisely, what really characterizes the revival of Ricardo is the prominence of two old concepts. First, the old classical concepts of *rent*, intended as a mere residual, or surplus, occurring to the resource's owner (and not as the strategic monopoly rent deriving from the constriction of production). Second, the concept resources' *appropriability* in the new theories of contracts, organizations and institutions, as opposed to the marginalist concept of *profit* obtained by substitution at the margin of the inputs. For this reason the obliged reference is Klein, Crawford and Alchian (1978). Yet, once again, this is just a further consequence of the reintroduction of the institution of asset ownership into the analysis, as we shall argue below. Note that it is not a case that contemporary theories of the firm in the managerial literature are mainly based on the concepts of resources, appropriability and rent (intended as residual surplus).

"good-old" allocative problem to the new (but even older) acquisition (resource ownership) problem.

This is certainly a consequence of the fact that the new incomplete contract frameworks developing from the Coasean "regulatory" legacy, originated in the partial equilibrium tradition, where the allocative process occurs on the background, and can be substantially neglected. However, any theory of ownership based on the concept of *capabilities intended as opportunities* cannot overlook the genuine economic problem deriving from the trade-off among competing opportunities in the way in which actors *employ* their goods and resources. Yet, in order to do that, an opportunity cost schedule is required. And opportunity cost schedules can only enter the picture in a general equilibrium analysis.

For the above reasons, before switching to a more accurate analysis of the economic theories of asset ownership, it is first necessary to give a brief look to the developments in the theories of institutions that are rooted in the standard *general equilibrium tradition*. This will provide a solid ground in order to better understand the relationship between transaction costs (market failures) and institutions. In effect, the essence of market failures in economics can only be properly understood in general competitive analysis. Only in a second step, once that a proper taxonomy of market failures (in general equilibrium) is defined, it will be possible to switch to the more recent theories of asset ownership, rooted in the partial equilibrium tradition. The final task will be to eventually merge the two in a unique coherent framework of asset ownership where both the distributive problem (linked to ownership rights) and the allocative problem (linked to use rights and capabilities) have equal weight and legitimacy.

In conclusion, the aim of the next chapter can be summarised as follows: to understand the mechanisms generating the decoupling between actors' subjective valuations and markets' objective valuations. A theory of ownership based on actors' subjective opportunity sets (in the employment of assets) will necessarily have to start from that.

CHAPTER 6

Property "Affected with Public Interest" and public duties in common law: understanding the legal foundations of the asset-service and private-public dichotomies and their implications for public regulation of private property

Chapter synopsis

Chapters two, three and four represent a general overview of the role of value with respect to the dichotomies between assets and services, and private and public domain, respectively. The chapters also provided policy and normative implications concerning the legitimacy to publicly regulate private properties. The key concept of these chapters was the one of use value, and its meaning represented by the act of private self-consumption of a property, or auto=consumption.

The reason why a correct understanding of the role of use value is fundamental in law and economics derives from the fact that it is not possible to understand the essence of common law, and some of its key founding concepts if the distinction between use value and exchange value (together with its economic implications) is not taken seriously. The present chapter shows precisely that the legal foundations of regulation, and of public control of private properties developed out of the key distinction between private self-employment of a property, and public use of a property through the provision of services to third parties.

In other words, the present chapter places the previous discussion into context, and gives it historical and legal relevance.

The chapter focusses on the two key concept of common law regulation: the one of public interest, and the one of public calling, from which the most important modern regulatory institution derives: the one of common carriage.

The aim of the chapter is twofold.

On the one hand, the chapter discusses the distinction between the concept of asset and the one of service in common law, showing that the original common law approach to public regulation cannot be properly understood if the distinction between the two is not preliminary address or taken into account. On the other hand, the chapter clarifies the meanings of the two concepts of public and private, and discusses their distinction. The chapter will show that a correct understanding of the distinction between assets and services becomes fundamental in order to properly grasp the meaning of the distinction between private and public in common law.

Consistent with the discussion developed in the previous chapters, the chapter shows that the private domain is always characterised by the ownership of the underlying physical property, or asset, while the public domain is always linked to the transaction and reward of services.

The chapter discusses and illustrates these "conjugate" dichotomies, and criticises the modern "neoclassical" and "failure-based" reinterpretations of the concepts. The chapter then, also represents an indirect critique of the modern deregulatory approach that has interpreted common law concepts in a neoclassical, and failure-based manner.

The chapter is divided into three main sections.

Section one discusses the concept of "public interest". The section shows that the neoclassical reinterpretations of the concepts are incomplete. A critical re-examination of the common law origins of the concept shows two fundamental aspects of the concept.

One, public interest is a concept defined in property law and, as such, it only characterises assets or resource, not a services, businesses, industries, operations or activities. Two, the test establishing "affectation with public interest" can only be conducted once that the property owner decides to employ his property to serve the public, but it can never be conducted in case of private auto-consumption of the property.

The chapter argues that contemporary (failure-based) interpretations of the concept have misinterpreted the original common law nature of "affectation with public interest" as they usually employ the concept to characterise an industry, or a business, never a property. This shortcoming, in turn, derives from the fact that the distinction between private and public is not framed in common law terms, and therefore as a distinction between *private self-consumption* of the services of the owned property, and *public consumption* of the same services. Conversely, a typical market-failure interpretation of public interest exclusively looks at the competitive characteristics of the market in which the services of the given property are sold. The logical causality is therefore reversed, from the final sale of services, to the property.

Sections two and three develop the distinction between private and public in greater depth, and focus on the concept of public calling.

The concept of public calling shares with the one of public interest its common law origins and for this reason, it also shares its approach to the public-private distinction.

Section two shows that, differently from modern interpretations, the status of public calling does not depend on the peculiar nature of a business, or on the competitive nature of the market served. Conversely, the distinction between private and public callings in common law distinguishes between the case in which a property is normally employed for private purposes, and is "held out" to serve third parties (the public) only salutary, and in an impromptu manner, and the case where properties are purposefully employed in order to serve the public and generate monetary compensations out of it.

As a result, the two duties of a public calling (the "hold out duty", and the duty to "serve with due care") should be intended as legal remedies in order to distinguish those properties or actors that should be held liable as they make their living out of serving the public (public callings), from properties and actors that should not held liable as the services provided are sporadic exceptions undertaken out the good will of the property owner. Thus, properties can be receive the status of public or private calling depending on whether their primary purpose is to serve the public, or to serve their private owner respectively.

The chapter remarks the important fact that this distinction relies on the distinction between the self-employment of a property for private self-subsistence, and the public employment of a property purposefully and exclusively employed to serve third parties. The imposition of the two common law duties derives from the awareness that only in the second case (normative) social value judgments can be legitimately derived.

Section three provides a critique of the "contractarian" (failure-based) approach to the distinction between public and private calling. It discusses the fact that the distinction between a private and a public calling cannot be interpreted in contractual terms. Public and private callings do not differ based on the type of contract that the business activity establishes. Conversely, they differ as while the former necessarily has to serve the public, the latter does not need to deal and to serve the public. As a result, while the duties of the former rely on social status (defined ex-ante), the obligations of the latter can only depend on the specific contract established ex-post, once that the property owner has decided to open his services to the public.

The difference between a private and a public calling is therefore not a distinction between different types of contracts, but it is a distinction between private contracts and social status: the obligations of the former can only causally follow *ex-post* a contingent decision to serve the public has been made, while the duties of the latter are established *ex-ante*, and they causally precede any request of service by a member of the public. The latter define the social status of the private endeavour, not only independently from the type of transaction, but even independently from whether there even is a transaction.

In sum, the chapter highlights a fundamental commonality between the conclusions developed in the previous chapters, and the approach adopted in common law: while the private sphere is univocally associated to the underlying property (property-based analysis), the public sphere is univocally characterised by the services provided by that property (the way in which the underlying physical support is employed). Public regulation can only develop its normative (legal) considerations based on the services transacted on the market, in the public sphere.

The essence of the legal constraints and obligations developed in common law should be interpreted in this way: as necessary legal tools whose role is to establish a univocal mapping between the *nature of the transacted service* and the *status* of the underlying private domain. Common law recognises that only in this case a univocal logical causality from the public sphere where services are valued and transacted to the private sphere where properties are operated becomes legitimate. The solution devised by common law is (not surprisingly) to rule out auto-consumption (private self-employment of properties), and to mandate "holding out" of the property to the general public under specific terms of service the property owner should conform to. In this way, the services transacted in the public domain can univocally identify (backward) the properties owned and appropriated in the private domain.

The chapter anticipates the essence of the common law origins of public regulation and public callings, which will be developed at greater depth in the following one: the recognition that the actions undertaken in the private domain can be normatively judged only if they conform to a pre-established public status that subject the private sphere to a public role, thus ensuring the existence of objective and univocal value judgments univocally linking private and public domains, while also merging the two into an indistinguishable "institutional monism". If this is not the case (in the absence of public duties), common law recognises that the public domain cannot legitimacy judge a private actor for how the latter ought to employ his own property, nor for the services provided. In the absence of public duties norming the behaviour of private actors the univocal "reversed causality" running from the service performed to the underlying private property cannot be established. In this case, common law recognises that the private actor cannot be charged for disservice, and cannot become liable for the services provided as there was no public duty norming what services ought to be provided to the public, and how they should have been provided, in the first place.

1. The origin of modern (price) regulation and "affectation with public interest" in common law: a historical reinterpretation of the relationship between property and services in the switch from *juris privati* to *juris publici*.

The present section provides a historical enquiry of the well-known expression "property affected with public interest." This represents the natural starting point of our analysis as the public regulation of tariffs has been officially legitimized by the sentence of the US Supreme Court in in *Munn v. Illinois* (1877) by relying on the common law concept of "public interest."

The literature on the topic is vast.[130] Yet, as opposed to the standard academic approach to the topic, the present section offers a slightly different angle on the topic. The present discussion does not focus on the legitimacy of the public authority to intervene on the process of rate-fixing (price regulation). Nor does it focus on the long-standing problem of access pricing, and on the conceptual problems raised by the centralised definition of prices.

Conversely, the section is exclusively interested in understanding the essence of the common law distinction between *juris privati* and *juris publici*, which characterises the legal foundation of "public interest". The remainder of the section focuses on the dichotomy

[130] The following is a non-exhaustive list of the references the present section relies on: Fairman (1953), Haberson (1930), Hamilton (1930), Hovenkamp (1988), J.L.A. (1932), Kitch and Bowler (1979), McAllister (1930), Miller (1971), Rabin (1986), Siegel (1984), Sorauf (1957), Wyman (1904).

between the concept of service and the one of asset (the property), while highlighting the fundamental conceptual challenges raised by the unclear and problematic interaction between the two from a legal point of view. The essence of the distinction between the two legal categories of *juris privati* and *juris publici* cannot be understood if their respective relationship with the two concepts of service and asset is not correctly understood first.

As the section will show, "affectation with public interest" applies to the given property, yet it depends on the specific way in which the latter is employed or, in other words, it depends on the specific service rendered by the given property.

The essence of the common law approach to "public interest" derives from the fact that, in order to understand the legal (public) status of a property, it is first necessary to understand how its respective owner ought to employ it: either to *serve the public* (public consumption of the property's services), or to simply *serve himself*, for an auto-consumption of his own property's services. The distinction between the two represents the essence of the first necessary and preliminary condition that has to be verified in order to legitimise the switch in the legal status of a property: from *juris privati*, to *juris publici*.

1.1. Public Interest in common law is defined by looking at the rights of things: juris publici vs. juris privati

The legitimacy of public control over rate setting was recognised by the Supreme Court in *Munn v. Illinois* leveraging on the concept of public interest, with the following words (cited in Fairman 1953: 675; emphasis in the original):

> "This brings us to inquire as to the principles upon which this power of regulation rests, in order that we may determine what is within and what without its operative effect. Looking, then, to the common law, from whence came the right which the Constitution protects, we find that when private property is "affected with a public interest, it ceases to be *juris privati* only." This was said by Lord Chief Justice Hale more than two hundred years ago, in his treatise *De Portibus Maris*, […] and has been accepted without objection as an essential element in the law of property ever since …"

Since then, public interest has been defined according to the common law tradition, and intended as the superimposition of *juris publici over juris privati*: whenever a private property is affected with public interest, it ceases to be *juris privati* only, and it becomes *juris publici*. This means that, given certain necessary conditions, as will be defined below, a thing, previously private, changes legal status, thus becoming *juris publici*.

This expression needs some further comments and clarifications.

The definition of public interest on which the legitimacy of public interventions rests comes from a treatise of the 17th century written by Sir Matthew Hale, the then Lord Chief Justice of the King Bench of England. Some historical descriptions can be found in Fairman (1953) and McAllister (1930). The treatise in question is entitled "*De Portibus Maris*" (literally: "on the ports of the sea"). In this and other writings, Lord Hale identifies three distinct branches of the law: *the rights and powers of the King* (which would then become "the State"), the *rights of persons*, and the *rights of things*. The public regulation of trade and commerce, as well as the identification and designation of places of public commerce, are prerogatives of the King (rights of the King). Yet, this branch of the law has nothing to do with the distinction between *juris privati* and *juris publici*.

The distinction of the two legal regimes applies to the third category of the law: *the rights of things*. The rights of things can be divided into *juris privati* and *juris publici* based on the location and the nature of the constraints shaping the way in which the thing (asset or resource) can be accessed, used and enjoyed by its owner.

The direct object to which the distinction between *juris privati* and *juris publici* applies, is therefore *the thing* or, the *physical property*: *some-thing* is defined as a *juris privati* or *juris publici*. The regulation of businesses, trades and activities all belong to a different department of common law, which is in principle unrelated with the rights of things and properties. As summarised by McAllister (1930: 761):

> "A further reference to the *Analysis* will enable us to make the definite statement that to Lord Hale the notion of *juris publici* not only had no relation to the regulation of trade and commerce, but also had no connection with the notion that a trade or occupation was 'common.'"

This is just a restatement of the fact that in common law, since the 17th century, the regulation of business and trade was conceptually distinguished from the law of property. While the former was under the prerogatives of the King, the law of property followed other logic. "Affectation with public interest" belonged to the latter, not to the former.

Three points are worth mentioning here.

First, the expression "affectation with public interest" applies to the thing that is the object of property. It is first fundamental to understand what the thing under consideration is, in order to derive some legal conclusion on the legal status (public vs. private).

Second, the same expression ("affectation with public interest") does not apply to the operation or management of a business, nor to the rendering of a service. This derives from the fact that the distinction between *public* and *private juris* does not apply to the operation of trade and commerce in the first place; it only applies to things, or physical objects. Thus, it is incorrect, at least in principle, to assert that a certain activity, business, industry, service or operation is "affected with public interest". The way in which the underlying thing (constituting the physical input) is linked to the services operated "over-the-top" (representing the immaterial outputs of the production process) represents a further logical and analytical step, and becomes legally relevant only later in the analysis. Yet, to establish a univocal linkage between the property and its employment is an undue step requiring some extra-justification.

Three, the definition of a certain occupation or business as "common" derives from the fact that *the things* used in order to operate that specific occupation is "affected with public interest", and is therefore a *juris publici*. As we shall see, although the specific legal status of a thing does transfer to the specific service rendered, or business operated, it still remains fundamental to highlight that this transfer has a very precise logical, analytical and temporal *causality* (directionality): form the rights of the property, to the rights of the business, and *not* vice versa.

In sum, a common law analysis of the foundations of public interest should start from the *rights of things*, while keeping them distinct from the rights of people and from the regulation of business and commerce. The legal status is originally endowed to the thing, and emanates from the thing. Accordingly, public interest, on which contemporary regulation is based, derives from, and is based upon, the legal status of the property, and not on the characteristics of the business or its productive operations.

1.2. Necessary conditions to understand when a property is affected with public interest in common law: from the property (asset) to its use (service)

The rights of things should therefore constitute the starting point of the common law approach to public interest and this implies that the given unit of analysis should be the thing, asset or resource employed. At this point, it becomes necessary to understand what the distinction between *juris publici* and *juris privati* means, under which circumstances it is possible to switch from one to the other, and what it entails for the notion of public interest.

In order to do so, it is still fundamental to go back to the original description provided by Lord Hale, as recalled by the U.S. Supreme Court in 1877.

Even though the distinction between *juris privati* and *juris publici* applies to the thing – meaning that the actual property considered is the unit of analysis, it has always been very clear that the legal status of this property exclusively depends on its employment or, on *the way in which the property is used by its owner*. This means two things. One, the legal status of the property depends on the specific use the latter is dedicated to, by its owner; two, the legal status of the property depends on whether the chosen employment is a public use or, conversely, a private use. Given a specific use chosen by the property owner, "affectation with public interest" can be mandated only as far as the property is devoted to *public use*.

It is then necessary to understand what public use means, and when it differs from private use.

In Lord Hale's words, the condition of public use can be summarised as follows (quoted in Fairman, 1953: 656):

> "If the king or subject have a publick wharf, unto which all persons that come to that port must come and unlade or lade their goods as for the purpose, because they are the wharfs only licensed by the queen, according to the statute of I. El. cap. II. or because there is no other wharf in that port, as it may fall out where a port is newly erected; in that case there cannot be taken arbitrary and excessive duties for cranage, wharfage, pesage, &c. neither can they be inhanced to an immoderate rate, but the duties must be reasonable and moderate, though settled by the king's license or charter. For now the wharf and crane and other conveniences are affected with a publick interest, and they cease to be juris privati only;"

Thus, if the property under consideration is devoted to serving the public (public use) so that it is used by the public under remuneration, while also becoming of fundamental importance to the public at large, meaning that the entire community will be greatly damaged or will be unable to carry on with its affairs without access to the property itself, the property has to be affected with public interest.

The logical connection from the wharf and the crane described by Lord Hale in the 17th century, to the case of grain elevators involved in the *Munn* sentence was apparent to Justice Bradley, as recalled by Fairman (1953: 656):

> "The doctrine 'is again asserted by Lord Hale' and 'applied to wharves and public accommodations erected thereon … Thus we see that the Common Law alone without any legislative interposition, recognizes the principle on which the legislation in question is founded. Surely, therefore, it is not going far to say that the subject is within the scope of legislative authority. The laws complained of do little more than carry out, and give practical effect to, the Common Law."

The interpretation provided by Justice Bradley convinced Chief Justice Waite, who stated the final sentence of the Supreme Court for the *Munn* case in the terms quoted above.

In sum, from the above overview, three fundamental points emerge. A given private property is affected with public interest whenever:

1. It is used and employed for third parties so to serve the general public (public use), and not exclusively its owner (definition of private use);
2. It affects the community at large, meaning that the specific service the property is devoted to becomes necessary and fundamental for the entire community (definition of the nature of the specific service the property is employed into);
3. It can be regarded as a "common charge", meaning that all members of the community have to pay in order to access or use the property so compensate the owner for making the property available to the public (definition of common charge).

Two fundamental remarks are in order at this point.

First of all, these three points can be regarded as the three *necessary conditions* for the emergence of public interest, and thus for the property to "cease to be *juris privati* only", and to acquire the new legal status of *juris publici*. They are all necessary to decree the "affectation with public interest". This means that, neither condition (nor a combination of two) is sufficient, alone, to decree the presence of public interest.

As a result, a property is affected with public interest (*juris publici*) whenever the property is open for the use of the public, whenever the whole public becomes dependent and reliant on it for its necessary activities, and whenever a certain monetary transaction (compensation) between the public and the property owner is involved. In Bradley's words (quoted in Fairman, 1953: 658; emphasis in the original):

> "Wherever a particular employment, or a business establishment becomes a matter of public consequence so as to affect the whole public and to become a '*common charge*,' it is subject to legislative regulation and control."

Secondly, these three points have to be read and interpreted *in sequence*, as they are listed above: from point one to point three. In other words, these are necessary conditions to be tested in a cumulative fashion. The level of the tariffs should be analysed only once it is ascertained that the specific way in which the property is employed is a public necessity. Similarly, the characteristic of "public necessity" should be verified only after it is first ascertained that the property is employed by its owner in a determined way, and that it is first devoted to public use.

This second passage is particularly important as it shows that, although the notion of public interest depends on the specific way in which the property is used, it nevertheless applies to the property itself. Consistent with the discussion above, the property must remain the starting unit of the analysis and, based on the way in which the given property is used, public interest can be ascertained for *that specific employment* of *the given property*, and *only once the property owner decides to devote it to public employment*.

1.3. Fundamental problems with the standard (monopoly-based) approach to regulation: properties are always employed to serve the public

The notion of public interest in common law relies on three conditions, and these form a sort of "cumulative" three-criteria test that, in its essence, is very similar to the one nowadays adopted by regulatory authorities in order to check whether a certain sector shall be regulated (ERG, 2008).[131]

[131] In order to understand whether a sector should be regulated the European Commission checks three distinct necessary conditions: (a) the presence of high and non-transitory entry barriers, dynamic of market structure, and the insufficiency of competition law alone. This test starts from the structural characteristics of the production process in order to derive normative policy implications (ERG, 2008).

A standard analysis of the concept of public interest (historical and analytical) usually focuses on *the last two* conditions highlighted above (2 and 3). The two summarise the two well-known and long-standing problems in regulatory economics.

On the one hand, condition two represents the essence of the standard contemporary approach to regulation as it requires a definition of the nature of public interest: what does it mean that a service affects the community at large? When is a certain service considered a necessity for society? The standard historical assessments of the birth of contemporary regulation starts from this second step, and presents the various possible interpretations provided. Usually, the two main interpretative approaches rely either on the concept of monopoly, or on the concept of public necessity, intended as "essential" or "universal" service. The first is usually identified as an "efficiency" explanation, while the second one is usually regarded as a "sociological explanation".[132] More generally, while the "efficiency explanation" of regulation relies on the concept of market failures, the sociological explanation relies on the concepts of necessities, rights, and fairness.[133]

On the other hand, the last (third) condition explicitly deals with the reasonable level of tariffs that can be charged by the property owner (*and* service provider). This third step corresponds to the well-known problem of access pricing, or pricing regulation in general. Even for this aspect, it is possible to distinguish two different approaches. The original "pricing" problem was framed in general and vague terms, and tariffs ought only to be "reasonable", thus lower than a certain threshold. It took several decades before that the pricing problem could shift from a general condition of "reasonable pricing" to a proper economic theory of optimal (efficient) pricing (for one of the earliest contributions, see Trebing, 1971).[134]

The theoretical framework offered by the theory of monopoly represents a powerful tool to merge the two aspects into one single theory: once that public interest is defined in function of the implications that production technologies have for the concentration of a market, then optimal pricing, production technology and market structure can all be interpreted as various aspects of the same conceptual framework. Ramsey pricing (Ramsey,

[132] The two historical references in this respect are Wyman (1904; 1911) and Burdick (1911). In the telecommunication context, see for instance the discussion provided by Nachbar (2008) or Yoo (2013).

[133] For a broad overview of the motivations requiring public regulation in a failure-based approach see Spulber (1989: chapter 1).

[134] The modern theory of access pricing is broad and diverse. Laffont and tirole (1994) is a comprehensive early contribution, while more recent overviews of the topic are Armstrong (2002), Laffont and tirole (2000), and Valletti and Estache (1999).

1927) represented the first technical attempt to provide an optimal solution to the problem of the regulation of tariffs (access pricing) from a monopoly-based perspective, thus based on the down-sloping elasticity of the demand, which represents the essence of the modern theory of monopoly (Chamberlin, 1962; Robinson, 1969).[135]

The problem with the standard treatment is twofold and the first shortcoming is a direct consequence of the other. First, the unit of analysis of a monopoly-approach is not the given property, but it is rather the given final service market in which competition occurs. In a partial-equilibrium approach to competition (such as the one adopted by modern theory of monopoly), we have no information on the nature of the assets and properties employed by the firms.[136] This first shortcoming is itself the consequence of a more fundamental limitation.

The fundamental problem with the standard treatment is that it simply overlooks, or rather it takes for granted, the first and preliminary condition determining "affectation with public interest" in Lord Hale's description: the decision of the property owner to employ his own property in order to serve third parties, rather than (or beside) himself. This first condition represents the distinction between private and public employment of the property.

Stated differently, the standard approach towards regulation, by ignoring the first necessary condition of the aforementioned "three-criteria test", enquires whether a property is affected with public interest while already assuming that the latter is necessarily employed in order to directly serve the public. The first necessary condition is usually disposed of "by assumption".

This first step in Lord-Hale's "three-criteria test" is far from being trivial or obvious; conversely, it is what really characterizes the essence of the common law approach to

---

[135] Since 1920s, the standard formula to establish the optimal access pricing was the Ramsey formula (Ramsey, 1927), which substantially corresponds to the well-known Lerner index for monopoly power. Thus the socially optimal mark-up (deriving from the price of good i (P_i) that maximizes the social welfare) is given by the equivalence $(L_i = (P_i - MC_i)/P_i = \theta/\varepsilon_i$. The left-hand side of the previous relation simply consists in the Lerner Index (L_i) denoting the degree of monopoly power (see Lerner, 1934), while in the left-hand side ε_i defines the elasticity of demand and θ is a Ramsey term which denotes the market conduct of the firm and derives from the shadow price of the constraint maximization problem (for an in-depth discussion on the interpretation of θ in industrial organization in general, and not only in the context of access pricing, the reader is referred to Genesove and Mullin (1998)). For a standard discussion on Ramsey pricing under complete information, see Armstrong (2002: 321-324), Armstrong et al. (1994: 51-60), Brown and Sibley (1986: 39-44), Mitchell and Vogelsang (1991: 43-61), Train (1991: chapter 4).

[136] In fact, this is identical of saying that the monopolistic approach to the theory of the firm is not really a theory of the firm as we now intend it, as a theory of asset ownership and vertical integration.

public interest and distinguishes it from market-failures interpretations typical of the antitrust approach (such as the monopoly approach). What is even more important, to assume that a property necessarily serves the public, rather than the owner, completely misses the meaning and the implication of "vertical integration". In order to better appreciate the role of the first necessary condition, it is necessary to reconsider the three steps leading to public interest and analyse them more carefully.

1.4. From a property-based approach to a service-based approach to the rights of things: Properties employed "in-house" always remain juris privati

Figure 6.1 graphically summarizes the logical steps.

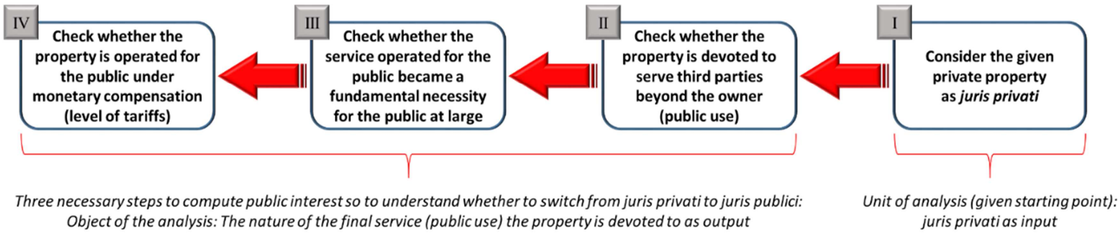

Figure 6. 1 – The definition of public interest in common law: logical, analytical and temporal steps from Juris privati to juris publici

The figure shows that the input of the analysis (the unit of the analysis) is the given property (step 1). Only after the owner decides to devote the property to public, rather than private, use (step 2) and only after the nature of the service provided (step 3), and the conditions under which this service is provided (step 4) are considered, is it possible to draw conclusions with respect to the desirable legal status of the given property under consideration.

Note that, although the figure shows four distinct steps (as described above), the three distinct "tests" corresponding to the three necessary conditions summarised above really correspond to the three arrows linking the four logical blocks. Thus, the first necessary condition corresponds to the first arrow linking step one with step two, the second necessary condition corresponds to the arrow bringing the analysis from step two to step three, and identically for the third arrow, corresponding to the third aforementioned condition.

The logical and analytical steps are summarised in figure 6.2 below.

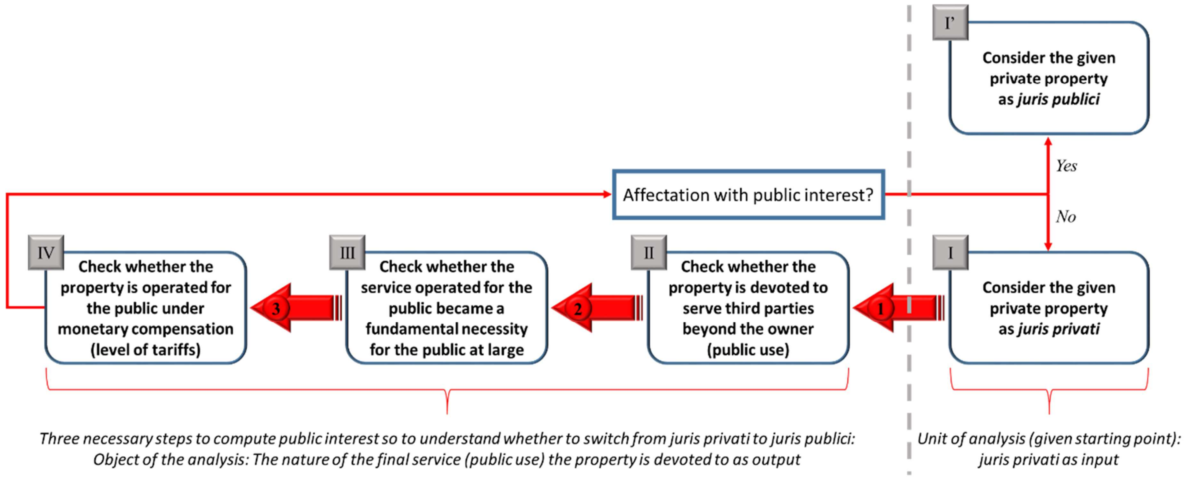

Figure 6. 2 – Public interest test in common law based on the rights of things: juris privati vs. juris publici

Figure 6.2 represents the complete overview of the logic behind the rights of property in common law.

Two points shall be noted from figure 6.2.

First, as already discussed, the figure shows that there is a feedback loop that brings the analysis back to the property, consistent with what has been already stressed, meaning that “affectation with public interest” in common law belongs to the law of property, not to the regulation of services. This means that the given property is, at the same time, the initial unit of analysis (the given entity from which the analysis starts and is conducted) and the final recipient of the decision concerning public interest.

Second, it is possible to note that the logical and analytical development is not homogeneous, as it shows a clear rupture between the first step and the second one. This rupture occurs in correspondence with the first arrow, which also corresponds to the first necessary condition mentioned above. As a result, two main stages can then be identified.

The first stage is the one where the sole property is involved and is considered. In this first stage the property is both the unit of analysis from which the analysis starts, and the object of the analysis or, in other words, the direct object of the predicate to which economic and legal concepts apply. For instance, the concept of public interest applies to the given property, as already remarked.

Starting from step 2 onwards, the object of the analysis to which legal concepts and predicates apply changes: *services* become the object of the analysis so that all steps after the first one (from two to four) apply *to the specific way* in which the property is employed and used or, in other words, to the service operated. For instance, the condition of necessity (step 3) necessarily applies to the specific service or employment considered. In the same vain, what is rewarded (step 4) is the specific service rendered. Differently from the first step, all steps and conditions which follow do not apply to the property itself, but to its particular employment or use.

As a result, it is possible to identify a service-based analysis and a property-based analysis, where the first necessary condition represents the dividing line between the two (figure 6.3). Depending on whether the first necessary condition is satisfied or not, the analysis will either focus on the service, or on the property. Thus, if the property is employed by its owner in order to serve third parties, then it is also possible to know how the property is actually employed *for* the public, or *by* the public, and therefore which type of service is rendered to the public. Conversely, if the property is not used in order to serve third parties, the first necessary condition of Lord Hale's test is not satisfied, and the property will certainly remain *juris privati*, where the latter characterizes, and applies to, the actual thing, and not to any of its peculiar employments (which now concerns only the property owner, being the sole recipient of the services operated by the property).

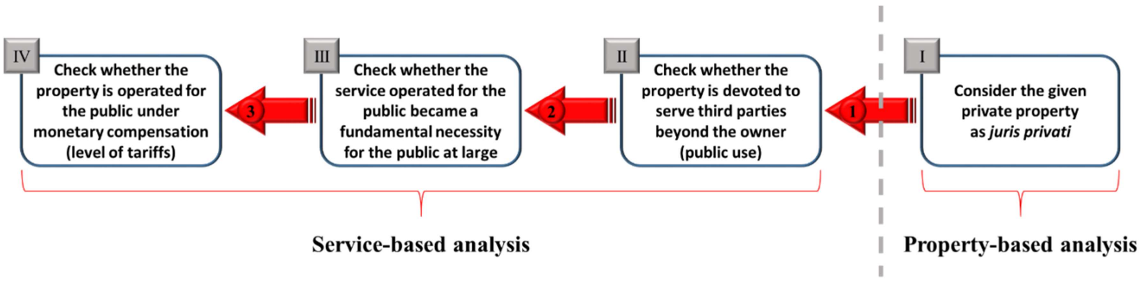

Figure 6. 3 – the two macro-phases of the common law approach towards public interest

The shift from step one to step two, triggered by the verification of the first necessary condition previously highlighted, constitutes the core of the common law analysis of public interest, and its most delicate passage. Most importantly, the switch from step one to step two also represents the key element to understanding how and why the common law approach to public interest departs from what will become the standard approach adopted by the economic and legal literature based on the monopoly explanation. This will in turn also become necessary in order to understand the essence of the duality (or departures) between the regulatory and antitrust approach to public control, asset ownership and institutions in general.

It is therefore necessary to complement the discussion of the common law approach to public interest, with the meaning and the implications of the first conditions as the differences between the monopoly approach adopted for the regulation of corporate charters and public service companies, and the common law approach to property theory derive from there.

1.5. "Private Use" of the property and the problematic interaction between property and services in interpreting the "right to decide how a property ought to be employed": the fundamental role of auto-consumption and self-subsistence in common law

If the switch from the property to the service is analysed, then two fundamental points emerge from Lord Hale's original discussion of the distinction between *juris privati and juris publici* applied to the law of property. The first point is preliminary in order to fully understand the implications of the second one. Moreover, both points have been recognised and retained by the US courts through the last quarter of the 19th century.

1.5.1. The problematic interaction between properties and services resolved in favour of services

First of all, as already remarked, the legal status of a property does not depend on its intrinsic nature: in principle, any property can be affected with public interest and, vice versa, any property considered as *juris publici* can be regarded as *juris privati* in different situations and regardless of the inherent nature of the property. Whether a property shall be regulated as *juris publici* depends on its specific use or employment, as recognised not only by Lord Hale, but also by the Supreme Court in *Munn* and by all the sentences since.

As summarised by Hamilton (1930: 1097; emphasis added):

> "He [chief judge Waite] declares, properly enough, that 'property' becomes 'clothed with a public interest *when used in a manner to make it of public consequence*, and affect the community at large.' He insists that, when '*one devotes his property to a use* in which the public has an interest, he, in effect, grants to the public an interest in that use.''

The previous excerpt clearly states that a given property shall not necessarily be affected with public interest merely based on its nature of structural features. Conversely, public interest (which still applies to the given property) depends on the manner in which this property is used and, even more importantly, *on the specific use the owner decides to devote its property to*.

In sum, "public interest" *applies to the underlying thing* (the given property), yet *it depends on the specific way in which the latter is employed* (the given service).

The relevance of the distinction between the given property and the (potentially arbitrary) employment in which the latter is used was constantly remarked in later sentences. Through the years and decades, the focus slowly shifted away from the property, towards the service in which the latter is employed. Thus, a decade later, in the *Munn* sentence regarding the *Georgia Railroad & Banking Co. v. Smith* (1888), the Court asserted the following (quoted in Fairman, 1953: 667; emphasis added):

> "the obligation … affect the property and employment with a public use; and *where property is thus affected, the business in which it is used is subject to legislative control* … This is not a new doctrine but old doctrine, always asserted whenever property or business is, by reason of special privileges received from the government, … affected with a public use. There have been differences of opinion among the judges of this court in some cases as to the circumstances or conditions under which some kinds of property or business may be properly held to be thus affected, as in Munn v. Illinois . . . ; but none as to the doctrine that when such use exists the business becomes subject to legislative control in all respects necessary to protect the public against danger, injustice, and oppression …"

In other words, although the unit of analysis is the given property under examination, the object of the regulatory intervention is the specific "business in which it is used". And only "when *such use* exists then the business [i.e. the property] becomes subject to legislative control …"

It should be noted that a clear distinction emerges between public interest (applied to the property) and public use (qualifying the specific employment). A property is affected with public interest as long as it is employed in a public use. While the *unit of analysis* the regulatory intervention applies to is the given property under consideration (the asset or resource), the *object* of the regulatory intervention is a specific use of the property or, in other words, the specific business in which this property is employed into. This distinction retraces the distinction between unit of analysis and object of the analysis for the so-called regulatory approach already highlighted above.

It should be noted that this delicate task of carefully distinguishing "the property" from "its employment" is full of perils. Already, in the above excerpt, the expression "*property or business*" appears, meaning that the two start to conflate and overlap.

The conceptual difficulties raised by the problematic interaction between property and services are solved once and for all by Justice McKenna in *German Alliance Insurance Co. v Kansas*, in 1914. Well aware of the fact that property and service are not two interchangeable synonyms, the Justice completely favoured the latter to the detriment of the former, by clearly stating that public interest can only be ascertained by looking at a specific employment of the property (the service), while completely discarding the role played by the underlying property (the thing). The Justice clearly asserted the following (cited in Hamilton, 1930: 1099):

> "In other words that which had been private property had from its use become, it was declared, of public concern, and the compensation to be charged for its use prescribed … it is the business which is the fundamental thing; property is but its instrument, the means of rendering the service which has become of public interest."

By that time the conceptual switch from the property to the service had been resolved in favour to the latter, by relegating the former to the background, in a completely passive role.

Yet the important message transpiring from all the above assertions is that all involved actors clearly recognised the fact that the relationship between the property and its services is not an easy one as any property can, at least in principle, be employed in multiple ways, and can be devoted to multiple uses. Moreover, what all parties recognised, even though somewhat implicitly, is that the one endowed with the rights to decide *how a certain property ought to be employed*, given the multiplicity of possible employments, is the property owner.

Yet, there are two different ways in which the right "to decide how a certain property ought to be employed" can be interpreted: across alternative *public employments*, or between *public employments* and *private ones*.

This last consideration leads to the second, and much more fundamental point highlighted by Lord Hale in his treatise on property law.

1.5.2. Exclusion and discrimination in public and in private uses: among the public vs. from the public

Above it was anticipated that the first necessary condition stated by Lord Hale in order to justify the potential "affectation with public interest" corresponds to the verification that the property under consideration is not employed by the owner "for himself" and for "his own needs", but that it is actually used as a means for commerce, thus in order to serve third parties.

In this respect, this is how Lord Hale originally introduces the concept of public interest in his "*De Portibus Maris*", as cited in Fairman (1953: 655; emphasis added):

> "'no man may set up a common ferry for all passengers, without a prescription time out of mind, or a charter from the king.' *He may make a ferry for his own use; but if he holds himself out to serve the public generally*, so that 'every man for his passage pays a toll, which is a common charge,' then the ferry 'is become a thing of public interest and use" and "ought to be under a public regulation.'"

The same concept is restated in McAllister (1930: 762; emphasis added) with respect to the same excerpt from Lord Hale:

> "Insofar as the small streams were used by the public, however, the king enjoyed certain prerogatives in them. For instance, *if a private person operated a ferry, not for the use of himself and his family, but for the use of all the king's subjects* that ferry became 'a thing of public interest and use.' Every such ferry must serve all who would use it, it must be kept in good order and could take only a reasonable toll."

The two "*ifs*" in the quoted passages above represent the key terms of both texts: only as far as the legitimate owner does not employ the given property for his own private purposes, the property can potentially become liable of public interest. The "if" clearly allows for the possibility that the owner may not use the property as a means for exchange.

This is the most profound and most fundamental interpretation of the nature of *juris privati* in common law applied to the law of property: a private property under *juris privati* confers first and foremost the right to decide whether a certain property shall be *held out* for the benefit of third parties (thus as an instrument for trade and exchange), or whether it shall be used as a mere means for self-subsistence or, in other words, auto-consumption.

This is the essence of the specific passage of the *Munn* sentence when it is asserted that *it is only when the owner "devotes his property to a use in which the public has an interest*" that "he, in effect, grants to the public an interest in that use." Stated differently, a property should not necessarily be intended by the part of its owner as an instrument of trade and commerce for the benefit of third parties (i.e. the public).

Thus, not only the owner of the property has, at least in principle, the right to decide among alternative employments. The most basic right deriving from the private ownership of a certain thing, which also characterises the essence of *juris privati* and distinguishes it from *juris publici* in common law, is yet another, and more fundamental one: the right conferred to the owner to exclude others (the public) from the use or enjoyment of the property, and therefore, ultimately, the right of the owner to use or enjoy the property *for himself* or, *for his own personal consumption* or *self-subsistence*. In common law, j*uris privati* fundamentally differs from *juris publici* as it confers to the owner the preliminary right to use the property for his own private needs in auto-consumption.

It should be noted that this "right to exclude" is very different from the right to exclude considered in standard antitrust law (Singer, 1996). In fact, it is preliminary to it.

On the one hand, the common law basic right to "use a property for your own self" has nothing to do with the standard antitrust debate on vertical foreclosure and leverage of monopoly power.[137] In a nutshell, contemporary antitrust law interprets discrimination and exclusion through vertical integration as a way to artificially restrict the production volume of a given service, so as to increase (leverage) monopoly power across stages. This is obviously a very different aim from the one pursued by the owner of a *juris privati*, which does not deal with the monopolistic volume of output, but is simply interested in the autonomous satisfaction of certain needs and wants through private properties.

On the other hand, exclusion in the interpretation originally provided by common law is not defined among potential competitors. In this latter case, the right to exclude would imply the right to *discriminate* among alternative users of the property (Werbach, 2007; Nachbar, 2008). Conversely, it is with respect to final users purchasing the services furnished by the property itself. Thus, the most basic right coming with the ownership of a certain property is not to exclude other competitors from a given trade, but it is not to trade at all in the first place. The difference between the two cases is the difference existing between using the property as a means for trade and commerce, and using the property as a means for auto-consumption.

[137] The contemporary approach to vertical relations and vertical restraints has replaced the old Chicago approach and has demonstrated that vertical leverage of monopoly power does indeed increase profits. For an overview on the topic of vertical foreclosure see Motta (2004: chapter 6), Rey and Tirole (2007), Rey and Vergé (2008), Riordan (2008). With respect to exclusive deals and exclusive contracts see Bernheim and Whinston (1998), Lafontaine and Slade (2008), Segal and Whinston (2000), Whinston (2006: chapter 4) are two main economic theories of vertical relations and vertical restraints.

It is easy to see how conditions two and three of the test to determine "affectation with public interest" only make sense if the property is preliminarily used in order to offer services to third parties under remuneration. It is true that Lord Hale's original interpretation of the public interest concept is used as a way to regulate the tariffs set to access the property. Yet, with no "holding out" of the property to anyone else, and with no monetary transaction occurring among the owner and the members of the public, there is no tariff to regulate in the first place. In sum, as auto-consumption and self-subsistence do not generate monetary transactions in the first place, they do not even generate tariffs to be regulated.

1.6. "Auto-consumption" vs. "public consumption" of the property as the key distinction between juris privati and juris publici in property law

The fundamental point that emerges from the original treatment provided by Lord Hale is the following: whether the given property shall be treated as a means for trade and commerce that raises tariffs from third parties, and not as a means for auto-consumption or self-subsistence for the exclusive enjoyment of the owner, is not the matter at stake, and cannot be treated as the object of the analysis of property law, according to common law. This decision shall be left to the owner himself. This means that the way in which the property is employed by its owner under *juris privati* shall be treated as the given data of the problem in common law, and not as the variable under examination.

As a result, given a property treated as *juris privati*, public control can be allowed only once it is ascertained that the property is used by its owner in order to serve third parties and not as a means for self-subsistence. This is just another way to say that the presence of some monetary compensation (the "public tariffs") is a necessary preliminary condition in order to proceed with the "affectation with public interest" so as to change the legal status of the property from *juris privati* to *juris publici*, as already noted above.

This was obviously the case for the *Munn* sentence. Grain elevators of the *Munn & Scott* were employed by third parties subject to the payment of a certain tariff.

Yet, it is this point that is a potential source of confusion, and represents the first source of misunderstanding in the history of regulation and property theory.

The *Munn* case officially introduced into the constitution of the United Stated the principle that the legislature can regulate the tariffs set by a private owner (Munn and Scott) for the access or use of its own property (grain elevators). *Munn* was a sentence on the regulation of tariffs or, in modern words, it was a sentence on price regulation. This was consistent with the fact that regulation until the second decade of the 20th century, corresponded to the sole regulation of pricing policies as no regulation of entry (regulation of structural monopoly as opposed to the regulation of legal monopoly) was conceived.[138]

Although it is obvious that the preliminary presence of some monetary transaction between the property owner and a third party (the public) represents the necessary condition for regulating access tariffs, not necessarily there must be some monetary transactions between the property owner and the general public. In other words, a market price and a monetary transaction are needed in order to regulate pricing policies and market transactions. Yet, at least in principle, this does not need to be. An inter-personal transaction between the property owner and a member of the public for the transaction and exchange of the services provided by the property might simply be missing.

This appears to be the essence of Lord Hale's statement that any owner "may make a ferry for his own use; *but if* he holds himself out to serve the public generally, so that 'every man for his passage pays a toll, which is a common charge, *then* the ferry is become a thing of public interest and use." The conjunctions "*if*" and "*then*" are highlighted precisely in order to show that only in the specific case in which the owner "holds himself out for the public", then the property might be potentially subject to public control.

Thus, this is how Justice Bradley summarised his thoughts in his private correspondence with Chief Justice Waite at the eve of the *Munn* sentence (cited in Fairman, 1953: 670-71; emphasis in the original):

> "The principle here involved is plainly this: *that whatever has the effect of a common charge* (that is, a constantly recurring charge on the public) *and is become of public interest and use, ought to be under public regulation, both as to the manner of its use, and the charges to be exacted*. The case of a ferry is merely one instance of the application of this principle."

These are just a reiteration of the three necessary conditions previously identified. The statement simply makes clear that in order for a certain property to be affected with *juris publici*, it is not only necessary that this property "became of public interest and use", but it is also necessary for this property to "have the effect of a common charge", meaning that

[138] See the first editorials commenting on the new trend in entry regulation of legal monopolies in Harvard Law Review (1920), Columbia Law Review (1924), and L.Y. (1926) in the Michigan Law Review.

the property must be held out for the use of third parties in the first instance, and its use and employment must be subject to the charge of a tariff (the market price) as a result.

In conclusion, “affectation with public interest” in common law is originally defined by first looking at whether the property under consideration is used for the private use and benefit of the owner or, alternatively, as a means of commerce. Only in this second case the property is operated for the benefits of third parties under monetary compensation, and only in this second case is a private property "affected with a public interest,” thus becoming *juris publici*, and “ceasing to be *juris privati* only."

This preliminary necessary condition was already clearly stated by Lord Hale in his treatise (on which both Justice Bradley and Chief Justice Waite rely on) (in Fairman, 1953: 670; emphasis added):

> “no man may set up a common ferry for all passengers, without a prescription time out of mind, or a charter from the king” he then goes on acknowledging that the same man “*may make a ferry for his own use or the use of his family, but not for the common use of all the king's subjects passing that way* …”

The US Supreme Court in *Munn* rediscovered Lord Hale’s treatise on the “rights of things” and applied the common law principle to the grain elevators used by grain shippers. In this light, the Supreme Court certainly had a good point in order to decree public regulation of the rates charged by private actors for the use of privately owned assets. Yet (and this is the key point) the fact that in the case of *Munn*, elevators were obviously used by their owners in order to serve the public and therefore make a monetary profit out of their use did not automatically mean that any property subject to *juris privati* necessarily has to be used in that way (in order to make monetary profits by serving third parties).

1.7. Public interest as “rights of things”: summing up the two key conditions explaining the switch from juris privati to juris publici in common law

In conclusion, the common law approach to the rights of things is based on the concept of “public interest”, while the latter in turn relies on the fundamental distinction between *juris privati* and *juris publici*. The switch from the former to the latter is not straightforward, and can only be justified if three necessary conditions contextually hold, where neither of these is alone sufficient to justify the change in the legal status of the property. Yet, two key points characterising the relationship between *juris privati* and *juris publici* clearly emerge from the previous discussion. The two are intimately linked, and one strictly follows the other.

First, there is a clear univocal causality in the common law approach from the given *juris privati* to, potentially, *juris publici*. It is not the other way around: there is no endowment of private properties following the decision to provide some given public service or necessity. It is important to underline this fact at this stage, as public franchises for the development of public service companies work precisely the other way around (from the public to the private). Yet, in common law the property is the given, and the legal status applies to it, it is not directly endowed to the property by other considerations. This means that first there is a property, and only afterwards someone decides to use it: t*he property exists independently from its functionality and potential employments*.

Two, the property does not necessarily have to be employed as a means to serve the public. It can also be employed privately, as an (intermediate) means for self-consumption of the property's services. Note that this point derives from the first one above as a property purposefully set up in order to provide a given service to the public will necessarily be employed as a means to serve the public. Thus, the fact that a property exists independently from the provision of a specific (given) service, also implies that the property owner might very well decide to employ it to serve his own needs, thus for his own auto-consumption and self-subsistence: *the property must not be necessarily employed "in-exchange" it can also be employed "in-use"*.

The second point represents the first necessary condition of Lord Hale's "test", while the first condition represents the logical, temporal and analytical causality of the framework.

Figure 6.4 summarizes the two key conditions implied by the common law approach to public interest as applied to property law.

As shown in figure 6.4, the two points represent two fundamental dimensions of economic analysis: point one defines the dimension of time, while point two characterises the dimension of value. While time flows one-way, the possible coexistence of self-consumption and third-party consumption retraces the classical duality in valuations, as first described by Adam Smith.

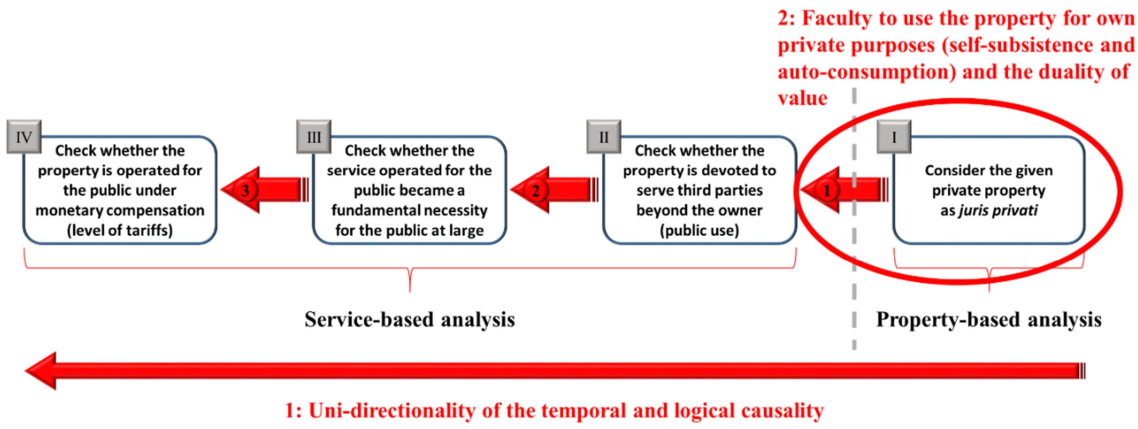

Figure 6. 4 – Two conditions of the common law tradition in public interest as applied to the law of property

These two considerations however need some more lengthy discussion that will be left for the next chapter.

The important point to highlight at this stage is that both conditions characterise an aspect (either in temporal or in value terms) of the key switch from the private dimension of juris privati to the public dimension of juris publici. It is the (unidirectional) transaction from step one to step two (as highlighted in the figure) that represents the key aspect worthy of further enquiry.

Lord Hale's discussion of public interest characterises only superficially the nature and the essence of this shift. It should be remembered that the aim of Lord Hale's treatment, on which the *Munn* sentence relied, was to characterise and justify "public interest". However, the concept of public interest is not limited to the shift from the private employment of the property to the public employment of the property. This shift only represents the first necessary but not sufficient condition to determine "affectation with public interest" and therefore *juris publici*.

Lord Hale's treatment is largely descriptive of the logic that has to be followed, but it leaves ample room for debate concerning the effective meanings of the various stages, and what the various conditions imply.

For this reason, in order to better enquire as to the meaning of the dichotomy between the private and the public employment of a property, it is necessary to look elsewhere.

The institution of common law that explicitly attempts to address the problems deriving from the unclear distinction between private and public is the one of public calling. The institution of public calling is much older than the one of public interest and, in a sense, it is possible to say that it represents its conceptual and legal foundations. I do not know whether it was Lord Hale's intention to base the legal dichotomy between *juris publici* and *juris privati* on the medieval institution of public calling. Certainly, the latter must have, at least indirectly, informed the former.

We now turn to the distinction between private and public calling, keeping in mind that even though this is a completely different issue and subject from the distinction between *juris privati* and *juris publici*, it can still shed some fundamental light on the way in which both aforementioned dimensions of the dichotomy between private and public consumption of a property (the temporal and the value dimension) ought to be interpreted, and on their very relevant legal and economic implications.

2. Public callings and "common businesses" in common law: the role of value and time in shaping the institution of common carriage

The previous section ended by highlighting two key conditions of Hale's "affectation with public interest": the univocal logical causality from the property to its employment (service), and the key distinction between the private domain and the public domain defined in function of the shift form a property-based analysis and a service-based analysis.

In order to understand the nature and the essence of the legal change from private to public status, it is necessary to resort to the concept of public calling. The institution of public calling represents the first legal definition of the distinction between public and private.

This section will provide an overview of the concept, and underline the key features of a public calling, while at the same time clarifying the essence of the distinction between public and private: whether and why some third parties (member of the public) is involved in, or affected by, the employment of a certain property.

2.1. From public interest to public calling: introductory remarks

The concept of public interest played a key role in constitutionally justifying rate regulation of private properties in the *Munn* sentence.

Beside Lord Hale's 17th century discussion of the meaning and the nature of public interest, there is also a second aspect of common law that has played a fundamental role in the development of contemporary legal and economic thinking on regulation: this second aspect includes the concepts of public calling, and common business, from which the institution of common carriage ultimately derives. A common carriage is simply a public calling, or "common undertaking", whose business it is "to carry".

There are probably no other concepts more relevant than those of public calling and common carriage in understanding the profound problems faced by courts and regulators in recent times, especially with respect to the fundamental issues of discrimination and refusal to deal.

As will be later illustrated, a public calling is just the underlying legal mechanism conferring the status of "common" to business, activities or, much more appropriately, to properties and assets. For this reason, it is justifiable to treat the two (common business and public calling) as synonyms although the underlying reasons why the two can be legitimately considered so, will require some further qualifications as will be discussed later on.

Although the concepts of public interest and public calling are very often confused and conflated, it is important to maintain their distinction as the concept of public interest in common law (as developed by Lord Hale in modern times) has nothing to do with the medieval institution of public calling. The concept of public calling was developed many centuries before Lord Hale's definition of public interest, and was aimed at addressing completely different issues and concerns. Yet, their common law roots become evident once the two concepts are examined in detail. These "common law" roots is what shall interest us here.

Although contemporary regulatory law is based on the common law concept of public interest, the concept of public calling is much older, and probably also occupies a more relevant place in the history of common law. In sum, the "common" status corresponds to the first step of Lord Hale's three-criteria step for public interest, and therefore corresponds to the key preliminary step dividing what is public from what is private. For this reason, the institution of public calling represents the essence and the core of the concept of public interest, although it is not a "sufficient condition" to determine "public interest". As already discussed above, public interest is a much more specific and restrictive concept than the preliminary legal distinction between private and public.

There is a substantial amount of literature on the topic.[139] Here we will briefly highlight the most important aspects emerging from the literature with a twofold aim in mind: to

[139] For in-depth discussions and analysis of the meaning and the origin of public callings and common businesses, the reader is referred to the following (certainly not exhaustive) list of references, on which the present discussion relies: Adler (1914), Ames (1888), Arterburn (1927),

highlight the common nature linking the common law concept of public interest with the one of public calling, and to also show the limitations of the contemporary approach to the concept of "common carriage", which (wrongly) interprets it as a synonym of public franchise for corporate charters, and as a legal remedy for market failures, in the same way in which public franchises are.

Thus, the current approach towards public callings and common carriage not only (incorrectly) conflates the concept of public interest with the one of public calling, but it also conflates the concept of public calling (itself incorrectly interpreted as a remedy to guarantee public interest) with the one of public franchise for public utility services, basing both on a market failure explanation (typically the monopoly explanation).

These misunderstandings derive from the incorrect interpretation of the term "*public*" in the two cases, and from a confused application of the fundamental distinction between public and private to the two concepts of assets (property) and services. As a result, it is the term *public* that represents the common denominator of all these expressions, and it is the confusing application of the term either to the service, or business, or to the underlying property that generates a great deal of this confusion.

As we show below, the term *public* in the expression "public calling" has to be interpreted in the same way in which public is interpreted in the expression "public interest", yet differently from the way in which public should be interpreted in the expression "public utility service". While in the first two expressions, the term public should be interpreted along standard common law lines, thus following a regulatory approach, in the last, the distinction between private and public should be reinterpreted in neoclassical terms, thus following a failure-based and antitrust-based approach.

It is the concept of value that, once again, becomes key in order to appreciate the distinction between the common law (regulatory) and the market-failure (antitrust) interpretations of the term "public". Only if value and time (intended as temporal and logical causality) are reconceptualised and reinterpreted, is it possible to understand the nature and essence of a public calling, and its shared roots with the common law concept of "affectation of public interest".

Basedow (1983), Beale (1897), Bogen (1996), Burdick (1911), Cherry (1999: chapters 2 & 4; 2015), Holmes (1879), Jeremy (1816), Payton (1981), Roberts (1920), Rossi (1998: 1242-1263), Simpson (1975), Wyman (1904; 1911: chapter 1).

2.2. Clarifying the original meaning of a public calling: a historical reconsideration of the dichotomy between public and private in tort law

2.2.1. The common law origins of a public calling: liability in performing a service and tort law

The standard way to define and characterise a public calling is to highlight the duties and obligations it is subject to. In sum, a business can be regarded a "public calling" if it is subject to two different duties or obligations: the duty to serve and the duty to perform "with due care".

On the one hand, the "duty to serve" implies that a public calling cannot refuse to serve any member of the public asking for the service. On the other hand, the duty to perform "with due care" implies that a public calling is liable for performing the requested service poorly and below the required, or expected, standards.

More generally, the institution of public calling derives its origins from the law of tort, and was originally defined in order to characterise the liabilities of the actor (businessman) providing the service: a public calling is liable for either not serving a member of the public if the latter receives damage from the fact that he has not been served (duty to serve), or for performing the requested service poorly and below the required standards, if a member of the public receives damage from the underperformance of the service (duty to perform with due care).

In truth, the obligation to perform "with due care" has historically been the sole original liability attached to a public calling (Bogen, 1996). What originally was characterised as "*common business*" in mediaeval times signalled the fact that they were liable for the service provided to the public. They also differed from *private* business in that the latter did not have any responsibility with respect to the service rendered. Thus, in a sentence delivered in 1440 the judges did not find the defendant guilty of having killed a horse with his medicament with the following remarks (quoted in Arterburn, 1927: 412):

> "You have not shown that he is a common surgeon to cure such horses, and so, although he killed your horse by his medicines, you shall have no action against him …"

In sum, only a "common" businessman was liable in front of the law in case he did not perform the service "with due care". This is, precisely, the first duty characterising any public calling: the duty to serve with due care.

As illustrated by Bogen (1996), the other obligation, the duty to serve, historically derives from the first and was defined in order to make the duty to serve with due care effective and to provide, so to say, a legal framework with which to enforce the liability associated to the "duty to serve with due care", in case the service performed was lacking and not satisfactory.

Common carriers are "common businesses"; as a result, they are subject to public callings' duties. Consistently, the institution of common carriage has always been characterised by the two aforementioned liabilities. The only difference is that in modern times the duty to serve has taken over and represents the fundamental obligation of common carriers, leaving the duty to serve with due care in the background.[140] Since the deregulatory process, this duty has become fundamental as both concepts of public necessity, linked to the topic of universal service, and discrimination, linked to the topic of open access and net neutrality, depend on it.

Beside the relative priority between the two basic duties of a public calling, there is also another, much more fundamental, aspect of contemporary common carriage that has changed since mediaeval times: the interpretations of the reasons why the institution of public calling was initially developed, its meaning, the rationale behind it, and the justification for the imposition of the two "common" duties on certain actors rather than others.

It is on this front that the original (mediaeval) common law interpretation, and the modern (neoclassical or failure-based) interpretation of public calling and common carriage fundamentally diverge. It is also on this aspect that we shall focus the remainder of our discussion. A proper clarification of the legal rationale justifying the two duties of a public calling is particularly necessary as two of the most urgent problems in contemporary antitrust and regulatory economics, discrimination and refusal to deal (i.e. open access), directly derive from them.

2.2.2. The original interpretation of public callings in common law: "common" defined as "business"

During the last two centuries, different interpretations have been advanced in order to understand where these duties come from, and what their rationale and the legitimacy were. Wyman's (1907) was the first to explicitly interpret public callings as legal responses

[140] For recent discussions of the concept of common carrier in the contemporary telecommunication sector see, among many others, Candeub (2004: 377-395), Cherry (1999), Deacon (2015), Noam (1994), Speta (2001), Yoo (2013).

to "virtually monopolistic" business and to officially inaugurate a market-failure approach to public callings and common carriages. This approach explicitly put public callings in strict connection with the regulation of public service corporations (which have always been regulated through public franchising).

The opening page of the preface to his famous treatise on the law of public callings (whose title "Public Service Corporations" is an explicit recognition of the substantial identity of public callings with corporate franchises), starts with a perfect restatement of the Coase theorem, where the role of *failures* in shaping the logical and analytical causality between economics and the law is very lucidly understood and very clearly described (Wyman, 1911: v; emphasis added):

> "The causes of the division [between private and public callings] are economic rather than strictly legal. Free competition, the very basis of the modern social organization, superseded almost completely medieval restrictions, but it has just come to be recognised that *the process of free competition fails* in some cases to secure the public good, and it has been at last admitted that *some control is necessary over such lines of industry as are affected with a public interest*. At this point the problem of public callings becomes a legal one."

The most striking feature of this passage is represented by the fact that "the law" is subordinated to the market mechanism, and it is the presence of some (competitive) failures that triggers the presence of legal concepts, such as that of public calling. This is striking if it is considered that this work on the economic foundations of the regulation of public utilities not only precedes any work in contemporary law and economics by nearly half century, but it also anticipates Pigou's *Wealth and Welfare* (published in 1912), namely the first ever proper treatise in welfare economics.

Nowadays, the "duty to serve" of common carriers is explained and interpreted along Wyman's lines, by looking at the competitive failures generated by the monopoly power held by infrastructure owners. In truth, even though monopoly power is not treated as the *exclusive* reason to justify the duties of a public calling, it is still true that nowadays "[e]ven sceptical commentators recognize that [monopoly power] has become the dominant, if not the sole, criterion for determining the scope of common carriage" (Yoo, 2013: 560).

This was also the essence of the deregulatory wave of the early 1980s. In 1980, for the first time the FCC associated the definition of common carriage to the presence of a market failure in monopolistic terms and distinguished between “dominant” and “non-dominant” carriers. The discerning element between the two was the presence of monopoly power (see Kearney and Merrill, 1998; Nichols, 1987). Only carriers with market power rested under public regulation, while the non-dominant ones were placed outside Title II of the 1934 Telecom Act.

Yet, the current interpretative approach to public calling is completely inconsistent with the standard and original interpretation of public callings that has been provided for many decades since Wyman. Since the publication of Wyman’s arguments, it has been recognised that the monopoly explanation proposed at the beginning of the 19th century (almost certainly under the influence of Marshallian Welfare economics) is completely unfounded and has to be rejected. As definitively concluded by Adler (1914: 148-49) more than a century ago:

> “No distinction based upon monopoly between a private and a common carrier prior to the year 1600 has been set forth, and no explanation has been offered as to why an inn- keeper should choose the monopoly (if any) of a common inn in preference to a private inn, free from the duties of a common innkeeper, and certainly no evidence is at hand of the existence of luxurious hotels where people of wealth made their homes in 1600, – the basis of the distinction at present made between common and private inns.”
>
> “When we consider the principle of monopoly as producing in the early days the supposed distinction between classes of callings, its failure is clearly apparent, for no evidence of any kind is offered that carriers were less numerous than butchers, or that innkeepers were fewer than carpenters, or barbers than weavers … Monopoly, therefore, cannot be accepted as an explanation of the distinction between public and private callings, either at present or in the distant past …”

Thus, monopoly cannot be a legitimate interpretation through which to understand the nature of public callings. This position was commonly accepted for several decades until the early 1980s, when the monopoly approach was rehabilitated in order to support and legitimize the deregulatory wave, although in those same years many other (unheard) authors were still trying to recall (although in vain) the profound difference between the motivations that originally justified the common law concept of public calling, and Wyman’s monopoly reinterpretation (Basedow, 1983; Cherry, 1999, & 2008; Dempsey, 1984; Nichols, 1987; Noam, 1997; O'Riordan, 1979; Rossi, 1998). As put by Payton (1981: 130-31) at the eve of the deregulatory wave:

> "Some modern commentators have attributed the duty to serve to the fact that the innkeeper, the smith, and the common carrier [as opposed to the other types of businesses] have a virtual monopoly vis-à-vis their individual customers. Although the concept of virtual monopoly may appeal to the modern mind because it makes imposition of the duty to serve economically rational, we should recognise that contemporaries would have been baffled by such explanation."

Not only is the monopoly approach not justifiable, but neither other types of explanations such as "affectation with public interest" can be used as alternative explanations in order to justify the "common status" of a business, as already mentioned. This is the case not only for mere temporal reasons, for a 17th century concept cannot be used in order to explain and justify a 14th century one, but also for logical and analytical reasons: a public calling is just a preliminary necessary yet not a sufficient condition in order to ascertain "public interest". In other words, although all properties affected with public interest necessarily have to be public callings, the reverse does not hold as *not all public callings are necessarily affected with public interest.* As a result, the concept of public interest cannot help in understanding the legal origins and rationale of a public calling.

Nor can we use a sort of "public franchise" argument, thus explaining the emergence of public callings with the pre-existing necessity to provide public services, as originally suggested by Burdick (1911). This conflation of the law of public callings with the regulation of public service companies not only constitutes another misleading and completely unfunded recent reinterpretation of the nature of public callings, but besides being temporally inconsistent (once again, a late-18th century concept would have been used to explain a 14th century one), it is also unhelpful as it is based on circular reasoning. As remarked by Bogen (1996: 52):

> "According to another theory, offering services to the public was an undertaking to serve all members of the public. But that is not so much a theory as a tautology. Innkeepers made no express promise to serve all comers, and an invitation to the public to request services does not necessarily imply a promise to provide them."

The question then remains: what motivated the emergence of a public calling in mediaeval common law?

The answer is so simple and so basic that it might just appear common non-sense at first. As convincingly recalled by Adler (1914), Arterburn (1927), Bogen (1996), Payton (1981), and many others, the institution of public calling was originally developed in order to unequivocally identify the nature and the type of a specific activity or, in other words, it was employed in order to simply characterise and identify the specific business under consideration. In fact, as recalled by Adler (1914: 151-52) the term "common" simply meant "business":

> "… the range of common employments must be greatly extended beyond that generally accepted, – and the subsequent shrinkage explained, – or that "common" does not have the significance usually ascribed to it, as characterizing and distinguishing one occupation from another. Not monopoly, or bailment, or necessity will be found differentiating all these employments from the few not somewhere called common. The list is so long and contains such different callings that we are led to the conclusion that the term "common" did not serve to distinguish one employment from another and that all occupations could be common."
>
> "What, then, did 'common' mean? Simply 'business' – business carrier, business tailor, business barber."

Adler's interpretation has been the one commonly accepted for several decades. The "common" status did not characterize any specific feature of one business vis-à-vis another one. "Common" was simply applied to any activity that could be defined clearly enough so to be clearly identified with a certain business, or profession. In other words, a "common business" was one that could be clearly identified and classified as a "publicly recognised" activity, and whose businessman's identity and profession could be clearly identified by the general public.

This point may appear as a meaningless and useless tautology as it leaves major questions still unsolved. In fact, what is the need and the rationale for defining a brand new institution, such as the "common status", in order to simply render any kind of business liable for its disservice? What is the need to distinguish between public and private businesses if they are both "just businesses"? And why does the liability for poor service (the original and primary duty of a public calling) have anything to do with the associated "duty to serve", which will be later interpreted as the fundamental pillar in justifying the public control of utilities and common carriers?

It should be noted that the same scepticism concerning the usefulness and meaningfulness of the concept of public interest that characterised later (neoclassical) economics can be applied to the concept of public calling. Both seem to possess tautological circular reasoning that potentially explains everything (any business, and any property) and therefore explain nothing.[141]

Some new perspectives on the matter are therefore required. In what follows we shall argue that, in the same way in which the common law concept of public interest acquires a very precise meaning when interpreted through non-neoclassical lenses, so does the common law concept of public calling, and for the exact same reasons: the dual nature of value.

2.3. Public and private callings in common law: preliminary characterizations and distinction

The difference between a common business and a private business under the law of public callings rests on whether a specific service represents the explicit and recurrent activity a certain business is purposefully *ex-ante* dedicated to (or a certain businessman is specialised into), as opposed to whether a certain service is just a sporadic, contingent, impromptu undertaking, also in the form of a "favour", provided only after a certain necessity arises. In this second case, the service is provided *ex-post* but with no ex-ante specific intention or willingness to purposefully undertake that specific trade, nor to perform that specific service for the public.

Thus, citing Bogen (1996: 73) once again:

> "The parties often went to the jury on whether the defendant was a 'common' innkeeper, i.e., whether the defendant was in the business of providing lodging to the general public. The best justification for making guests responsible for their own goods was that the host was not in the business of innkeeping."

So that (Bogen, 1996: 74):

> "The courts ultimately accepted the principle that hosts' liability depended on whether they were in the regular business of providing lodging."

[141] Thus for instance, Nachbar (2008: 84) summarises the problem as follows: "While there was considerable dispute throughout the era over the criteria for determining whether a particular business was public, it was common ground that publicness provided its own justification for imposing special duties on businesses deemed to be public … If "public" were self-defining, that would be the end of the matter, but it's not self-defining. The heart of the inquiry, then, is to identify exactly what it is that makes a business 'public.'" See also Yoo (2013: 553-559).

As a result, a certain activity was common for the simple reason of being a declared business for the provision of a certain service (innkeeping in this case) to the public.

Thus, if we go back to the case of the horse who died after careless treatment had been provided by the part of a "*non-common* surgeon", the motivations supporting the aforementioned judge's remarks in finding the defendant not guilty, were the following (quoted in Arterburn, 1927: 412; emphasis in the original):

> "Perhaps he applied his medicine *de son bon gré*, and afterwards your horse died; now, since he did it *de son bon gré* you shall not have an action …"

Or, in other words, as the defendant that spontaneously decided to help cure the horse did it voluntary and spontaneously ("*de son bon gré"* thus in the form of an undue favour), with no clear competence on the matter, he could not be found guilty of the action. For this reason, the defendant was not liable of the death of the horse. This is how liability and the first duty of public callings (to serve with due care) has to be linked.

The same concept is expressed in Jeremy (1816: 7; emphasis in the original):

> "A mandatary not being by profession skilful in the business undertaken shall not be liable to an action if he performs his commission *bona fide*, and to the best of his ability."

The two expressions "*de son bon gré"* and "*bona fide*" should be interpreted in the same way: a private citizen, being a non-*common* businessman, can trade and provide services to third parties a*t his own discretion*, and with no clear expectations on the type of service provided, nor on the outcome. For this reason, the terms of trade are not legally binding, making the private actor not liable for the final result. This characterised the original common law distinction between private and common.

In Arterburn (1927: 418) words:

> "The distinction made was between those engaged in a trade or 'common' occupation, and those unskilled in any trade. If one chose to allow one of the latter class to come into contact with his person or goods, he assumed the risk of injury, without any remedy for damages."

As a result, far from being an institution denoting some form of market power, some peculiarity of the competitive or productive environment, or some necessity, the distinction between private and public calling (or private and public business) really has to be intended as a “distinction between persons who performed an act sporadically and those who performed it as a regular basis” (Bogen, 1996: 74).

In sum, as we shall see in the next paragraphs, the distinction between private and common really distinguished the various *identities* of the actors engaged in transactions with the public. In Arterburn’s (1927: 418-19; emphasis added) words:

> “The writer thinks there is no basis for any classification of business into public or private. There is neither logic nor practicability in such an arrangement, nor is there any reason historically for such a division. The word ‘common’ meant simply ‘business.’ That is, one engaged in work as a trade in which he made his living. It seems then, the words "common carrier" meant one in the business of carrying as a trade, *as distinguished from one who did it as a casual act or by special agreement*.”

3. Public vs. private calling as a dichotomy between status in tort law and terms of trade in contract law

This last section clarifies a widespread misunderstanding on the nature of public calling that has developed in the last decades: the “contractarian approach” to a public calling. A contractarian approach to public calling adopts a failure-based approach and interprets the distinction between a private and a private calling in function of the type of contract (type of the contractual relation) established.

This cannot be correct. A public calling should be interpreted in function of the social status conferred ex-ante, and not in function of the type of contract established ex-post in function of a specific type of transaction. While the status of a property or actor characterises the ex-ante purpose of the activity undertaken, contracts only regulate the terms of a contingent transactions (thus ex-post).

3.1. Misconceptions in the way in which the dichotomy between public and private calling has been reinterpreted in more recent times: from status to contract in the “contractarian” approach

Although it is important to clarify the fact that in common law the original status of “common businesses” has nothing to do with the characteristics of the market in which services are provided (in the same way in which “affectation with public interest” has nothing to do with market structure), the previous clarifications are still not enough as they are still open to another, more subtle, type of misunderstanding.

The aforementioned distinction between public and private can still be interpreted along two different lines.

On the one hand, the distinction between private calling and public calling is usually interpreted as a distinction between contract law and tort law respectively. The law of public callings is usually seen as the original incubator of the English Law of contract (Ames, 1888; Simpson, 1975).

The well-known interpretation of the public-private distinction adopted by this "contractarian" approach can be (very briefly) summarised as follows.

Only private businesses can modify their liabilities and duties by stipulating bilateral contracts with "privately chosen" and "selected" third parties. Public callings cannot stipulate private and bilateral contracts with any specific third party as their status impedes them from selecting the specific third parties they deal with, and the specific duties and liabilities associated to the service provided. While the former would be a violation of the duty to serve all indifferently, the second would be a violation of the duty to serve "with due care". As privately stipulated contracts can potentially undermine the duties characterising the status of a common business, common law explicitly denies public callings the possibility to subject themselves to contract law. Thus, while the terms of transactions and trade undertaken by a private business are subject to contract law, terms of transaction and trade of common (public) businesses are subject to the law of tort.

In sum, the standard interpretation of the public-private distinction provided by the literature can be interpreted as a dichotomy between duties imposed on the public calling according to *tort law*, and duties imposed on the private calling according to the *law of contracts*. In legal terms, the difference between the two is characterised by the fact that in tort law the compensation of individuals occurs "for losses which they have suffered in respect of all their legally recognised interests, rather than one interest only" (Prosser, 1975: 6). The distinction is between a multiplicity of generic interests (plural) versus a single, specific and contingent interest (singular).

Stated differently, the fundamental point to remark is that duties and obligations of a public calling in (common) tort law derive from its *status*, and not from the stipulated terms of trade as defined in contract law. The distinction between tort law and contract law therefore reflects the distinction between a *social status* and a *bilateral contract*.

Unfortunately, this fundamental aspect of the distinction between tort and contract law has usually been misinterpreted and not correctly appreciated, as very clearly discussed by Cherry (1999: chapter 4). The current literature has often interpreted the distinction between a private and a public calling as a distinction between *two types* of contracts, and not as a distinction between *status and contract*, as the following passage (among many others) exemplifies (Stone, 1991: 29): "a common calling was one in which a person was in business –selling to customers– as opposed to being in the exclusive employ or service of another person." This type of interpretation cannot be correct as it does not accurately reflect the original essence of a public calling.

More generally, the recent literature, adopting a standard "failure-based approach" has tended to reinterpret the status of private calling as one with the faculty of stipulating exclusive contracts with specific customers or, in other words, to *discriminate* among potential customers, differently from the case of a public calling, characterised by the fact that businessmen in this second case cannot discretionarily choose and select customers "from the general public".

The tendency to interpret the distinction between public and private duties as a distinction between duties deriving from *alternative types of contracts* can be interpreted as a "contractarian approach" to the public-private dichotomy of common law, as it exclusively focuses on the nature and the type of the contract. This contractarian approach is common to all antitrust-based (failure-based) approaches. On the one hand it is consistent with a monopoly explanation of public callings: contemporary (monopoly-based) antitrust law focusses on the effects of exclusive (specific) contracts on monopoly power, as it reinterprets the faculty to select customers and to set up exclusive or special contracts as an attempt to foreclose markets, or leverage monopoly power across stages of the chain (Lafontaine and Slade, 2008; Rey and Tirole, 2007; Rey and Vergé, 2008; Riordan, 2008). On the other hand "efficiency-based" theories, such as TCE (Williamson, 1985: chapters 2 & 3), focus on the advances inherent in specific or exclusive contracts deriving from productive inefficiencies along the chain. However, regardless of the different interpretative approaches, all antitrust approaches are linked by this peculiar "contractarian" interpretation of the public-private distinction.

Yet, as clearly put by Cherry (1999: 35), "[t]he primary problem with the regulatory contract theory, as applied by the courts to uphold the telephone companies' limited liability tariff provisions, is that it is simply factually wrong." Public callings do not derive their duties from the fact that they cannot stipulate exclusive contracts, as it is usually asserted, but from the fact that in order to operate they cannot (or do not need to) stipulate contracts *tout court*. It is really the *absence* of any contractual form that characterises a public calling, and that originally motivated its emergence. The contractarian approach, which adopts a genuine "antitrust approach" to asset ownership, reflects an unfortunate (19^{th} century) conflation of the theory of public callings (to be intended in tort law), with the theory of public service utilities and public franchises, to be interpreted in contract law (see the interesting discussion in Cherry, 1999: chapter 4). (In the same way in which the 19^{th} century reinterpretation of public interest reflects an unfortunate conflation of the original common law concept with the theory of public service utilities and public franchises.)

For this reason, the dichotomy between a public and a private calling shall not be interpreted by looking at the differences between the types of contracts stipulated, but rather by trying to understand the fundamental difference between the concepts of *status* in common (tort) law, and the concept of a *contract*. Stated differently, the real question to ask should become: what confers social status independently from contractual relations? Or, how should social status be interpreted independently form contractual terms? And what distinguishes a businessman's social status from its faculty to stipulate contracts?

In sum, in order to properly appreciate the distinction between private and public calling it is fundamental to keep *status* (the key concept in common tort law) and *contracts* (the key concept in contract law) well-distinct so that the former can be explained and interpreted separately and independently from the latter.

For this reason, a second interpretative solution is needed in order to appreciate the distinction between a public and a private calling. It is at this point that the clear parallelism with the concept of "public interest" emerges, and the role of value becomes key.

3.2. The original distinction between private and public calling in common law: employment of the property "to serve one's self" (private self-consumption) vs. "to serve the public" (public consumption)

The distinction between public and private has to be intended as a distinction between *status* and *contract*. Yet, no clarity has been done over this distinction. What does the dichotomy between contract law and tort law mean? What is this distinction based on?

In order to appreciate what the distinction between public and private ultimately means it is necessary to go back to the origin of the common carrier concept.

As recalled by Bogen (1996), common carrier status was first used in order to define the business of innkeepers (see also Adler, 1914, and Jeremy, 1816). This is how Bogen (1996: 54-55) describes the process leading to the creation of "common inns" in mediaeval England:

> "Travelers in the Middle Ages relied to a great extent on private hospitality –the wealthy staying at the castles of other noblemen or in religious houses, the poor in servant's quarters or with other local villagers. Monasteries established guest houses to provide accommodation for travellers. Some of the guest houses became public inns. Inns flourished on the continent in the twelfth century, and by the thirteenth century, as written records demonstrate, both reputable and disreputable inns were also well established in England."

> "The rise of trade and the breakdown of the feudal order put more people on the roads and created a demand for places to stay … The inn normally had a stable for the horses and provided food and shelter for the merchants, messengers, carriers, small landowners, and other members of the middle class when they travelled. By Chaucer's time [late 14th century], innkeeping was a thriving business, and the law was forced to deal with the questions that it raised."

The above passage is sufficient to understand the real nature of a public calling *vis-à-vis* a private calling: *a private calling is characterised by the fact that a given property is primarily and prominently used in order to satisfy the needs and the necessities of its respective owner*.

In this case, a "private hospitality" is characterised by the fact the castles and monasteries were primarily intended and employed by their respective owners as their own private dwellings, and they were therefore used by their respective owners in order to address private purposes. In other words, this simply means that these properties were primarily and purposefully used for self-subsistence, or auto-consumption.

Castles and monasteries could originally offer *private hospitality* in the sense that their primary and prominent role was to provide subsistence to their owners, in order to address their own necessities, and they were *not* employed as means for exchange, trade and commerce in the first place. This is how the term "private" should be interpreted: they could offer *private hospitality* in the sense that their original employment was *a private one*. For this reason, the undertaking derives its status from the original purpose of its "natural employment": a private one.

Thus, inns could offer private hospitality as long as they were not ex-ante dedicated to the satisfaction of third parties' necessities by means of market transactions (Adler, 1914: 153):

> "In medieval times merchant strangers were arbitrarily assigned to 'hosts' who were responsible for their good behavior. Travelling nobles often lodged in the religious houses without invitation, but these houses were not common inns."

In the same vein, the concept of "public" should be interpreted in the specular way: a calling is "public" as it simply serves the public. Once again, the public status of a calling derives from its nature and from the fact that it is purposefully and exclusively used in order to transact with the public, and serve the public. If an "inn" is systematically and primarily employed by its owner, not to satisfy private (personal) needs, but rather in order to purposefully address the needs and requests of third parties of the public, then it qualifies as a "common inn", meaning that that specific inn is i*n the business of inn keeping*.

It is now clear how the term public should be interpreted, and the reasons why the two terms "public" and "common" can be considered as synonyms. A "common calling" is characterised by the fact that it *serves the public*, rather than its direct owner (see Arterburn, 1927: 419 and footnotes herein).

As stated by Adler (1914: 154; emphasis added):

> "It is scarcely open to dispute, therefore, that all trades mentioned in the books were on occasion called 'common,' that the private was distinguished from *the public exercise of a trade*, and it is difficult to see how 'common,' in the earliest uses as applied to all trades, can have any significance other than the one at the present time limited to a few trades, - that is, profession to serve all … *it follows that 'common' means, and meant in this connection, 'open to public service*'."

Thus, the expression "public calling" in common law has nothing to do with the fact that a certain property is owned or controlled by the governmental or legislative authority, as is usually thought, and incorrectly asserted.

This confusion on the correct interpretation of the word "public" certainly derives from the fact that, very often public callings (such as common carriers) are *also* public franchises. In the case of corporate charters, the term public really means that the properties are operated on behalf of the public authority, and under governmental control. A public calling is still characterised by privately owned and controlled assets and properties, *in the same way in which private callings do*.

In sum, what distinguishes private from public callings is not the nature of the ownership of the underlying input (asset), but rather the way in which the underlying assets are operated, and the nature of the specific employment the underlying property is dedicated to: *the same privately owned infrastructure or assets can be either a private or a public calling depending on whether it is primarily and purposefully employed in order to serve the owner or the public*.

Two remarks are in order at this point.

First, it is easy to see how the switch from a private to a public inn perfectly retraces the aforementioned description of Lord Hale's switch from *juris privati* to *juris publici*: in medieval tort law inns become *common* (callings) as soon as they are purposefully employed in order to *serve the public* rather than the property owner; while they remain *private* (callings) as long as they are employed in order to (almost exclusively) serve the owner. If the owner is also willing to salutary open his own property (*juris privati*) to the use of other third parties in an impromptu manner, he can do so. Yet, as the service is provided to the public by the employment of a property originally employed for private purposes, the owner cannot be held responsible (liable) for any damaged or lost item affecting the guest. Under the regime of private callings, there is no duty to serve "with due care", for the simple reason that there is no duty to serve the general public in the first place.

The distinction between public and private calling perfectly retraces Lord Hale's first necessary condition of the three-criteria test for the definition of "affectation with public interest". This is not surprising as Lord Hale's public interest is genuinely rooted in common law, and the institution of public calling was developed in English common law (four centuries earlier). Yet, this is just further proof that the most appropriate common law distinction between public and private should be interpreted by looking at the distinction between "to serve one's self" in order to address private needs and "to serve the public" in order to address public (third parties) needs.

Second, it should be noted that what really distinguishes a private calling from a public calling is the role of market transactions undertaken in order to serve third parties, and relative monetary compensations.

A public calling is characterised by the fact that its employment necessarily always involves a market transaction, and that its services are always enjoyed subject to a monetary compensation. This is just another way (but not the only, nor the most important, way) to understand why “common” in tort law should be simply intended as a synonym for “business”.

As a result, what links together the term “common”, “public” and “business”, all three perfectly interchangeable synonyms,[142] and distinguishes them from “private properties or undertaking” is *the necessary presence of a market transaction, deriving from the fact that the ultimate aim motivating the employment and the ownership of the property corresponds to the purposeful pursuit of some monetary compensation in exchange for the service provided*.

In Arterburn’s (1927: 420) words:

> “Common as used in ‘common carrier’ is generally assumed to mean ‘public,’ when in reality the original use of the adjective was merely to distinguish those who carried as a trade and those who carried as occasional acts”
>
> “Carrying on a business is not a casual act, but a habitual and *common* practice of rendering service *for compensation*.”

This passage is key. There is no “common” or “public” undertaking without a market transaction, nor is there a common, or public undertaking without a purposeful pursue of monetary compensations received in exchange from the use of the underlying property. A public calling (common business) is characterised by the fact that whenever a property is employed, and a service is enjoyed, a monetary transaction is undertaken, and a compensation is generated.

This point is clearly put by Adler (1914: 156; emphasis in the original):

[142] “The only possible meaning in "common" here is "public," and "public" can signify here nothing but serving the public, in the business of.” (Adler, 1914: 154; footnote 75).

> "The carrier, like every other business man, purports to serve and to deal with the public. Business is impersonal; … A man is engaged in business when he solicits the favor of and undertakes to deal with persons *indifferently* for profit."

From which, a fundamental conclusion follows (Adler, 1914: 158; emphasis added):

> "Under a true interpretation of the common law all business is public, and the phrase 'private business' is a contradiction in terms. *Whatever is private is not business, and that which is business is not private*."

In other words, in common (tort) law any service undertaken in order to serve the public upon monetary compensation is a *business*, and in this capacity it is "common" and, thus a "public" undertaking. The adjective "private" in common law can only be reserved for those undertakings aimed at private self-subsistence and auto-consumption. From which it follows that the expression "*private business*" is "a contradiction in terms"; something can be regarded as private if and only if it is *not* a business activity.

This clarification is also perfectly in line with the common law interpretation of public interest: as already mentioned above; only when some "common charge" is applied in order to employ the property, "affectation with public interest" might apply. However, there can be no public interest if there are no charges to regulate in the first place.

3.3. Public vs. private calling as a dichotomy between (monetary) compensated and uncompensated employments of a given property

The above conclusion of Adler summarizes the essence of the original distinction between a private and a public calling in common law: the status of "private" depends on the fact that a property is not necessarily employed on the market in order serve the public, *so to* purposefully generate monetary returns out of its use. This is not the case for a public calling, where the presence of a monetary transaction compensating the service rendered to third parties represents the necessary and ex-ante declared purpose of the undertaking.

For this reason, a public calling cannot be characterised by the nature of the specific business considered or, alternatively, by the characteristics of the final market in which the service is performed, as asserted by the so-called "monopoly approach". For the same reason, the two concepts of "public calling" and "public interest" are very different ones: as already recalling, the public calling status is a necessary precondition to "public interest", but it does not exhaust it.

At the same time, the distinction between public and private calling cannot even be characterised by the nature of the contractual relationship between the "businessman" and the "client" (user or "guest"), as asserted by the "efficiency approach" of TCE, and by any "contractarian" approach in general. TCE in particular developed its own theory of institutions (and asset ownership) based on the "contractarian" distinction between spot "classical" contracts and more dedicated, long term, contracts, such as the "neoclassical" or the "relational" ones (Williamson, 1985: chapter 3, esp. figure 3-2 at page 79).

Yet, the distinction between "public calling" and "private calling" does not distinguish between different types of businesses, and this means that it does not even distinguish between different *types of market transactions*, consistent with what has been discussed in the previous section. Rather, it distinguishes between the presence of a public business, from the absence of any business-oriented undertaking and this means that the distinction is between the recurrent and necessary presence of a market transaction, and the absence of any monetary market transaction *tout court*. In the same vein, this means that the public-private dichotomy in common law does not even distinguish among different *types of contracts*, as it rather distinguishes between the presence of a (bilateral and contingent) contract, and the absence of any type of (bilateral) contract.

A more appropriate interpretation of the dichotomy between the notion of *status* proposed by the law of tort and the notion of "*negligence*" (or *contract*) based on contract law (see Cherry, 1999: 11–13; and references herein) must be based on the dichotomy between auto-consumption for self-subsistence on the one hand, and "holding out to the public" for third parties' consumption on the other hand, where the critical dividing line between the two is given by the presence (or absence) of a market (monetary) transaction between two parties. This is how the common law approach based on the concept of status differs from the neoclassical contractarian approach: the former understands an actors' status by explicitly allowing for the employment and consumption of a property's services even outside a (compensated) market transaction, the latter does not as it requires that any property is employed only subject to a monetary compensation.

For this reason, the dichotomy between status and contract, tort and negligence, and therefore between common (public) and private respectively, cannot be based on the standard failure-based explanation of the antitrust approach, regardless of whether the failure is interpreted in monopoly terms, or in contractarian-efficiency-incentives terms (as already remarked, all these three terms can be used in order to define contemporary approaches to asset ownership and institutions).

Figure 6.5 provides a general framework where the previous discussion on public interest and the present discussion on public calling are brought together.

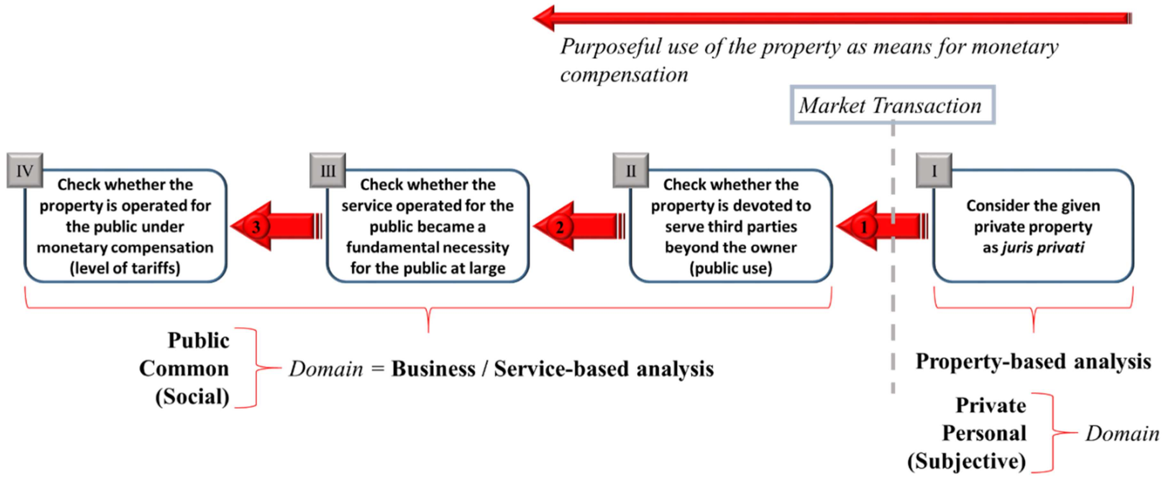

Figure 6. 5 – The private/public distinction in common law as a distinction between monetary (compensated) and non-monetary (non-compensated) employment of the property

As already noted, the distinction between "common" and "private" in tort law precisely corresponds to the first step of the "public interest test" discussed in the previous section. In common law, a property is affected with public interest" if and only if it is preliminarily employed in order to serve the public, rather than for personal self-subsistence.

Although only common undertakings can be potentially "affected with public interest", it obviously does not follow that any property must necessarily be affected with public interest as soon as it is employed in order to extract monetary profits: other necessary conditions must be respect for a *juris privati* to be treated as *juris publici*, as discussed above. Yet, the characteristic of being a business, and therefore the presence of a monetary transaction (or compensation) is the first preliminary and necessary condition in order to declare "affectation with public interest".

4. Concluding remarks

The chapter has shown that the dichotomy between public and private in common law can only be understood in light of the distinction between assets and services. Moreover, the chapter has also showed that the distinction between assets and services in common law can only become meaningful is the private employment (self-consumption) of the property is distinguished from its public employment. While the private domain must also follow a property-based analysis, the public domain is characterised by a service-based analysis.

From the discussion here provided one aspect emerges clearly: the common law approach can only be adopted if dualism in value is retained, and if the public-private dichotomy is explained in function of the dualism between exchange and use value. Moreover, the chapter shows a striking similarity between the duties of a public calling and the conditions (assumptions) imposed on the theory of value in order to reach monistic and objective valuations, so to justify value judgments.

This is indeed the case. The next chapter will put everything together, and will show that the common law foundations to public control and regulation can be perfectly understood by simply looking at how neoclassical economics “solved” the original dualism in value: the duties imposed by common law in order to regulate private properties have precisely the same effect, and the same intent, of the two assumptions of neoclassical value theory: to legitimise and justify social value judgments with respect to the way in which private properties ought to be employed.

CHAPTER 7

A new perspective on the regulation of private assets: the role of value and time in understanding the relationship between the public and the private dimension

Chapter synopsis

In this chapter, three important aspects of the public-private dichotomy in common law are discussed in order to clarify the essence of the institution of "common carriage", to highlight its close relationship with the common law concept of "public interest", and to finally understand the essence of the public-private dichotomy in common (tort) law, which is also a dichotomy between (public) *status* and (private) *contracts*. Implications will also be derived with respect to the contemporary theory of the firm, asset ownership and institutions in general.

This will be necessary in order to better understand the essence of the deregulatory process, and its conceptual flaws, especially with respect to the topics of unbundling, mandatory access to property, and discrimination.

The chapter links the concepts developed in chapters 4 and 5 with the discussion on public and private callings developed in the previous chapter. The aim of the chapter is to show that the two public duties of common businesses (and therefore common carriages) can only be understood if a dualistic framework in value (and institutions) is adopted, consistent with the original classical, and early Austrian approach to value theory. As discussed above, dualism in value necessarily also implies some temporal causality between the two domains: private and public. Only if value is a monism, and the institutional environment collapses to a singularity, the causality between the underlying asset and the service operated does not matter.

The chapter is organised around two main sections. While section one focusses on the role of value, section two mainly focusses on the role and the meaning of the presence of a univocal logical and temporal causality between the two units of analysis: services and assets.

The main message of section one can simply be summarised as follows: the definition of a public status through the imposition of the two public duties creates a univocal mapping between the specific service performed, and the nature of the underlying private property. The ultimate implication of this univocal mapping is to reduce the original dualism in value to a monism.

The way in which it does so is by establishing a univocal mapping between the private identity of the asset owners and their public identity. This generates an important implication: while the private identity of an economic actor relies on the set of properties that they possess, the public identity of an economic actor transacting services on the market is defined according to set of services that they is allowed to provide.

Section two focusses on the logical causality between services and properties. It especially focuses on the distinction between status, that defines duties ex-ante and is "general" or, transaction-independent, and contracts, that define obligations ex-post, and are transaction-specific. The section highlights the parallelism between the evolution of common law in contract law, and the development of the TCE framework.

The legal distinction operated starting from the late 1950s between common carriers contract-carriers and reflects (and was certainly the inspiration for) Williamson's distinction between general purpose assets, and service-specific assets, respectively.

1. Assets and services: interpreting common law duties in value terms

1.1. Private callings, public callings and the duality of value: the dynamics of use and exchange value

The first important consideration concerns the way in which the dichotomy between private and public should be interpreted in common law and, more generally, in law and economics.

As already remarked, the switch from the private to the public domain cannot be understood in failure terms. Neither market structure (competitive) considerations, nor "contractarian" or "transactional" failures, are by any means useful in order to understand what ought to be public and what shall remain private. Conversely, consistent with what has been previously asserted with respect to the switch from *juris privati* to *juris publici*, the real factor giving meaning to the dichotomy between private and public is represented by the dual nature of value.

Private callings can be defined as such only if self-consumption is allowed, meaning that private value-in-use must remain independent and distinct from (public) value-in-exchange. In this case, the property employed also remains *juris privati*, consistent with the discussion developed above. Thus, *juris privati* necessarily defines and characterises private callings, as in both cases the property is used as mere means for self-subsistence, or as means for auto-consumption. On the other hand, whenever the property is used as a means for exchange, thus as a support in order to generate transactions, and therefore monetary compensations, a "business" is also created. A business, by definition, is always common, thus public, as already discussed above. In this second case, the value of a property used in order to do business with third parties is necessarily "in exchange".

We therefore reach the same conclusion already stated above: while value-in-exchange is necessarily a social, collective, thus public and common category, value-in-use necessarily belongs to the private sphere, meaning that it is personal and subjective. Moreover, this distinction between private value- in-exchange and public value-in-use reflects another ontological dichotomy, the one between services (transacted to serve the public), and assets (appropriated in the private sphere).

In sum, in the same way in which the distinction between *juris privati* and *juris publici* can only be appreciated and understood by recognising the dualism of value, the much older (mediaeval) distinction between private callings and public callings can only be understood if the duality between value-in-use and value-in-exchange is retained and accepted. In other words, value in common law must be treated as a dualism, not as a monism: the common law *dualism in juridical status* (private and public) can only be correctly interpreted by adopting a *dualism in value*.

Figure 7.1 summarizes the above considerations.

From the figure above it is possible to see that the duality in legal status (private vs. public) can be interpreted in a twofold manner: either as a dualism in value, as already mentioned, or as a dualism in object of analysis. In fact, the public domain, characterised by value-in-exchange, overlaps with that part of the analytical and logical flow corresponding to the service-based analysis, while the private domain overlaps with the first step, where the object of the analysis is the property. It should be noted that what is exchanged or transacted in the public domain are services, meaning that value-in-exchange applies to a certain *service*, not to the underlying property. This is consistent with basic microeconomic analysis where the object exchanged and "valued" is the service: prices, in microeconomics, evaluate resources' services, certainly not resources. Conversely, value-in-use is a characteristic of a certain owned property.

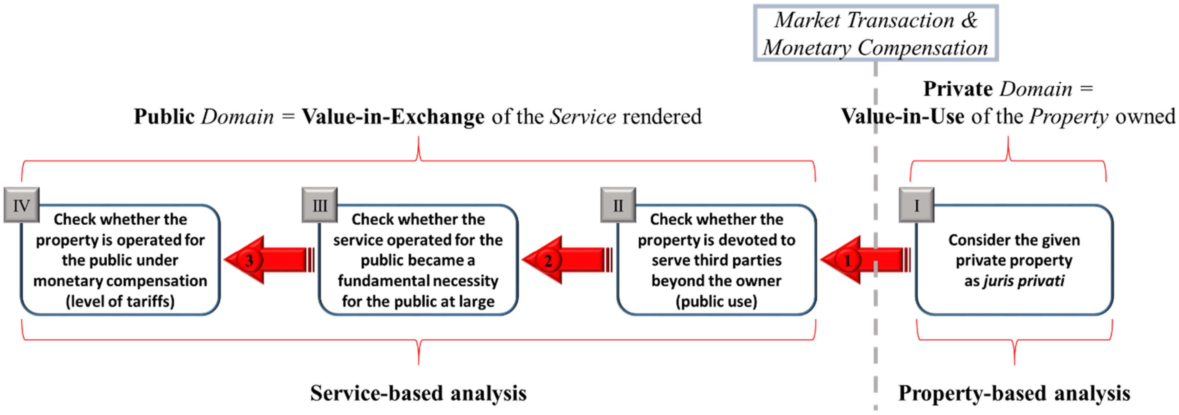

Figure 7. 1 - The juridical duality between public and. private understood as a duality in value (in-exchange vs. in-use) and in object of analysis (service vs. property)

Note that this second dichotomy in unit of analysis is consistent with the way in which the terms "common" or "public" are defined in common law: while *juris privati* always characterises the legal status of a certain *property* (thus the underlying "thing", or asset), the terms "public" or "common" are synonyms of business, as already highlighted above. Yet, a certain business is not characterised nor defined by the underlying property owned as inputs, but rather by the nature of the services sold as output on the final market. The unit of analysis characterising a business is necessarily the service, and this makes the service the unit of analysis characterising any endeavour in the public domain.

Leaving the issue of the duality in the object of analysis for later considerations (section 1.2 below), we now turn to two important clarifications in the relationship between the common law dichotomy between public and private, and the respective duality in value.

1.5.3. The duality of value in common law: ambiguities between necessary and sufficient conditions

First of all, it should be noted that the duality between value-in-use and value-in-exchange is not interpreted in the same identical way by the two common law approaches discussed thus far: Lord Hale's discussion of public interest on the one hand, and the English common (tort) law distinction between private and public calling on the other.

While the presence of a market transaction in the first case is a *sufficient condition* in order to trigger the potential "affectation with public interest" of a certain property, it only remains a *necessary condition*, yet not sufficient, as in the case of a pubic calling.

As already seen, in the case of Lord Hale's treatment of public interest, a property remains *juris privati* as long as it is used *exclusively* by its respective owner (or their family and close relatives). Yet, as soon as any member of the public becomes reliant on it, and must pay a compensation to enjoy its services, the same asset can potentially be affected with public interest (provided other conditions are also verified). On the other hand, the case of a public calling in tort law is specular. As discussed above, it is still possible for a private calling to provide occasional and sporadic services to members of the public, provided that the public transaction occurs *ex-post* (only after a certain request for service is made). This means that a calling can still remain private as long as the property (i.e. the inn) is not purposefully and *exclusively* ex-ante dedicated to serve the public.

As a result, in the case of Lord Hale's "affectation with public interest" any property can *become juris publici* as long as it serves members of the public, meaning that it *is not exclusively used for private purposes*. On the other hand, in the case of mediaeval tort law any property or undertaking can remain private as long as it is also and primarily employed for private purposes, meaning that it is *not exclusively used for common (public) purposes*. The difference between the two cases reflects the fundamental difference between a sufficient and a necessary condition and it can be better appreciated and understood if the logical and temporal causality of the events is taken into consideration (section 1.2 below).

This inconsistency can be quite serious as it changes the way in which the nature of value should be treated, and the relationship between (subjective) use value and (objective) exchange value should be intended. In the case of pure monism in value, no conceptual problem arises, as the two approaches remain identical. In the case of some dualism, whenever subjective value in use and social value in exchange coexist, the two common law approaches do not lead to the same conclusions.

In the case of public interest, the mere presence of exchange value beside some value in use is alone sufficient to potentially trigger the new legal status of *juris publici*. Thus, the concomitant presence of the dual nature of value is sufficient for a certain property to be potentially treated as an asset subject to public control. In this case *juris privati* can only be retained if value reduces to a monism, when value completely collapses into use value alone. On the other hand, in the case of tort law the status of "private calling" can be retained as long as the property under consideration can *also* generate some value in use for its owner, granted that private use remains the primary purpose an asset is dedicated to. In this second case, a private calling can allow for a dual nature of value, while a public calling is characterised by monism in value, where value completely collapses to value-in-exchange.

It should be noted that, among the two, the way in which the duality in value is treated by tort law is the one consistent with the assumptions and the approach adopted by modern neoclassical economics, where only services are exchanged in the social dimension, and value in use completely disappears and conflates into value-in-exchange. The latter also remains the sole dimension of value. This is consistent with the fact that in neoclassical microeconomics, value can be computed in the public dimension (in the form of services' prices) if and only if all vestiges of subjectivity in valuations are ruled out (by assumption).

Conversely, Lord Hale's treatment of public interest is the one posing the most serious conceptual and analytical problems as a cohabitation between public value-in-exchange and private value-in-use of the property seems to be sufficient to switch to *juris publici*. However, the simultaneous presence of "self-service" and "public service" would severely undermine the basic assumptions on which modern welfare economics rests (duality and monotonicity of economic functions).

It therefore seems that while in medieval tort a property operates in the public domain only under the condition of structural separation (vertical disintegration), in Lord Hale's treatment complete structural separation is not necessary in order to switch from the private domain of juris privati to the public domain of juris publici.

1.5.4. The division of labour, monism in value, and the switch from a common law to a contractual approach to public callings

The second consideration worthy of mention concerns the dynamic aspect of the dual nature of value, consistent with the discussion developed above with respect to the distinction between *juris privati* and *juris publici*. In order to make proper sense of the various ways in which the distinction between common and private callings has been interpreted through the centuries, it is necessary to take into account the dynamic nature of value.

As mentioned above, the original (mediaeval) interpretation of a public calling was simply that of "business", to be intended as "purposefully serving the public", thus for public use. Originally, since the 13th century, practically a*ny* business could hold the status of "common" as long as it was, precisely, a business. Yet, the public-private dichotomy has never been a clear-cut and static concept. Two things have to be remarked upon here: one, the precise legal meaning and application of the dichotomy changed over time, and two, the conceptual foundations and the analytical interpretation of the dichotomy changed greatly through the centuries and especially in the last few decades, as already anticipated above.

Although the original medieval interpretation of the institution of public calling reflected the interpretation provided in the present section, things started to change at the end of the 15th century (Adler, 1914: 153-58, Cherry, 1999: 12-14, Payton, 1981: 130-133; Wyman, 1911: chapter 1). More specifically, this led to a gradual exemptions of the scope of businesses required to respect original public callings' duties.[143] This deviation created a rupture in legal regimes: while some businesses remained under mediaeval common (tort) law, many more shifted towards contract law.

This process reached its complete maturation by the very early 18th century, when only two businesses remained under the original legal framework: "At the opening of the eighteenth century, then, the English common law had fixed on common carriers and innkeepers as a special class of businesses with extraordinary duties to the public enforceable at common law" (Payton, 1981: 132). All other types of businesses shifted to contract law. Thus, a dual legal regime has eventually been established, where the dividing line is now merely based on the *nature* of the business considered. This dual legal regime has lasted until the deregulatory wave of the 1980s-90s (when even the last vestige of common law for carriers was dismantled).

This historical passage completely corrupted the original definition of "public calling", yet it offers an understanding of the origins of the contemporary "contractarian" interpretation of the public-private dichotomy (see Cherry, 1999: 48-51).

For the above reasons, there is a general agreement among commentators on the fact that the original common law interpretation of public calling, which eventually only applied to common carriers and innkeepers, just represents an outdated medieval heritage, and that what fostered this trigger was the expansion of trade and the market, beginning in the contemporary epoch.[144] Thus, there is little doubt among all scholars that even for the

---

[143] According to Wyman (1911: 17): "the early cases which were just under discussion are illustrations of this course of events. Barber, surgeon, smith and tailor are no longer in common calling because the situation in the modern times does not require it; but innkeeper, carrier, ferryman and wharfinger are still in that classification, since even in modern business the conditions require them to be so treated."

[144] "The laws governing common carriers, blacksmiths, and innkeepers were medieval survivals recrafted to the requirements of a world that was breaking away from its medieval past." (Payton, 1981: 132).

fundamental switch in legal regimes, from the medieval law of tort to the more modern law of contracts, it was Adam Smith's "*expansion of the market*" that represented the main and most important factor of change (Adler, 1914: 157, Stone, 1991: 25, Wyman, 1911: 2-3).

As is well-known, once that market expands, the division of labour follows. As a result, the changing nature of the public-private dichotomy, and the rationale that justified the gradual shift from public callings regulated under tort law to private callings regulated under contract law, was also consistently interpreted by looking at Adam Smith's "division of labour".

The problem here derives from the fact that the "division of labour" itself can be interpreted either in antitrust (failures) terms, consistent with a contractarian approach, or in regulatory (status) terms, in line with the original common law approach to the public-private dichotomy. Thus, on the one hand the standard and accepted explanation links Smith's "division of labour" with the genuine antitrust concepts of scale economies, market structure and competition (Chandler, 1977; Coase, 1988b; Stigler, 1951) or, in other words, with the "conditions of trade" associated to, and deriving from, a certain size of the market, as standard industrial organization teaches (Bain, 1956, 1959; Schmalensee, 1989; Sutton, 2007). This is the essence of Wyman's (1904) original reinterpretation of "*the law of the public callings as a solution to the trust problem*", as is the title of his 1904 paper, where antirust remedies (the trust problem) and regulatory practices (the public calling) have been explicitly merged under the same interpretative framework for the first time.

Yet, as already discussed above, the monopoly (or market structure) interpretation of the institution of public calling is unfounded and misleading, in the same way as any market failure interpretation.

For this reason, it would be a complete logical fallacy to try to explain the change in legal meaning of the public-private dichotomy using a market failure (monopoly) interpretation of the division of labour, if the monopoly explanation does not fit the phenomena that one aims to explain in the first place.

The advantage of the "value approach" to the public-private dichotomy derives precisely form the fact that it can still explain the historical changes in the way in which public and private callings have been interpreted in function of the dynamic expansion of the market, and therefore of the growing division of labour, while at the same time being faithful to the original regulatory interpretation of medieval common law, which was *not* based on failure-based concepts.

Thus, consistent with what has been anticipated above, it is clear to see why the original dichotomy between public and private has changed its meaning through the centuries as trade slowly started to develop, once the dualism between (private) value-in-use and (common) value-in-exchange is recognised. Building on Adam Smith's original intuition (as discussed in chapter 2 above), the story can be summarised as follows.

On the one hand, in the "early and rude" medieval eve, properties and assets were nearly exclusively used for self-substance and personal auto-consumption or, for the self-subsistence of an enlarged family or social group defined by informal social ties. The prominence of self-subsistence corresponds to an epoch where the private sphere based on subjective value-in-use largely prevails over the public sphere where formal market transactions generated by trade and commerce relied on social value-in-exchange. As demand grew and trade developed, from the 15th century onwards, more and more undertakings and properties could be *purposefully* and *exclusively* dedicated to serve the public, thus to formal trade and commerce. With time, more and more properties started to be endowed with the "common" status. With the increased division of labour, the "exchange" nature of value became more prominent with respect to the "use" nature of value.

As soon as nearly all productive properties and undertakings became completely dedicated to serve the public through trade, rather than to serve their respective owners in private self-subsistence, more properties were affected with the status of "business" in the proper sense. At this point, when nearly everything could be regarded as "business", the original common law distinction between public and private callings ceased to make sense. This is true in part because in a world with widespread division of labour and advanced trade relations nearly any profession can potentially be labelled as "common". In this case, the status of "common" becomes redundant and meaningless.

It is at this point, when value has eventually became a monism and the value of a property conflates into the value-in-exchange of the service operated (its "business"), that the public-private dichotomy acquires a new, contractarian, meaning. Yet, only in this last stage, when value collapses into a monism, and private value-in-use substantially disappears, that the public-private dichotomy can be used to distinguish *different types of businesses*, and therefore *different types of contractual relationships*, or *different types of* *market transactions*, and not to distinguish a business activity (public exchange) from a *non-business one* (private self-subsistence), as the original dualism in value implied. This extreme case is, of course, the case of completely deverticalised businesses as well. The latter is, in fact, the default situation from where any contractarian approach begins its analysis of the public-private dichotomy).

This approach is in line with Adler's (1914: 157-58) reinterpretation of the nature of the mechanism that dynamically shifted the meaning of "common" through time:

> "But with the inventions of Arkwright, the writings of Adam Smith, and the spread of the idea of free trade, a great change took place in business conditions toward the close of the eighteenth century. In ordinary trades there ceased to be any need for a distinction between the common and the private exercise of a trade. With the repeal of the Statute of Apprentices in 1814 the distinction made in such a statement as 'To make a man of a trade, he must be apprentice to him who did openly, commonly, and by publick profession sell, and not privately by stealth,' would cease to be necessary and would be gradually dropped as meaningless."

In sum, it is not because of structural, competitive or transactional reasons that the distinction between public and private eventually "would cease to be necessary and would be gradually dropped as meaningless." Most importantly, it is not necessarily because of the increase in competition, and the advent of perfectly competitive markets, as Wyman (1904, 1911) first suggested, that the public calling status changed its nature and legal meaning.[145] Rather, it was just because the very socio-economic conditions on which that distinction originally was based eventually disappeared, that the public-private distinction also ceased to have any useful meaning.

This interpretation is in line with the original spirit of a public calling that, it should be recalled, was not to regulate market structure or competitive conditions in a certain market, but it was rather to provide a legal framework on which normative judgments of the actions of private actors could be based, so to have a legal foundations in order to make them liable in case of disservice or damage to any member of the public.

---

[145] "In the early part of the nineteenth century a combination of economic factors brought about in the business world as near an approach to condition of freedom in competition as can never happen in a world limited by time and space. Naturally enough with such individual freedom of action laissez faire became the accepted policy for dealing with the business world as the occasions for the application of the principles of law regulating public callings become fewer.2 (Wyman, 1911: 22).

The above discussion highlights two important points. One, the economic interpretation of the legal regime, as well as the epistemological relationship between law and economics, depends on the way in which value is conceived and treated. This means that the contractarian approach, which distinguishes the legal duties of different businesses only by looking at the different contractual arraignments, rather than at their different legal status, can only be valid once the preliminary necessary condition of monism in value is resolved and ascertained. Two, the above considerations highlight that the same legal concepts can naturally and spontaneously change meanings through time, in function of the changes occurring in the external socio-cultural and technological context.

Both these points were already made with respect to the dichotomy between *juris publici* and *juris privati*. Even in that case it was remarked that: (i) the shift from value-in-use to value-in-exchange represents the first necessary condition that has to be satisfied in order to ascertain whether a property is "affected with public interest" and that (ii) the balance between *juris privati* and *juris publici* directly depends on the changing habits of the property owners.

It should be noted that a transition from the original dual regime of tort law to the new contractarian approach can also be reframed in the following way: *a duality (heterogeneity) in legal regimes is supported by a duality in the nature of value, while monism in value necessarily leads to a monism (homogeneity) in legal jurisprudence*.

This conclusion, which holds for both "affectation with public interest" and the concept of public calling, can be also reinterpreted as a dichotomy between (absolute) status and (bilateral) contract.[146]

It is now necessary to turn to this distinction, as it will shed some light on two other fundamental aspects of the concept of public calling: the unit of analysis (the distinction between services and assets) and the logical and temporal causality, or directionality (given by the distinction between *ex-ante* and *ex-post* decisions). Both are key in order to understand how the public-private dichotomy interacts with the concept of asset ownership.

This leads to the second chapter of the present discussion.

[146] Note, moreover, that the same holds with respect to the duality in epistemological approaches: while the regulatory approach deals with the various statuses deriving from heterogeneity in valuations, the antitrust approach develops form a monism in value, and only focusses on the contractarian failures deriving from the various types of contracts, consistent with what has been discussed above.

1.2. From property to services: from the ex-ante private sphere to ex-post revelation in the social sphere and back

As discussed above, the duties of a public calling, and therefore of common carriers, derive their legitimacy from tort law, not from contract law. This in turn means that a common undertaking is bound to the "duty to serve all indiscriminately" and to the "duty to serve with care", because of its special (legal or juridical) status, and not because of specific contractual (yet contingent) terms stipulated with the transacting counterpart.

As recalled by Cherry (1999: 45):

> "the development of the common law led to confusion between the spheres of tort and contract law, and contributed to confusion regarding the legal basis of the duties and liabilities of common carriers and bailees to their customers."
>
> "As with the other obligations imposed on public callings, the basis for the duty to serve was that of *status* under tort law and not under contract law" (p. 51; emphasis in the original).

This difference is important as the meaning of the "common duties" changes according to whether they are interpreted in tort or in contract terms.

If we go back to the medieval origins of a public calling, it is possible to note that the original aim and rationale of the duties of a public calling was to make possible that a certain social actor could be clearly *identified* by the other members of a society, so to make him *recognisable to the public*. In other words, a public calling conferred *status* because it provided an actor with a *public identity*, beside his original *private* identity. After all, a "status" can simply be intended as a "public identity", and it is possible to interpret the role played by a "status" precisely as the one to transform an actor's identity from its original inherent *private nature*, to a *public nature*. In fact, status only makes sense if it has a "public" dimension.

This is how the term "public" should be understood and how it acquires its proper meaning in tort (common) law. Figure 7.2 summarizes the concept.

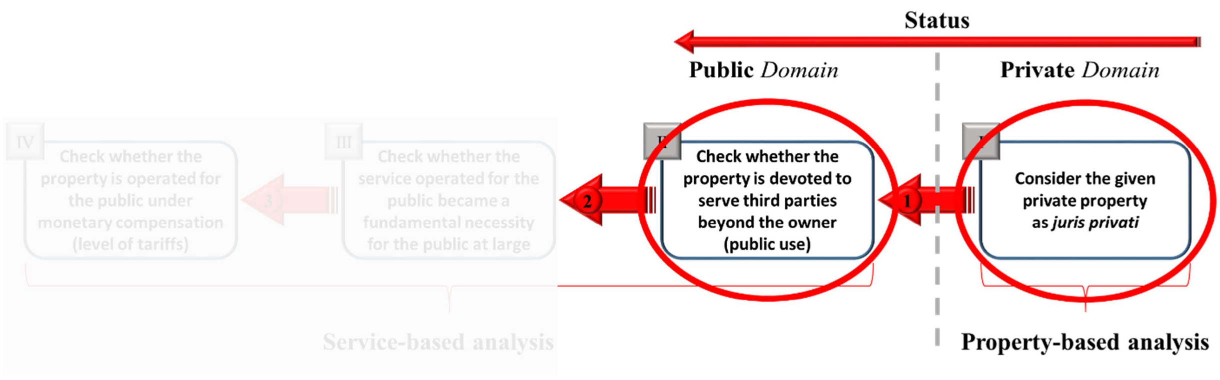

Figure 7. 2 - The role of status in common law: to "extract" an actors' identity from the private domain and place it in the public domain in order to "reveal" it to the public

The way in which tort law ensures that an actor's identity becomes of the public domain is to *univocally* link the identity of the actor to his *profession* or, in other words, to *what he does for living*. In Arterburn's (1927: 418-19; emphasis added) words, "The word 'common' meant simply 'business.' That is, *one engaged in work as a trade in which he made his living*."

The definition of some behavioural obligations, represented by the two duties of the public calling (to serve all indifferently and to serve with due care) becomes the necessary means in order to ensure that the private identity of the actor turns into a social, or public, identity. The status of "common" does precisely that:

1. It precisely identifies the specific service (or business) that the actor endowed with the status is supposed to undertake, and
2. It forces the actor to do precisely that as a means for living or, in other words, it makes sure that that specific business becomes his own profession or, "the way in which he makes his living".

These two points correspond to the two duties of a public calling as highlighted. On the one hand, by identifying the nature and the essence of the final service rendered it clarifies the terms of service, and makes the businessman liable for potential disservice in case he does not match the specific service requirements (duty to serve with due care). On the other hand, by forcing the actor to serve all indiscriminately, and by preventing the "refusal to deal", the "common status" ensures that the actor only makes his means for living by exclusively trading with the public, and by selling the services of his profession to the general public (duty to serve). When an actor is forced to "make his living" exclusively through the profession he is specialised in, living through self-subsistence, by employing their properties in auto-consumption is clearly ruled out.

This is the essence of the tort law approach to the two duties of a public calling, and this is the way in which the three synonyms "public", "business", and "common" should be linked together and understood in common law.

This particular condition is particularly interesting as the *identity* of a "common businessman" is completely and univocally characterised by his own profession or, *by the specific type of service* (or capability) that he sells and delivers to the public. The nature of this specific service is the one identified by the duties of the public calling. For this reason it becomes possible to identify any "common businessman" by simply appealing to him is profession and nothing more, by saying that a certain actor is "a surgeon", "a tailor", "a smith", "a baker", "an innkeeper", "a carrier", and so forth. Note that, consistent with the original common law tradition, there is no need to specify the adjective "common": as long as someone is appealed and identified through his own business or profession, this means that the latter is already "common", otherwise it would not be possible to appeal him in this way in the first place. In this sense, "common" and "business" are synonyms, with no need for further qualification.

On the other hand, it is possible to understand the meaning of "private" in function of status and identity, according to the original common law interpretation of the public-private dichotomy: a "private transaction" is one where the *identity* of the counterpart offering the service, and engaged in the transaction, is not univocally and unequivocally identified by the business she is performing, and therefore by the specific service she is rendering. A "random" person offering to cure someone else's sick horse, can do it "*de son bon gré*", as we saw in the previous chapter. Yet, if he cannot be identified in function of the specific service he is performing, this means that his "real", underlying identity remains "private", thus *not socially revealed by the specific transaction*. The service occasionally rendered by a "private" (anonymous) citizen does not identify him, nor does it reveal his social status.

As we saw above, it is for this reason that, whenever the identity of the counterpart of the transaction cannot be socially identified, the actor performing the service could not be subject to the duties and obligations characterising a proper (common) "surgeon" in tort law. For, the sole fact that the actor performing the service could be appealed a "surgeon", would already mean that that private actor could otherwise be identified by the service performed. For this reason, as "[y]ou have not shown that he is a common surgeon to cure such horses, and so, although he killed your horse by his medicines, you shall have no action against him without an assumpsit" (Arterburn, 1927: 412), as already mentioned above.

The same holds of course for any other profession, such as the only two that were still in the 18th century characterised by their common law duties: innkeepers and carriers. Yet, the fact that a "private citizen" offers shelter to a pilgrim does not make him an "innkeeper" as, in principle, his private identity can be *any,* independently from the contingent service that he is furnishing. This is true unless "hosting pilgrims" is what he explicitly does for a living (i.e. he is i*n the business of inn keeping*) so that his identity could be univocally and unmistakably publicly revealed through his profession. In this last case, the "common" duties apply, and he will be liable because of his (public) status (i.e. identity).

Figure 7.3 summarizes the way in which the dichotomy between status (tort) and contract (negligence) should be interpreted in common law: in function of the locus and the source of someone's identity.

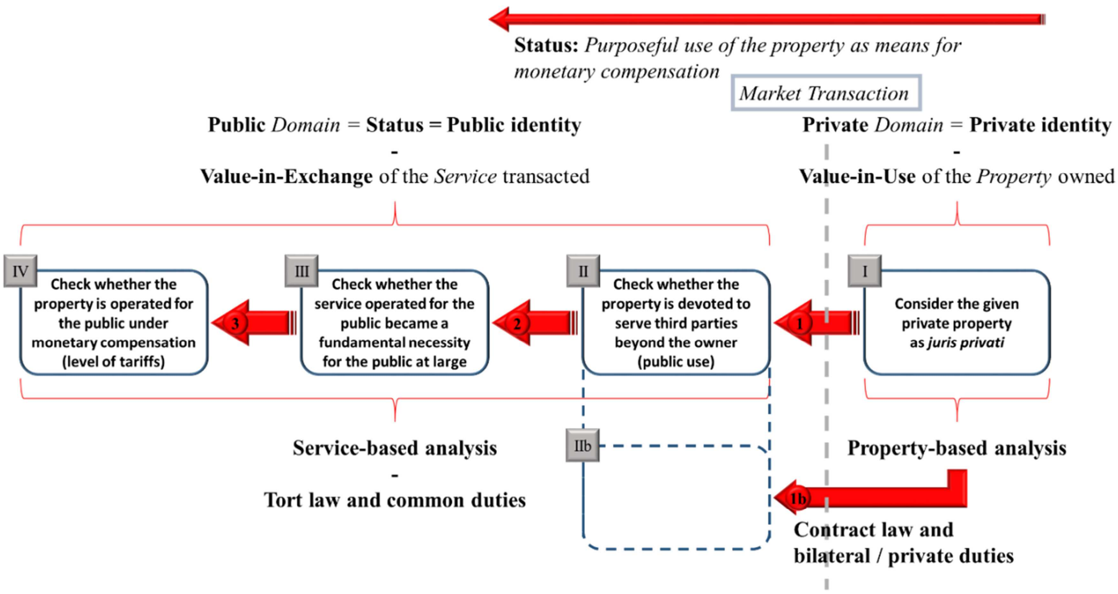

Figure 7. 3 - General recap of the public – private dichotomy in common law

Figure 7.3 provides a complete overview of the logical argument developed thus far. Before discussing some fundamental points already emerging from the figure, it is first necessary to highlight that a new element appears. In fact, from the figure, a third case can be noted beside the "standard" two already mentioned. The first case is the one where mere auto-consumption of the property occurs (stage I). This is the realm of *juris private*, as noted above. The second is the case where a "public status" is established, meaning that the property is exclusively and purposefully used in order to serve the public, as means for public exchange (stage II, occurring through condition 1). In this case the property can potentially be regarded as *juris publici*.

The third case is the new one, and corresponds to the one of private employment of the property, where the property owner however also provides impromptu "private" services under special (personal) request. This case corresponds to the second arrow (1bis) leading to step IIbis. This is the case where the dichotomy between *juris private* and *juris publici* does not match with the dichotomy between private and public calling. In the former case the opening of the property to the use of the public represents a sufficient condition to switch to *juris publici* (provided the other two conditions are also satisfied). In the latter, the mere fact of opening the property to public use is not sufficient alone to switch to the new "common status", as long as the property is also primarily used for private purposes. For this reason, in this second case the private (bilateral) trade is still regulated under contract law, rather than tort law.

Now that this first point has been clarified, it is time to highlight some important considerations emerging from the previous discussion with respect to unit of analysis and temporal and logical causality.

1.3. Unit of analysis and the public-private dichotomy: private identities are characterised by the properties owned, while public identities (status) are characterised by the services rendered to the public

First of all, it is important to note how the switch from a private to a public calling entails a switch in units of analysis.

On the one hand, the unit of analysis in the case of a private calling is the property employed or, much more generally, the productive resource available for either self-subsistence or private trade. Thus, in the case of the innkeeper, or the carrier, the unit of analysis is the inn, or the property used for transportation.[147]

A given service can never be the unit of analysis of a private calling for two reasons. In the case of employment of the property for mere self-substance (pure autarky of the single owner), what characterises and identifies a property owner in a social context is the property that he uses and nothing more as, no social exchange of services is even present.

[147] Note that, in the case of the doctor or the baker, the unit of analysis would be the productive power of the actor himself, thus its productive capacity in labour terms. Here we overlook this second case and focus on the case where the unit of analysis is the actual physical property employed, as the cases of interest of the present work are the assets, in the form of inn keepings or common carriage, even though the reasoning would remain the same.

On the other hand, in the salutary event in which a private calling decides to offer a service to a member of the public, there is still no ex-ante social obligation on the owner telling him *whether* to serve a member of the public asking for a service (no mandatory duty to serve), and *how* the property shall be employed in order to serve a specific member of the public, nor the specific nature of the service to provide (no duty to serve with due care).

As a result, a private calling is never characterised by a specific service, business, or operation, and this derives from the fact that there are no social duties super-imposed on the property owner forcing him to use the property in a certain way, and telling him *the way in which the property ought to be used*. For this reason it is not possible, nor legitimate, to univocally associate any specific service to the given property. In the case of a private calling, the situation is the one characterised by figure 7.4..

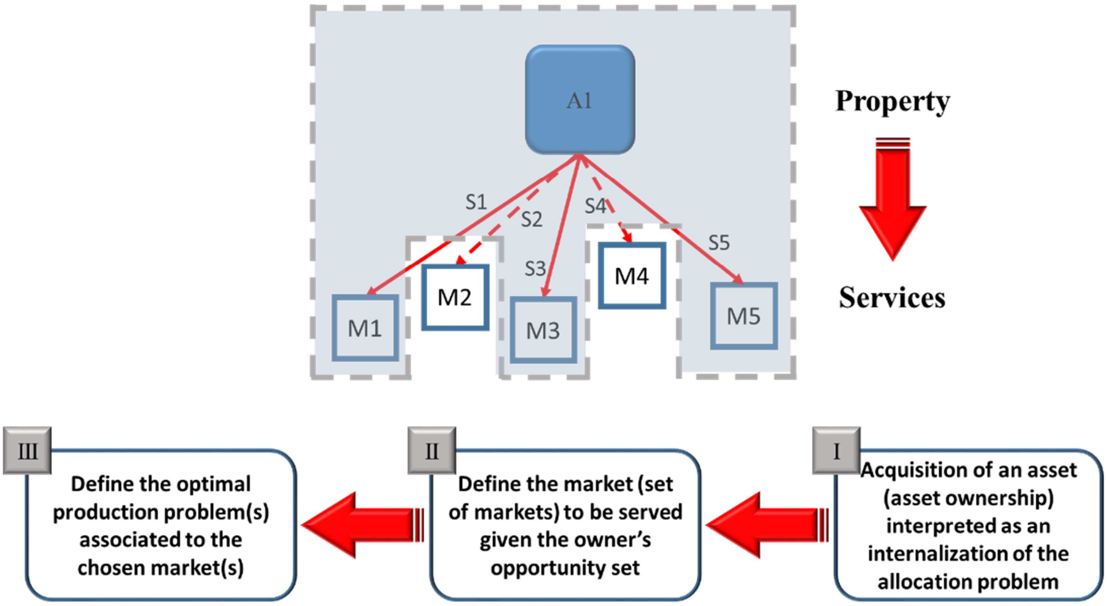

Figure 7. 4 - The unit of analysis of a private calling is the given asset (property) as no public duty is super-imposed on the property owner telling him how he ought to employ his own property

In sum, the private dimension is always characterised by the given property, while it is not possible to derive any conclusion with respect to the *service operated* and therefore with respect to the *business undertaken* as long as we remain in the private domain. The essence of a private calling consists of the fact that the "social given" is the given property, never the service operated.

This is consistent with the approach adopted by Lord Hale in his discussion on the "public interest". Even in that case, *juris privati* applied to "the thing" or, the given property, and could be extended to a certain service, or business, only in later steps of the analysis.

Things are very different in the case of a public calling and given the previous discussion it is now possible to understand why.

The essence of a public calling relies on the fact that the "social given" is the specific business to be performed, or in other words, the specific service that has to be rendered. The reason why this shift in the unit of analysis occurs derives from the fact that *social* constraints are now imposed on the private property with respect to the way in which the latter ought to be employed. These social constraints represent the crux of the matter and correspond to the two fundamental duties imposed over a public calling. The duty to serve with due care in particular (the original duty of public callings) is the one clarifying which service has to be provided and under which specific terms and conditions.

In other words, it is the imposition of social constraints on how the property ought to be employed that allows the switch in the unit of analysis from the property used as an input to the service rendered as the output of a specific business operated.

In conclusion, the public dimension is always characterised by the given services rendered by (superimposed on) the property owners, meaning that the unit of analysis of the public domain is the given service, not the property. This assertion can be interpreted in two alternative ways. On the one hand, a public calling is, by definition, characterised by *the way in which a property ought to be used*, or by *the specific business the various actors ought to operate*. On the other hand, services belong to the public domain as, by definition, they represent those activities that place the private property in relationship with the public, provided that these services are not direct to the property owner himself (in this case, it would be self-subsistence or auto-consumption of services).

It should be noted that this decoupling in the unit of analysis between private and public in common law perfectly retraces the shift in unit of analysis occurring between *juris privati* (which applies to the property) and *juris publici* (defined in function of the specific service performed by the property), as discussed in chapter 6 above.

Two further points have to be highlighted with respect to the change in the unit of analysis in conclusion of the section.

One, the same dichotomy in the unit of analysis is also reflected in the way in which the fundamental concept of *identity* should be intended. There is no univocal way to define identity. Above, it was asserted that it is possible to distinguish between a private and a public identity, where the latter can also be intended as the "social status" of an actor. It is now possible to understand what this distinction implies and what it actually means: *while the private identity of an actor is defined by the given set of resources or assets that he owns, the public (social) identity of an actor is defined by the given service (set of services) that he renders and transact with the public*. In other words, the private identity of an actor is shaped by what he owns, while the public identity of an actor is shaped by what he does (with what he owns). Status, being a public identity, is only defined by what somebody does or, better, ought to do in the public sphere.

Two, although it is true that in order to switch to the public domain it is necessary to first know the "common status" of an actor, the opposite also holds: it is not possible to confer any public status, nor to identify any public calling if the specific service to be rendered is not clearly and precisely identified in the first place. Any public status is based on the rights and duties linked to the specific activity operated, or service rendered. This means that a service defined in a vague and ambiguous manner will also necessarily provide vague and ambiguous social duties, and therefore a vague and ambiguous social status. In order to define a public calling it is first necessary to clearly define how a resource ought to be employed in the sense of: (i) what has to be rendered to the public (duty to serve) and (ii) under which conditions (duty to serve with due care).

This issue has historically represented one of the most severe problems and limitations of public regulation since the early 19th century. The problems of non-price discrimination and net neutrality in modern telecommunication policy both arise from the fact that in a converged environment (whenever it is not easy to distinguish among the various services simultaneously provided to the public by a given property) the clear definition of the "common business" (the specific business characterising the public calling) is far from an easy task. Stated differently, in a converged environment it becomes very hard to establish a univocal causality running from the underlying property to the services operated over, and by, it.

This problem of the correct identification of the specific business subject to public calling duties became central during the late 1960s in the telecommunication industry, soon after the advent of digital computer processing. During the first computer inquiry (concluded in 1971), the FCC distinguished between *carriage or primary services* (thus services covered by common duties consistent with the common law legacy) and *incidental services*. Only the former characterise the "common business" (thus the public identity) of the carrier, while "incidental services" lie outside the publicly recognised *profession*, or expertise, of the public calling.[148]

In truth, the distinction between common and incidental services operated by a public calling is much older than that. Not only was the distinction already adopted in railway regulation during the 19th century, but it went back all the way to the medieval origin of the institution of public calling itself. Bogen (1996: 67; emphasis added) recalls that with respect to inn keeping, already in 1365, "[a]ny obligation concerning goods a traveller left in his room was *incidental* to providing a room and unlikely to be the subject of any express undertaking." Consistent with the interpretation provided by the FCC various centuries later, already in medieval times, incidental services were not covered by common duties, thus absolute liability did not apply to them.

2. Logical and temporal causality and the public-private dichotomy: reversing the causality from services to assets to publicly reveal private identities

The second main consideration to develop concerns the temporal and logical causality of the process and of the events in tort and contract law.

Above, a fundamental distinction between tort and contract law (status or contract, or absolute liability and negligence) was discussed, the shift in the unit of analysis from the bare property (belonging to the private domain) to the services operated over or by it (belonging to the public domain). In truth, the switch in unit of analysis is just a manifestation of a much more fundamental mechanism: the reversal in the logical and temporal flow of causality. When the unit of analysis is reversed from the property to the services, it means that instead of moving from the physical asset to its employment, we are moving backwards, from its (visible) employment to the underlying property.

This causal inversion has many manifestations, and can be interpreted in various ways as it is at the same time temporal, logical, and analytical.

[148] For early discussions on the legal problems posed by the unclear interdependence between the underlying transmission infrastructure and the services operated over it, see Berman (1974), Berman and Oettinger (1975), and Mathison and Walker (1970). For overviews of the computer inquiries in the telecommunication industry see Benjamin et al. (2001: chapter 21), Blevins (2016), Bonica (1979), Cannon (2003), Frieden (1981, 1987), Kellog et a. (1992: chapter 11), Samet (1978). The present considerations are in line with Yoo (2013: 547) assertions stating that "common carriage regimes work best for commodities for which product quality does not vary." This is correct as a variation in product quality can be interpreted as a variation in the nature of the potential set of services provided, and therefore as a break-down of the univocal mapping between assets (infrastructures) and services.

2.1. The logical causality in common law is from the private property to the publicly exchanged service

First of all, the switch in the unit of analysis from given assets to given services implies a *logical* inversion from the private domain as unit of analysis and starting point, to the public domain a unit of analysis and starting point. In other words, the change in the unit of analysis from a property to a service and vice versa matters in order to understand *which domain informs the other*. On the one hand, if the property is the unit of analysis, then the service becomes the direct object of the analysis, and what occurs in the public domain depends on the decisions taken in (originating from) the private domain (the locus where the property belongs to). On the other hand, if the services exchanged with the public become the unit of analysis (the logical starting point), the underlying property now becomes the object of the analysis and the outcome of the economic and social reasoning. In this second case what happens in the public domain informs the decisions taken in the private domain.

This symmetry between unit of analysis (the starting given) and object of the analysis (the direct outcome) is one of the key pillars of the distinction between the antitrust (failure-based) approach and the regulatory approach in common law. Yet, this also entails a symmetry in the domain (or locus) of the analysis: public (the exchanged or transacted services) versus private (the owned property).

The fact that the reversed causality between property and services can be reinterpreted as a reversed logical causality between the private and the public domain may seem hard to interpret at first. In truth, the public-private logical reversal conceals the essence of the distinction between the common law concept of public calling, and the much more recent concept of public service utility: public service companies are always defined by taking the public service or necessity to make available to the public as the unit of analysis (in the public domain), so to analyse the property private financiers should invest into in turn (private domain). This is exactly the opposite of the approach adopted by common law, which never focusses on a specific service, but always starts its legal enquiry from the property in the private sphere (*juris privati*) and from there it derives implications for the public sphere. This is always true in common law unless some extra-considerations or additional institutions allowing the logical inversion is defined, such as the one of status, to which we now turn.

2.2. Public status in common law allows us to reveal private identities and to invert the analytical causality backwards: from publicly traded services to privately owned properties

Secondly, the possibility to invert the unit of analysis (the given input) with the object of the analysis (the generated output) has one fundamental *analytical* advantage: it allows us to overcome the problems posed by the presence of private identities and private activities, and it allows us to retrieve the latter by merely looking at the public identity, or activities, of any individual or firms. Stated differently, the reason why the imposition of a public status is so important from an analytical point of view derives from the fact that, once that an actor is endowed with public privileges and constrained by public duties, his hidden (private) subjectivity is (forcefully) *revealed* to the public. And this derives from the fact that social duties and social obligations define *what ought to be done*, or what *it is expected* to be done.

For instance, a common business is completely defined by the specific service that it provides to the public and this means that every single actor dealing with that specific business already knows in advance how the latter ought to behave, and which service it ought to provide. The same holds when it comes to the profession of a certain person.

Social revelation of private identities or activities is the fundamental mechanism on which contemporary (normative) microeconomics is based (Samuelson, 1938a, 1938b, 1947). The mechanism of revelation has precisely this meaning: it allows an analytical inversion from the outputs (public services) to the inputs (private properties) and vice versa, generating a univocal mapping between the two. This mechanism is also called the duality assumption in economics, where the duality is, precisely, between inputs and outputs. The duality assumption, or revelation mechanism, is precisely what guarantees that markets where *services are publicly exchanged* always provide the necessary and sufficient signals to infer normative conclusions on how underlying *private properties and assets* (resources) ought to be employed in a social system.

Yet, the above discussion shows that this is precisely the aim of the common law institution of public calling, and to guarantee revelation and duality is precisely what the status conferred by the public calling does: it constrains the private sphere by super-imposing social duties on the way in which a private property *ought to* be employed, so that the observation of the (revealed) services provided to the public is enough to understand and infer the private identity of the businessman (firm), and the nature of the underlying properties or assets employed.

This leads to three important concluding remarks on the topic.

2.2.1. *The reverse causality between services and assets can only be legitimate if some form of social regulatory institution norming actors' behaviours is already in place*

One, every time that the analytical flow goes from the services performed to the inputs used, this necessarily means that some social institution is in place that guarantees that this reverse causality is legitimate and can be implemented. This can be the status of a public calling, characterised by the social duties superimposed on the "common business", or it can be the institution of *juris publici*, as described above, which deals with the public control of service pricing.

2.2.2. *A univocal social revelation of private identities can only occur if any property owner is identified by their public profession (what he must do for living)*

Two, by looking at the meaning of a public calling, it is also possible to perceive a clear parallelism between the role of the public duties, and the role of the (Walrasian) vNAD constraints on actors' behaviours discussed above. As a result, it is possible to understand what really ensures the duality between inputs and outputs in microeconomics, so that normative conclusions can be merely based on the services exchanged, rather than on the characteristics of the various resources owned: the fact that production functions (regardless of whether they are firms or single businessmen) are defined in function of a given "*profession*". The latter is defined according to the specific (given) service rendered to the public or, alternatively, according to the "business", or the industry in which the "profession" operates. As already asserted, to define an economic actor in function of its *social status*, thus in function of his *profession*, is already a super-imposition of a social identity over a private identity. Normally, it would not be possible to infer what the "real" underlying private identity of an actor would be by simply looking at a service publicly exchanged, for the simple reason that it is not possible to know whether this service is provided in an impromptu and occasional manner, or if it is provided regularly as the purposeful way to make a living.

The above considerations closely retrace the essence of new-new-welfare economics (Reiter, 1977), which acknowledges the fact that under information asymmetries the private identities of actors are unknown and cannot be "revealed" by a transaction.[149] On the other hand, it also retraces the managerial distinction between industry and a

[149] This is also the essence of incentives-based regulation (Laffont and Tirole, 1993).

resource-based analysis of firms' behaviours (Barney and Clark, 2007; Conner, 1991; Dierickx and Cool, 1989; Prahalad and Hamel, 1990; Wernerfelt, 1984), the former focussing on the services a firm sells, the latter on the properties it operates, or on the competences detained.

In truth, the problem deriving from the fact that the observed public transaction cannot univocally reveal the private identity of the property owner was already understood (and solved) many centuries ago through the institution of a public calling. In medieval times it was already perfectly understood that it would be incorrect to define the owner of a house as an "innkeeper" only because he occasionally decides to host people impromptu in a private fashion. The institution of public calling was developed precisely in order to solve this specific problem: the private identity of a property owner can be neglected by establishing public duties imposing that "*to host people*" is the specific business or activity that that specific house owner must undertake to make a living, so that the service provided to the public can now fully characterise the identity of the private owner, and therefore also judge him based on his behaviour (make him liable).

Thus, the simple fact that actors in microeconomic analysis are defined based on what they (are expected to) do vis-à-vis the public (their public identity), rather than on what they originally own (independently and regardless of what they eventually decide to do: their original private identity) is already symptomatic of the fact that the vNAD model is *not* institution-free and does not occur in an institutional void. Some hidden and not declared regulatory institution is always super-imposed on the exchange mechanism in order to make sure that certain actors must necessarily undertake a certain business in order to make a living, and must univocally identify with a certain profession, industry or business (all service-based) *independently from their privately owned property*. This is indeed the case as section 5 will discuss.

2.2.3. A univocal social revelation of private identities implies a switch from the subjective opportunity costs sustained in order to operate a given property, to the objective production costs sustained to provide a given service

The third and last remark is a further development of the previous one, and helps to clarify it.

It has been asserted that what distinguishes a private from a public calling is the fact that only the latter can be defined and identified based on the specific business (in the case of a firm), or profession (in the case of a single actor) operated. This means that the "cost" experienced by a public calling in operating a certain business, or profession, will always depend on the specific activity or service rendered, and will always be independent from the private identity of the actor. In other words, the "cost" sustained by a public calling will always be objectively defined by its public status, and will never depend on the private subjectivity of the actor the "common status" applies to.

This represents a further analytical benefit of imposing social status on actors. It not only allows us to revert the logical flow of the analysis, from the public to the private sphere, and to reveal the private identity of an actor based on what the interactions that he establishes with the public, independently from its original initial conditions, it also allows us to overcome the subjectivity of costs, and to make them objective, by linking the concept of cost not to the private identity of the property owner (which would lead to private and subjective costs), but to the resulting service or activity provided to the public. The fact that costs are not associated to the identity of the "private calling" anymore, but to the specific profession identified by the "public calling" make them objective, rather than private, thus subjective (as they would originally be).

Thus, in the same way in which the status is a social category that is defined based on what someone does, rather than owns, the "economic cost" deriving from the exercise of a certain status should be interpreted as the "*cost in providing a certain service*", or as the costs of operating a certain business, and not as the "*cost in operating a certain property*". This is the difference that distinguishes the cost sustained in providing a certain output (service), as opposed to the cost sustained in operating and employing a certain input (asset).

This third considerations brings the analysis back to the distinction between (objective) production costs and (subjective) opportunity costs, developed in chapter 4. While the cost that applies to a private calling (juris privati) can be defined as the *subjective opportunity costs sustained in order to operate a certain given property*, the cost that applies to a public calling can be defined as the *objective production costs sustained in order to provide a certain service*. While the concept of "opportunity cost" can only be defined with respect to the employment of a given resource's constrained production capacity (in input), the concept of "production cost" can only be defined with respect to the provision of a given service (in output).

It should be remarked that, consistent with the previous discussion, the latter are public, and objective costs as they apply to services that are only defined in the public sphere, while the former are private and subjective costs, as they apply to the operation of assets that can only be defined in the private sphere, and across subjective alternatives (given by the options of the opportunity-cost schedule).

In conclusion, the "social revelation" of one's private identity corresponds to the "objectification" of one's subjective costs, and into the transformation of private costs into social costs. This is precisely the essence of the Coase theorem, already discussed above (Stigler, 1966: 113): "The Coase theorem thus asserts that under perfect competition private and social costs will be equal".

The problem with the Coase theorem is that it neglects the fundamental fact that the univocal identification of an actor's identity through a specific public profession, business or industry is already itself the outcome of a social regulatory superstructure, norming how actors ought to behave, which services ought to be provided and how, and therefore which costs ought to be sustained.

2.3. Temporal causality and problems deriving from the ex-ante vs. ex-post reversibility

The third consideration of the distinction between status in tort (common) law and contracts concerns the opposite flow in the *temporal* causality of events.

As already mentioned, tort law is defined by the fact that the compensation of individuals occurs "for losses which they have suffered in respect of all their legally recognised interests, rather than one interest only" (Prosser, 1975: 6).

In tort law, when duties derive from a certain legal status, the obligations of an actor derive from, and apply to, their own public identity. Contracts do not define public identities, but rather contingent relationships with respect to one specific undertaking, independently from the real identity of the counterpart. This difference is also reflected in the two underlying reasons according to which a certain actor can be sued: for absolute liability (in case of public identity or status) or for negligence (private identity or contract). The former is generic and absolute (common), while the second is contingent and relational (specific).

Another way to put the distinction between contingent negligence and absolute tort is to note that while in tort law the status of a person confers rights and duties *ex-ante*, irrespective of the specific undertaking or relationship, in contract law, contracts only define rights and duties in function of a specific and contingent undertaking, and *purposefully for that specific undertaking*. Contracts differ from status as they occur *ex-post* a decision has been taken, and in function of a specific transaction. In other words, the status is generic, and comes *first* and preliminary to any economic decision or transaction, contracts are contingent, thus come *after* and in function of a specific economic decision or transaction.

The definition of a public identity through status reverses the temporal causality of events.

It should be noted that this reversed causality derives from the fact that a contract applies to a specific transaction, thus it regulates the performance of a specific *service*. From there, the specific private identity of the undertaker must be retrieved, together with their private duties. The realm of contract law is therefore service-based, and for this reason the temporal causality of the events proceeds backwards, consistent with the logical and analytical directionality of any service-based approach: from the specific activity to be exchanged (the service or business) to the underlying private identity of the actors involved. Conversely, tort law follows the other temporal succession of events: the specific services to be exchanged derive (logically and temporally) from the identity of the parties, which are established *ex-ante*.

The flow of temporal causality becomes key in contractarian (failure-based) approaches to public callings and especially the distinction between *ex-ante* and *ex-post* in economic events.

2.3.1. Temporal causality: some considerations on the role of legal status vis-à-vis bilateral contracts

A key aspect of any contractarian approach is the "breach of a contract", occurring whenever the ex-ante terms of bilateral contacts are not respected ex-post by the transacting parties. Contractual breaches are paradigmatic of all failures based on uncertainty and asymmetric information and are the key failures in the modern economics of contracts (Bajari and Tadelis, 2001; Hart and Moore, 1999; Hart and Tirole, 1988; Maskin and Moore, 1999; Maskin and Tirole, 1999; Segal, 1999; Tirole, 1999). Yet, two points are worth mentioning here.

First, any breach of a contract is service-specific, and this means that it only becomes meaningful if the specific obligations and duties of a party are defined in function of the given service to be performed. There is no breach of a contract if there is no specific service to perform which is first defined in advance. This problem that closely retraces the problem arising when "incidental services" are introduced next to "primary" carrier services, has plagued the regulatory debate in the last twenty years (see Sidak and Spulber, 1996; 1997). The rise of new, unregulated, "incidental services" next to the legacy, regulated, "carrier services" in the telecommunication industry has challenged the meaning of a breach of a contract as regulators have experienced through the three computer enquires since the 1970s (see section 1.3 above).

Second, any ex-post breach of the contract clearly highlights the frictions emerging between the private sphere of the property owner, and the public sphere of the transacted services. The standard failures deriving from the ex-post breach of a contract arise from the fact that the *temporal* reversal of the flow of events (from the specific public service to the private identity of the actor) is not backed by a contextual logical and analytical inversion of the events that ensures that the private identity of the transacting parties really matches, and is consistent with, the specific contractual relationship established. Every time that the *directionality* of the events is reverted, so that we go from the publicly exchanged service to the private identity of the property owners, it is first necessary to make sure that the subjective dimensions of the private identity of the counterparts (in terms of costs and values) are first replaced by their social "duals" through the imposition of public duties, and the contextual recognition of a public status. If this is not the case, contractual failures deriving from the inconsistency of the framework will necessarily arise

The exact aim of a public calling was to ensure that private (bilateral) exchanges could always be backed by public institutions that could ensure and legitimise the causal inversion of the logical and temporal causality was exactly. The definition of a public status makes sure that it is always possible to causally and temporally revert the logical flow of events, so that the *ex-post* performance of a service can always be traced back to the specific expectations generated *ex-ante* with respect to the way in which a certain actor *ought to* behave.

Not surprisingly, the original problems deriving from the breach of private (bilateral and contingent) promises is what originally fostered the development of public status in the form of a public calling first, and the common contract law in turn (Ames, 1888; Simpson, 1975).

2.3.2. *Status and contracts in the deregulation from common carriers to contract carriers: the common law roots of Williamson's TCE*

In a market-failure (contractarian) explanation on the one hand a private calling is associated to a situation of bilateral contract (leading to specific relationships, hold-ups, or specific assets, in Williamson's terminology), while a public calling is associated to a situation of general (or generic) contracts. As explained by the current literature in the economics of contracts, in the first case the identity of the counterpart matters, while the second case characterizes a situation where the identities of the transacting parties do not matter. This is how the private-public dichotomy in status is interpreted in contractarian terms, when assets can only be used "publicly" as a "common business", and the choice is only among alternative *compensated* transactions.

Not only the whole economic literature on incentives and contracts developed form this peculiar reinterpretation of the public-private dichotomy, but also the legal literature adopted this "contractarian" interpretative approach to distinguish public from private callings: one of the first deregulatory acts was the *Above 890* decision of the FCC (taken in 1959), where a new type of *private carrier* was created beside the standard institution of *common carrier*. The decision was soon followed, in 1971, by another very similar decision taken by the FCC, the *Specialised Common Carrier* decision, where the previous notion of private carrier was expanded and extended into a new entity: the *specialised common carrier*. By the early 1970s, three different public status could be identified: private, specialised and common carriage.

The legal rationale that distinguishes common carriers from private and specialised carriers perfectly reflects the "contractarian" interpretation of the distinction between public and private calling: while common carriers can only set up general contracts in order to serve the general public with no distinction nor discrimination among the various transacting parties, nor among the various services provided (in conformity with the "duty to serve" contradistinguishing all public callings), specialised carriers can set up specific contracts with specific clients in order to serve specialised or, specific, services. While the identity of the transacting parties in the case of "common business" (i.e. common carriage) does not matter, the identity of the transacting parties in the case of "private and specialised business" does matter. It is not a case that specialised carriers are also referred to as "contract carrier" in order to distinguish them from common carriers, which cannot sign specific contracts as they are obliged to respect the duty to serve of public callings.

The "contractarian" re-interpretation of the public-private dichotomy in legal jurisprudence precedes the re-interpretation provided by the theoretical literature in economics. It is not a case that Williamson's distinction of generic (common) assets from specific assets (which were only defined in the late 1970s: Williamson, 1979) closely retraces the FCC legal distinction between, respectively, common carriers and specialised carriers. There is little doubt that TCE has been heavily inspired by the distinction operated by the FCC more than a decade earlier, and that it just reconceptualised the contractarian dichotomy between *common and specialised carriers* (i.e. common and private callings) into the new contractarian dichotomy between *generic and specific assets*, leading to a new taxonomy of contractual relationships (from the so-called “classical contract” for “common business”, or public callings, to neoclassical or relational contracts in the case of “specialised carriers”, all the way down to internal governance for extreme forms of “specialised businesses”, such as private callings, or carrier).

In sum, starting from the late 1950s, the public-private dichotomy started to be reinterpreted in the following (contractarian) way: while *common* carriers serve the general public under generic contractual terms according to their tort law status, private or specialised carriers can set up specific relationships with selected customers through the stipulation of a customer-specific contract tailored on the specific identity of the counterpart, and protected not by tort law anymore, but by contract law. Williamson’s theory of the firm (TCE), relying on the interaction between the nature of the contracts and the nature of the property (asset), eventually built on that. It should be noted that this reinterpretation gained ground much before the deregulatory process of the early 1980s, which was mainly focussed on the role of monopoly and competitive failures, rather than on the role of contractual or transaction failures.

Yet, as already mentioned, the contractarian reinterpretation of the distinction between public (common) and private is simply incorrect and misleading. The only appropriate way to interpret the distinction between public and private is to frame it in value terms, by retaining the classical duality of value. Few authors have underlined this conceptual flaw of the contemporary legal and economic literature, Cherry (1999; 2006, 2008, 2015) being one of the very few notable exceptions. Yet, no author has tried to make sense of this inconsistency by proving an alternative interpretation to the contractarian one.

Moreover, it should be noted that two of TCE’s key concepts, opportunism and hold-up can simply be interpreted as a reversal of the causality flow: while causality originally goes from the private sphere of the actor (asset appropriation) to the public sphere of the transaction of a service, the effect of opportunism is to reverse the original causality flow backward, thus from the specific service that should have been transacted to the private (subjective) spheres of the transacting parties.

Yet, this break-down of the contractual terms (negligence) occurs only because no ex-ante social duty has been established so that no subjective identity can be “enacted” ex-post.

3. Concluding remarks

The aim of the present chapter was to show that, as long as the public regulation of private assets is interpreted through the lenses of a monistic framework, the real meaning of the two public calling duties cannot be properly understood.

The interpretation of the duty to "serve with due care" and of the duty to "hold out to the public" in a monistic framework would lead to the wrong conclusion that the two public duties have been imposed in function of the competitive conditions of the markets in which the various carriers compete, or in function of the different natures of the various transactions undertaken by the carriers.

The first approach reflects a standard monopolistic explanation of the common carrier status. This was the approach adopted by the Federal Communication Commission in 1980, when it distinguished between dominant and non-dominant carriers. The second approach reflects a contractarian approach to the common carrier status. This approach was adopted by the Federal Communication Commission in the 1960s, and it led to the distinction between "general purpose" carriers (standard common carriage), and specialised or contract carriers. While the first approach was clearly affected by the developments in antitrust law, the second inspired the rise of a contractarian approach to public regulation that first emerged in the 1970s (Goldberg, 1976, Williamson, 1976).

Both approaches are, however, market-failure interpretations of the common carriage duties.

CHAPTER 8

Public callings, private ownership and institutions: why the regulatory approach can never be replaced by a pure (failure-based) antitrust approach

The argument developed in the previous chapters can be summarised as in figure 8.1.

The various dichotomies explored throughout the work highlight a fracture between two specular approaches that, for convenience, we have relabelled "antitrust" and regulatory". The two approaches represent the duality of any socio-economic problem.

The important point to highlight is that the two sides of this duality correspond to the two legitimate approaches in the case of the two social domains: the private and the public domain, where both are defined looking at asset ownership, thus given a certain distribution of properties among actors.

The way in which figure 8.1 shall be interpreted is the following one.

The dualism between the regulatory and the antirust approach retraces the way in which value theory is modelled. The temporal, logical and analytical causal flow between the two determines which one can be determined in function of the other, and therefore which side causally determines which.

Consistent with what has been argued in the previous chapters, the natural causality flow is from the private domain where physical properties are appropriated, to the public domain where services are publicly transacted. Any inversion in the logical, temporal and analytical causality between the two requires the definition of some legal (institutional) framework that ensures that the reversed path, from public services to private properties can be legitimately undertaken. This institutional remedies can be either statuses, or contracts, or public duties, and so forth.

If no legal institution is first devised, it is not legitimate to retrieve and extrapolate any social conclusion by merely observing the way in which services are publicly transacted:

some form of legal regulation is always necessary in order to move from the private to the public domain *given private properties* (*juris privati*).

The antitrust approach runs backward: from the public to the private domains. What does this mean? There are two ways to interpret this reversed causality between the two spheres depending on the way in which the term “public” is interpreted in function of value.

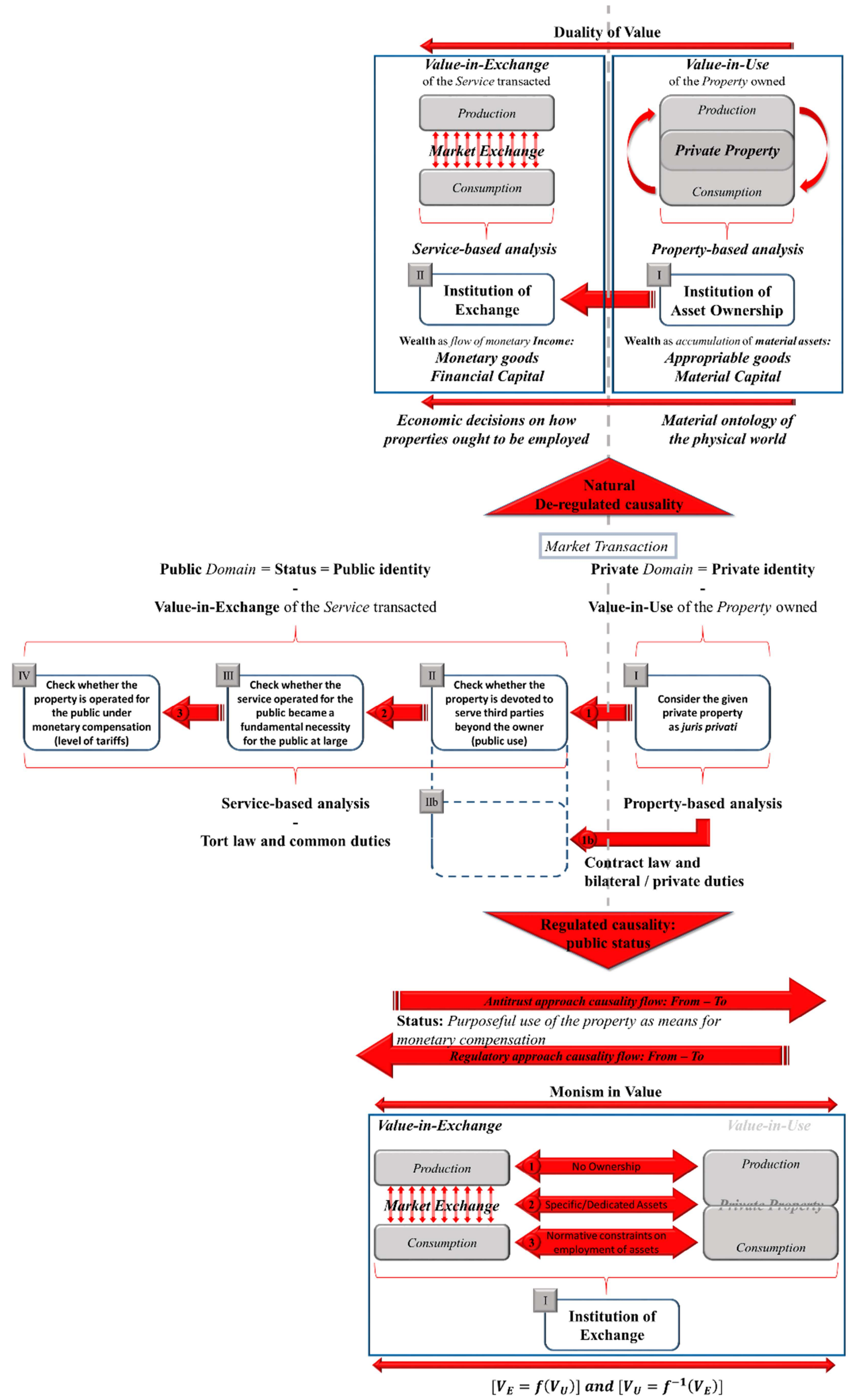

Figure 8. 1 - Overall recap: Private domain and public domain in the dichotomy between Private property and publicly transacted services under monism and dualism in value

1. The original meaning of the antitrust approach according to a pure failure-based interpretation

The first way to interpret the essence of the antitrust approach is the mainstream "neoclassical", or failure-based approach: given the social (monetary) transactions of the various actors purchasing and remunerating the exchanged services, if markets do not fail it is always possible to retrieve the private identities of the various actors, and their subjective characteristics. If there are failures, then social legal institutions are needed in order to operate this backward process. Yet, the essence of the antitrust approach, in line with any market failure approach, is that if markets do not fail in transacting and pricing services, the observation of the social interactions will always be sufficient in order to retrieve the private sphere.

This is the standard way in which the reverted causality of the antitrust approach should be approached. The well-known implication of this approach is that, if there are no failures whatever, the private sphere becomes irrelevant and, as the private domain is defined by the private possession of properties, in an antitrust approach the subjective identities of the actors operating the assets become irrelevant together with the institution of asset ownership (together with all other legal institutions and rights).

It should be noted that, while in the regulatory approach some legal institution is always needed in order to super-impose the public dimension on the private dimension, this does not occur in the antitrust approach. In this second case, if failures are absent the public dimension can always rule over the private dimension with no need for any "extra-help" from any other (legal) institution beside the market (pricing) mechanism. This is consistent with the failure-based nature of the antitrust approach.

Thus, given the presence of privately owned assets, and given the social exchange of services among actors through market transactions according to a failure-based approach, the essence of the difference between the two sides of the dual problem (regulatory and antitrust approaches) can be summarised as follows.

In the *regulatory approach,* some public legal institution is *always* needed (no matter what) in order to decide how the given (privately owned) properties ought to be employed by their respective owners, in order to achieve a given desired outcome in the public transaction of services among actors. Note that this is in line with the essence of the regulatory approach as whenever the starting point (unit) of the analysis is a *given* allocative outcome of services, in order to go from the publicly exchanged services to the employment of privately owned properties, some social legal institution is always needed.

In the *antitrust approach*, if the mechanism allocating and pricing the various services is working "failure-less", then some public legal institution is *never* needed to decide how the given (privately owned) properties ought to be employed by their respective owners in order to achieve a given desired efficient outcome in the public transaction of services among actors. This means that if there are no failures it is always possible to retrieve the private dimension of asset ownership given the desired public exchange and allocation of services no matter what. In this case, an antitrust approach must be necessarily turned into a regulatory approach if and only if some failure is present.

The meaning of the term "public" in the two cases differs. In the regulatory case, the public dimension is always represented by the legal framework of the social system super-imposed on the behaviour of property owners. It is this legal superstructure that ensures that the backward reversal from the domain where services are transacted to the domain where properties owned can be legitimately undertaken. In the antitrust approach, the nature of the public dimension varies depending on whether failures are present. In the ideal case of no failures, "public" does not define the legal framework, but rather the bare exchange (market) mechanism. The latter remains the necessary and sufficient institution that ensures the reversal from the publicly transacted services to the privately owned properties.

2. Fundamental inconsistencies inherent in the pure antitrust interpretation of relationship between the public and the private sphere

The present section will show that the pure (failure-based) version of the antitrust-approach, where the social sphere can be super-imposed over the private sphere by solely relying on perfectly efficient (and cost-less) markets is an illusion based on a misunderstanding of the assumptions behind contemporary neoclassical economics. The key (misinterpreted) concept is value.

A pure antitrust approach can hold if, and only if, subjectivity in actors' valuations on how their own private properties *ought to be employed* are ruled out *by assumption*.

In order to understand what subjective valuations are, and where they come from, it is sufficient to go back to the essence of the legal status of *juris privati* in the "rights of things", and its tort law counterpart of *private calling*.

2.1. The meaning and the nature of compensated transactions: integrability, duality, and public calling duties

Private clings and public callings are distinguished by the presence of a "dividing line" separating the private sphere from the public sphere. This dividing line represented in figure 8.1 above represents the presence of a market transaction *compensating* the property owner for service rendered to the *public*.

It is for this reason that the term "common" in tort law is simply used as a mere synonym of the term "business": the exclusive purpose of any business is to *serve the public* in exchange for some monetary *compensation*. On the other hand, it should be remembered that the presence of some monetary compensation for the use of a private property is necessary in order to potentially trigger "affectation with public interest" for the regulation of tariffs. More precisely, the presence of some monetary transaction compensating for the enjoyment of any service rendered by any privately owned property is also the first necessary condition that has to hold for a property to "cease to be *juris privati* only", and to become *juris publici* or, "affected with public interest".

This means that, as already recalled, *juris privati* legally characterises a property that is used for private purposes. This is the case where no monetary compensation occurs as, obviously, no property owner compensates himself for using his own property. More generally, this legal status applies to all those cases where the property can be employed by persons that do not need to remunerate the property owner such as, for instance, the case where a property is employed by the owner "*for his own use or the use of his family*", as already recognised by Lord Hale.

This means that the private employment of a property (auto-consumption) corresponds to an "uncompensated consumption" of a service. Consistent, subjective valuations in the employment of a property are generated every time that properties are employed in-house for the owners' own purposes and needs, consistent with the reinterpretation of the classical dual nature of value, and Adam Smith's dynamic model of the division of labour.

Yet, so-called "uncompensated transactions" (such as externalities or bilateral exchanges compromised by information asymmetries between the two parties) always generate failures in the allocation (or transaction) process. Nearly all types of failures (externalities, public goods, information asymmetries, adaptations to unforeseen circumstances under uncertainty) can be reinterpreted as "uncompensated transactions" or, in other words, as "*uncompensated services' allocation*". Any non-compensated transaction ultimately undermines the legitimacy of the normative conclusions based on the pure market exchanges and monetary transactions. This is why, if an antitrust approach is adopted, some extra-legal framework is always needed in order to guarantee the switch from the public to the private sphere whenever uncompensated allocations occur or, in other words, whenever markets fail.

In conclusion, subjective valuations on how a property ought to be employed will always be present as long as uncompensated "private employment" of the property in auto-consumption for self-subsistence is allowed.

Yet, subjective valuations can be ruled out by superimposing very precise social constraints on actors' behaviours with respect to the way in which they ought to use their property. These social constraints are, precisely, the ones of a public calling:

1. Any property owner has to publicly ex-ante identify himself with a specific business that can only serve a specific service (or set of services), and
2. Any property owner is not only bound to always and exclusively "hold his property out to the public", but he is also obliged to "make his living" always and exclusively by selling the services of his own property to the public, and therefore he always has to serve any member of the public requiring services prior to monetary compensation, no matter what.

These two conditions ensures that the dividing line between private and public is always crossed on the left, as the private employment of a property necessarily becomes a public one (figure 8.2). Yet, as soon as we move from the private domain to the public domain, subjective valuations emerging from the self-employment of a property disappear: subjective value-in-use conflates into value-in-exchange and value collapses to a monism.

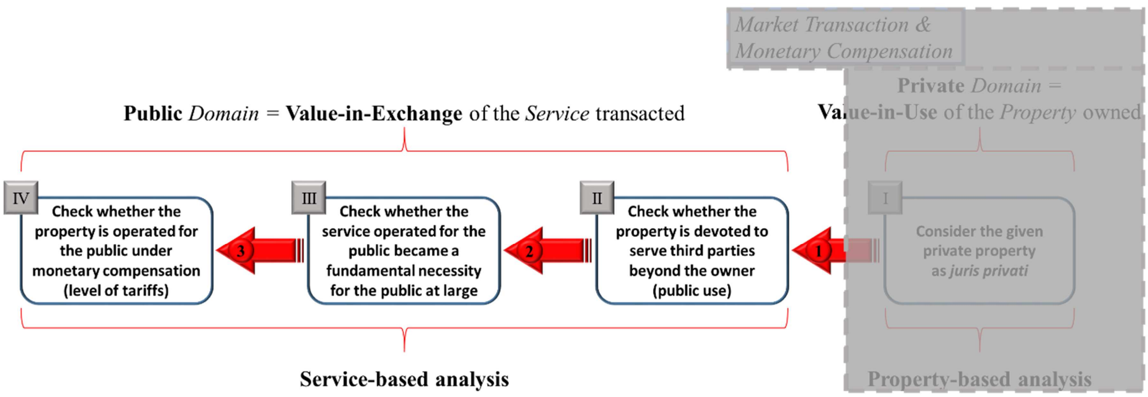

Figure 8. 2 - Whenever the private domain is ruled out by assumption, the duality of value reduces to a monism where only value in exchange is present

It was already extensively discussed how these two social constraints represent, respectively, the duty to serve with due care, and the duty to hold out to the public indifferently.

What has not been remarked upon is that these two social constraints also correspond to the two basic normative behavioural assumptions of microeconomic theory: *convexity* and *non-station (monotonicity)*. The former makes sure that actors do not change their respective social identities by changing the service operated, and therefore the market served. The latter makes sure that all wealth privately possessed by actors is employed in "exchange", to serve the public, against monetary compensation, and never internally, "in-house".

More generally, the *normative* content of contemporary institutional analysis, which guarantees the switch from the public to the private domain granted that allocative (pricing) failures are absent, derives from the settlement of the integrability problem (which occurred in a preliminary way in 1950, and in a definitive way in the early 1970s).[150]

The integrability problem can be alternatively reinterpreted in three completely equivalent ways:[151]

- As Samuelson's Weak Axiom of Revealed Preferences (WARP) and Houthakker's Strong Axiom of Revealed Preferences (SARP) if the focus is on the nature of the adaptation (or re-allocation) process;
- As the von Neumann-Dantzing-Kantorovic-Koopmans' *duality condition* if the focus is on the relationship between the capacity constraint and the maximization problem;

[150] Hands (2006) is the sole author I am aware of that explicitly attempted a critical reconstruction of the integrability debate in institutional terms along similar lines.

[151] The standard reference on the integrability debate is Samuelson (1950). One of the clearest treatments of the topic remains, in my opinion, Wold's (1944: 109-117). Alternatively, the most comprehensive treatment is provided in Chipman et al. (1971), where the chapters of Hurwicz (1971), Hurwicz and Uzawa (1971) and Chipman (1971) are worth of note. For an historical treatment of the integrability problem, see Mirowski (1989), and the response articles of Carlson (1997) and Hands (1993). Hands (2006) represents the most exhaustive historical summary on the topic. Concerning the topic of duality in linear programming, the reader is referred to the standard textbook of Dorfman et al. (1958), to the path-breaking monograph edited by Koopmans (1951), and to the other, more affordable, classic treatment of Koopmans (1957). The more technical topic of duality in microeconomic analysis is treated by Cornes (1992), Diewert (1982) and Krishna and Sonneschein (1990). A treatment of duality specific for production theory is provided by McFadden (1978). Afriat (1980) is the most exhaustive treatment of integrability and duality applied to consumer theory and demand functions. With respect to rationality, see Debreu (1954) and Arrow (1959). For a general, comprehensive treatment, see Mas-Colell et al. (1995: chapters 2 and 3).

- As Arrow's *rationality conditions* defined as reflexivity, completeness and transitivity in economic choices if the focus is on the way in which actors make their personal valuations concerning the way in which properties ought to be allocated.

No matter which specific interpretative approach is adopted, the essence of all three frameworks is the same: they ensure that the two "social duties" of "public identity" and "mandatory holding out" strictly hold. What changes, among the three, is the type of approach adopted.

First, Samuelson's *WARP and SARP* ensures that an actor's private identity is revealed to the public by means of monetary transactions, thus making sure that an actor's monetary compensations can, alone, replace one's private identity with his public (revealed) one. Second, the von Neumann-Dantzing-Kantorovic-Koopmans *duality condition* ensures that property owners will only employ their properties so that the services transacted and purchased in the market can, alone, reveal to the public their private valuation frameworks with respect to the ways in which their properties ought to be used. Three, Arrow's rationality conditions ensures that the publicly revealed identity of the property owners is not perturbed by changes in the patters in which they decide to employ their own properties.

As a result, the normative assumptions established by the contemporary ordinalist framework in microeconomics fulfil precisely the same function that the two duties of a public calling perform in common law: they associate each actor to one specific and socially recognised profession (duty to serve properly given a specific service), and they ensure that each actor is completely identified with that profession, meaning that he can only sell the services of his property according to the precepts of this profession for living (mandatory duty to serve the entire public generally).

The above discussion can be summarised in one very simple way:

> *The normative assumptions behind contemporary welfare economics describe a world where only public callings operate, while private callings are necessarily absent*.

As a result, the necessary conclusion is that:

> *In order derive normative conclusions on how (non-specific) privately owned properties ought to be employed across alternative services, it is first necessary to*

superimpose public calling duties on these properties, and to regulate their operations.

Once public calling duties are imposed, we trace the social system back to the framework illustrated by normative welfare economics (vNAD) and, starting from there, an efficiency-based analysis based on competitive failures is legitimate and can be undertaken.

This explains why in this specific case, a failure-based approach to ownership and institutions finally makes sense, and why the absence of failures becomes, alone, a sufficient condition in order to invert the logical causality backwards, from the publicly transacted services to the privately owned properties.

This also explains why asset ownership is irrelevant in a failure-based approach:[152] very simply, properties are never treated as *juris privati*, and this rules private sphere out by assumption. It should not be surprising that, when only public callings can operate, private ownership can obviously be ignored: the role of a public calling is precisely to replace an actor's private identity with a generic public status. This, however, becomes a tautology, rather than an interesting and meaningful result as it would lead to the following conclusion: given that private properties are always publicly regulated as *juris publici*, then private property becomes irrelevant if market failures are absent. And this is obviously true for the simple reason that there was no *juris privati* in the first place.

It is important to remark that, if assets are service-specific, meaning that infrastructure can only provide one single service, the failure-based approach is perfectly legitimate, and consistent with the common law approach.

What is however most important is yet another point, that brings the discussion back to where we left it: in the same way in which a failure-based interpretation of institutions is founded on a logical paradox, the antitrust approach to (de)-regulation is also contradictory, and as such has to be rejected.

In sum, to assume that asset ownership and (legal) institutions do not matter as long as all property owners behave according to socially imposed duties constraining their behaviour is inconsistent. In fact, the presence of socially recognised duties constraining the way in which property owners can employ their own assets already implies the presence of some regulatory (extra-market) institution in the first place.

[152] As summarised by Laffont and Tirole (1993: 642): "the ownership structure does not matter if complete contracts can be written. Therefore, if we want to distinguish between public enterprises and regulated firms, we must point at some contract incompleteness."

2.2. Convex and "rational" employment of private properties: the necessary role of public duties

This same observation could also be reached through other paths.

First, the same conclusions can be appreciated by simply noting that actors in microeconomics are defined by well-established and exogenously given "primitives" (consumption and production functions). However, a primitive in the theory of production defines the "identity" of the actor, or the "nature of the business" performed in order to "make a living". In other words, primitives in microeconomics define actors' public identities or their public status.

Actors' "primitives" in microeconomics must have very precise properties: continuous, convex, and continuously differentiable of class 2. Primitives reflecting these characteristics are also called "smooth" (Debreu, 1972). "Smooth primitives" are also necessarily: (i) rational, they also necessarily (ii) reveal private identities to the public, and (iii) they can also always be inverted into a dual. In other words, "smooth curves" satisfy the integrability conditions, allowing for a constant inversion of the causality flow from the public domain of services' allocation to the private domain of property employment.

The problem with this interpretation of failures is apparent: it does not explain why primitives must be "smooth", and what specific conditions ensure that "smoothness" is the desirable shape of the curves. The easy answer is that actors endowed with the public status of a public calling employ their own properties in a way that is consistent with the shape of "smooth primitives".

The typical failure-based approach that conceives failures as a deviation from the "smoothness" shape of curves does not realise that this specific shape of the curves is itself the result of pre-existing super-imposed social duties on actors' behaviours, constraining the way in which they ought to employ their own property.

As a result, to assert, as any antitrust-approach does, that no legal (extra-market) institution is necessary in order to retrieve private identities from the observation of social allocation of services, *as long as "primitives are smooth"*, is a contradiction in terms: the smoothness of the curves itself is the most visible proof of the fact that some pre-existent legal framework is already in place before and beyond the allocative (pricing) mechanism, telling actors how they ought to employ their properties, and how they ought to interact with the general public

In sum, the presence of “smooth curves” is already the signal that actors behave consistent to a very precise public role, by respecting the duties and obligations inherent in their public status.

From the previous remarks an important caveat follows.

Above it was mentioned that whenever curves are not smooth, they can be at least “artificially rationalised” through some social intervention. The concept of rationality is then clearly linked to the shape of the curves, and therefore to the social obligations and duties super-imposed on actors, so that they can behave as if they were following a “smooth curve”. From these considerations, it is possible to appreciate the real meaning of the term “rationality” in normative welfare economics in light of the nature of the duties of a public calling: actors are considered to act “*rationally*” as long as they employ their own properties accordingly with the public expectations generated by their public identity, and therefore, as long as they act in line with the duties attached to their public status.

If the two duties of a public calling are reconsidered, and if the meaning of the normative assumptions in microeconomics are re-examined, this assertion simply means that actors act rationally as long as they always employ their owned property (representing their endowed “wealth” in welfare economics): (i) in line with the service (set or services) that has been ex-ante decided to be offered to the general public, and (ii) and never for their own private self-subsistence, or for the private purposes of their own family or closed circle of relatives, but rather always in order to make a profit out of it, so to receive monetary compensation for the “holding out” of the property to the general public subject to a market tariff.

This is the actual meaning of the term “rationality” that, of course, only makes sense if a preliminary social legal framework establishes with clarity and precision how a certain property *ought to* be employed by its respective owner. A behaviour can only be rational in function of a given public status, but there can be no rationality in the private domain, where there is no super-imposed constrain on how privately owned properties ought to be employed.

2.3. Social costs, private costs and the fundamental inconsistency of the “Coase theorem”.

Second, the same observations could have already been reached if the Coase theorem is analysed: the retrieval of private costs from social costs can only be legitimate if public identities are super-imposed on private actors. Yet, to say that legal institutions do not matter in a world where actors have publicly revealed identities exclusively defined by their own profession (business operated, or service rendered to the public) as Coase does through his 1960 article, is contradictory as only the legal setting can ensure that a random person performing a certain service can actually be defined and socially identified in function of that specific service. Only in that case, can a private citizen indeed be recognised as "a surgeon", and a certain company offering to move and transport goods is clearly recognisable as a (common) carrier.

Conversely, social costs and private costs do not need to be identical for private callings, if actors can decide to operate resources for their own self-subsistence, while trading services occasionally, whenever they decide to "hold their property out" to the public in the specific way they *contingently* decide to. In this case, actors' private identities cannot be defined by the impromptu service performed under private request.

This was perfectly understood by legal scholars and practitioners in the 14th century. The definition of public calling, and its clear distinction from the private calling in function of the liability of the person publicly performing a certain service, was precisely the way in which medieval tort law managed to solve the (unsolvable) issues deriving from the duality of value, and the discrepancy between the private sphere of asset ownership, and the public sphere where services are transacted (and the underlying assets or resources allocated).

For this reason, although it is unfortunate to see that an analytical problem that was already very well understood in medieval tort law was forgotten and completely overlooked by modern commentators, this should not be too surprising as the essence of medieval common cannot be appreciated as long as the dichotomy between private (*juris privati*) and public (*juris publici*) is not reinterpreted in value terms, in function of the dichotomy between *private value-in-use* and *social value-in-exchange*.

3. Misplaced faith in the competitive panacea: any social mechanism willing to prescribe how juris privati ought to be employed must necessarily adopt a regulatory approach even if no failure undermines the competitive process

In conclusion, no private callings (defined by the fact that actors can deliberately decide whether to undertake a certain service impromptu or to just employ their own properties for their own private purposes) are present in the Arrow-Debreu model, which is therefore much more a framework describing a situation of perfect regulation, rather than a situation of perfectly decentralised private callings.

This conclusion should not be too surprising considering that the mathematical normative assumptions constraining actors’ behaviours have been clearly developed through the decades with the explicit intent to demonstrate that subjectivity in actors’ private valuations could be overcome, so that central planning and public control could be (at least theoretically) implemented. The willingness to demonstrate that normative conclusions on how *socially available* resources ought to be employed and allocated could be derived by merely looking at the social sphere of services’ exchange, while neglecting the private sphere of asset ownership (and treat it as irrelevant) has generated the illusion that what was framed was in effect a system characterised by *juris privati* and *private callings*.

In truth, nothing can be further away from the truth: a failure-based world, where only the mere performance of the competitive market can be used in order to derive normative conclusions on how resources ought to be employed, corresponds to a world where all asset owners are publicly regulated, certainly not to a world where public regulation is absent. This generated the illusions that it is possible to regulate private properties by simply deregulating, which is clearly an unsustainable conclusion. Some form of legal superstructure must necessarily always be there in order to operate an inversion of the temporal, logical and analytical causality: from the social allocation of transacted services, to the private ownership of assets and resources.

This makes the pure failure approach of any antitrust framework simply invalid, and brings all frameworks back to the scar of the regulatory approach, which can be summarised as follows:

> *It is never possible to induce how privately owned assets ought to be employed by their respective owners, unless some legal framework constraining private actors’ employment (allocation) of their properties is introduced.*

Or, in other words:

> *It is never possible to derive any normative conclusion on how privately owned assets ought to be employed unless they are either service-specific, or subject to public callings duties*

If and only if the above conditions hold, then it is possible to invert the causal flow, and to go from the public dimension of publicly allocated services, to the private dimension of privately owned properties. Yet, some non-market (legal) institution is always preliminary needed to make sure that those duties are defined, applied, and enforced.

References

Adler, E.A. (1914),'Business Jurisprudence', *Harvard Law Review,* 28*(2):135-162.*

Afriat, S. N. (1980), *Demand Functions and Slutsky Matrix*, Princeton, Princeton University Press.

Alchian, A. A. (1968), 'Cost' in D. Sills (Ed.), *International Encyclopedia of Social Science*, Vol. 3, pp: 404- 415.

Alchian, A. A. (1977), *Some Economics of Property Rights,* in A.A. Alchian (ed.), Economic Forces at Work, pp: 127-150.

Alchian, A.A. and Demsetz, H (1972), 'Production, Information Costs, and Economic Organization', *American Economic Review*, 62(5): 777-795.

Allen, D. (2000), 'Transaction Cost', in B. Bouckaert and G. Geest (eds.), *Encyclopedia of Law and Economics*, Volume 1, Cheltenham: Edward Elgar, pp. 893–926.

Allen, R.G.D. (1932), 'The foundation of a Mathematical Theory of Exchange', *Economica,* 36: 197-226.

Alter, M. (1982), 'Carl Menger and *Homo Oeconomicus*: Some Thoughts on Austrian Theory and Methodology'*,* Journal of Economic Issues, 16(1):149-160.

Alter, M. (1990),'What do we know about Menger?' in Caldwell (Ed.) *Carl Menger and his legacy in economics*, pp: 313-348.

Ames, J.B. (1888), **'**The History of Assumpsit I & II', *Harvard Law Review*, 2(1): 1-19 & 2(2): 53-69.

Armstrong, M. (2002), 'The Theory of Access Pricing and Interconnection', in M. Cave, S.K. Mjumdar and I. Vogelsang (Eds.) *Handbook of Telecommunication Economics* (Volume 1), Amsterdam: Elsevier, pp: 295-384.

Armstrong, M, Cowan, S. and Vickers, J. (1994), *Regulatory Reform: Economic Analysis and British Experience*, Cambridge (Mass.): MIT Press.

Arrow, K. (1959), 'Rational Choice Functions and Orderings', *Economica,* 26(102): 121-127.

Arrow, K. (1969), 'The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Non-market Allocation', *Joint Economic Committee of Congress 1969.*

Arrow, K. and Debreu, G. (1954), 'Existence of an Equilibrium for a Competitive Economy', *Econometrica,* 22(3): 265-290*.*

Arrow, K. and Hahn F.H. (1971), *General Competitive Analysis*, Amsterdam: North-Holland.

Arterburn, N.F. (1927), 'The origin and first test of public callings', *University of Pennsylvania Law Review and American Law Register,* 75,(5): 411-428.

Aspromourgos, T. (2009), *The Science of Wealth, Adam Smith and the Framing of Political Economy*, New York: Routledge.

Bain, J.S. (1956), *Barriers to New Competition*, Cambridge: Harvard University Press.

Bain, J.S. (1959), *Industrial Organization*, New York: John Wiley e Sons.

Bajari, P. and Tadelis, S. (2001), 'Incentives versus Transaction Costs: A Theory of Procurement Contracts', *RAND Journal of Economics*, 32(3): 387–407.

Barney, J. and Clark, D. (2007), *Resource-Based Theory*, Oxford: Oxford University Press.

Barone, E. (1896), 'Studi sulla distribuzione: I. La prima approssimazione sintetica', *Giornale degli Economisti*, 12: 107-155.

Barzel,Y. (1997), *Economic Analysis of Property Rights* (2nd ed.), New York: Cambridge University Press.

Barzel, Y. and Kochin, A. (1992), 'Ronald Coase on the Nature of Social Cost as a Key to the Problem of the Firm'*, Scandinavian Journal of Economics*, 94(1): 19-31.

Basedow, J. (1983), 'Common carriers: continuity and disintegration in United States transportation law: Part 1'. *Transportation Law Journal* 13(1): 1-42.

Beale, J.H. (1897), 'The Carrier's Liability: Its History', *Harvard Law Review*, 11(3), 158-168.

Benjamin, S.M., Litchmn, D.G., Shelanski, H.A. (2001), *Telecommunications Law and Policy*, Durham: Carolina Academic Press.

Berman, P.J. (1974), 'Computer or Communications? Allocation of Functions and the Role of the federal Communications commission', *Federal Communications Bar Journal*, 27: 161-230.

Berman, P.J. Oettinger, A.G. (1975), 'Changing Functions and Facilities: The Politics of Information Resources', *Federal Communications Bar Journal*, 28: 227-273.

Bharadwaj, K. (1978), 'The subversion of classical analysis: Alfred Marshall's early writing on value', *Cambridge Journal of Economics,* 2(3): 253-271.

Blaug, M., (1958), *Ricardian Economics,* Yale University Press.

Blaug,M. (1990), 'On the Historiography of Economics', *Journal of the History of Economic Thought,* 12: 27-37.

Blaug, M. (1996), *Economic Theory in Retrospect* (5th ed.), Cambridge: Cambridge University Press.

Blevins, J. (2016), 'The FCC and the "Pre-Internet"', *Indiana Law Journal*, 91: 1309-1362.

Bonica, J. (1979), 'A Re-examination of the Computer Inquiry', *Journal of Computers, Technology and Law*, 7: 23-75.

Bogen, D.S. (1996), 'The innkeeper's Tale: The legal Development of a Public Calling', *Utah Law Review,* 51(1): 51-92.

Böhm-Bawerk, E. (1894a), 'The Ultimate Standard of Value', *Annals of the American Academy of Political and Social Science*, 5: 1-60.

Böhm-Bawerk, E. (1894b), 'One Word More on the Ultimate Standard of Value', *Economic Journal*, 4(15): 719-724.

Border,K.C. (1985), *Fixed point theorems with applications to economics and game theory*, New York: Cambridge University Press.

Bowley, M. (1972), 'The predecessors of Jevons, the revolution that wasn't ', *The Manchester School of Economics and Social Studies,* 40*(1):* 9-29.

Brennan, T. J. (1993), 'Copyright property and the right to deny', *Chicago Kent Law Review*, 68: 675-714.

Brock, G.W. (1981), *The Telecommunications Industry: The Dynamics of Market Structure,* Harvard: Harvard University Press.

Brown, S.J. and Sibley, D.S. (1986), *The theory of public utility pricing*, Cambridge: Cambridge University Press.

Buchanan, D.H. (1929), 'The Historical Approach to Rent and Price Theory', *Economica,* 26: 123-155.

Buchanan, J. M. (1969), *Cost and Choice*, Chicago: Chicago University Press.

Buchanan, J. M. (1973), 'Introduction: L.S.E. cost theory in retrospect', in James Buchanan and G.F. Thirlby (ed.), *L.S.E. Essays on Costs*, New York: New York University Press, pp: 3-16.

Buchanan, J.M. (2008), 'Opportunity Cost", in in S.N. Durlauf and L.E. Blume *The New Palgrave Dictionary of Economics* (2nd Ed.), Palgrave Macmillan, 2008.

Buchanan ,J. M. and Stubblebine, W.C. (1962), 'Externality', *Economica*, 29(116): 371-384.

Buchanan, J. M. and G. Thirlby, (1973), *L.S.E. Essays on Costs*, New York: New York University Press.

Burdick, (1911), 'The origin of the peculiar duties of public service companies', *Columbia Law Review*, 11(6): 514-764.

Burt, R.S. (1992), *Structural Holes: The Social Structure of Competition,* Cambridge: Harvard University Press*.*

Carlson, M.J. (1997), 'Mirowski's Thesis and the 'Integrability Problem' in Neoclassical Economics, *Journal of Economics Issues,* 31(3): 741-760.

Calabresi, G. (1968), 'Transaction Costs, Resource Allocation and Liability Rules-A Comment', *The Journal of Law & Economics,* 11(1):67-73.

Calabresi, G. (1961), 'Some thoughts on risk distribution and the law of torts', *Yale Law Journal,* 70(4): 499-553.

Calabresi, G. and Melamed, AD. (1972), 'Property rules, liability rules, and inalienability: one view of the cathedral', *Harvard Law Review*, 85(6): 1089-1128.

Caldwell, B. (1990), *Carl Menger and his legacy in economics,* London: Duke University Press.

Candeub, A. (2004), 'Network Interconnection and Takings*,' Syracuse Law Review*, 54: 369-433.

Cannan, E. (1903), *A History of the theories of production and distribution in English political economy from 1776 to 1848,* London: P.S King & Son.

Cannon, R. (2003), 'The legacy of the federal communications commission's computer inquires', *Federal Communications Law Journal*, 55(2): 167-206.

Carlton, D.W. (2001),'A general Analysis of Exclusionary Conduct and Refusal to Deal, Why Aspen and Kodak are Misguided', *Antitrust Law Journal*, 68(3): 659-683.

Cassells, (1935), 'A Re interpretation of Ricardo on Value', *The Quarterly journal of Economics, 49(3):518-532.*

Chamberlin, E.H. (1962 [1933]), *The Theory of Monopolistic Competition* (8th Ed.), Cambridge (Mass.): Harvard University Press.

Chandler, A.D. (1977), *The Visible Hand: The Managerial Revolution in American Business,* Cambridge: The Belknap Press of Harvard University Press.

Cheung, S.N.S. (1969), 'Transaction costs, risk aversion, and the choice of contractual arrangements', *Journal of Law and Economics,* 12: 23-42*.*

Chipman, J., Hurwicz, L., Richter, M. (1971), *Preferences, Utility, and Demand (Eds.),* New York: Harcourt Brace Jovanovich.

Chipman, J., (1971), 'Introduction to Part II', in J. Chipman, L. Hurwicz, M. Richter, (Eds.), *Preferences, Utility, and Demand*, New York: Harcourt Brace Jovanovich, p: 321-332.

Cherry, B.A. (1999), *The Crisis in Telecommunications Carrier Liability: Historical Regulatory Flaws and Recommended Reform*, Springer.

Cherry, B. A. (2006), 'Misusing network neutrality to eliminate common carriage threatens free speech and the postal system', *Northern Kentucky Law Review*, 33: 482-511.

Cherry, B.A. (2008), 'Back to the Future: How Transportation Deregulatory Policies Foreshadow Evolution of Communications Policies', *The Information Society*, 24(5):273-291.

Cherry, B.A. (2015), 'Historical Distortion: How Misuse of 'Public Utility and 'Natural Monopoly' Misdirects U.S. Telecommunications Policy Development, ITS LA Conference 2015.

Clark, J. B. (1899), *The distribution of wealth: a theory of wages, Interest and Profits*, New York: MacMillan.

Coase, R. H. (1937), 'The Nature of the Firm', *Economica*, 4(16): 386–405.

Coase, R. (1973 [1938]), ‘Business organization and the accountant’, in J. Buchanan and G.F. Thirlby (Eds.), *L.S.E. Essays on Costs*, pp: 97-132.

Coase, R.H. (1946), ‘The Marginal Cost Controversy’, *Economica*, 13(51): 169–182.

Coase, R.H. (1947), ‘The Marginal Cost Controversy: Some Further Comments’, *Economica*, 14(54): 150-153.

Coase, R.H. (1959), ‘The Federal Communications Commissions’, *Journal of Law and Economics*, 2: 1-40.

Coase, R.H. (1960), ‘The Problem of Social Cost’, *Journal of Law and Economics*, 3: 1–44.

Coase, R.H. (1964), ‘The Regulated Industries – Discussion’, *American Economic Review P&P*, 54(3): 194–197.

Coase, R.H. (1970), ‘The Theory of Public Utility Pricing and Its Application’, *The Bell Journal of Economics and Management Science,* 1(1): 113–128.

Coase, R.H. (1972a), ‘Industrial Organization: A Proposal for Research’, in V. Fuchs (Ed.), *Policy Issues and Research Opportunities in Industrial Organization*, New York: National Bureau of Economics Research, pp. 59–73.

Coase, R.H. (1972b), ‘Durability and Monopoly’, *Journal of Law and Economics*, 15(1): 143–149.

Coase, R.H. (1973 [1938]), ‘Business Administration and the Accountant’, in J. Buchanan and G. Thirlby (Eds.) *L.S.E. Essays on Costs*, New York: New York University Press, pp: 95–132.

Coase,R.H. (1982), ‘Economics at LSE in the 1930’s: A Personal View’, *Atlantic Economic Journal*, 31-34.

Coase, R.H. (1988a), *The Firm, the Market, and the Law*, Chicago: Chicago University Press.

Coase, R.H. (1988b), ‘The Nature of the Firm: Origin, Meaning, Influence’, *Journal of Law, Economics, and Organization*, 4(1): 3–47.

Coase, R.H. (1990), ‘Accounting and the Theory of the Firm’, *Journal of Accounting and Economics*, 12(1): 3–13.

Coase, R.H. (1992), ‘The Institutional Structure of Production’, *American Economic Review*, 82(4): 713–719.

Coase, R.H. (1993), 'Law and Economics at Chicago', *Journal of Law and Economics*, 36(1): 239.

Coase, R.H. (1994), *Essays on Economics and Economists*, Chicago: University of Chicago Press.

Cornes, R.H. (1992), *Duality and Modern Economics*, Cambridge (New York: Cambridge University Press.

Dantzing, G.B. (1963), *Linear Programming and Extensions*, Princeton: Princeton University Press.

Davenport, H.J. (1894), 'The formula of sacrifice', *Journal of political Economy*, 2(4): 561-573.

Deacon, D. (2015), 'Common Carrier Essentialism and the Emerging Common Law of Internet regulation', *Administrative Law Review*, 67(1): 133-183.

Debreu, G. (1954), 'Representation of a preference ordering by a numerical function', in RM Thrall, CH Coombs and RL Davis (Eds.), *Decision Processes*, p: 105-110.

Debreu, G. (1959), *Theory of Value,* Cowles Foundation for Research in Economics, New York: Wiley.

Debreu, G. (1972), 'Smooth preferences', *Econometrica,* 40(4): 603-615.

deMarchi, N.B.(1972), 'Mill and Cairnes and the Emergence of Marginalism in England', *History of Political Economy*, 4(3):344-263.

Dempsey, P.S. (1984), 'The Interstate Commerce Commission--Disintegration of an American Legal Institution', *American University Law Review*, 34(1): 1-51.

Demsetz,H. (1967), 'Toward a property rights', *American Economic Review*, 62: 347–359.

Demsetz,H. (1988), 'The theory of the firm revisited', *Journal of Law, Economics, & Organization*, 4(1): 141-161.

Diewert, W.E. (1982), 'Duality Approaches to Microeconomic theory', in K.J. Arrow and M.D. lntriligator (Ed.) *Handbook of Mathematical Economics,* New York North-Holland, Volume II, pp: 535-591.

Dobb, M. (1973), *Theories of value and distribution since Adam Smith,* London: Cambridge University Press.

Dorfman R., Samuelson, P.A., Solow R.M. (1958), *Linear Programming and Economic Analysis*, New York: Dover.

Edelberg, V., (1933), 'The Ricardian Theory of Profits', *Economica,* 39: 51-74.

Edgeworth, F.Y. (1892), 'Review of Böhm-Bawerk, Positive Theory of Capital', *Economic Journal,* 2: 328-36.

Edgeworth, F.Y. (1894), 'Professor Böhm-Bawerk on the Ultimate Standard of Value', *Economic Journal, 4(15): 518-521.*

Eggerston, T. (1990), *Economic Behavior and Institutions*, Cambridge: Cambridge University Press.

ERG (2008), '*ERG Report on guidance on the application of the 3 criteria test'*.

Fairman, C. (1953), 'The so-called Granger Cases, Lord Hale and Justice Bradley', *Stanford Law Review,* 5(4): 587-679.

Farrell, J. (1987), 'Information and the Coase theorem', *Journal of Economic Perspectives,* 1(2):113-129.

Frieden, R. M. (1981), 'The Computer Inquiries: Mapping the Communications/Information Processing Terrain,' *Federal Communications Law Journal*, 33(1): 55-115.

Frieden, R. M. (1987), 'The Third Computer Inquiry: A Deregulatory Dilemma', *Federal Communications Law Journal*, 38(3): 383-410.

Frischmann, B., and Hogendoorn, C. (2015), 'The Marginal Cost Controversy', *Journal of Economics Perspectives*, 29(1): 193-206.

Foss, K. and Foss, N. (2001), 'Assets, Attributes and Ownerships', *International Journal of the Economics of Business*, 8(1): 19-37.

Foss, K. and Foss, N. (2001), 'Understanding Opportunity Discovery and Sustainable

Advantage: the role of Transaction Costs and Property Rights', *Strategic Entrepreneurship Journal*, 2: 191-207.

Furubotn, E.G. and Richter, R. (2005), *Institutions and Economic Theory. The Contribution of the New Institutional Economics* (2nd Ed.), Ann Arbor: University of Michigan Press.

Furubotn, E.G. and Pejovich, S. (1972),' Property rights and economic theory: a survey of recent literature', *Journal of Economic Literature*, 10(4): 1137-1162.

Genesove, D. and Mullin, W.P. (1998), Testing Static Oligopoly Models: conduct and cost in the sugar industry, 1890-1914', *RAND Journal of Economics*, 29(2): 355-377.

Gibbons, R. (2005), 'Four formal(izable) theories of the firm?', *Journal of Economic Behavior & Organization*, 58-(2): 200-245.

Goldberg,V.P. (1976), 'Regulation and Administered Contracts', *The Bell Journal of Economics,* 7(2): 426-448.

Granovetter, M. (1985), 'Economic Action and Social Structure: The Problem of Embeddedness', *American Journal of Sociology*, 91(3): 481-510.

Green, D.I. (1894), 'Pain-Cost and Opportunity-Cost', *Quarterly Journal of Economics,* 8(2): 218-229

Groenewengen,P. (2003), 'English Marginalism: Jevons, Marshall, and Pigou' in J.Samuels, E Biddle, J Davis (Eds.), *A Companion to the History of Economic Thought*, pp.246-261.

Grossman, S.J. and Hart, O.D. (1986), 'The costs and benefits of ownership: A theory of vertical and lateral integration', *Journal of Political Economy,* 94(4): 691-719.

Haberler, G. (1929), 'The Theory of Comparative Cost Once More', The Quarterly Journal of Economics, 43(2): 376-381.

Haberler, G. (1936), *The Theory of International Trade: With its Applications to Commercial Policy*, London: Hodge.

Haberler, G. (1950), 'Some Problems in the Pure Theory of International Trade', *Economic Journal*, 60(238): 223–240.

Hands, W. (1993), 'More Lights on Integrability, Symmetry and Utility, as potential Energy in Mirowski's Critical History', *History of Political Economy,* 25: 118-130*.*

Hands, W. (1993), 'More Lights on Integrability, Symmetry and Utility, as potential Energy in Mirowski's Critical History', *History of Political Economy*, 25: 118-130.

Hands, W. (2006), 'Integrability, Razionalizability, and Path-Dependency in the History of Demand Theory', in Mirowski and D.Wade Hands, (Eds.), Agreement on Demand: Consumer Theory in the Twentieth Century, Duke University Press, pp: 151-185.

Hands, W. (2006), ‘Integrability, Razionalizability, and Path-Dependency in the History of Demand Theory’, in Mirowski and D.Wade Hands, (Eds.), *Agreement on Demand: Consumer Theory in the Twentieth Century*, Duke University Press, p: 151-185.

Hamilton, W.A. (1930), ‘Affectation with public interest’, *The Yale Law Journal*, 39(8):1089-1112.

Hayek, F. (1934), ‘Carl Menger’, *Economica*, 1(4): 393-420.

Hayek, F. (1935), *Collectivist Economic Planning*, London: Routledge.

Hayek, F. (1937), ‘Economics and Knowledge’, *Economica*, 4(13): 33–54.

Hayek, F. (1940), ‘Socialist Calculation: The Competitive ‘Solution’, *Economica*, 7(26): 125–149.

Hayek, F. (1945), ‘The Use of Knowledge in Society’, *American Economic Review*, 35(4): 520–530.

Hart, O. (1989), ’An Economist's Perspective on the Theory of the Firm’, *Columbia Law Review*, 89(7): 1757-1774.

Hart, O. and Moore, J. (1990), ‘Property Rights and the Nature of the Firm’, *Journal of Political Economy*, 98(6): 1119-1158.

Hart, O., and Moore, J. (1999), ‘Foundations of Incomplete Contracts’, *Review of Economic Studies*, 66(1): 115-138

Hart, O. and Tirole, J. (1988), ‘Contract Renegotiation and Coasian Dynamics’, *Review of Economic Studies*, 55(4): 509-540.

Heilbroner, R. (2008), ‘Wealth’, *The New Palgrave Dictionary of Economics*.

Hicks,J. (1932), *The theory of Wages*, MacMillan and Co.

Hicks,J. (1934), ’Leon Walras’, *Econometrica*, 2(4):338-348.

Hicks, J. (1939), ‘The Foundations of Welfare Economics’, *Economic Journal*, 49(196): 696–712.

Hicks, J. (1956), *A Revision of Demand Theory*, New York: Oxford University Press.

Hicks, J. and Weber (1973), *Carl Menger and the Austrian school of economics,* Oxford: Clarendon Press.

Hodgson, G.M. (1988), *Economics and Institutions. A Manifesto for Modern Institutional Economics*, Cambridge, Polity Press.

Hodgson, G.M. (2014), 'What is capital? Economists and sociologists have changed its meaning: should it be changed back?', *Cambridge Journal of Economics, 38 (5): 1063-1086.*

Hodgson, G.M. (2015), 'Much of the 'economics of property rights' devalues property and legal rights', *Journal of Institutional Economics*, 11(4): 683-709.

Hollander, J.H. (1904), 'The development of Ricardo's theory of value,' *The Quarterly Journal of Economics*, 18(4):455-491.

Holmes, O. W. (1879), 'Common carriers and the common law', The American law review, 13:609-631.

Horwitz, S. (2003), 'The Austrian Marginalists: Menger, Böhm-Bawerk, and Wieser', in W.

Samuels, J. Biddle and J. Davis (Eds.), *A Companion to the History of Economic Though*, Oxford: Blackwell, pp. 262–277.

Hotelling, H. (1932), 'Edgeworth's Taxation Paradox and the Nature of Demand and Supply Functions', *Journal of Political Economy,* 40(5): 577-616.

Hotelling, H. (1938), 'The General Welfare in Relation to Problems of Taxation and of Railway and Utility rates', *Econometrica*, 6(3): 242–269.

Hovenkamp, H. (1990), 'Positivism in Law and Economics', California Law Review, 78(4): 815-852.

Hovenkamp, H. (2005), 'Exclusion and the Sherman Act', *Chicago Law Review, 72(1):147-164.*

Hurwicz, L. (1971), 'On the problem of Integrability of Demand Functions', in J. Chipman, L. Hurwicz, M. Richter, M. and Sonneschien (Eds.), *Preferences, Utility, and Demand*, New York: Harcourt Brace Jovanovich, p: 174-214.

Hurwicz, L. and Uzawa, H., (1971), 'On the Integrability of Demand Functions', in J. Chipman, L. Hurwicz, M. Richter, M. and Sonneschien (Eds.), *Preferences, Utility, and Demand*, New York: Harcourt Brace Jovanovich, p:174-214.

Hutchison, T. W. (1953), *A review of economic Doctrines 1870-1929*, London: ClarendonPress.

Jeremy, H. (1816), *The Law of Carriers, Inn-Keepers, Warehousemen, and Other Depositories of Goods for Hire*, London: I. Riley.

Jaffé, W. (1976), 'Menger, Jevons and Walras De-Homogenized', *Economic Inquiry*, 14(4): 511–524.

Jevons, W.S. (1965 [1871]), *The Theory of Political Economy* (5th Ed.), New York: Sentry Press.

Jolls, C., Sunstein, C.R., Thaler, R. (1998), 'A Behavioral Approach to Law and Economics', *Stanford Law Review*, 50(5): 1471-1550.

Kahneman, D., Knetsch, J. and Thaler, R. (1990), 'Experimental Test of the Endowment Effect and the Coase Theorem', *Journal of Political Economy*, 98(6): 1325-1348.

Kahneman, D., Knetsch, J. and Thaler, R. (1991), 'The Endowment Effect, Loss Aversion, and Status Quo Bias', *Journal of Economic Perspectives*, 5(1): 193-206.

Kaldor, N. (1939), 'Welfare Propositions of Economics and Interpersonal Comparisons of Utility', *Economic Journal*, 49(195): 549–552.

Kaldor, N. (1956), 'Alternative Theories of Distribution', The *Review of Economic Studies*, 823(2):83-100.

Kearney,J.D. and Merrill, T.W,(1998), 'The great Transformation of regulated industries law *', Columbia Law Review', 98(6):1323-1409.*

Kellog, M., Thorne, J., and Huber, P. (1992). *Federal Telecommunications Law*, Boston: Little, Brown.

Kitch, E.W. and Bowler C.A., (1979), 'The facts of Munn v. Illinois', *The Supreme Court Review*, 313-343.

Klaes, M. (2000), 'The Birth of the Concept of Transaction Cost: Issues and Controversies', *Industrial and Corporate Change*, 9(4): 567–593.

Klein, B., Crawford, R.G., Alchian, A.A. (1978), 'Vertical Integration, Appropriable Rents, and the Competitive Contracting Process', *The Journal of Law & Economics*, 21(2): 297-326.

Knight, F. (1924), 'Some Fallacies in the Interpretation of Social Cost', *Quarterly Journal of Economics,* 38(4): 582-606.

Knight, F (1928), A suggestion for simplifying the statement of the general theory of price, *Journal of Political Economy*, 36(3): 353-370.

Knight, F. (1935a), 'The Ricardian Theory of Production and Distribution I, II', *The Canadian Journal of Economics and Political Science,* 1(1): 3-25.

Knight, F. (1935b), 'The Ricardian Theory of Production and Distribution III', *The Canadian Journal of Economics and Political Science,* 1(2): 171-196.

Knight, F. (1935c), *The Ethics of Competition*, London : Allen & Unwin Knight, F. (1967a [1935]), 'Notes on Utility and Cost', in F. Knight (ed.), *The Economic Organization,* New York: Augustus M. Kelley Publishers, p: 122-149.

Knight, F. (1967b [1935]), 'Suggestion for the Reconstruction of Cost Theory', in F. Knight (ed.), *The Economic Organization,* New York: Augustus M. Kelley Publishers, p: 150-179.

Koopmans T. C. (1951), *Activity, Analysis of Production and Allocation* (Ed.), New York: J.W.& Sons.

Koopmans,T.C. (1957), *Three Essays on the State of Economic Science*, New York: McGraw-Hill.

Krishna, V., Sonneschein, H. (1990), 'Duality in Consumer Theory', in J. Chipman, D.

McFadden and M. Richter (Eds.), *Preferences, Uncertainty and Optimality*, Westview Press,p: 44-55.

Laffont, J.J. (1988), *Fundamentals of Public Economics*, Cambridge (Mass): MIT Press.

Laffont, J.J. (2000), *Incentives and Political Economy,* New York: Oxford University Press.

Laffont, J.J. (2002), 'Public economics yesterday, today and tomorrow', *Journal of Public Economics*, 86(3):327-334.

Laffont, J.J. and Martimort, D., (2001), *The Theory of Incentives: The Principal-Agent Model,* Princeton University Press.

Laffont, J.J. and Tirole (1993), *A Theory of Incentives in Procurement and Regulation*, Cambridge (Mass): MIT Press.

Laffont, J.J. and Tirole J.(1994),'Access Pricing and Competition', *European Economic Review,* 38(9):1673-1710.

Laffont, JJ. And Tirole, J. (2000), *Competition in Telecommunications*, Cambridge, Mass: MIT University Press.

Lafontaine, F. and Slade, M. (2008), 'Exclusive Contracts and Vertical Restraints: Empirical Evidence and Public Policy', in P. Buccirossi (Ed.), *Handbook of Antitrust Economics*, Cambridge, Mass: MIT University Press, pp: 391-414.

Lerner, A. (1934), 'The Concept of Monopoly and the Measurement of Monopoly Power', *Review of Economic Studies*, 1(3): 157–175.

Lerner, A. (1944), *The Economics of Control: Principles of Welfare Economics*, New York: Macmillan.

Lueck, D. and Miceli, TJ. (2007), 'Property Law', in Polinsky,M and Shavell,S. (eds.), *Handbook of Law and Economics*, Vol 1, p: 183-257.

Marx, K. (1995 [1867]), *Capital*, Volume I, Oxford: Oxford University Press.

Mas-Colell, A. (1985), *The Theory of General Economic Equilibrium. A Differentiable Approach*, Cambridge: Cambridge University Press.

Mas-Colell, A., Whinston, M. and Green, J. (1995), *Microeconomic Theory*, New York: Oxford University Press.

Maskin, E. and Moore, J. (1999), 'Implementation and Renegotiation', *Review of Economic Studies*, 66(1): 39-56

Maskin, E. and Tirole, J. (1999), 'Unforeseen Contingencies and Incomplete Contracts', *Review of Economic Studies*, 66(1): 83-114.

Mathison, S.T. and Walker, P.M. (1970), *Computers and Telecommunications Issues in Public Policy*, Englewood Cliffs: Prentice Hall.

McAllister, B.P. (1930), 'Lord Hale and Business Affected with a Public Interest', *Harvard Law Review* 43(5): 759-791.

McCulloch, J. (1977), 'The Austrian Theory of the Marginal Use and of Ordinal Marginal Utility', *Zeitschrift fur Nationalo Konomie*, 37(3): 249–280.

McFadden, D. (1978), 'Cost revenue and profit function', in M. Fuss and D. McFadden (Eds.), *Production Economics: A Dual Approach to Theory and Applications,* New York: North-Holland, Vol. 1, pp: 3-109.

Meade, J.E (1944), ‘Price and output policy of state enterprise’, *The Economic Journal,* 54:321-339.

Medema, S. (1994a), *Ronald H. Coase*, London: Mcmillan.

Medema, S. (1994b), ‘The Myth of two Coases’, *Journal of Economic Issues,* 28(1): 208–217.

Medema, S. (2011), ‘A case of Mistaken Identity: George Stigler, “The Problem of Social cost” and the Coase theorem’, *European Journal Law and Economics*, 31(1): 11–38.

Medema,S. and Zerbe (2000), ’The Coase Theorem’, in B. Bouckaert and G. Geest (Eds.), *Encyclopedia of Law and Economics*, Cheltenham: Edward Elgar, Volume 1, pp: 836–891.

Meek, R. L. (1973), *Studies in the Labor Theory of Value,* New York Monthly Review Press.

Meese, A. J. (2005), ‘Property Aspen and Refusal to deal’, *Antitrust Law Journal*, 73(1): 81-114.

Menger, C. (1981 [1871]), *Principles of Economics*, New York: NYU Press.

Mercuro, N. Medema, S.G. (1997), *Economics and the Law. From Posner to Post-Modernism*, Princeton: Princeton University Press.

Milgrom, P. and Roberts, J. (1992), *Economics, Organizations & Management*, Upper Saddle River: Prentice Hall.

Mill, J.S., (1848), *Principles of Political Economy with some of their applications to Social Philosophy*, London: John Parker.

Mirowski, P. (1989), ‘*More Heat than Light. Economics as Social Physics, Physics as Nature’s Economics’*, Cambridge: Cambridge University Press.

Mirowski, P. (2002), *‘Economics Becomes a Cyborg Science’,* Cambridge: Cambridge University Press.

Mirowski, P. and D. W. Hands (2006), *Agreement on Demand: Consumer Theory in the Twentieth Century* (Eds.), Durham: Duke University Press.

Mitchell, B. And Vogelsang, I. (1991), *Telecommunications Pricing*: Theory and Practice, Cambridge: Cambridge University Press.

Montani, G. (1987), ‘Scarcity’, In the *New Palgrave Dictionary of Economics*, vol. 4, pp: 253-254.

Moscati, I. (2007), ‘History of consumer demand theory 1871-1971: A Neo-Kantian rational reconstruction’, *The European Journal of the History of Economic Thought*, 14(1): 119-156.

Moscati, I. (2013a), ‘How Cardinal Utility Entered Economic Analysis: 1909–1944’, *European Journal of the History of Economic Thought*, 20(6): 906–939.

Moscati, I. (2013b), ‘Were Jevons, Menger, and Walras Really Cardinalists? On the Notion of Measurement in Utility Theory, Psychology, Mathematics, and Other Disciplines, 1870–1910’, *History of Political Economy*, 45(3): 373–414.

Moscati, I. (2013c), ‘Austrian Debates on Utility Measurement from Menger to Hayek’, *CESMEP Working Paper 6/2013* (November 2013): 1–42.

Motta, M. (2004), *Competition Policy. Theory and Practice*, Cambridge: Cambridge University Press.

Nachbar, T. B. (2008), ‘The Public Network’, *Common Law Conspectus*, 17: 67-115.

Nelson, R. Winter, S. (1982), *An Evolutionary Theory of economic Change*, Cambridge, Mass: Belknap Press of Harvard University Press.

Nichols, P. (1987), ‘Redefining “Common Carrier”: The FCC’s Attempt at Deregulation by Redefinition’, Duke Law Journal, 1987(3): 501-520.

von Neumann, J.V. (1945 [1933]), ‘A Model of General Economic Equilibrium‘, *Review of Economic Studies*, p: 1-9.

Noam, E.M. (1994), ‘Beyond liberalization. The impending doom of common carriage’, 18(6):435-452.

Noam, E. (1997), ‘Will Universal Service and Common Carriage Survive the Telecommunications Act of 1996?’, *Columbia Law Review*, 97(4): 955-975

O’Brien, D. (1990), ‘Lionel Robbins and the Austrian Connection’, in B. Caldwell (ed.), *Carl Menger and his Legacy in Economics*, Durham: Duke University Press, pp. 155–184.

O’Brien, D. P. (1988), ‘Classical Reassessments’, in W.Thweatt (Ed.) *Classical Political Economy,* Boston: Luwer, pp: 179-220.

O’Brien, D. (2004), *The Classical Economists Revisited*, Princeton: Princeton University Press.

O'Riordan, K.L. (1979), 'An examination of the application of common carrier regulation to entities providing new telecommunications services', *Case Western Law Review*, 29: 577-602.

Payton, S. (1981),'The Duty of a Public Utility to Serve in the Presence of New Competition, in Sichel and Thomas G. Gies, (eds.), *Applications of Economic Principles in Public Utility Industries*, University of Michigan Publisher, pp:121-152.

Pasinetti, L. (1960), 'A Mathematical Formulation of the Ricardian System', *The Review of Economic Studies*, 27(2): 78-98.

Peach, T. (1993), *Interpreting Ricardo*, New York: Cambridge University Press.

Peach, T. (2009), 'Adam Smith and the Labor Theory of (Real) Value: A Reconsideration', History of Political Economy, 41(2): 383-406.

Popovsky, M. (2006), 'Defining exclusionary conduct: Section 2, the rule of reason, and the unifying principle underlying antitrust rules, *Antitrust Law Journal,* 73:435-482.

Prahalad, C ; Hamel, G. (1990), 'The Core Competence of the Corporation', *Harvard Business Review*, May/Jun 68(3): 79

Prosser, W. L. (1975), *Handbook of the Law of Torts,* (4th Ed). St. Paul, MN: West Publishing Co.

Ramsey, F. (1927), 'A Contribution to the Theory of Taxation', *Economic Journal*, 37(1): 47-61.

Rey, P. and Tirole, J. (2007), 'A Primer on Foreclosure', in M. Armstrong and R. Porter (Eds.), *Handbook of Industrial Organization*, vol.3, chapter 33, Burlington: Elsevier Science, pp: 2145-2220

Rey, P. and Vergé, T. (2008), 'Economics of Vertical Restraints', in P. Buccirossi (Ed.), *Handbook of Antitrust Economics*, Cambridge, Mass: MIT University Press, pp: 353-390.

Ricardo, D. (1911 [1821]), *The Principles of Political Economy and Taxation (3rd Ed.),* London: J.M. Dent & sons.

Riordan, M.H. (2008), 'Competitive effects of vertical integration', in in P. Buccirossi (Ed.), *Handbook of Antitrust Economics*, Cambridge, Mass.: MIT University Press, pp: 145-182.

Robinson, J. (1934), 'Euler's Theorem and the Problem of Distribution', *The economic journal,* 44(175):398-414.

Robinson, J. (1969 [1933]), *The Economics of Imperfect Competition* (2nd ed.), Edinburgh: MacMillan St Martins Press.

Robbins, L., (1930), 'On a certain ambiguity in the conception of stationary equilibrium', *Economic Journal*, 40(158): 194-214.

Robbins, L. (1933), 'Introduction', in Robbins (Ed.) *The Common sense of Political Economy by Philip Wicksteed*', London: Routledge, pp: v-xxiii.

Robbins, L. (1934), 'Remarks upon certain aspects of the theory of costs'*, The Economic Journal'*, 44:1-18.

Robbins, L. (1935 [1932]), *An Essay on the Nature and Significance of Economic Science* (2nd ed.), London: Macmillan.

Robbins, (1958), *Robert Torrens and the evolution of classical economics,* London: Millan St Martins Press.

Robbins, L., (1970), *The Evolution of Modern Economic Theory,* London: Macmillan.

Robbins, L. (1981), 'Economics and Political Economy', *The American Economic Review*, 71: 1-10.

Rossi, J.(1998), 'The common Law 'Duty to Serve'and Protection of Consumers in an Age of competitive Retail Public Utility Restructuring, *Vanderblit Law Review, 5: 1233-1321.*

Ruggles, N. (1949a), 'The Welfare Basis of the Marginal Cost Pricing Principle', *The Review of Economic Studies,* 17(1): 29-46.

Ruggles,N. (1949b), 'Recent Developments in the Theory of Marginal Cost Pricing', *The Review of Economic Studies,* 17(2): 107-126.

Rutherford, M. (1994), *Institutions in economics. The old and the new institutionalism*, New York: Cambridge University Press.

Samet, H. (1978), 'Computers and Communications: The FCC Dilemma in Determining What to Regulate', Depaul Law Review, 28: 71-103.

Samuelson, P.A. (1938a), 'A Note on the Pure Theory of Consumer's Behavior', *Economica*, 5(17): 61-71.

Samuelson, P.A. (1938b), ’The Empirical Implications of Utility Analysis*, Economica*, 6(4): 344-356.

Samuelson, P.A. (1947), *Foundations of Economic Analysis*, Cambridge, MA: Harvard University Press.

Samuelson, P.A. (1950), ‘Evaluation of Real National Income’, *Oxford Economic Papers*, 2(1):1–29.

Samuelson, P.A. (1959), ‘A Modern Treatment of the Ricardian Economy: The pricing of goods and of Labor and Land Services’, *Quarterly Journal of Economics*, 73(1): 1-35.

Samuelson, P.A. (1995), ‘Some Uneasiness with the Coase Theorem’, *Japan and the World Economy*, 7(1): 1-7.

Schefold, B. (1999), ‘Use value and the ‘commercial knowledge of commodities’, *Value, Distribution and Capital*, New York: Routledge.

Schoen, F.X. (2005),’Exclusionary conduct after Trinko’, *New York University Law Review,* 80: 1625-1664.

Schumpeter, J. (1954), *History of Economic Analysis*, New York: Oxford University Press.

Scitovsky, T. (1964), ’A survey of Some Theories of income Distribution’, *University of California and Berkeley*, Princeton University Press, 15-51.

Segal, Y. and M. Whinston (2000),’Exclusive contracts and protection of investments’, *Journal* of economics, 31(4):603-633.

Senior, N.W. (1936),’Report on the State of Agricolture’, *Quarterly Review*, 466-504.

Sidak, J.G and Spulber, D.F. (1996), ‘Deregulatory takings and breach of the regulatory contract’, *New York university Law Review*, 71: 851.

Sidak, J.G and Spulber, D.F. (1997), *Deregulatory Takings and the Regulatory Contract. The Competitive Transformation of Networks Industries in the United States*, Cambridge: Cambridge University Press.

Simpson ,A.W.(1975), *A History of the Common Law of contract. The rise of the Action of Assumpsit*, New York: Oxford University Press.

Sinha, A. (2010), *Theories of Value from Adam Smith to Piero Sraffa,* Routledge, New Delhi.

Singer,(1996), ‘No Right to Exclude: Public Accommodations and Private Property ’,*Northwestern University Law Review*, pp 1283-1480.

Skinner, A. (2008), ’Smith Adam’ (1723-1790), *The New Palagrave Dictionary of Economics,* 2nd Edition, v. 7, pp. 559-61.

Smith, A. (1991 [1776]), *The Wealth of Nations*, London : Everyman's Library

Sotiropoulos, D. (2009), ’Why John Stewart Mill should not be enlisted among neoclassical economists’, *Journal History of Economic Though,* 16(3):455-473.

Speta, J.B. (2001), ‘A common carrier approach to Internet interconnection’, *Federal Communications Law Journal, 54(2): 225-279*.

Spulber, D. (1989), *Regulation and Markets*, Cambridge (Mass): MIT Press.

Sraffa, P. (1951), *The Works and Correspondence of David Ricardo*, Cambridge: Cambridge University Press.

Starret, D.A. (1972), ‘Fundamental non convexities in the theory of externalities’, *Journal of Economic Theory*, 4(2): 180-199.

Stigler, G. (1937),’The economics of Carl Menger’, *Journal of Political Economy*, 45(2):229-250*.*

Stigler, G., (1941), *Production and Distribution Theories,* New York: Macmillan.

Stigler, G., (1951), the division of labour is limited by the extend of the market..

Stigler, G., (1952), ‘The Ricardian Theory of Value and Distribution’, *The Journal of Political Economy,* 60(3): 187-207.

Stigler, G., (1958), ‘Ricardo and the 93% Labor Theory of Value’, *The American Economic Review*, 48(3): 357-367.

Stigler, G. (1966), *The Theory of Price* (3rd ed.), New York: Macmillan.

Streissler, E. (1972), ‘To what Extent was the Austrian School Marginalist?’, *History of Political Economy*, 4(2): 426–441.

Stone, A. (1991), *Public service liberalism: Telecommunications and transitions in public policy*, Princeton, N.J: Princeton University Press

Sutton, J. (2007), 'Market Structure: Theory and Evidence', in M. Armstrong and R. Porter (Eds.), *Handbook of Industrial Organization*, vol.3, chapter 35, Burlington: Elsevier Science, pp: 2301-2368.

Swedberg, R. (2005), *The Handbook of Economic Sociology*, Princeton: Princeton University Press.

Tadelis, S. and O. Williamson (2013), 'Transaction Cost Economics', in R. Gibbons and J. Roberts (eds.), *The Handbook of Organizational Economics*, Princeton: Princeton University Press, pp. 159–189.

Thirlby, G.F. (1946), 'The Subjective Theory of Value and Accounting "Cost"', *Economica*, 13(49): 32-49.

Thirlby, G.F. (1947), 'The Marginal Cost Controversy: A Note on Mr. Coase's Model', *Economica*, 14(53): 48-53.

Thirlby, G.F. (1960), 'Economists Cost Rules and Equilibrium Theory', *Economica*, 27(106): 148-157.

Tirole, J. (1999), 'Incomplete contracts: Where do we stand?', *Econometrica*, 67(4): 741–781.

Tirole, J. (1988), *The Theory of Industrial Organization*, Cambridge (Mass.): MIT Press.

Torrens, R. (1921), 'An Essay on the External Corn Trade', London.

Train, K. (1991), *Optimal Regulation*, Cambridge (Mass.): MIT Press.

Trebing, H.M. (1971), *Essays on public utility pricing and regulation*. Institute of Public Utilities, Division of Research, Michigan State University.

Valletti, T. and Estache, (1999), 'The theory of Access pricing: an overview for infrastructure regulators', *World Bank Policy Resource working Paper 2097.*

Vickers, (2010), 'Competition Policy and Property Rights', *The Economic Journal*, 120(544):375-392.

Vives , X. (1999), *Oligopoly Pricing. Old Ideas and New Tools*, Cambridge (Mass): MIT Press.

Walras, L., (1954 [1900]), *Elements of Pure Economics* (4th and Definitive Ed.), London: Allen & Unwin.

Walker, D.A. (2003), 'Early General Equilibrium Economics: Walras, Pareto, and Cassel', in J.Samuels, E Biddle, J Davis (eds.), *A Companion to the History of Economic Thought*, pp.278-293.

Werbach, K. (2007), 'Only Connect', 22 Berkeley Tech', *Law Journal,* 1233: 1261-62.

Wernerfelt, B. (1984), 'A Resource-Based View of the Firm', *Strategic Management Journal*, 5(2): 171-180.

West, E.G. (1988),'Developments in the Literature on Adam Smith: an evaluative survey', *Classical Political Economy*, 13-44.

Whinston, M.D. (2006), *Lectures on Antitrust Economics*, Cambridge (Mass): MIT Press.

Whish, R. and Bailey, D. (2012), *Competition Law*, Oxford: Oxford University Press.

Wicksell, K. (1934), *Lectures on Political Economy (Vol.1)*, London: George Routledge and sons, ltd.

Wicksteed, PH. (1970 [1888]), *The Alphabet of Economic Science: Elements of the Theory of Value or Worth*, New York: August Kelley.

Wicksteed, P.H. (1991 [1896]), *The Co-ordination of the Laws of Distribution*, in I. Steedman, Worchester: Billing and Sons.

Wieser, F. (1930 [1889]), *Natural Value*, New York: Stechert & Co.

Williamson, O.E. (1971), 'The vertical integration of production: market failure considerations', *The American Economic Review* 61(2):112-123.

Williamson, O.E. (1974), 'The economics of antitrust: Transaction cost considerations', *University of Pennsylvania Law Review*, 122(6): 1439-1496.

Williamson, O.E. (1975), *Markets and Hierarchies,* New York: Free Press.

Williamson, O.E. (1976), 'Franchise bidding for natural monopolies-in general and with respect to CATV', *The Bell Journal of Economics*, 7(1): 73-104.

Williamson, O.E. (1979), 'Transaction-cost economics: the governance of contractual relations', *The Journal of Law & Economics,* 22(2): 233-261.

Williamson, O.E. (1980), 'The Organization of Work: A Comparative Institutional Assessment', *Journal of Economic Behaviour and Organization.*

Williamson, O.E. (1985a), *The Economic Institutions of Capitalism*, New York: Free Press.

Williamson, O.E. (1985b), 'Asset Specificity and Economic Organization', *International Journal of Economic Organization 3,* 365-378.

Williamson, O.E. (1988), 'Technology and Transaction Cost Economics: A Reply', *Journal of Economic Behavior and Organization*, 10(3): 355–363.

Williamson, O.E. (1994), 'Evaluating Coase', *Journal of Economic Perspectives*, 8(2): 201-209.

Williamson, O.E. (2000), 'The New Institutional Economics: Taking Stock, Looking Ahead', *Journal of Economic Literature*, 38(3): 595–613.

Williamson, O.E. (2010), 'Transaction Cost Economics: The Natural progression', *American Economic Review*, 100(3): 673–690.

Wiseman, J. (1953), 'Uncertainty, Costs, and Collectivist Economic Planning', *Economica,* 20(78): 118-128.

Wiseman, J. (1957), 'The theory of public utility price: an empty box', *Oxford Economic Papers*, 9(1): 56-74.

Wiseman, j. (1959), 'The theory of public utility price: A Further Note', *Oxford Economic Papers,* 11(1): 88-97.

Wold, H. (1944), 'A synthesis of pure demand analysis', *Scandinavian Actuarial Journal*, 1944(1-2): 69-120.

Wyman, B. (1904), 'The Law of the Public Callings as a solution of the Trust problem', *Harvard law review*, 17(3):156-173.

Wyman,B. (1911), *Public service corporations and all others engaged in public employment, Washinghton: Beard Books.*

Yoo, C.S. (2013), 'Is there a role for common carriage in an Internet-based world?', *Houston Law Review*, 51: 545-672.

Young, A.A. (1912), 'Jevons' "Theory of Political Economy"', *American Economic Review*, 2(3): 576-589.