ESSAYS ON GROWTH:
IMPERFECT COMPETITION, LABOUR SUPPLY AND LOCAL PUBLIC GOODS

Gilles Duranton

Submitted for the Degree of Ph.D.

London School of Economics and Political Science

University of London

1997

ABSTRACT

This dissertation is primarily concerned with two major issues: 1/ Is growth really sustainable in the "long-run"? 2/ What are the consequences for growth of imperfect competition?

Chapter 1 explores a simple model of endogenous growth in an overlapping generations framework when labour supply is made endogenous. If leisure and consumption are substitutes, the economy experiences multiple equilibrium paths. If leisure and consumption are complements, then production remains bounded, although endogenous growth is possible and socially desirable.

In chapter 2, instead of assuming that local public goods only affect the utility of consumers as is usual, we assume that they are purely productive. The implications of this assumption are analysed within standard dynamic growth models where all factors are mobile. We show that the decentralisation of the first-best is more demanding than usually.

In chapter 3, we propose a strategic model of imperfect competition with endogenous growth and endogenous market structure. Assuming increasing returns at the firm level and heterogeneity on the labour market, short-run efficiency can be maximised under monopoly. However, in the long-run competition can generate growth through a distribution effect, whereas a monopoly leads to a no-growth steady-state.

In chapter 4, the evolution of industries is viewed as a cumulative purposeful cost-reduction process subject to spillovers in a differentiated oligopoly. The long-run outcome depends primarily on spillovers. When they are weak, firms dig their niche over time and keep investing. On the contrary, if spillovers are strong and if the diffusion function of spillovers is concave, firms use ever more similar technologies. This involves less and less investment and thus a fall in the growth rate of productivity.

In the final chapter, we propose a two-sector economy where products are substitutes. The innovation function is random, product specific and the probability of success is increasing in R&D investment. The successful innovator is in a temporary monopoly which provides an incentive for R&D. When a product has a relatively lower marginal cost, the monopoly profits are larger because of its bigger market size. Consequently, R&D investment in this product increases. So does the probability of a new cost-reducing innovation. This simple feed-back effect implies a divergence in the investment and development patterns.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

First of all, this dissertation owes a lot to Charlie Bean. His careful readings of previous versions of the various chapters have been immensely helpful to transform half-formed ideas into coherent chapters. I also must thank him for his warm welcomes and his efforts to understand the ideas expressed in this dissertation, however weird they may have seemed in the beginning.

Over the years numerous discussions with many people have helped me to shape my ideas and then sharpen them. I must mention first Etienne Wasmer and Toni Haniotis for all their informal input as well as for their detailed comments on some parts of this work. Stéphane Déo also deserves a special mention for his co-autorship on Chapter 2. Danny Quah, Kevin Roberts, Jacques Thisse, Richard Freeman, Vernon Henderson, David Wildasin, Peter Funk, Hylke Vandenbussche, Jan Eeckhout, Claudio Michelacci, Joël Maurice, Olivier Jeanne, Gianmarco Ottaviano, Donata Hoesch and Adam Lubanski all provided valuable suggestions and / or detailed comments on some of the chapters.

I must also thank seminar and conference participants at LSE (many times), Toulouse, Florence, Tokyo, Bonn, Louvain, Zurich, Barcelona, Paris, etc... Eventually I am grateful to Siwan Kemp who proof read the whole dissertation.

From a more down to the earth point of view, financial support from the French Ministry of Research and Technology, ESRC, LSE (through a teaching assistantship) and the European Union (through a TMR grant is gratefully acknowledged). Tony Venables has also been instrumental in my coming to LSE. I thank the Centre for Economic Performance at LSE and CERAS in Paris for their hospitality. Finally, I am also grateful to the French Ministry of Defence which let me accomplish my military obligations as a Ph.D. student at LSE.

"Le Progrès est l'injustice que chaque génération commet à l'égard de celle qui l'a précédée", E.M. Cioran, *De l'Inconvénient d'Etre Né*, Gallimard 1977, Paris.

INTRODUCTION

This dissertation is primarily concerned with two major issues:

1/ Is growth really sustainable in the "long-run"?

2/ What are the consequences for growth of imperfect competition?

In order to motivate the chapters that follow, I will first give a short history of the theory of growth and then try to make the meaning of the two questions above more explicit.[1]

Theories of growth: a satellite overview

Discussions about economic growth are a fairly recent phenomenon. There are probably many reasons for this, be they religious or ideological. A simple answer is maybe that, prior to the industrial revolution, rates of growth were so low that progress was barely noticeable in a life-span. Traditionally, economists celebrate the role of Adam Smith (Smith [1776]) who first highlighted the division of labour as a way to increase production. However, Adam Smith did not consider explicitly long-run economic evolution and the

[1] Of course, a complete answer to those two questions (if any can be proposed now) would require much more than a couple of years of thinking and five chapters of a Ph.D. dissertation!

possibility of ever-increasing production. The next decisive step was probably accomplished by Karl Marx (Marx and Engels [1967]) who first described a process of evolution that would, in the long-run and after a capitalist and a socialist stage, reach an era where goods would be available according to one's need.

Since then, the nature of economic evolution has been at the forefront of all sorts of reflections, except maybe that of economic theorists (Schumpeter being a notable exception). All types of theories relying on wide variety of arguments were proposed. None was really accepted by economic theorists.

A major reason was probably because of the inability of most proposed theories to replicate empirical facts or because their implications were not testable. The first full-fledged "model" of growth that was logically consistent and empirically relevant for many stylised facts was developed by Solow in 1956. But here, and without diminishing the giant step forward made by Solow, one can speak only of "model" and not of "theory" since Solow aimed primarily at a description of the growth process (Solow [1957]).[2] In other words, the "neo-classical" model of growth does not explain *why* the total productivity of factor increases, but it is a very powerful tool to describe *how* the growth process takes place[3]. It proved very useful in decomposing growth into various components (capital, labour and technology) and in understanding how the evolution of technology, tastes and labour influence capital accumulation and production. However, the model (wisely?) avoids the major issue: *Why* do economies grow in the long-run?

A second reason has probably much to do (retrospectively) with the general advancement of the rest of the economic discipline. Until the late 70s, formal analysis progressed essentially by using assumptions of constant returns to scale in a static framework. These two assumptions have many methodological advantages. First, they provide a reference framework that is now well understood. Second, they may have an important explanatory power for a wide range of phenomena. Third, they avoid the contradictory implications of naive or straightforward use of increasing returns. Theoretically, we can interpret the

---

[2] If by "model" we mean an answer to the question "how" and by "theory" and answer to "why".
[3] However, the neo-classical model of growth is not comfortable with the qualitative aspects of economic evolution since it captures only 2 of the 6 broad stylised facts proposed by Kuznets [1973].

Solow model as a proof that economic growth is impossible in the long-run without increasing returns (at least asymptotically) or exogenous change. The same argument applies in spatial economics. Without increasing returns, there would not be any economic agglomeration (Starrett [1978], see Krugman [1995] for longer developments on the analogy between the two fields).

The next step towards the development of a formal theory of growth was taken by Romer [1986] and [1990] (as well as Lucas [1988] and Grossman and Helpman [1991]). There has been much debate concerning the "real" insights brought by Romer in the theory of growth. It seems to me that there are two main technical insights. The first one was to propose a simple way of solving what seemed a complex dynamic programming exercise and to show the existence and uniqueness of the solution. The second one was the development of a model where some competition was taking place in a world of increasing returns. From a modelling point of view, his 1990 model keeps all (or nearly all) the features of the Solow model, allows for unbounded growth and does not imply convergence. A simplified presentation can be the following. The consumer good can be produced using intermediate goods (blueprints) and labour. Firms operate with constant returns to scale and an increasing number of intermediate goods increases the productivity of firms. Labour can also be also to produce new blueprints. The new blueprints give rise to patents and enable the production of new intermediate goods. The two important elements are that blueprints are imperfect substitutes so that and that the production of new blueprints is a linear function of existing blueprints. Moreover, the use of the blueprint to produce final goods gives rise to a rent that accrues to the innovator. But the new blueprints can be used freely by other researchers to invent new products. Truly, one can find many predecessors to the framework proposed by Romer. It is true that his modelling bears some resemblance to what can be found in Marshall [1890], Young [1928], Domar [1947], Harrod [1948], Arrow [1962], Uzawa [1965] or Shell [1966]. What makes his research important is that his work appealed to many other researchers. Furthermore, Romer was able to spell out the implications of his models in terms of economic policy in a simple way.

Romer's two seminal contributions display with a tension between the necessity of a major market imperfection for sustained growth to be possible (i.e., increasing returns or some sort of externality) on the one hand, and a desire to stay as close as possible to an Arrow-Debreu framework on the other hand. As can be seen from numerous syntheses (Barro and Sala-i-Martin [1995] or Mankiw [1995]), most of the students of the growth process followed Romer in their approaches and used either a competitive model with an externality or a monopolistic competition model à la Dixit and Stiglitz [1977]. This has two main advantages. First it enables a better understanding of many growth models, since comparison with a well-understood benchmark can easily be made. Second, it also enables incorporation into the analysis of other real life market imperfections likely to affect the growth process (for instance when focusing on financial intermediation or international trade to name just two possible directions). The rest of this introduction discusses successively the way the new growth theory uses increasing returns and how it deals with competition. At the same time, it will also introduce the following chapters of this dissertation.

The sustainability of growth

The new "theories of growth" have not been without their critics. Empirically, cases against them have been made by Mankiw, Romer and Weil [1992] and Jones [1995a,b]. Theoretically, most critiques of endogenous growth models usually focus on the razor edge aspect of the production function of the accumulative factor. Indeed, most specifications in existing models are linear as pointed out by Solow [1994]. A usual form is:

$$\dot{H} = f(L)\overline{H}^{\delta} \text{ with } \delta = 1,$$

where H is a reproducible factor and L a non-reproducible factor. It is true that if $\delta < 1$, endogenous growth is not sustainable and if $\delta > 1$, the growth rate is explosive (which can be interpreted as another form of non-sustainability). Thus the model only "works" (i.e., gives a bounded and strictly positive rate of growth) for a set of parameters of measure zero on the technological side (i.e., $\delta = 1$).

Firstly, note that Romer [1986]'s original demonstration is more general than the linear case since it only requires $\dot{H}$ to be a concave function of H for which the ratio $\dot{H}/H$ has a strictly positive lower bound. Moreover, it is possible to propose some formulations for the accumulation equation for which an explosive growth rate is potentially possible ($\delta > 1$), but where it is ruled out by some convex adjustment costs in the production function of the final good. Parente [1994] in this respect goes even further since he proposes a model where all the knowledge already exists (and it is the only accumulative factor). In his model, technologies are indexed by their productivity parameter A. For a given technology, depending on the accumulated production, firms have a bounded level of expertise (learning-by-doing). If they want to use a more productive technology, they incur a fixed cost and a loss of expertise which increases in the ratio of productivity levels between the new technology and the old one. What is obtained in equilibrium (which can be first-best) is a situation where firms regularly upgrade their technologies by a given finite proportion. The same kind of result can be obtained also with a capital vintage model à la Chari and Hopenhayn [1991].

So the conclusion is that the assumptions concerning the accumulation equation are not as specific as they may seem at first. They can be relaxed. However, the policy implications might be quite substantially affected. In standard models, the emphasis is mainly on the under-provision of R&D in a competitive equilibrium. Knowledge is viewed as the "engine" of growth (which indeed it ultimately is). However, when using an increasing returns to scale function for the generation of knowledge (i.e., of degree more than one in the reproducible factor), adaptation costs can limit the growth rate and act as the limiting constraint on growth as in Parente [1994] and make R&D a secondary variable. This vision is supported by Jovanovic [1995] who claims that the generation of knowledge in modern economies receives around 0.5% of the resources, whereas other activities to sustain growth (education, training) require around 25% of the resources. Crafts [1995] also takes a similar view.

Besides, an overwhelming part of the attention has been given to the reproducible factors, whereas non-reproducible factors have been somewhat neglected (Hahn [1989]). The question of the sustainability of growth with respect to natural resources is not tackled here since it seems obvious that if some exhaustible factors are necessary for production, then growth is not sustainable in the long-run. Although this problem may not be theoretically very interesting, it may be empirically crucial. The focus must then be on "non-exhaustible" and non-reproducible factors, i.e., labour. Then, if we take a "long-run" view, which was the original focus of endogenous growth theory, demand factors such as labour cannot be neglected as underlined forcefully from different perspectives by Kuznets [1959], Blanchard [1994] or Fogel [1994].

A first small strand of literature has focused on the formation of the labour force, exploring the links between growth and fertility (e.g., Barro and Becker [1989] or Becker, Murphy and Tamura [1990]). The issue of labour supply has been even more neglected. Usually, in endogenous growth models, labour supply is supposed to be exogenous and constant. When made endogenous, it is under the assumption of a zero elasticity of labour supply with respect to long-run income (King, Plosser and Rebelo [1988] or Jones, Manuelli and Rossi [1994] among others). On the contrary, popular theories of growth relate economic development to the attitude towards work assuming implicitly a long-run flexibility of labour supply.

Chapter 1 therefore executes an analysis of the growth process with a non-trivial labour supply decision. The production function exhibits increasing returns so that sustained growth is possible and people live for two periods. When consumption and leisure are substitute, which is empirically relevant for low levels of incomes, two equilibria are possible. If the initial level of accumulated capital is below a given threshold, people asymptotically do not work or consume. On the contrary, if the initial level of capital is above the threshold, people asymptotically give up all their leisure for more and more consumption.

When leisure and consumption are complements, which is probably a good working assumption for developed economies (Pencavel [1986]), it is shown that the economy converges towards a no-growth steady-state. When a specific utility function is used (i.e., a CES), this result holds even if explosive growth is possible (i.e., with a production function homogenous of degree more than one in capital). This no-growth steady-state is shown to be inefficient, whereas a Pigouvian tax on investment can improve the welfare and generate growth. In effect, the growth process meets its limits on the demand; rather than supply-side.

Another important and yet not often tackled real-life change, brought about by the growth process, is an increased mobility of factors. A great deal of work has been devoted to the analysis of the locational changes induced by this increased mobility (see Fujita and Thisse [1996]). However, very few efforts have been made to assess the ability for states or more generally any kind of jurisdictions, to insure their financing when all factors become mobile. Indeed, the importance of public goods and infrastructure for production seems difficult to deny. The analysis of infrastructure in a growth framework was first undertaken by Barro [1990] and Barro and Sala-i-Martin [1992]. The production function they use at the firm level is:

$$Y = G^{\beta} k^{1-\beta} l^{\beta}$$

where G is the amount of accumulated public infrastructure. What is assumed by this production function is that public capital is a necessary input for both the production of consumer goods and that of reproducible factors (since there is only one sector in the model). What is considered in our analysis is an "open-economy" case where at the level of each jurisdiction, we get:

$$Y = G^{\beta} K^{1-\beta} L^{\beta}$$

A benevolent government then maximises the utility of the representative agent in the economy. A second-best steady-state is reached when the fiscal distortions created to finance the public goods are balanced by the efficiency gain of more infrastructure. In equilibrium, public and private capital grow at the same rate. In studies focusing only on the taxation side, King and Rebelo [1990] and Jones, Manuelli and Rossi [1993] show that the welfare loss of taxation can be much larger in a growth framework than in a standard static framework.

However, the effect of taxation on factors that are mobile between different fiscal jurisdictions is absent in their analysis. Given perfect mobility of all productive factors, will the jurisdictions still be able to provide necessary productive public goods? Consideration of this issue is the objective of Chapter 2. Technically it uses a framework close to Barro [1990] and considers the effects of tax competition. It is shown that the financing of public goods will still be possible, and indeed even optimal within some institutional settings. The idea, as in Tiebout [1956], is to rely on perfect mobility of factors and also to use the land market as a way of collecting taxes. Unfortunately, it appears that complete taxation of the differential land rent as with public consumer goods (i.e., the Henry George rule, see Vickrey [1977]) does not lead to the first-best. The residential taxes that are required are more complex (although they can be decentralised if the jurisdictions have enough degrees of liberty).

Growth and imperfect competition

What the short overview above also made clear was that the new theory of growth, for consistency reasons, must rely on market imperfections (imperfect competition and/or market incompleteness). This is generally made either in a competitive setting with externalities or via monopolistic competition and a large number of firms. Thus, the exploration of the inter-relation between market structure and growth has been rather neglected by analysts of the new growth theories. However, this inter-relation between market structure and growth has always been recognised as a key issue and the topic has received a considerable amount of attention from policy makers as well as from the informal literature. To my knowledge, the number of formal studies focusing on this issue is very small[4]. One exception is Smulders and Van de Klundert [1995]. They consider a small number of firms in a monopolistic competition framework. They can invest in R&D to improve their product. A higher market power reduces the free-rider problem arising because of R&D spillovers across firms and increases the appropriability of investments. But, the monopoly power creates its usual dead-weight loss. The problem of the model is that the market structure is fixed exogenously and can never change. Moreover, this paper does not consider strategic interactions among firms although the focus is on imperfect

[4] As stated by Stiglitz [1994]: "Rather, it is that the standard Arrow-Debreu model (the competitive paradigm) not only does not include (endogenous) changes in the technology but its framework is fundamentally inconsistent with incorporating technological change".

competition.

A second perspective is given by Schumpeterian competition. This strand of literature is based on the arguments developed by Schumpeter [1911] where the successful entrepreneur gains a competitive edge over his or her competitors and is rewarded by monopoly profits (the seminal paper combining the new growth theory with Schumpeterian ideas is Aghion and Howitt [1992]). This monopoly rent is only temporary as another entrepreneur will eventually innovate to reap the monopoly rent for himself. From our perspective, the major drawback of Aghion, Harris and Vickers [1995] or Aghion, Dewatripont and Rey [1995] is that again the market structure is ignored or modelled without strategic interactions (the focus of these two papers however is different from ours). The second problem of these neo-Schumpeterian studies is that they primarily conclude that monopolies are a strong factor of growth. This, of course, clashes with the usual intuitions of economists concerning competition (be it casual empiricism or more serious econometric studies like Nickell [1996]). It is only with major amendments that neo-Schumpeterian models favour competition (Aghion, Dewatripont and Rey [1995]).

The main reason for this difficulty of considering imperfect competition probably lies in the basic contradiction in straightforward extensions of the basic framework between the micro-modelling of the market structure and the technical requirements for growth to occur. The fundamental problem is that the accumulation equation is again of the form:

$$\dot{H} = f(L)\overline{H}$$

That is, new knowledge ($\dot{H}$) is an increasing function of non-accumulative resources (e.g., skilled labour) and a linear function of existing knowledge ($\overline{H}$). At an industry level, it seems reasonable to think that productivity gains are created within the sector (maybe by the adaptation of some knowledge generated elsewhere). Since, in many sectors, the number of competing firms is small, one can disaggregate:

$$\overline{H} = \frac{1}{n}\sum_{n} H_i$$

Then, if n is small, it seems natural to think that firms are able to internalise those dynamic gains (at least partly). This assumption is not usual in the endogenous growth framework but it is a major feature in the industrial organisation literature[5] (see Tirole [1988] or Kamien and Schwartz [1982] for related empirical evidence). The implication of this assumption in a standard endogenous growth setting nonetheless is that, any degree of dynamic appropriability, however small, implies a monopoly, at least in the sector where the reproducible factor is accumulated.

In Chapter 3, we explore the connection between market structure and growth. As in much of the existing literature, we assume increasing returns at the firm level. The overlapping-generations set-up implies that labour is a short-run factor, whose allocation is decided after that of capital. This sequencing makes it impossible for one firm to secure all the available capital when savers cannot co-ordinate. Then, due to the presence of labour market heterogeneity, firms, exploiting a differentiated technology have no incentive to hire all the labour. In other words, this imperfect factor mobility makes it impossible for a monopoly to dominate the economy when the market is sufficiently large, despite there being increasing returns at the firm level. We can then show that within a period the First Welfare Theorem does not hold because of increasing returns. Indeed, productive efficiency is maximised under monopoly. But in the long run, the monopoly generates a monopoly rent which is not re-injected into the economy and growth may not take place. By contrast, when the market structure is competitive, sustained positive growth may occur. So competition is sub-optimal, but generates growth, whereas a monopolistic market structure implies a particular Pareto-optimal dynamic path that leads to a no-growth steady-state.

In Chapter 4, we also tackle the issue of imperfect competition and growth but this time with a focus on purposeful (and incremental) innovation and spillovers. In the existing literature, it is assumed either that knowledge is purely firm- (or product-) specific (it is the case in the neo-Schumpeterian literature, see Grossman and Helpman [1991, chapter 4]) or that there exists a common body of knowledge that everyone can use. We study a differentiated oligopoly model à la Hotelling where cost-reducing investments are possible

---

[5] But the problem in this literature is that the time horizon of the model is limited.

and where the cost reductions can spillover partly to competitors. The intensity of spillovers is decreasing in the distance between firms in the product space. The location of firms is itself endogenous so that firms locate, invest and sell and then may change their location, invest again, etc...

Depending on the shape of the diffusion function for spillovers, two polar situations are possible. It is shown that products can remain highly differentiated when spillovers are low. In this case, firms keep investing at the same rate. On the contrary, for a diffusion function concave in distance and strong spillovers, products tend to acquire increasingly similar attributes until they are so close to each other that, given the spillovers, there is no more incentive to perform R&D.

Finally, chapter 5 also tackles the problem of imperfect competition and growth but this time with a focus on Schumpeterian style or drastic, instead of incremental, innovation. It is shown that Schumpeterian competition acts as a self-reinforcing mechanism. If products are substitutes, a small advantage in the beginning for one product induces more R&D invested in this product and implies asymptotically that eventually most or all the research is engaged only in one direction. Consequently, it is not so much that a monopoly is good for growth, but growth which creates monopolies by reducing diversity.

CHAPTER 1

ENDOGENOUS LABOUR SUPPLY, GROWTH AND OVERLAPPING GENERATIONS

Abstract : This chapter explores a simple model of endogenous growth in an overlapping generations framework when labour supply is made endogenous. The following results are obtained:

- If leisure and consumption are substitutes, the economy experiences multiple equilibrium paths (either high growth and high labour supply or no growth and low labour supply).
- If the demand for leisure is inelastic, then the economy enjoys steady growth as in standard models.
- If leisure and consumption are complements, then production remains bounded, although endogenous growth is possible and socially desirable.

In the light of our earlier discussion, technological and other limitations on the supply side can hardly be viewed as an important factor. [...] A long term rise in real income per capita would make leisure an increasingly preferred good as is clearly evidenced by the marked reduction in the working week in freely organized non authoritarian advanced countries. [...] The pressure on the demand side for further increase is likely to slacken. Kuznets [1959]

I. INTRODUCTION

Although modern theories of growth tend to focus primarily on technological change, human capital, knowledge and capital accumulation, many popular explanations often relate the wealth of nations to the supply of labour. These popular arguments are used to account for two different phenomena: the persistence of under-development in the Third World and the productivity slowdown in the developed countries. Some countries, supposedly, remain poor because their populations are "lazy" (*i.e.*, a level effect) and more subtly, rich countries do not grow as fast as they used to because of "declining efforts" of the labour force (*i.e.*, a growth rate effect). What credit should we give to these popular arguments? The answer on the face of it is: Not much. Of course, one can always introduce a cultural parameter to explain differences in labour supply but the causality that goes from laziness to poverty is probably at best spurious. The most obvious counter-example is that this type of view attributed the stagnation of China until 1950 to the spirit of Confucianism (refusal of innovation, laziness, etc.). Yet now the same Confucianist motives are used to explain the astounding growth in China (hard-working people, emphasis on savings, respect for authority, etc)!

However, labour, which still receives a large share of income in most countries, is subject to important time-series and cross-section variations. Historically, Blanchard [1994] underlines that before the industrial revolution people enjoyed a relatively light work load in Europe, labouring only 100-150 days a year. This pattern was spatially and temporally widespread. With a reduction of population, and a resultant rise in wages in fifteenth-century England and the Netherlands, however, they worked only some 80-100 days. Later, with a higher population depressing wage rates, the peasants were forced to deploy their labour time in commercial and industrial pursuits, and accordingly had to work harder. Modern patterns of labour and leisure emerged only with the Industrial Revolution. The new norm was set around 300 days of ten hours a year. It continuously decreased since then. One should also mention that potential labour supply has sharply increased since 1750, as underlined by Fogel [1994]. Consequently the fraction of labour effectively supplied has strongly decreased. More precisely, Maddison [1991] provides us with some data concerning labour supply in developed countries over the last century:

Annual hours worked per person	Germany	UK	USA
1870	2,941	2,984	2,964
1950	2,316	1,958	1,867
1973	1,804	1,688	1,717
1989	1,607	1,552	1,604

Table 1

Among others, this decline of working hours is also emphasised by Pencavel [1986] who provides significant additional evidence on shorter week hours, longer vacations and earlier retirement. So our first stylised fact is a steady decrease of labour supply over time as incomes grow in developed countries. Note also that the growth rate has been declining over the last 20 years, possibly as a consequence of lower working hours. One may wonder how they might evolve in the future.

If we now look at a cross-section of countries (Smith-Morris [1990]), we can see that:

Country	Working Hours in Manufacturing (per week, 1985-87)	Growth rate of GDP per capita (%, 1980-88)
South-Korea	54.0	7.9
Singapore	49.2	5.3
Malaysia	45.6	2.2
Kenya	41.0	-0.5

Table 2

Our second set of facts shows that among developing countries, those enjoying high growth also tend to work more, whereas persistent poverty is associated with a low work load. Of course, our sample is by no means large. However, it is representative of a wide pattern. How can this duality be explained? Do we need to invoke an exogenous parameter or is it the consequence of basic economic forces?

Given their emphasis on technology, most endogenous growth models assume constant and exogenous labour supply and do not offer many insights or deal with this issue.[1] The only causality runs from labour supply to output and growth through the production function. When endogenous labour supply is assumed in the growth literature, it is invariably under the hypothesis of a zero wage elasticity of leisure demand. Then, it is immediate that individual labour supply will remain constant over time, as capital (or knowledge) is accumulated, which, as we have seen, is counter-factual. Only changes in the value of technological parameters or in preferences can lead to a change in the supply of labour. However convenient it might be, this assumption is not sustainable in the long-run, when one focuses on issues dealing with growth and labour. If labour supply is endogenous there is no a priori reason why the *elasticity of labour supply with respect to the wage should precisely be zero*. This zero elasticity is only a particular case in a general model. If anything goes, according to the empirical facts given by the survey of Pencavel [1986], the income elasticity of leisure demand is positive in the long-run and negative in the short-run in developed countries. Our objective is then to study the behaviour of growth models when labour supply is made endogenous and see whether we can generate stylised facts similar to those above.

We consider an overlapping-generations (OLG) model where labour is an essential input and where, depending on the amount of labour supplied, growth can be either positive or negative. Our reasons for choosing an OLG structure are multiple. First, the Infinitely-Lived-Agents (ILA) with exponential discounting assumption does not receive much empirical support (Wilhelm [1996]). Second, ILA models with endogenous growth do not allow us to replicate some of the stylised facts described above (see the discussion below). Third, ILA models are often restricted to the analysis of balanced growth paths for heavily specified models. Here, our OLG structure enables us to obtain results for more general

[1] There exists a sizeable literature dealing with the endogenous evolution of fertility and its interactions with growth (e.g., Becker, Murphy and Tamura [1990]). We focus here on the determination of the *individual* labour supply and we abstract from the evolution of the population.

paths than balanced growth paths using more general utility functions. We are also able to derive some results for production functions that cannot be studied under the standard ILA approach.

Our results show that, when leisure and consumption are (gross) complements, we find asymptotically that endogenous *growth does not occur, although it is both possible and socially desirable*. Complementarity between leisure and consumption thus implies that people work too little and that "demand" limits to growth are met (in the sense of Kuznets [1959]). The increase in wealth induces a rise in wages. Given the complementarity assumption, workers will substitute part of their consumption for increased leisure. Working hours will then decrease until production remains constant over time. As usual in an OLG structure, welfare is not maximised in the long-run since part of the work of the current generation benefits future generations (longer working hours result in higher savings and consequently higher wages at the next period, but this is not internalised by the current generation in the absence of inter-generational altruism). In models with fixed labour supply, there is no possible Pareto improvement since higher future consumption is at the cost of a lower utility now. With variable labour supply, however, there is room for Pareto improvement. Increasing working hours can benefit all generations, including the first one. The disutility of a heavier work load can be compensated at the next period by higher returns to savings due to the harder work that will be performed by the next generation. This Pareto improvement also generates a positive growth rate. This cannot happen in a competitive equilibrium, because there is no enforceable contract that can link future generations with the current one.

This result might support the view that labour supply should be constrained directly or indirectly through a tax mechanism, because people are not working hard enough. This "rigidity" of labour supply is crucial to maintain economic expansion over time. Our argument then goes against the traditional laisser-faire argument saying that labour supply should be made as unconstrained as possible.[2] So, under complementarity of leisure and consumption, our model gives some support to the popular view on the evolution of labour

[2] Leisure is a good whose consumption obviously enters the utility function of agents. If labour supply is constrained, the equilibrium cannot be in general Pareto-optimal in any Arrow-Debreu framework. The first welfare theorem requires unconstrained labour supply. Moreover, subsidising investment as in traditional endogenous growth models would only make things worse.

supply and its consequences over time in developed countries. What we also show is that this evolution can be explained by using a simple assumption receiving wide empirical support concerning the demand for leisure.

Under substitutability, which may constitute a good working assumption for low-income countries, multiple equilibria are possible. The model converges either towards a poverty trap with no work or towards a high growth - heavy work equilibrium. The intuition for the result is the following. In the high growth case, as capital is accumulated, the successive generations experience an age of augmented expectations concerning the labour supply of the following generations. Then, the current generation works hard, since it expects that the next one will supply a higher quantity of labour and so will induce high yields for its savings. On the contrary, the poverty trap is entered after an age of diminished expectations concerning the labour supply of the following generations. Expectations of low yields induce a substitution of work for leisure. The next generation will then face very low wages and low expectations, so that it effectively has even less incentive to supply labour heavily. Thus the economy is stuck in a poverty trap. So people are not poor because they are lazy, rather they do not work because they are poor. The popular causality is thus reversed. In such a poverty trap, only strong government intervention, sacrificing the welfare of the current generations, can set the economy on a growth path again.

The rest of the chapter is organised as follows. Sections II to IV explore respectively the case of zero, positive and negative income elasticity of leisure demand in a simple general equilibrium model with endogenous growth. The last section contains some final remarks.

II. INELASTIC LABOUR SUPPLY

To introduce the notation and structure of the model, consider first a very simple succession of overlapping generations of agents in a production economy. Individuals live for two periods: they work when they are young and are retired when old. At the end of their youth, they get a wage. This wage is saved and used as productive capital by the next generation. Before dying, old individuals consume the interest and the principal of their savings. They leave no bequest and we assume that the population of each generation remains constant, normalised to one. Labour supply is endogenous and maximum labour supply is equal to one. For simplicity we restrict our attention to an environment, where the individual born in t works in period t and consumes only in period $t+1$. Of course, there is a disutility of work or, conversely, the individual is happier with increased leisure:

$$\begin{aligned} &U_t = U(l_t, c_{t+1}), \\ &U_l > 0,\ U_c > 0,\ U_{ll} < 0,\ U_{cc} < 0,\ U_{lc} > 0, \end{aligned} \tag{1}$$

with l_t being the individual leisure time ($l_t \leq 1$). There is only one good in this economy. It can be used either as an investment or as a consumption good. The production function at the firm level is standard, assumes that labour is an essential input and it is homogenous of degree one in capital (k_t) and labour:

$$y_t = A_t k_t^{\beta} (1 - l_t)^{1-\beta}. \tag{2}$$

There is an externality at the aggregate level. Following Romer [1986], we suppose that the aggregate production function shows linear returns in the accumulated factor where upper-case letters denote economy-wide aggregates:[3]

$$A_t = AK_t^{1-\beta} \quad => \quad Y_t = AK_t(1 - L_t)^{1-\beta}. \tag{3}$$

The economy is assumed to be perfectly competitive. Capital depreciates fully after one period.[4] It is accumulated according to:

$$K_{t+1} = w_t(1 - L_t). \tag{4}$$

The budget constraint is :

[3] As usual, the accumulated factor is capital in its broadest sense, be it knowledge, human or physical capital. A more complete discussion on that issue is provided further in this paper.
[4] Our results are robust to less extreme assumptions concerning depreciation. They hold as long as the depreciation rate is strictly superior to zero.

$$c_{t+1} = w_t r_{t+1}(1 - l_t). \tag{5}$$

With competitive factor markets, we obtain:

$$w_t = (1-\beta)AK_t(1-L_t)^{-\beta} \qquad \text{and} \qquad r_{t+1} = \beta A(1-L_{t+1})^{1-\beta}. \tag{6}$$

Whence: $$c_{t+1} = \frac{\beta(1-\beta)A^2 K_t (1-L_{t+1})^{1-\beta}}{(1-L_t)^{\beta}}(1-l_t). \tag{7}$$

And thus at the aggregate level:

$= w_t \beta A (1 - L$

$$C_{t+1} = Z_t(1-L_t)^{1-\beta} \qquad \text{with } Z_t = \beta(1-\beta)A^2 K_t (1-L_{t+1})^{1-\beta}. \tag{8}$$

Moreover, given (1) and (8) we can derive the existence of a simple demand function $l(Z_t)$ for aggregate leisure. We also define the "zero-growth labour supply" L^* such that $K_{t+1}/K_t = 1$, we find:

$$L^* = 1 - ((1-\beta)A)^{1/(\beta-1)}. \tag{9}$$

In this section, we also make an additional assumption concerning leisure demand: the income elasticity of leisure demand is zero. The alternative cases of positive or negative elasticities will be explored in the next two sections. Assumption (a0) is then:

(a0) Inelasticity of leisure demand: $\partial l(Z_t)/\partial Z_t = 0 \quad \forall Z_t$.

Young individuals born in t maximise their utility given the values of the different parameters and variables in period t and the expected values of the parameters and variables in period t+1. Namely, if we assume the stability of the production function and tastes, individuals optimise over l_t given L_t, K_t and $E(L_{t+1}/(L_t, K_t))$, that is the expected value of L_{t+1} given L_t and K_t. A temporary equilibrium is defined by $l_t = L_t$ and

$L_{t+1} = E\left(L_{t+1} / \left(L_t, K_t\right)\right)$. The consumer program can be written:

$$\underset{l_t}{Max}\ U\left(l_t, c_{t+1}\right), \tag{10}$$

subject to: $\left(1 - l_t\right) w_t r_{t+1} = c_{t+1}$.

Due to (a0), we get the simple solution: $l_t = L_t = L$. (11)

The specific feature of our assumption concerning the utility function (the income effect exactly offsets the substitution effect) implies that the choice of the effort today is independent of the expected returns of the savings. The dynamic behaviour of the capital stock is then almost trivial:

$$\frac{K_{t+1}}{K_t} = \left(1 - \beta\right) A \left(1 - L\right)^{1-\beta}. \tag{12}$$

For instance, we can use the following utility function as an illustrative case: $U = ln\left(l_t\right) + x.ln\left(c_{t+1}\right)$. We obtain $K_{t+1}/K_t = \left(1 - \beta\right) A \left(x/\left(1 + x\right)\right)^{1-\beta}$ and $l_t = L_t = 1/\left(1 + x\right)$. Thus, depending on the value of *x*, the economy will experience steady negative growth (low *x*), a steady-state or steady positive growth (high *x*). See Figure 1 for an illustration.

As a conclusion, we can see that the inelasticity-of-labour-supply assumption is very convenient because it allows us to neglect future events and because it implies a constant share of leisure. However this simplification is obtained with an important loss of generality. Nonetheless, except for Eriksson [1996] in a continuous time framework (but in his case labour supply does not affect accumulation) or Hahn [1989] for neo-classical growth models, papers dealing with growth and endogenous labour supply usually adopt this assumption. The fact is that labour supply in these papers is just an ingredient not their specific object of study. See Benhabib and Perli [1994], Jones, Manuelli, and Rossi [1993] or King, Plosser, and Rebelo [1988] for examples of such treatments.

III. SUBSTITUTABILITY BETWEEN LEISURE AND CONSUMPTION

A first justification for this section is that substitutability between leisure and consumption should be considered as a theoretical possibility. Moreover, it is empirically relevant for low levels of income as assumed by traditional labour supply curves.[5] We make the following assumptions:

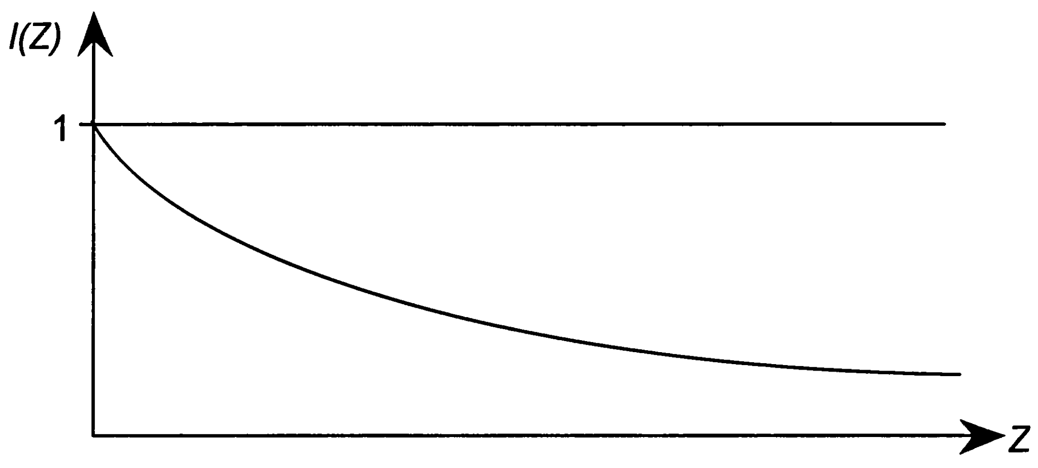
(a1) Substitutability: $\partial l(Z_t)/\partial Z_t < 0$.

(a2) Convexity: $\partial^2 l(Z_t)/\partial Z_t^2 > 0$.

(a3) Continuity: $l(Z)$ is continuous.

(a4) Boundary conditions: $\lim_{Z \to 0} l(Z) = 1$ and $\lim_{Z \to +\infty} l(Z) = 0$.

Assumption (a1) is the main assumption. It states sure that an increase in wage increases the supply of labour. Assumption (a2) ensures that the labour supply curve must be concave (or convex demand for leisure). Assumption (a3) implies that the demand function is well defined. Note that the boundary conditions in (a4) could be relaxed. We just need labour supply to be able to create positive and negative growth. The use of 0 and 1 will just make the proofs easier. All these assumptions can be summarised on a simple graph:



Graph 1

[5] It seems also that, empirically in the "short-term", leisure and consumption are substitute (Pencavel [1986]). Yet this short-term assumption is apparently contradicting our OLG structure. We can nonetheless put forward the following argument: by the word generation we do not mean literally half a lifetime but the horizon of predictable future. Then "generations" are assumed to last around 5 years. And 5 years can be set as a limit for predictable future, as well as for the time during which consumption and leisure are substitute. The OLG structure can then be used as a modelling device capturing the idea of myopic agents who do not consider their future beyond the horizon of predictable future.

Using (6), we can write:

(P1) $\partial L_{t+1}\left(L_t,\ K_t\right)/\partial L_t > 0$ and $\partial L_{t+1}\left(L_t,\ K_t\right)/\partial K_t > 0$.

(P2) $\partial^2 L_{t+1}\left(L_t,\ K_t\right)/\partial L_t^2 < 0$.

(P3) $L_{t+1}\left(1,\ K_t\right) = 1$ and for K_t sufficiently large: $\partial L_{t+1}\left(L_t,\ K_t\right)/\partial L_t\Big|_{L_t=1} < 1$.

(P4) $\lim\limits_{L_t \to 0} L_{t+1}\left(L_t, K_t\right) = -\infty$ and $\forall \varepsilon > 0,\ \exists\ K_t$ such as $L_{t+1}\left(\varepsilon,\ K_t\right) = 0$.

That is, there is a map $L_{t+1}\left(L_t,\ K_t\right)$ that is increasing in L_t and in K_t.[6] If the labour supplied in the next period is high, the interest rate is also high. Young individuals are then willing to substitute some of their current leisure time for more consumption tomorrow. For a high enough K, due to (P3), we can define the fixed point $\underline{L}\left(K_t\right)$ of $L_{t+1}\left(L_t,\ K_t\right)$ such that $\underline{L}\left(K_t\right) = L_{t+1}\left(\underline{L}\left(K_t\right), K_t\right)$ $\left(\text{and } L_t \neq 1\right)$. Note that in general this fixed point is not a rational expectation equilibrium given that it varies with K. We also define $L_0\left(K_t\right)$ such that $L_{t+1}\left(L_0\left(K_t\right), K_t\right) = 0$. Figure 2 gives a simple diagrammatic illustration of some typical situations. In short $\underline{L}\left(K_t\right)$ is the demand for leisure in t such that the same demand is expected in t+1. $L_0\left(K_t\right)$ is the demand for leisure if zero leisure demand is expected in t+1. Of course, since $L_{t+1}\left(L_t,\ K_t\right)$ is an expected leisure demand, it can be negative for some values of L (in that case, it means simply that the expectations are not consistent with the rational expectation assumption).

In other words and more formally, property (P1) establishes substitutability, whereas properties (P2)-(P4) ensure the existence and uniqueness of $\underline{L}\left(K_t\right)$ for K sufficiently large and that $L_{t+1}\left(L_t,\ K_t\right)$ is well-behaved to avoid trivial issues (or intractability if labour supply is not concave). For example the CES utility function

[6] We analyse the dynamic mapping $L_{t+1}\left(L_t, K_t\right)$ instead of the more traditional $L_t\left(L_{t+1}, K_t\right)$ since many of the proofs are not backward looking but forward looking whereby the equilibrium path is derived by contradiction.

$U_t = l_t^{1-\sigma}/(1-\sigma) + x\,c_{t+1}^{1-\sigma}/(1-\sigma)$ with $x > 0$ and $0 < \sigma < 1$ satisfies assumptions (a1) to (a4) (and consequently (P1)-(P4)).

The interesting feature when leisure demand is elastic is that the dynamic behaviour of the economy is determined by both its past behaviour and its expected future behaviour. The equilibrium l_t (= L_t) depends on L_{t+1} and K_t. In turn L_{t+1} is determined by L_{t+2}. Then L_t depends on the whole sequence of $\{L_{t+1},\ L_{t+2}\ldots\}$. On the contrary K_t is a state variable, resulting of all past actions and initial conditions. Since the relation is strongly non-linear, there is no hope of generating a complete analytical solution of our problem. Nonetheless some results can be derived:

Proposition 1 *Under (a1)- (a4), there exist a trivial steady-state and a non-trivial one.*

Proof If the economy expects $L_{t+1} = 1$, (P3) implies that no labour is supplied today (*i.e.*, $L_t = 1$). Then $K_{t+1} = 0$ and the economy is trapped in this equilibrium forever. This is an equilibrium since this is self-fulfilling. We call this equilibrium the trivial steady-state. We can also define K^* such that $L_{t+1}\left(L^*, K^*\right) = L^*$. The existence and uniqueness of K^* directly stem from the intermediate value theorem with (a1)-(a4). We can check easily that $\left(L^*, K^*\right)$ is a steady-state since $K_{t+1}\left(L^*, K^*\right) = K^*$ and since demanding L^* is self-fulfilling. According to the definition of L^*, no other level of leisure demand can imply a steady-state. Any other level of capital except for K^*, makes the demand L^* inconsistent with rational expectations. ■

Proposition 2 *Under (a1)-(a4), there exists at least one equilibrium path with sustained growth for a large enough initial K. For this equilibrium path, demand for leisure tends to zero.*

Proof in Appendix 1. ■

The argument is as follows: we can show that if L_t is in a neighbourhood left of $\underline{L}\left(K_t\right)$, the underlying expectations would not be consistent (they would imply L_t above L^* and thus negative growth in a finite time). There exists also a neighbourhood right of $L_0(K_t)$ such that, if L_t is in that neighbourhood, the underlying expectation should become inferior to $L_0(K_t)$, which is again inconsistent (expected L is negative in finite time). Since no point can belong to both neighbourhoods, there is at least one equilibrium path with sustained growth. Moreover, since L_t is right of $L_0(K_t)$ and left of $\underline{L}(K_t)$, when K becomes very large, L is driven to zero. Any equilibrium with growth then implies that leisure time should converge to zero. Individuals are asymptotically working all the time.

Proposition 3 *Under (a1)-(a4), the equilibrium path with growth is unique.*

Proof in Appendix 2. ■

This last result builds on the slope of $L_{t+1}(L_t, K_t)$. If K becomes very large, the slope of the curve becomes very steep. Then in case of two (or more) equilibrium paths with growth for a given level of capital, the amounts of leisure demanded should be diverging. This is ruled out since growth drives leisure demand to zero asymptotically.

Proposition 4 *Under (a1)-(a4), there exists at least one equilibrium path with negative growth. Negative growth is the only outcome when K is less than K^*.*

Proof The first part of the proposition is straightforward: the trivial steady-state is always possible. The second part is also quite easy to understand. In case of sustained growth, working backwards, clearly we need $L_t < L^* \ \forall t$ and $\{L_t\}$ is decreasing. Then if we reverse the time scale $\{L_{-t}\}$ is increasing and bounded above. Then it must converge toward a level lower than L^* (if it was converging towards $L > L^*$, then growth would be impossible in the first stages). From Proposition 3, the lowest level of capital consistent with that assertion $\left(\text{i.e., } \underline{L}(K_t) = L^*\right)$ is K^*. By contraposition, $K < K^*$ implies negative growth. ■

Corollary 1 *The non-trivial steady-state is unstable.*

This corollary stems directly from Propositions 1 and 4. ■

If the economy is at $\left(L^*,K^*\right)$, a small perturbation in K drives the economy out of the steady-state for ever. Consequently, the economy tends either to the trivial steady-state (which is reached in case of negative growth) or with the unique positive growth path.

The most important result of this section is that growth, although possible, is not automatically given. If the economy starts at a sufficiently low level of accumulated capital, growth will not occur and the trivial steady-state with no capital will eventually be reached. The trivial steady-state can be interpreted as a poverty trap.[7] Even above the threshold, the low level equilibrium can never be ruled out. Rational expectations can lead to multiple equilibria that can be Pareto ranked.

The only equilibrium path that involves growth also implies a sustained decrease of leisure time towards zero. The intuition of the proposition is the following: assuming a sufficiently high level of capital, the current generation works hard because it expects the next generation to work very hard as well and, in so doing, to offer the future retired people high yields for their savings. The next generation has the same beliefs. Since the amount of accumulated capital is higher, wages are higher. So the incentive to substitute leisure for work is even stronger. In that case, expectations are self-fulfilling and labour supply increases over time. Note that upper-corner solutions for labour supply are impossible because of the infinite marginal utility of leisure (as assumed implicitly with the demand function), when leisure demand is close to zero. Technically, this results simply from the substitutability assumption that implies a rise in labour supply when wages are higher.

On the contrary, expecting a low return for their savings, people will mainly demand

[7] If the depreciation rate was strictly positive or if the minimum labour supply was above zero, then the low steady-state would imply a strictly positive level of capital. Moreover it must be noted that this poverty trap does not stem from restrictions on the production function as in Azariadis and Drazen [1990] (i.e., an exogenous non-convexity in the returns of the production function.) but it results from the structure of tastes, initial conditions and expectations. Of course, the model can also be "growth-constrained" if the minimum labour supply is high enough to generate growth whatever the expectations or if the depreciation rate is zero.

leisure, so that capital at the next period will be scarce and wages will be low. This will make work unattractive for the next generation and will confirm previous expectations. The simplest case is given by the trivial equilibrium. Work is not supplied today with the expectation that it will not be supplied tomorrow. If no work is supplied, no capital is accumulated and there is effectively no point in supplying any work at the next period.[8] This co-ordination failure arises because of the OLG structure, not because of the externality in the aggregate production function which merely enables growth. In this second case, the economy is then trapped in a low production equilibrium (production is zero) and there is no possibility to escape it in a laisser-faire economy. This argument can act as a rationale for government intervention (justified only by a Utilitarian and not a Pareto criterion) in low income countries, for which the substitutability assumption might be valid. The point here is that this intervention may not create too many distortions since it suffices for the government to intervene for a finite number of periods. Once the high growth equilibrium path is triggered, no more intervention is required. The length of this period is ultimately an empirical question. The Korean and Japanese experiences in the post-war period may suggest however that it can stay reasonable.

Propositions 1 to 4 hold under well-behaved utility functions exhibiting substitutability.[9] If individuals consume when young as well as when old, results get more complicated. However it seems natural to focus on the assumption of complementarity or multiplicative separability of the two consumption levels. As a consequence, both consumption levels are "tied" in the sense that pessimistic expectations about future returns will not induce people to substitute future consumption for present consumption but consumption for leisure.[10]

[8] An interpretation in terms of product multiplicity can be also developed (with l figuring some index of consumption of goods produced with constant returns such as agricultural products or manufactured goods produced with a traditional technology).
[9] They are also robust to the formation of expectations. The rational expectation assumption might look a bit strong. If one thinks of simple possible rules of thumb, one may supply L_t such that $E\left(L_{t+1}/l_t,K_t\right)=L_t$. The individual is rational but unable to forecast the changes of economic conditions due to the accumulation of capital. With this rule of thumb, we can see that $L_t=\underline{L}\left(K_t\right)$. Then we have three different possible situations. If $K=K^*$, then the economy stays at the (unstable) steady-state. If $K>K^*$, then $L_t<L^*$. The economy keeps growing: $K_{t+1}<K_t$ and $L_{t+1}<L_t$. Capital grows unbounded, whereas leisure tends to zero. If $K<K^*$, then $L_t > L^*$. The economy experiences negative growth with labour supply and capital going to zero. Our rule of thumb then yields results with the same flavour as the model with rational expectations.
[10] The consumption during the youth is positively related to the future consumption as typically implied by the FOCs. Consequently future consumption still determines the demand for leisure.

The essence of the results then should not be modified by this refinement of the model.

Finally, note that sociological theories of growth often emphasise the importance of culture in the process of development. Cultural values are supposed to govern the attitude towards work. Here, we simply assume, that the economic returns are higher when one works more. If we accept here the idea that consumption and leisure are substitutes when people are poor, there exist multiple equilibria. Consequently, poverty may not be the result of cultural values, but a matter of self-fulfilling expectations and initial conditions. Nonetheless, sociological explanations may have a part to play. To escape the trivial (or low) steady-state, Weber [1930] emphasises some form of irrationality with respect to our specification for utility. In his introduction, he underlines that *(With Protestantism) man is dominated by the making of money, by acquisition as the ultimate purpose of his life.* [...] *Economic acquisition is no longer subordinated to man as the means for the satisfaction of material needs.*

IV. COMPLEMENTARITY BETWEEN LEISURE AND CONSUMPTION

The case for which leisure and consumption are substitutes is empirically weakly relevant for rich countries since most studies surveyed by Pencavel [1986] show a long-run complementarity between those two goods (the wage elasticity of labour supply is negative in all cases; however it is close to zero). Another reason to justify this complementarity assumption is that *time is the ultimate scarce resource as consumption goods become abundant*. Even with a longer life-expectancy, our lifetimes still remain bounded, whereas there is no reason for consumption to be bounded in the future. This view states that time is valued for its own sake. Becker [1965] takes a more materialistic approach (where time has no value as such) and underlines that utility is derived from both the amount of consumption and the time devoted to consumption with a complementarity between the two. Both approaches are equivalent in our model.

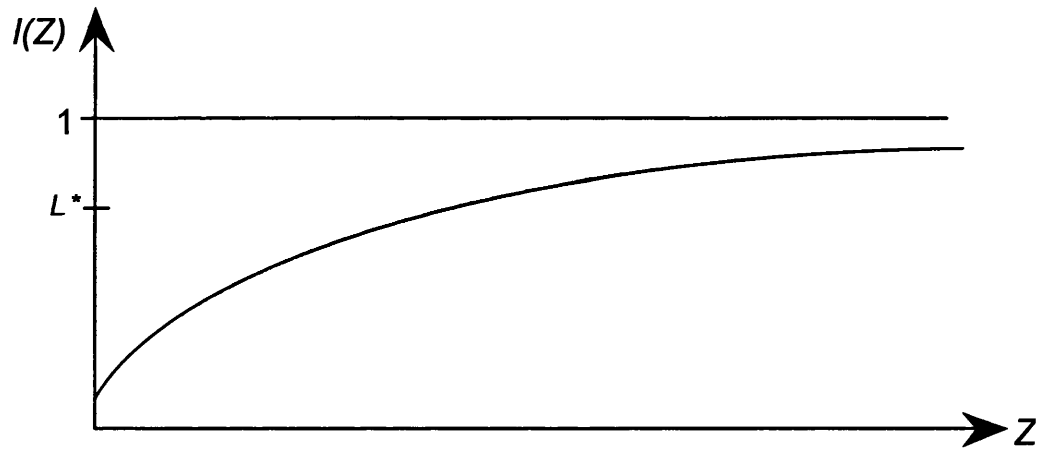
We make the following assumptions:

(a5) Complementarity: $\partial l(Z_t)/\partial Z_t > 0$.

(a6) Continuity: $l(Z)$ is continuous.

(a7) Boundary conditions: $\lim_{Z\to+\infty} l(Z) \in (L^*,\ 1)$ and $\lim_{Z\to 0} l(Z) < L^*$.

Assumption (a5) states that leisure demand is increasing with the net return to labour. Assumption (a7) is just here to ensure that leisure demand can be low enough to imply positive growth or high enough to yield negative growth. So, apart from complementarity, we just specify that demand is continuous and can change sufficiently to create positive or negative growth. As previously, these assumptions can be summarised on a simple graph:



Graph 2

There is a new map $L_{t+1}(L_t,\ K_t)$. From (a5-a7), it has the following properties:

(P5): $\partial L_{t+1}(L_t,\ K_t)/\partial L_t < 0$ and $\partial L_{t+1}(L_t,\ K_t)/\partial K_t > 0$.

(P6): $L_{t+1}(L_t,\ K_t)$ is continuous.

(P7): $L_{t+1}(0,\ K_t) \geq 1$ and $\lim_{L_t \to Max\ L} L_{t+1}(L_t, K_t) = -\infty$.

These properties ensure the existence of $\underline{L}(K)$ and $L_0(K)$ such that $\underline{L}(K) < L_0(K)$ (refer to Figure 3 for a diagrammatic exposition). For example the CES utility function

$U_t = l_t^{1-\sigma}/(1-\sigma) + x c_{t+1}^{1-\sigma}/(1-\sigma)$ with $x > 0$ and $\sigma > 1$ satisfies assumptions (a5) to (a7). Our basic result for this section is then quite easy to establish:

Proposition 5 *Under (a5)-(a7), sustained positive or negative growth are impossible.*

The proof is by contradiction: suppose positive growth is possible, then we need $L < L^*$ for capital to keep growing. Consequently K would become superior to K^* and then $L_{t+1}(L_t,\ K_t)$ would become such that $\underline{L}(K)$ and $L_0(K)$ are superior to L^*. For capital to continue to grow, we still need $L_t < L^*$. This implies that $L_{t+1} > L^*$ is expected. Then, when these rational expectations are fulfilled, it means $K_{t+2} < K_{t+1}$. Thus sustained growth is impossible and the same argument runs with negative growth. The economy cannot experience explosive cycles either, since the growth rate is bounded by $(1-\beta)A$. If we are below the threshold for which $L_0(K) < L^*$, capital increases without cycles. ■

The corollary of this result is that the economy in the long-run is either at the steady-state (L^*, K^*) or experiencing small scale fluctuations within some given bands around the steady-state.[11] In other words, the Romer assumption coupled with a positive income elasticity of leisure demand imply a Solow-type outcome. The intuition is that, when wealth increases, individuals tend to work less due to an increased demand for leisure. Working hours then decrease, until production remains constant over time. Unsurprisingly, depending on the slope of $L_{t+1}(L_t,\ K_t)$ at (L^*, K^*), the steady-state is either locally stable (if the slope is superior to -1) or unstable. Moreover our results can be extended for a larger class of production functions, which can be studied only in an OLG framework:

Proposition 5' *Sustained positive or negative growth is impossible with the production function* $y_t = AK_t^{1-\beta+\mu} k_t^{\beta}(1-l_t)^{1-\beta}$ *and with* $U_t = l_t^{1-\sigma}/(1-\sigma) + x c_{t+1}^{1-\sigma}/(1-\sigma)$ $(x > 0$ *and* $\sigma > 1)$.

---

[11] This result is robust to the formation of expectations like in the previous section. If one expects $L_{t+1}=L_t$, when $K_t>K^*$, one gets $L_{t+1}>L^*$ and hence negative growth. When $K_t<K^*$, one gets $L_{t+1}<L^*$ and hence positive growth.

Proof in Appendix 3. ■

With fixed labour supply, this production function should yield explosive rates of growth. Here, we show that the no-growth result can be extended for potentially explosive returns in the case of a CES utility function. In other words, again effort decreases until it becomes insufficient to generate growth.

A possible objection to these two results could be raised using the fact that our engine of growth is very specific. In particular, do our results hold with what is usually referred to in endogenous growth models as "accumulation of knowledge" or "technological progress" (Romer [1990] and many subsequent papers)? These more sophisticated models of endogenous growth use a constant returns to scale production function for the consumption good, along with another sector (R&D) for which the creation of new knowledge is a linear function of existing knowledge and an increasing function of labour employed in this sector. Technically, both frameworks are the same, yielding the same kind of results (see Grossman and Helpman [1994]). The only differences rest with the treatment of the competitive assumption. So our conjecture is that in such models, growth would imply a decrease of the labour supplied in the R&D sector (growth should reduce the incentive to supply labour in the manufacturing sector, which in turn reduces the value of patents). Thus this would drive growth to end. If instead of knowledge, human capital was accumulated the results should also remain valid as long as the demand for leisure increases in the income.

This stagnation due to Propositions 5 and 5' would not be very worrying if this situation was Pareto-optimal (and it would be so in an ILA framework with a CES utility function after correcting for the externality in the production function). One can easily imagine a situation with growth still being possible, but the agent preferring to work less and still enjoying a high income. But this is not the case.

Proposition 6 *When the economy is at its steady-state, it is possible to improve the welfare of all future generations including the current one. This welfare improvement implies positive growth.*

See Appendix 4 for a formal proof. ■

Then, although endogenous growth is possible and socially profitable, it does not take place. A Pareto-improvement is possible because a small increase in the labour supply of the future generations (i.e., higher consumption) compensates for a small increase in the labour supply of the current generation. Yet the current utility decreases due to more work today, but it also increases due to more consumption tomorrow (higher interest rate given by the supplementary work of the next generation).

In contrast to many endogenous growth models, the inefficiency does not come from insufficient investment in R&D or in physical capital, but from sub-optimal labour supply. Typically, in models with exogenous labour supply, the unregulated growth rate is equal to the optimal growth rate minus a constant which depends on the difference between private and social returns. Here, whatever the "potential" growth rate, our complementarity between consumption and labour drives the growth process to a complete standstill. To restore efficiency, it may be possible to tax investment or savings in order to induce workers to work more (instead of subsidising it in models with constrained labour supply). It may also be possible to directly compel people to work sufficiently so that perpetual growth should occur. This may be very hard to implement since it is possible to extend the model realistically with every generation voting on taxes or working hours. This voting process should bring us back to Proposition 5.

V. CONCLUDING REMARKS

The perspective on growth in this chapter is quite different that of the existing growth literatures, both neo-classical or endogenous. Their typical questions are: under what technological assumptions does growth happen? (or what are the engines of growth and what should we do to propel growth?). Here, we explore what are the conditions on the "demand" side under which growth occurs, given that it is technically possible on the supply side. Although barely tackled, a better understanding of the determination of "effort" seems to be a major issue in the growth and development issues. In the framework developed above, it is shown that obtaining positive growth is demanding concerning labour supply since it precludes leisure and consumption being gross complements.

This assumption, however, does make sense empirically. We face here an apparent contradiction. Although the model predicts that growth should stop, it still occurs. It might simply be that we are converging to a steady-state we have not reached yet, but that eventually demand limits to growth will appear. Yet, many other candidate explanations can be at work to explain the *ex-post* rigidity of labour supply, but most of them must be ruled out regarding our case. Traditional labour economics stresses the importance of labour externalities, indivisibilities and increasing returns to labour effectively supplied. But except from extreme forms of indivisibilities (individual returns are zero if one works less than $24\left(1-L^*\right)$ hours per day), this may not be sufficient to reverse Propositions 4 and 5.

One possible way to explain the persistence of high levels of labour supply is the existence of some "oligarchic wealth". Typically the only engine of growth usually found in economic models is greediness for consumption. Given the consumerist objective, people work, accumulate, learn or trade indefinitely to maximise their intertemporal utility assuming that wealth is "democratic", i.e., that all goods can be multiplied. For instance one may assume that total social reward within a group is fixed (that is, there can be only one winner) and that honours or positions are sought for their own sake (more generally, there are goods in limited supply whatever the state of development). Then unlike the quest for consumption, the quest for social reward (for instance being the richest man in the village, the most quoted economist, or owning a painting by Vermeer) is endless.

welfare.

Cole, Mailath and Postlewaite [1992] proposed a matching model where people are interested in physical attributes and thus have an incentive to work in order to be able to mate with the best partners. In other papers, Corneo and Jeanne [1994] or Fershtman, Murphy and Weiss [1996] introduced directly a concern for social status in the utility function. Then, the pursuit of relative utility is never exhausted since they are no decreasing marginal returns. This type of simple argument is consistent with a deeper observation of the empirical data provided by Pencavel [1986]. The average working time may be declining, but many highly qualified and wealthy people still work 60 hours a week, very often at a high cost for their family or their health. The elite may be part of a "rat race" from which growth might be only a side-effect.

As a final point, this chapter has explored the behaviour of a standard growth model with endogenous labour supply. We showed that with a negative income elasticity of labour supply, the economy converges either towards a level equilibrium or towards a high growth and high labour supply situation, thus replicating the situation in many developing countries. On the contrary, when consumption and leisure are complement, demand for leisure should increase as the economy grows richer until a no-growth steady-state is reached. Further work should focus on the generality of this result. Moreover, new arguments able to restore the possibility of positive growth should enrich our understanding of the growth process.

APPENDIX 1 PROOF OF PROPOSITION 2

Because of (P1), $\forall K > K^*$ and $L \in [\underline{L}(K), 1]$, we observe $L_{t+1} > L_t$. As long as $L \in [\underline{L}(K), L^*]$, $L_{t+1}(L_t)$ is shifting to the left since $L < L^*$. But $\{L_t\}$ is increasing and eventually becomes larger than L^*. Then capital starts to decrease and falls below K^*. When $L > L^*$ and $K < K^*$, growth is not possible anymore: to get sustained growth (*i.e.*, growth for all future periods), we need the necessary condition $L < \underline{L}(K)$ to be satisfied. It means that leisure time must decline over time since $\partial \underline{L}(K)/\partial K < 0$.

The second step is to show that there is a neighbourhood left of $\underline{L}(K)$ such that sustained growth is impossible for any L in this neighbourhood. If $K_{t+1} > K_t$, note that $\underline{L}(K_{t+1}) < L_{t+1}(\underline{L}_t, K_t) = \underline{L}(K_t)$ since $\underline{L}(K)$ is decreasing in K due to (a2)-(a4). By continuity, there is a neighbourhood $I_1(K)$, left of $\underline{L}(K)$ such that for all $x \in I_1(K)$, we have $\underline{L}(K_{t+1}(x)) < L_{t+1}(x, K_t) \leq x \leq \underline{L}(K)$. Since $L_{t+1} > \underline{L}(K_t)$, this implies $L_{t+2} > L_{t+1}$ and a continuous decline of labour supply until the capital accumulated starts to decrease (*i.e.*, until $L > L^*$) as we demonstrated in the first step.

Recursively we have $I_2(K)$, left of $I_1(K)$ and such that $L_{t+2}(L_{t+1}, K_{t+1}) < L_{t+1}(x, K_t) < x$ and $L_{t+3} > L_{t+2}$. In the same way and due to the continuity of $L_{t+1}(.)$, we can define $I_n(K)$ for any integer n>0. Then we have:

$$I(K) \equiv \bigcup_{i=1}^{i=+\infty} I_i(K). \tag{A1}$$

Thus, to get sustained growth, we need L_t be left of $I(K_t)$. Otherwise expectations would be dynamically inconsistent with the growth hypothesis (L would end up being superior to L^*, see step 1).

For the third step, we can show that we also have a neighbourhood $J_1(K)$, right of $L_0(K)$ such that it would be dynamically inconsistent to demand $L_t \in J_1(K)$. In fact, we would get $L_{t+1} < L_0(K_{t+1})$. If we demand $L_0(K)$, we will need $L_{t+1} = 0$ to fulfil our expectations. This is impossible (this is ruled out by (A4)). The existence of $J_1(K)$ is established by continuity. We set $J_2(K)$ such that L_{t+1} is feasible and $L_{t+2} < L_0(K_{t+2})$. Then as previously:

$$J(K) \equiv \bigcup_{i=1}^{i=+\infty} J_i(K). \qquad \text{(A2)}$$

So for sustained growth, a necessary condition is that any L_t must be left of *I(K)* and right of *J(K)*. If leisure demand on the right of *I(K)*, the path of labour supply is not decreasing enough, whereas if it is on the left of *J(K)*, it is decreasing too fast to be dynamically consistent. A growth path will exist if, for any $K > K^*$, there is at least one sequence of leisure demand not inconsistent with rational expectations. In particular, if there is any L_t left of $I(K_t)$ and right of $J(K_t)$, we will observe L_{t+n} left of $I(K_{t+n})$ and right of $J(K_{t+n})$ for any positive integer *n* (this means that expected leisure demand is always consistent with rational expectations for all future periods). Then one can check that this is sufficient for growth to occur since: $L_t < L^* \Rightarrow K_{t+1} > K_t$.

So for the fourth step, it remains to show that, no point can belong to both neighbourhoods, *I(K)* and *J(K)*. Suppose there is $L_t \in I(K)$ and $L_t \in J(K)$, then there must be $\tau \geq t$ such that $L_{\tau+1} > L_\tau$ and $L_{\tau+1} < L_0(K_{\tau+1})$. These last two conditions imply $L_0(K_{\tau+1}) > L_\tau$. By definition, we also have $L_0(K_\tau) > L_0(K_{\tau+1})$. Then we must observe $L_\tau > L_0(K_\tau)$. This leads to a contradiction since we need $L_\tau < L_0(K_\tau)$, otherwise it would mean $L_{\tau+1} = 0$. So, there always exists at least one feasible perfect foresight equilibrium with growth. This path is such that leisure is decreasing and capital is increasing. ■

APPENDIX 2 PROOF OF PROPOSITION 3

Step 1 Suppose that for a given level of capital K_τ sufficiently large, there are two possible levels of leisure demand L_τ and L_τ' consistent with rational expectations and growth. We define θ such that $\theta(K) \equiv \partial L_{\tau+1}/L_\tau \big|_{L_\tau=\underline{L}}$. Due to the concavity of $L_{t+1}(L_t,\ K_t)$ and the boundary assumptions, θ increases with K and can be arbitrarily large when K is large enough. Then, for K large enough, we obtain $L_{\tau+1} - L_{\tau+1}' > \theta\left(L_\tau - L_\tau'\right)$ and the accumulation equation implies that:

$$\frac{K_{\tau+1}' - K_{\tau+1}}{K_{\tau+1}'} = \frac{K_\tau\left(1-L_\tau'\right)^{1-\beta} - K_\tau\left(1-L_\tau\right)^{1-\beta}}{K_\tau\left(1-L_\tau'\right)^{1-\beta}} \quad \Rightarrow \quad \frac{K_{\tau+1}' - K_{\tau+1}}{K_{\tau+1}'} < \frac{L_\tau - L_\tau'}{1-L_\tau'}. \tag{A3}$$

Then if $\theta > 1 + 1/\left(1-\underline{L}(K_\tau)\right)$, we find that $\left(K_{\tau+1}' - K_{\tau+1}\right)/K_{\tau+1}' < L_{\tau+1} - L_{\tau+1}'$.

Step 2 Assume that for any $t > \tau$, the following property holds:

$$L_t - L_t' > \left(K_t' - K_t\right)/K_t'. \tag{A4}$$

Now we can define $L_{t+1}''\left(K_t, L_t'\right)$ the underlying leisure demand induced by K_t and L_t'. Note that $L_{t+1}''\left(K_t,\ L_t'\right) > L_{t+1}'\left(K_t',\ L_t'\right)$. Since $K_t > K_\tau$, we get the result that $L_{t+1} - L_{t+1}'' > \theta\left(L_t - L_t'\right)$.

Step 3 Using the equilibrium condition $l_t = L_t$, we can write:

$$L_{t+1}' - L_{t+1}'' = \left[\frac{l^{-1}\left(L_t'\right)}{\beta(1-\beta)A^2K_t}\right]^{1/(1-\beta)} - \left[\frac{l^{-1}\left(L_t'\right)}{\beta(1-\beta)A^2K_t'}\right]^{1/(1-\beta)}. \tag{A5}$$

From (a1) and (a4): $L_t < 1$; that is $\left[l^{-1}\left(L_t'\right)/\left(\beta(1-\beta)A^2 K_t'\right)\right]^{1/(1-\beta)} < 1$. Then (A5) implies: $L_{t+1}' - L_{t+1}'' < 1 - \left[K_t / K_t'\right]^{1/(1-\beta)}$. Using (A4), this induces $L_{t+1}' - L_{t+1}'' < L_t - L_t'$. Combining this last inequality with the result $L_{t+1} - L_{t+1}'' > \theta\left(L_t - L_t'\right)$ obtained above gives us $L_{t+1} - L_{t+1}' > (\theta - 1)\left(L_t - L_t'\right)$.

Step 4 Given the saving behaviour, we have:

$$\frac{K_{t+1}' - K_{t+1}}{K_{t+1}'} = \frac{K_t'\left(1 - L_t'\right)^{1-\beta} - K_t\left(1 - L_t\right)^{1-\beta}}{K_t'\left(1 - L_t'\right)^{1-\beta}}. \tag{A6}$$

Using (A4), we find $\left(K_{t+1}' - K_{t+1}\right)/K_{t+1}' < \left(1 + K_t / \left(K_t'\left(1 - L_t'\right)\right)\right)\left(L_t - L_t'\right)$ after manipulation. Now using the result of step 3, we find:

$$\frac{K_{t+1}' - K_{t+1}}{K_{t+1}'} < \left(1 + \frac{K_t}{K_t'\left(1 - L_t'\right)}\right) \Big/ (\theta - 1)\left(L_{t+1} - L_{t+1}'\right). \tag{A7}$$

So for $\theta > 2 + 1/\left(1 - \underline{L}\left(K_\tau\right)\right)$, we find $L_{t+1} - L_{t+1}' > \left(K_{t+1}' - K_{t+1}\right)/K_{t+1}'$. Then recursively, the property holds for all $t > \tau$. So the distance between L_{t+n} and L_{t+n}' is increasing geometrically, which obviously contradicts that the fact L_t should converge to zero in all cases. The equilibrium path with growth must then be unique for K large enough.

Step 5 Now it remains to show that uniqueness is also true for any other $K > K^*$. Assume a temporary equilibrium $\left(L_t, K_t\right)$ on the growth path such as the slope of $L_{t+1}(.,.)$ in L_t is superior to $2 + 1/\left(1 - L_t\right)$. Because of the monotonicity of $L_{t+1}(.,.)$, for a unique K_{t-1} there exists a unique L_{t-1} such that $\left(L_t, K_t\right)$ will be reached in t.

Moreover, it is straightforward that, for $K'_{t-1} > K_{t-1}$, the associated leisure demand such that (L_t, K_t) is attained in t is $L'_{t-1} < L_{t-1}$. We can write that $K_t = (1-\beta)AK_{t-1}\left(1 - L_{t-1}(K_{t-1})\right)^{1-\beta}$. Because of this last equation K_t is increasing with K_{t-1}. Consequently, there exists a unique K_{t-1} such that K_t will be reached at the next period. Of course, $K_t > K_{t-1}$ and $L_t > L_{t-1}$. Recursively, the temporary equilibrium with positive growth is unique for each level of capital above K^*. ■

APPENDIX 3 PROPOSITION 5'

We assume $U_t = l_t^{1-\sigma}/(1-\sigma) + x\,c_{t+1}^{1-\sigma}/(1-\sigma)$ (with $x > 0$ and $\sigma > 1$) and the aggregate production function to be $Y_t = AK_t^{1+\mu}(1-L_t)^{1-\beta}$. In this case, the labour supply that leaves initial conditions constant is not independent of *K* anymore. We can define $L^*(K_t)$ such that $K_{t+1}/K_t = 1$, that is $(1-\beta)AK_t^{\mu}\left(1 - L^*(K_t)\right)^{1-\beta} = 1$. In order for sustained positive growth to be possible, we need $L_t < L^*(K_t)$ and the expectations are such that $L_{t+1} < L^*(K_{t+1})$. Individual maximisation implies:

$$L_{t+1} = 1 - \left(\left(\frac{l_t}{1-l_t}\right)^{\sigma/(1-\sigma)} \frac{(1-l_t)}{x^{1/(1-\sigma)}\beta A\left((1-\beta)AK_t^{1+\mu}(1-L_t)^{1-\beta}\right)^{1+\mu}}\right)^{1/(1-\beta)} . \qquad \text{(A8)}$$

The temporary equilibrium is then such that $l_t = L_t$, so we find:

$$L_{t+1} = 1 - \left(\left(\frac{L_t}{1-L_t}\right)^{\sigma/(1-\sigma)} \frac{(1-L_t)^{1-(1-\beta)(1+\mu)}}{x^{1/(1-\sigma)}\beta A^{2+\mu}(1-\beta)^{1+\mu} K_t^{(1+\mu)(1+\mu)}}\right)^{1/(1-\beta)} . \qquad \text{(A9)}$$

Using the definition of $L^*(K_{t+1})$, we get:

$$L^*(K_{t+1}) = 1 - \left(\left(1-\beta\right)^{1+\mu} A^{1+\mu} K_t^{\mu(1+\mu)} \left(1-L_t\right)^{\mu(1-\beta)}\right)^{-1/(1-\beta)}. \tag{A10}$$

The consistency of expectations with growth requires $L_{t+1} < L^*(K_{t+1})$. From (A8) and (A9), this condition writes $x^{1/(1-\sigma)}\beta A K_t < \left(L_t/(1-L_t)\right)^{\sigma/(1-\sigma)}(1-L_t)^{\beta}$. Growth in period t also requires $L_t < L^*(K_t)$, which implies $(1-\beta)AK_t^{\mu}(1-L_t)^{1-\beta} > 1$. Combining the two previous inequalities, we obtain:

$$\left(\frac{\sigma}{\sigma-1} + \frac{1-\beta}{\mu} + \beta\right) ln(1-L_t) > ln\frac{x^{1/(1-\sigma)}\beta A^{1-1/\mu}}{(1-\beta)^{1/\mu}} + \frac{\sigma}{\sigma-1} ln(L_t). \tag{A11}$$

Since $\lim\limits_{K_t \to +\infty} L_t = 1$, we need $\sigma/(\sigma-1) + (1-\beta)/\mu + \beta < 0$ which is a contradiction. ■

APPENDIX 4 PROOF OF PROPOSITION 6

At the steady-state (L^*, K^*), the utility of each generation is:

$$U^* = U\left(L^*,\ (1-\beta)AK^*\left(1-L^*\right)^{1-\beta}\right) \quad \text{with} \quad L^* = 1 - \left((1-\beta)A\right)^{-1/(1-\beta)} \tag{A12}$$

and L^* satisfies:

$$\frac{dU\left(l,\ (1-\beta)\beta A^2\left(1-L^*\right)^{-\beta}(1-l)K^*\left(1-L^*\right)^{1-\beta}\right)}{dl} = 0, \tag{A13}$$

$$\Rightarrow \begin{array}{l} U_l\left(L^*, (1-\beta)AK^*\left(1-L^*\right)^{1-\beta}\right) - (1-\beta)\beta A^2 K^*\left(1-L^*\right)^{1-2\beta} \times \\ U_c\left(L^*, (1-\beta)AK^*\left(1-L^*\right)^{1-\beta}\right) = 0. \end{array} \tag{A14}$$

The authority can fix L^{**} forever. If all excess production for $\left(1-L^{**}\right)>\left(1-L^{*}\right)$ is given to the young generation, we observe then:

$$\Delta U=\Delta L\times U_l\left(L^{**},(1-\beta)AK^{*}\left(1-L^{**}\right)^{1-\beta}\right)-\Delta C\times U_c\left(L^{**},(1-\beta)AK^{*}\left(1-L^{**}\right)^{1-\beta}\right).\text{(A15)}$$

For $L^{**}=L^{*}$, it implies:

$$\Delta U=\Delta L\begin{pmatrix} U_l\left(L^{*},(1-\beta)AK^{*}\left(1-L^{*}\right)^{1-\beta}\right)-(1-\beta)A^2K^{*}\left(1-L^{*}\right)^{1-2\beta}\times \\ U_c\left(L^{*},(1-\beta)AK^{*}\left(1-L^{*}\right)^{1-\beta}\right)\end{pmatrix}. \qquad \text{(A16)}$$

In this case, using (A14), one can check that the second term of the *rhs* of this expression is strictly positive. Consequently by continuity there exists $L^{**}<L^{*}$ such that we can set $L=L^{**}$. In that case, the generation born in *t-1* receives the same income, whereas the generation born in *t* is better-off. All the generations born after *t* are also better off since they work the same and consume more than the generation born in t. ■

THE CASE OF UNITARY ELASTICTITY OF SUBSTITUTION BETWEEN LEISURE AND CONSUMPTION

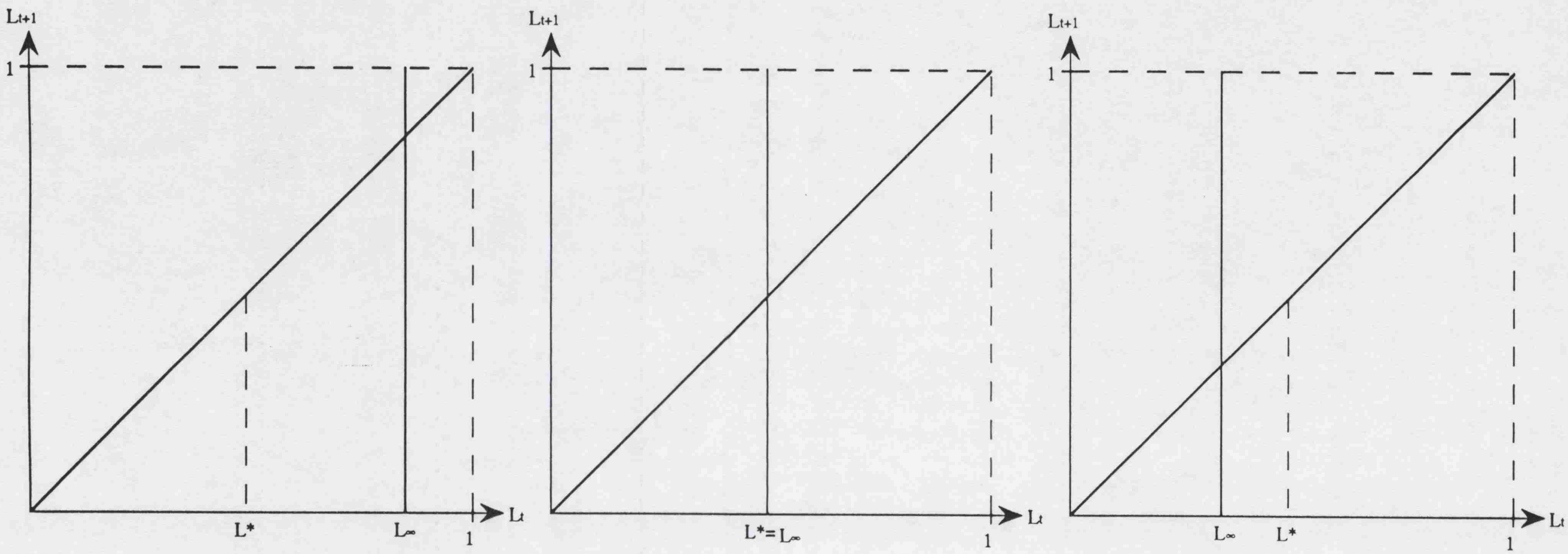


Figure 1-a

Recession

Figure 1-b

Stagnation

Figure 1-c

Growth

SUBSTITUTABILITY BETWEEN LEISURE AND CONSUMPTION

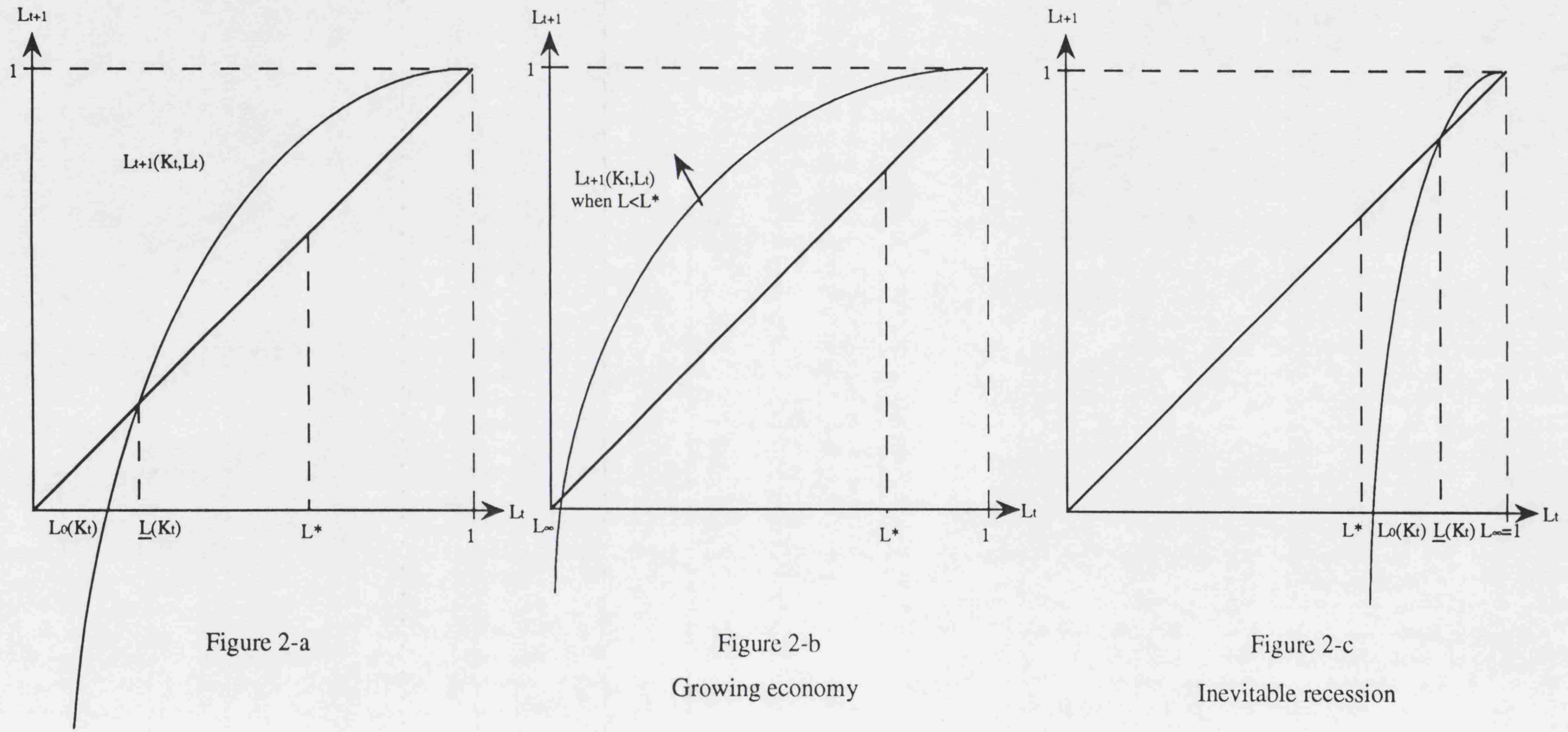


Figure 2-a

Figure 2-b

Growing economy

Figure 2-c

Inevitable recession

COMPLEMENTARITY BETWEEN LEISURE AND CONSUMPTION

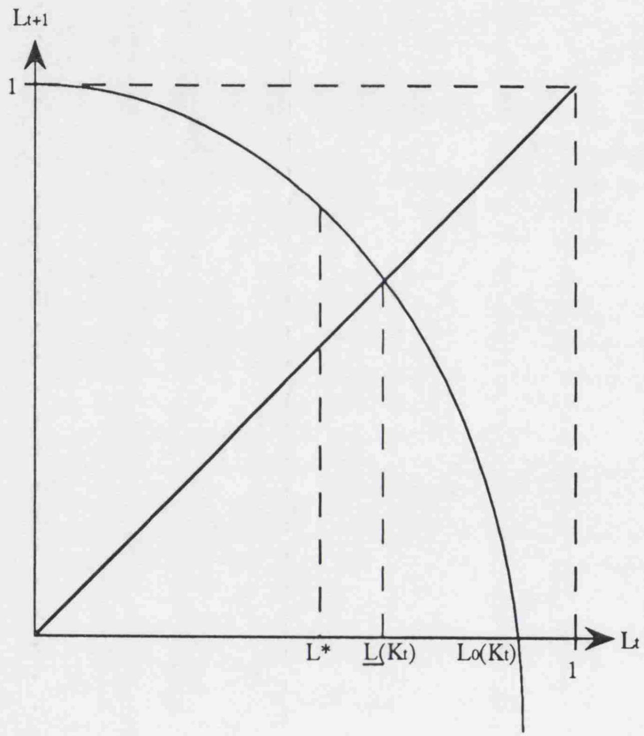


Figure 3

CHAPTER 2

FINANCING PRODUCTIVE LOCAL PUBLIC GOODS

Abstract: Public economics typically assumes that local public goods only affect the utility of consumers. We assume on the contrary that local public goods are purely productive. The implications of this assumption are analysed within standard dynamic growth models. Investment in the public good enhances productivity only in the jurisdiction where it takes place. Capital as well as people are perfectly mobile. After characterising the first-best equilibrium, we show that its decentralisation is more demanding than with local public consumer goods. In particular, efficient decentralisation cannot be obtained with competing land developers providing the public good through a simple land capitalisation scheme.

I. INTRODUCTION

How can an economy achieve optimal provision of local public goods (LPG)? Traditional public economics, originating from Samuelson [1954], stresses the difficulty of the provision of public goods in a general context. Basically, a decentralised scheme cannot be implemented because of a standard free-rider problem. Moreover, the first-order conditions for the central planner to achieve first-best require him to observe the consumers' marginal rates of substitution. So, even with a benevolent planner, the optimum is difficult to reach. However, Tiebout [1956] devised a clever alternative solution. His answer to the pivotal question raised above was to rely on the local aspects of some public goods and to use competition among the jurisdictions. In his original paper, assuming an optimal number of jurisdictions led by profit-maximising developers, people, by voting with their feet, can choose between different levels of provision for the public good and the associated head-taxes. Then the competitive equilibrium is a first-best situation.

Yet the assumption of an *ex-ante* optimal number of jurisdictions and the possibility of head-taxes are rather restrictive. However, the subsequent literature (see Wildasin [1987] or Mieszkowski and Zodrow [1989] for complete surveys on this issue) managed to relax them elegantly.[1] The Tiebout idea relies on the strong analogy between fiscal competition and the competition to supply private goods. Consequently free-entry of developers leads to the first-best just as free-entry of producers can achieve the First Welfare Theorem for private goods. As for the head-tax, the idea is to replace it by using the land market. Implementation of the first-best just requires the developer to be able to take advantage of the differential land rent in her jurisdiction since public spending is capitalised in the land value. This differential land rent stems from the agents' competition for space.[2]

The profit function of each developer in the most direct case is equivalent to total differential land rent (TDR) minus expenses for the public good (G). Then an immediate implication of the zero-profit condition is $TDR = G$. This result is known as the Henry

[1] To insure existence, some restrictions on the utility functions are still necessary. See Fujita [1989] for developments.

[2] By differential land rent we mean the share of land rent created by the action of the developer. In the paper, without any land development, land rent is equal to zero so that total land rent and differential land rent are equal.

George Theorem (George [1879]). Then the only informational requirement to implement the first-best in a decentralised way is the observation of the land market. This result appears in Flatters, Henderson and Mieszkowski [1974], Vickrey [1977] and Arnott and Stiglitz [1979] among others. Of course it relies on strong assumptions. It is not valid with:
- Imperfect competition, see Scotchmer [1986];
- Imperfect geography (if the land-rent is not well-defined at the border of the city), see Arnott and Stiglitz [1979] and Pines [1991];
- Congestion, see Scotchmer [1986] and Fujita [1989];
- Imperfect taxation (if only a property tax is available instead of a land tax), see Mieszkowski [1972] or Hoyt [1991];
- Imperfect mobility (if agents cannot vote with their feet at zero cost).

Despite all this, Tiebout's idea seems to be empirically relevant: Oates [1969], Edel and Sclar [1974], Hamilton [1976], Meadows [1976] or more recently Wassmer [1993] all draw favourable conclusions and show that mobility influences the mix of public goods offered at the local level. Moreover, some important projects are explicitly funded by land capitalisation schemes. One can think for instance of the railways in Tokyo (Kanemoto [1984], Kanemoto and Kiyono [1993], Midgley [1994]) or the public transports in Hongkong (Midgley [1994]).

Strange as it might seem, most potential applications of the capitalisation hypothesis (i.e., public good finance through the differential land rent) concern infrastructures (which are "productive public goods"), whereas the theoretical analysis focuses entirely on "public consumer goods" (which directly enter the utility function).[3] For infrastructures, it is possible that user-charges can provide convenient, albeit not always simple, pricing schemes; a general synthesis on these aspects is provided by Laffont and Tirole [1993]. In particular, perfect spatial discriminatory pricing is seldom available, although most infrastructure have an important geographic dimension. For road networks, airports or even electricity or water distribution, location matters. In other words, *public infrastructures are a crucial part of the production process at the local level* (see World Bank [1994]).

[3] The concepts of productive public good, public capital and infrastructure are equivalent in the analysis that follows.

Apart from user-charges, another alternative to finance public infrastructure is given by capital taxation at the local level (Wildasin [1989]). The literature on this matter shows the existence of a fiscal externality leading jurisdictions to implement sub-optimal levels of capital taxation because of factor mobility. This literature usually takes public expenditures as given and ignores any linkage between taxation and the marginal productivity of capital; that is, it again assumes implicitly that public goods are consumed rather than productive. However, the fiscal externality is likely to persist and even worsen with productive local public goods. So, as it seems difficult to finance productive public goods with either user-charges or capital taxation, our aim in what follows is to assess to what extent the land market can contribute to the financing of productive local public goods as it does for public consumer goods.

Another possible perspective is to consider that many local public consumer goods also have a productive aspect. Indeed, only purely recreational public goods such as parks and museums can be considered as pure public consumer goods. Even in these extreme cases, it may be argued that they have some productive aspects: museums can improve education, whose productive role is obvious.[4] The quality of leisure also has an impact on production. A final motivation is that public economics typically recommends that local expenditure should be financed through the taxation of the differential land rent. However land taxes represent only a small fraction of local public finances (see Prud'Homme [1987] and Henderson [1995] for evidence). We offer here a new explanation for this stylised fact.

For public consumer goods, a partial equilibrium analysis is sufficient. We just need to consider an exogenous income, which generates a demand for the public good mediated by the land markets. Then, the problem is to see how the public good can be financed through the land market. On the contrary, in the case of productive public goods, the initial income generates a demand for private goods, for land and for savings. Using the demand for land, it is possible to finance local public goods but the story does not end here, since the savings and the amount of public good determine future production. A dynamic general equilibrium analysis is thus required. In other words, the analysis of productive public

---

[4] Education presents some specific features that justify a separate analysis with different assumptions. For instance, see Bénabou [1993].

goods is inherently dynamic, because public capital can be accumulated and influences further production.

Moreover, assume for convenience that private producers operate with constant returns to scale. Then the introduction of public capital which is assumed to increase the marginal productivity of private capital, implies increasing returns at the aggregate level. For these reasons, our model is dynamic and also embodies increasing returns. Using such a framework, we explore whether the Classical results due to Tiebout and George are still valid when the productive aspect of public goods is taken into account. It is shown below that, although Tiebout-style (i.e., first-best) results can be obtained, they are very demanding and rely on strong assumptions. The Henry George Theorem does not hold in the usual sense. Simple capitalisation schemes do not work because the marginal product of public investment in one jurisdiction benefits both the workers who live in the jurisdiction and also capital holders who need not live where their savings are invested. As a consequence, land rent capitalises public investment only up to the increase in wages it causes multiplied by the share of housing in expenditures. On the contrary, increased public investments in one particular jurisdiction raise the marginal return to capital. Thus, it increases the financial income of everyone in the economy and it follows that the demand for land in all the jurisdictions goes up.

In Section II, we propose a simple framework of a competitive production economy with productive public goods. In Section III, we analyse various decentralised frameworks. Finally the last section ends the analysis with some considerations concerning the provision of local public goods.

II. ECONOMIES WITH PRODUCTIVE PUBLIC GOODS: THE FIRST-BEST

We consider a large economy composed of S "islands", each of surface unity. The total population size N is such that $N < SL^*$, where $1/L^*$ is the minimal amount of land to be consumed per capita in order to enjoy positive utility. A given island may or may not be populated. Individuals are infinitely-lived (or finitely-lived with a dynastic utility function) and supply continuously one unit of labour inelastically. their utility function is:

$$
\begin{aligned}
& U = \int_0^{+\infty} \frac{(u_t)^{1-\sigma}}{1-\sigma} e^{-\rho t} dt, \\
& \text{with} \quad u_t = z_t s_t^x \qquad \text{if } s_t \geq 1/L^*, \\
& \qquad\quad\; u_t = 0 \qquad\quad \text{if } s_t < 1/L^*, \\
& \qquad\quad\; \sigma > 1.
\end{aligned} \tag{1}
$$

Instantaneous utility is thus a Cobb-Douglas function of z, the per capita consumption of the final good and s, the per capita quantity of land (which is a proxy for housing consumption). There is a single good, which can be consumed, used as public investment or transformed into private capital. The production function in each island is:[5]

$$
Y_{i,t} = AG_{i,t}^{\beta} K_{i,t}^{1-\alpha} L_{i,t}^{\alpha} \quad \text{with} \quad \alpha \geq \beta > x, \ i = 1, \ldots, S. \tag{2}
$$

This specification makes public capital a necessary input. In this respect we follow the empirical literature (Aschauer [1989] or Gramlich [1994]). Moreover, we do not consider the case of publicly provided private goods. Note that if land consumption is equally divided between inhabitants of the same island i, we observe $s_i = 1/L_i$ since we suppose that people should live and work on the same island. For the sake of simplicity we ignore depreciation and any unit of capital (K), labour (L) or public spending (G) can be used only in one island.[6]

Capital and labour are perfectly mobile across islands. The good can be transformed into capital and shipped elsewhere, but then it cannot be consumed anymore. In short, what consumers eat must come from their residential area. This assumption may seem rather unappealing but it makes sense for non-tradable goods like most personal services. Furthermore it leads to considerable technical simplification.[7] Note also that only

---

[5] As usual, dynamic results can be obtained only for heavily specified models. Our instantaneous utility function may seem *ad-hoc*. It allows however to consider the inferior good aspect of land. A Stone-Geary specification would be more elegant but it is not tractable in a dynamic context.

[6] There is no overlap. See Hochman, Pines and Thisse [1995] for the implications of this problem.

[7] In particular, it will simplify our first-best land use (land is used with the same intensity wherever it is used). Otherwise, we may be led to more complex patterns of land use (the density in inhabited islands need

consumers use space. This simplifying assumption, which is relaxed below, allows us to neglect the issue of competition for space between different categories of agents (producers and consumers). Finally, the analysis that follows concentrates on steady-state behaviours of our economy.[8]

A tendency to agglomeration is driven by the increasing returns present in the production function. Reducing the number of populated (or developed) islands increases production. Production becomes very large when only one island is developed. On the contrary, the taste for land acts as a dispersion force. The balance between agglomeration and dispersion is stated in the following lemma:

Lemma 1 *The first-best where individuals are treated symmetrically is such that N/L^* islands are populated by L^* inhabitants each.*[9]

See proof in Appendix 1. ■

The intuition for this lemma is that as long as the marginal productivity of public capital multiplied by the marginal utility of consumption is higher than the marginal utility of the consumption of space, it is socially worthwhile to reduce the consumption of space to increase production in order to avoid the spatial dilution of public capital.

Proposition 1 : *The first-best is such that mobile factors are symmetrically allocated across existing jurisdictions. If $\beta<\alpha$, the long-run level of capital accumulated is $\overline{K}=\left((1-\alpha)^{1-\beta}A\beta^{\beta}L^{*\alpha}/\rho\right)^{1/(\alpha-\beta)}$. If $\beta=\alpha$, public capital, private capital and consumption all grow at the rate $\left((1-\alpha)^{1-\alpha}\alpha^{\alpha}AL^{*\alpha}-\rho\right)/\sigma$.*

---

not be the same everywhere at the optimum). This "asymmetric" first best would render decentralisation more problematic. This assumption is relaxed in Duranton [1997].

[8] In the model we consider, the analysis of the transition dynamic is either straightforward since the problem reduces to a neo-classical model of growth when $\beta<\alpha$ or trivial since there is no transition dynamics when $\beta=\alpha$.

[9] This assumption is here to avoid a first-best for which consumers would not be treated symmetrically as it sometimes happens in public or urban economics. See Mirrlees [1972] for an example where utilitarian welfare leads to unequal treatment.

Proof Thanks to Lemma 1, we can restrict our attention to symmetric situations with constant consumption of land. It is then possible to write the social planner's program independently of land consumption.

$$\underset{z,G,K}{Max}\ U = \int_0^{+\infty} \frac{z_t^{1-\sigma}}{1-\sigma} e^{-\rho t} dt, \qquad \text{(SP)}$$

Subject to the "representative island" budget constraint: $Y_t = L^* z_t + \dot{G}_t + \dot{K}_t$,

and to the transversality condition: $\lim_{t \to +\infty} e^{-\rho t} \lambda_t K_t = 0$. (3)

Since $\sigma > 1$, intertemporal utility is always well defined. Thus this transversality condition does not play an important role in our analysis. The first-order conditions of (SP) are:

$$\begin{cases} z^{-\sigma} e^{-\rho t} = \lambda \\ \lambda(1-\alpha) A G^{\beta} K^{-\alpha} L^{*\alpha} = -\dot{\lambda} \\ \lambda \beta A G^{\beta-1} K^{1-\alpha} L^{*\alpha} = -\dot{\lambda} \end{cases} \qquad (4)$$

From this, we can state:

$$G = \frac{\beta K}{1-\alpha} \qquad (5)$$

and $$\frac{\dot{z}}{z} = \frac{(1-\alpha)^{1-\beta} A \beta^{\beta} K^{\beta-\alpha} L^{*\alpha} - \rho}{\sigma}. \qquad (6)$$

Depending on the value of β, we distinguish two different cases:

- $\beta < \alpha$ (Solow case). There is a long-run steady state for which $\dot{z}/z = 0$:

$$\Rightarrow \overline{K}=\left(\frac{(1-\alpha)^{1-\beta} A\beta^{\beta} L^{*\alpha}}{\rho}\right)^{1/(\alpha-\beta)} . \tag{7}$$

In this case, we have overall increasing returns but decreasing returns to the reproducible factors. Consequently the economy converges towards its long-run equilibrium level $\overline{K}$.

- $\beta=\alpha$ (Barro-Romer case). From the previous equations, we get:

$$\Rightarrow \frac{\dot{z}}{z}=\frac{\dot{K}}{K}=\frac{\dot{G}}{G}=\frac{(1-\alpha)^{1-\alpha}\alpha^{\alpha} AL^{*\alpha}-\rho}{\sigma}. \tag{8}$$

Here, given that returns to the reproducible factors are precisely one, endogenous steady growth is possible. The equilibrium growth path is such that consumption, public and private capital are all increasing at the same rate. ■

The initial program involved a dynamic optimisation over two choice variables with two state variables. However, thanks to Lemma 1, the consumption of space remains constant which allows us to get rid of one choice variable. Then, the marginal productivity of public capital should be equal to that of private capital which reduces our problem to a standard dynamic programming problem. Depending on the value of β, two different results can be obtained.[10] These two different cases underline two different conceptions of the role of public spending in the growth process. In the first case, growth is taken as exogenous (to the model), that is to say infrastructures do not act as the engine of growth (although they may intervene crucially in the productive process). This vision is defended for instance by Gramlich [1994]. It states that infrastructures do matter for production to occur, but that they do not influence the long-run growth rate. The other conception states that public spending can be considered as a major engine of growth. Theoretically, this vision was put forward by Barro [1990] and Barro and Sala-i-Martin [1992]. Empirically, this conception is defended by De Long and Summers[1991] and [1993]. We do not want to intervene here in the controversy on this matter. We use both specifications and show that our results are similar for both cases.

[10] As usual in growth models, the case $\beta>\alpha$ does not make much sense economically since it implies explosive growth.

It can be shown easily that the first-best solution in these two programs can be decentralised, given the provision of public goods. In the Solow case with no exogenous growth, the government just needs to provide the optimal amount of public infrastructure in each jurisdiction as in equation (5). Firms then face a well-behaved production function with constant returns to scale in private inputs. The steady-state is such that no further public spending is required after it is reached. In either the exogenous or endogenous growth case, public spending can be financed through a tax on production.[11] Barro [1990] shows that the government's optimal rate of taxation is also the one that maximises the growth rate. In a competitive equilibrium, consumers receive an interest rate r_t for their savings and a wage w_t for their work. Each factor is paid its marginal product and firms make zero profits. Moreover consumers have to pay a rent R_t for each unit of land (and this rent can be re-injected through public land property). Consequently, the competitive equilibrium is efficient. Our concern in what follows is then to explore whether it is also possible to decentralise the provision of public infrastructure and to explore the implications of factor mobility.

III. DECENTRALISATION OF THE FIRST-BEST

Some possible institutional regimes

We suppose now that the development of islands is undertaken by competing developers. The competitive process is the following. On any undeveloped island, a developer can create a jurisdiction where she will provide the public good. Consumers then make their location and saving decisions with both labour and capital being perfectly mobile.

Within this broad framework, decentralisation can be established through various institutional regimes. The first alternative we consider is one in which the developer has complete control over everything in her jurisdiction (*Regime 1*). All possible fiscal and financial instruments are available and can be implemented. In this case, therefore, the

[11] The efficiency of this tax rests on three strong assumptions, which are i/ observability of private production, ii/ no distortion of labour supply and iii/ immobility of factors. Financing infrastructure through the land market avoids these shortcomings.

decentralisation of production is not complete due to the possible intervention of developers in the production process. *Regime 1* can be viewed as a benchmark case, but it may also be applicable to the US where land developers can have extensive powers (see Henderson and Slade [1993]). It also strongly reminds us of traditional factory-towns.

We also consider a second regime where the jurisdiction can either can implement a residential tax or set the price of land as a local monopoly (*Regime 2*). This is the case in countries with a decentralised system of government like Switzerland. With this regime, jurisdictions can also issue debts. Decentralisation takes place in this framework at two levels since, on the one hand, the provision of public goods is left to private developers and, on the other hand, production is realised within the jurisdictions by independent private producers.

Finally, and this may be more relevant to describe the countries with a centralised system of government (where local public goods remain locally provided), we consider an environment in which the land developer cannot use any fiscal instrument and faces a competitive land market on the supply as well as on the demand-side (*Regime 3*). Consumers bid on an individually optimal quantity of land, taking its price as given. Although the developer cannot set land rent directly in her jurisdiction, she can manipulate it indirectly by public expenditure actions. Besides, since the Henry George Theorem is often valid with public consumer goods, the natural temptation is to restrict the jurisdictions to use only land rent to finance the productive public goods. If this arrangement enables us to attain first-best, then its practical implementation should be simpler than in the two previous cases since it would just require land prices to be observed.

Our aim is thus to see how we can replicate the first-best equilibrium within the least demanding framework. We may prefer *a priori* a situation for which the developer should not be allowed to intervene too much in the local economy with her role restricted to the provision of the public good. Our analysis starts from the most demanding framework, that is the one for which the developer can intervene extensively in the local economy (fixing land rent, wages, public spending, capital investments, borrowing on the capital markets...). Then we examine the effects of various restrictions. Before going into the detailed analysis of these institutions, it is useful to state the following lemma.

Lemma 2: *Whatever the regime, any competitive equilibrium is such that the area developed per consumer is equal to* $s = 1/L^*$.

Proof: An equilibrium must be such that no developer should have an incentive to change her strategy and no consumer should have an incentive to move. For regimes 1 and 2, in case an individual should move, suppose that all the surplus created is distributed to the consumers within the jurisdiction and consumed immediately. If this raises the level of instantaneous utility, then the situation cannot be in equilibrium. It is a sufficient condition for an increase in total utility since it does not modify the state variable (a similar argument could be made using increasing profits). If $L < L^*$, then one can calculate that $\Delta u/\Delta L = \alpha Y/L^{x+2} - xz/L^{x+1}$. Since $\alpha > x$ and $Y \geq Lz$, then $\Delta u/\Delta L > 0$. For Regime 3, a similar proof can be made by comparing the increase in wage with the rise in land rent. ■

Regime 1

In that case, we get the following proposition, which unsurprisingly states that if developers have enough degrees of freedom, the first-best can be reached.

Proposition 2 : *The equilibrium under Regime 1 is first-best.*

Proof: Developers maximise the sum of their discounted profits, but due to perfect factor mobility this objective is equivalent to the maximisation of instantaneous utility at each point in time. Thus, each developer faces the problem:

$$\begin{aligned} &Max\, \pi_t = Y_t + R_t - w_t L_t - r_t K_t - r_t G_t, \\ &s.t. \quad U_t \geq \overline{U}. \end{aligned} \qquad \text{(DP1)}$$

The developer can maximise profits without any restriction on the instruments she uses. Her only constraint is to offer a level of utility as high as can be obtained in other islands.

Given the concavity of the utility function, each developer will divide her land equally among her consumers. Wages and rents are redundant since labour supply is inelastic and land is equally divided. Without loss of generality we can therefore normalise R=0. Our program becomes:

$$\begin{aligned} &\underset{\hat{w},G,K,L}{Max} \quad \pi_t = Y_t - \hat{w}_t L_t - r_t K_t - r_t G_t \\ &s.t. \quad U_t = \overline{U} \end{aligned} \tag{9}$$

where $\hat{w}$ is the wage distributed when land rent is normalised to zero. Before solving the developers' program, we must consider the consumers' program:

$$\underset{z,i}{Max} \; U = \int_0^{+\infty} \frac{\left(z_t s_{i,t}^x\right)^{1-\sigma}}{1-\sigma} e^{-\rho t} dt, \tag{CP1}$$

$$\text{s.t.} \quad \hat{w}_{i,t} + r_t Z_t = z_t + \dot{Z}_t$$

$$\text{and} \quad s_i \geq 1/L^*,$$

where Z is individual wealth. Due to Lemma 2, the available lots all have the same size, constant over time. Consequently, it is possible to take the term in s in the utility function out of the integral. The first-order conditions for (CP1) are then:

$$\begin{cases} z^{-\sigma} e^{-\rho t} = \lambda, \\ \lambda r = -\dot{\lambda}. \end{cases} \tag{10}$$

Then, we find that:

$$\hat{w}_t + \left(\frac{r(\sigma-1)+\rho}{\sigma}\right) Z_t = z_t. \tag{11}$$

Returning to the developers' program (DP1) and using (11), the first-order conditions are:

$$\begin{cases} \dfrac{\partial Y}{\partial K} = r, & \dfrac{\partial Y}{\partial L} - \hat{w} = -x\mu \dfrac{\hat{w} + \left(\dfrac{r(\sigma - 1) + \rho}{\sigma}\right) Z}{L^{x+1}}, \\ \dfrac{\partial Y}{\partial G} = r, & -L = \mu \dfrac{1}{L^x}, \end{cases} \tag{12}$$

where μ is the Lagrange multiplier of (DP1). Using Lemma 2 and the zero-profit condition, we obtain in the Solow case:

$$\begin{cases} (1 - \alpha + \beta) Y = (G + K) r, \\ G = \dfrac{\beta K}{1 - \alpha}, \\ \hat{w} = \dfrac{\alpha - \beta}{L} Y, \\ r = (1 - \alpha)^{1-\beta} \beta^{\beta} A K^{\beta - \alpha} L^{*\alpha}. \end{cases} \tag{13}$$

Consumers then save until the steady-state is reached. This steady-state is such that $r = \rho$. We find the long-run level of capital is $\overline{K}$ as in equation (7). Consequently, our decentralisation scheme enables us to implement the first-best in the Solow case. As for the Barro-Romer case, the first-order and zero-profit conditions read:

$$\begin{cases} Y = (G + K) r, \\ G = \dfrac{\alpha K}{1 - \alpha}, \\ \hat{w} = 0, \\ r = (1 - \alpha)^{1-\beta} \beta^{\beta} A L^{*\alpha}. \end{cases} \tag{14}$$

From this we can easily obtain that the growth rate is the same as in equation (8). Thus, it is also possible to decentralise our first-best solution if we allow for endogenous growth. ■

Note that heavy intervention is required in our "factory-island" since the developer is responsible for both the wage policy and the spatial allocation of consumers. Moreover the Henry George Theorem becomes meaningless since the first-best solution can be obtained for any level of land rent. To get more intuition concerning the working of the model, we need to explore the two other regimes.

Regime 2

If the developer can intervene on the land market through a residential tax, we can now show that the first-best equilibrium can be decentralised without her direct intervention in the production process.

Proposition 3 : *The equilibrium in Regime 2 is first-best. Developers set a (steady-state) residential tax equal to:*

- $T_t = \left(1-(x+1)(\alpha-\beta)\right)w_t - x\rho Z_t$ *in the Solow case.*
- $T_t = w_t - x\dfrac{r(\sigma-1)+\rho}{\sigma}Z_t$ *in the Barro-Romer case.*

Proof The consumers' program is:

$$\underset{z,s,i}{Max}\ U = \int_0^{+\infty} \frac{\left(z_t s_{i,t}^x\right)^{1-\sigma}}{1-\sigma} e^{-\rho t}\, dt, \tag{CP2}$$

s.t. $w_{i,t} + r_t Z_t - \dot{Z}_t - T_{i,t} = R_{i,t}s_{i,t} + z_t = E\left(w_{i,t}\right)$

and $s_{i,t} \geq 1/L^*$.

$E\left(w_t\right)$ is the expenditure of the consumer at date t and $T_{i,t}$ is the taxation in jurisdiction i. The wage w is offered by independent competing producers. After simplifications, and given that the equilibrium is such that s is constant over time, the first-order conditions can be written:

$$\begin{cases} \dfrac{\dot{z}}{z} = \dfrac{r-\rho}{\sigma}, \\ Rs = xz, \\ s = 1/L^*. \end{cases} \tag{15}$$

Because of the mobility of the productive factors, the developers' program reduces to:

$$\max_{G,T,L} \ \pi_t = R_t + L_t T_t - r_t G_t, \qquad \text{(DP2)}$$
$$s.t. \quad U_t = \overline{U},$$

where T is the residential tax. Instead of solving this program, note that we can write:

$$w_t - T_t - R_t s_t = \hat{w}_t. \qquad (16)$$

We can then re-write the developer's program:

$$\max_{G,\hat{w},L} \ \pi_t = w_t L_t - \hat{w}_t L_t - r_t G_t, \qquad (17)$$
$$s.t. \quad U_t = \overline{U}.$$

Moreover competing producers within each jurisdiction will induce:

$$w_t L_t = \beta Y_t \quad \text{and} \quad r_t K_t = (1-\beta) Y_t. \qquad (18)$$

Equations (17) and (18) are equivalent to (DP1) after using the first-order condition for K (remember that equation (12) states that $r_t K_t = (1-\beta)Y_t$ and is thus equivalent to (18)). Then, using the zero-profit condition and (15), we obtain (13) again. This means that the first-best can be obtained if we replace the developer's direct intervention in production by a residential tax. This result is not very surprising. Due to capital mobility, the first-order condition with respect to K is the same if the developer decentralises production. Direct wage setting is replaced by a residential tax. This tax can be characterised easily, in the Solow case, using (13), (15) and the steady-state condition for which $\dot{Z} = 0$. We get:

$$T_t = \big(1 - (x+1)(\alpha - \beta)\big) w_t - x\rho Z_t. \qquad (19)$$

In the Barro-Romer situation, we have:

$$T_t = w_t - x\frac{r(\sigma - 1) + \rho}{\sigma} Z_t, \qquad (20)$$

with r as in (14). ■

This residential tax may seem at first difficult to implement since it requires knowledge of all the parameters in the model (as in most public good provision problems). Note that our tax has two components, one related to wealth and the other to the wage, but since we are in a representative agent framework and at the steady-state, where all endogenous variables change proportionally, this tax could be calculated differently. Specifically, because of the symmetry across individuals, we just need the developer to be able to observe the land rent and to set a local income tax to maximise her total profit. Note also that the taxes in Proposition 3 can be interpreted partly as "congestion" taxes. The idea is that any marginal consumer in a jurisdiction lowers the average productivity of labour due to decreasing returns. On the contrary, the developer benefits from his arrival since he also generates a positive pecuniary externality for the returns to private and public capital and since his demand for land increases land rent. The rest of the intuition is left for the next sub-section (i.e., why this tax is not equal to zero).

Regime 3

Since the first-best requires a non-zero tax/subsidy, it is impossible to obtain efficiency under *Regime 3*. However analysis of *Regime 3* is revealing because we can explore the inefficiencies resulting from a more constrained regime. This also strengthens the intuition concerning our previous results.

Proposition 4 *Regime 3 leads to under-investment in public and private capital. The equilibrium is such that jurisdictions may run positive profits. Profit increases with the preference for land when x is small.*

See proof in Appendix 2. ■

Consumers want to live where they receive the highest wages. Since land is scarce, competition for land creates a land rent that enables the financing of the local public goods. In the traditional case of public consumer goods, there is a direct relation between public expenditures and the utility of consumers. In the case of productive public goods, the relation is not as direct: more public expenditures induce higher production and only a share of the marginal production is received by workers (i.e., a fraction α). And the marginal increase of wages generated by the marginal public investment is used only partially for housing expenditures (the share of housing is $x/(x+1)$ in the total expenditures). So the marginal value of public investment is capitalised only through the share of wages in the production multiplied by the share of housing in consumer expenditures.

Alternatively, note that in the case of public consumer goods, efficiency is defined by the equalisation of marginal rates of substitution. Here, on the contrary, efficiency is defined by the equalisation of marginal rates of return. The problem is that the developer's marginal revenue for public investment is given by the marginal increase of land rent, which is driven by a preference parameter. Thus, we introduce a preference parameter in a "technological relation".

On the other hand, land rent in a given jurisdiction partly capitalises public investments made in all other jurisdictions. This happens because the demand for land (and thus land rent) depends not only on current income, but also on total wealth. So higher public investments in jurisdiction j imply a higher interest rate for private capital in the economy, and thus a higher wealth for consumers in jurisdiction i. So land rent in jurisdiction i increases after investments in public goods are made in j (as well as in any other jurisdictions). If this effect is sufficiently strong, land developers make positive profits. (No zero profit condition can be at work here since there is no instrument available for the developer to redistribute the rent.)

Land as factor of production

Until now we have assumed that land was only used for housing purposes. It may be worth looking at what happens when land is (only) a factor of production. More specifically, suppose that now utility is:

$$U = \int_0^{+\infty} \frac{(z_t)^{1-\sigma}}{1-\sigma} e^{-\rho t} dt \qquad \text{with} \quad \sigma > 1, \tag{21}$$

but he individual production function is:

$$y_t = AG_t^{\beta} k_t^{1-\alpha-\gamma} s_t^{\gamma} l_t^{\alpha} \quad \text{with} \quad \gamma > \beta \quad \text{and} \quad \alpha > 0. \tag{22}$$

So in each jurisdiction:

$$Y_{i,t} = AG_{i,t}^{\beta} K_{i,t}^{1-\alpha-\gamma} L_{i,t}^{\alpha}. \tag{23}$$

Since $\gamma > \beta$, the first-best is such that all land is developed since the marginal productivity of land is superior to the marginal productivity of the local public good (and the optimal population per island is $N/S = L^{**}$). It can be shown easily that the first-best level of public good is:

$$\overline{G} = \left(\frac{A\beta^{(\alpha+\gamma)}(1-\alpha-\gamma)^{(1-\alpha-\gamma)}}{\rho} L^{**\alpha} \right)^{1/(\alpha+\gamma-\beta)} \qquad \text{with} \quad r = \rho. \tag{24}$$

But this level of accumulated public capital cannot be reached using only land rent. Indeed, at the market equilibrium, each factor is paid at its marginal productivity. Thus, total differential land rent equals:

$$R_t = \gamma AG_t^{\beta} K_t^{1-\alpha-\gamma} L_t^{\alpha}. \tag{25}$$

We can also show in the long-run that free-entry drives profits to zero (all the land rent is devoted to the financing of the public good) and that the population in each island is

$L = L^{**}$. Then in the steady-state:

$$\overline{\overline{R}} = \rho\overline{\overline{G}} \quad \Leftrightarrow \quad \overline{\overline{G}} = \gamma\left((1-\alpha-\gamma)^{(1-\alpha-\gamma)}\frac{\beta^{\beta}}{\rho}AL^{**^{\alpha}}\right)^{1/(\alpha+\gamma-\beta)} \tag{26}$$

$$\Rightarrow \overline{\overline{R}} = \frac{\gamma}{\beta}\overline{G} \tag{27}$$

Consequently, straightforward land capitalisation does not induce first-best and again the Henry George Theorem does not hold. Another instrument is therefore necessary to decentralise the first-best. The reason is that at the market equilibrium all factors are paid at their marginal productivity when we assume homogeneity of degree one for private factors. Then there is no reason for the marginal productivity of land to be equal to the marginal productivity of the public good. In other words, the land rent, which is equal to the share of land in the production function, does not provide in general an optimal level of infrastructure since it is not equal to the share of public capital.

IV. CONCLUDING REMARKS ON THE PROVISION OF LOCAL PUBLIC GOODS

We have shown in this chapter that considering the productive aspects of local public goods seriously complicates the decentralisation of the provision of infrastructures. The first-best can be decentralised, but the local authority must be able to intervene at least through a residential tax (indexed on wealth and the wage). With heterogeneous agents, the required residential tax could be even more difficult to implement; other instruments may be necessary, and the second-best taxation should be indexed on variables that are clearly difficult to observe at the local level (e.g., wealth in the upper-tail of the distribution).

Without any tax, direct internalisation of the land rent is not efficient. Marginal local public investments are capitalised in local land values only through the share of housing multiplied by the share of wages in local production. Conversely, land values within one jurisdiction also capitalise public investments made in all the economy because of the mobility of capital (higher interest rates induce a higher demand for land).

However, despite our apparently rather negative results, Tiebout and George's ideas may deserve more attention than they are presently given.[12] Our interpretation of existing land capitalisation schemes is that they constitute useful additional instruments (capitalisation occurs up to the share of wage multiplied by that of housing). For instance the Hongkong underground network was partly financed through a land capitalisation scheme (see Midgley [1994] for more details). Before the construction of the network, the operator was able to buy large parcels surrounding the future stations. Now the operator receives a majority of its income through the fares, but the profit made on land operations (development of shopping areas, commercial real estate and residential buildings) is by no means negligible, at around 15% of the construction cost. As a consequence the Hongkong underground is the only profitable underground network in the world.

[12] Respectively introduce competition at the local level and use information available on the land market.

APPENDIX 1 PROOF OF LEMMA 1

After some straightforward manipulations (i.e., broadly similar to the ones performed through equations (2) to (6)), the optimal consumption path for a fixed amount of developed land is:

$$z = Y - g(K+G) \quad \text{with} \quad g = \frac{(1-\alpha)^{1-\beta} A\beta^{\beta} K^{\beta-\alpha} s^{-\alpha} - \rho}{\sigma}, \tag{A1}$$

where g is the growth rate of the economy. Then we can write:

$$z = \frac{AG^{\beta}K^{1-\alpha}}{s^{\beta}} - \frac{(1-\alpha)^{1-\beta} A\beta^{\beta} K^{\beta-\alpha} s^{-\alpha} - \rho}{\sigma}\left(\frac{1-\alpha+\beta}{1-\alpha}\right)K. \tag{A2}$$

This leads to:

$$u = s^{x-\alpha} AK^{1-\alpha+\beta}\left(\left(\frac{\beta}{1-\alpha}\right)^{\beta} s^{\alpha-\beta} - \frac{(1-\alpha)^{1-\beta} A\beta^{\beta}}{\sigma}\left(\frac{1-\alpha+\beta}{1-\alpha}\right) + \frac{\rho}{s^{\alpha}\sigma AK^{\beta-\alpha}}\left(\frac{1-\alpha+\beta}{1-\alpha}\right)\right) \tag{A3}$$

After simplifications, one can check that for any $z > 0$:

$$x < \beta \quad \Rightarrow \quad \frac{\partial u}{\partial s} \leq 0. \tag{A4}$$

So for a given accumulation path (i.e., some existing G, K and savings at each date), from $x < \beta$, a typical consumer can always increase his instantaneous utility by consuming less space until the constraint $s \geq 1/L^{*}$ is binding. Suppose now that there exists an asymmetric equilibrium (with some jurisdictions that are less populated). Then, at the equal-treatment first-best, we require instantaneous utility to be the same for everybody. By assumption, we also need production to take place in the jurisdiction where one lives. Whatever c, the share of production consumed, we find that instantaneous utility is equal to

$$u = cs^{x-\beta} AK^{1-\alpha+\beta}\left(\frac{\beta}{1-\alpha}\right)^{\beta}. \tag{A5}$$

Then, whatever c, K and G, we find that the *per capita* consumption of land that maximises instantaneous utility is $s = 1/L^*$ in all the jurisdictions. The first-best is then symmetric and $s = 1/L^*$. ■

APPENDIX 2 PROOF OF PROPOSITION 4

The consumers' program is:

$$\underset{z,s}{Max}\ U = \int_0^{+\infty} \frac{\left(z_t s_t^x\right)^{1-\sigma}}{1-\sigma} e^{-\rho t} dt \tag{CP3}$$

s.t. $w_t + r_t Z_t - \dot{Z}_t = R_t s_t + z_t = E(w_t)$,

and $s_t \geq 1/L^*$.

After simplifications, the first-order conditions can be expressed by:

$$\begin{cases} R = \dfrac{E(w)}{s(1+1/x)}, \\ \dot{Z} = Z\dfrac{r-\rho}{\sigma}. \end{cases} \tag{A6}$$

As for the developers' program, we have:

$$\begin{aligned} &\underset{G,L}{Max}\ \ \pi_t = R_t - r_t G_t \\ &s.t.\ \ U_t = \overline{U}. \end{aligned} \tag{DP3}$$

Using (A6), the first-order conditions of the developers are:

$$\begin{cases} L^{x+1} = -\lambda(x+1), \\ \dfrac{\partial w}{\partial G} \cdot \dfrac{L}{(1+1/x)} - r = \lambda \dfrac{\partial w}{\partial G} \cdot \dfrac{1}{L^x(1+1/x)}, \end{cases} \tag{A7}$$

$$\Rightarrow \; r = \frac{\partial w}{\partial G} L \frac{x(x+2)}{(x+1)^2}. \tag{A8}$$

Moreover, equalisation of the marginal rates of productivity between private and public investment implies:

$$K = \frac{1-\alpha}{\beta} G. \tag{A9}$$

Wages offered by private firms are then:

$$w = \alpha A K^{1-\alpha} G^{\beta} L^{\alpha-1}. \tag{A10}$$

If we plug the derivative of (A10) in (A8), we find after simplification:

$$r = \alpha\beta^{\beta}(1-\alpha)^{1-\alpha} A G^{\beta-\alpha} L^{\alpha} \frac{x(x+2)}{(x+1)^2}. \tag{A11}$$

Of course, as in the previous cases, the equilibrium is such that $L = L^*$ (Lemma 2).

• $\beta < \alpha$ (Solow case). The steady-state is reached when $r = \rho$. The long-run level of accumulated capital is then:

$$\overline{\overline{K}} = \left(\frac{\alpha.(1-\alpha)^{1-\beta} A \beta^{\beta} L^{*\alpha} \dfrac{x(x+2)}{(x+1)^2}}{\rho} \right)^{1/(\alpha-\beta)}. \tag{A12}$$

Using equation (7) and after some manipulations, we can observe that $\overline{\overline{K}} < \overline{K}$. Note also that free-entry does not drive profits to zero because of the increasing returns (see Lemma 2 again). The steady-state is such that all instantaneous income is spent. Consequently, the steady-state land rent is:

$$\overline{\overline{R}} = \frac{\overline{\overline{Y}}}{L^*(1 + 1/x)}. \tag{A13}$$

Then, the long-run level of profit is given by:

$$\overline{\overline{\pi}} = A\left(\frac{\beta}{1-\alpha}\right)^{\beta} \overline{\overline{K}}^{1+\beta-\alpha} \left(\frac{x}{x+1} L^* - \alpha\beta\right). \tag{A14}$$

Profits can be either positive or negative. This stems essentially from the passive role played by the developers (they just maximise over *G*). We do not really bother about negative profits and we can assume that they can be financed through lump-sum transfers. Note that profits increase with the preference for land.

- $\beta = \alpha$ (Barro-Romer case). Using (A11) and the symmetric equilibrium condition $L = L^*$, we obtain a growth rate equal to:

$$\frac{\dot{z}}{z} = \frac{\alpha\beta^{\beta}(1-\alpha)^{1-\alpha} A G^{\beta-\alpha} L^{*\alpha} \dfrac{x(x+2)}{(x+1)^2} - \rho}{\sigma}. \tag{A15}$$

By a straightforward comparison with equation (6), we can see that, due to the terms in α and *x*, the growth rate is below the optimal one. After simplification, the profit is expressed by:

$$\pi = \frac{xK}{x+1}\left(\begin{array}{l} A\dfrac{\alpha^{\alpha}}{(1-\alpha)^{\alpha} L^{*1-\alpha}} - \dfrac{\alpha^{2+\alpha}}{(1-\alpha)^{\alpha}} AL^{*\alpha} \dfrac{x(x+2)}{(x+1)^2} \\ -\dfrac{1}{(1-\alpha)L^{*}} \dfrac{\alpha^{1+\alpha}(1-\alpha)^{1-\alpha} AL^{*\alpha} \dfrac{x(x+2)}{(x+1)^2} - \rho}{\sigma} \end{array}\right). \tag{A16}$$

Then, for $x = 0$, we get $\pi|_{x=0} = 0$ and:

$$\left.\frac{\partial \pi}{\partial x}\right|_{x=0} = \frac{K}{4}\left(A\frac{\alpha^{\alpha}}{(1-\alpha)^{\alpha} L^{*1-\alpha}} + \frac{\rho}{(1-\alpha)L^{*}\sigma}\right) \geq 0, \tag{A17}$$

which ends the proof. ■

CHAPTER 3

GROWTH AND IMPERFECT COMPETITION ON FACTOR MARKETS: INCREASING RETURNS AND DISTRIBUTION

Abstract: Although seldom modelled outside the monopolistic competition framework, market incompleteness and imperfect competition are central to the new growth theories. We propose here a strategic model of imperfect competition with endogenous growth and endogenous market structure where we focus on labour market issues. For growth to be possible, we assume increasing returns at the firm level. Due to heterogeneity on the labour market, the market structure is not degenerate. Then, because of increasing returns, short-run efficiency is maximised under monopoly and free entry implies too many firms in the market. However, in the long-run competition can generate growth through a distribution effect, whereas a monopoly leads to a zero-growth steady-state. Thus there is a trade-off between static and dynamic efficiency. This trade-off implies the existence of a growth-maximising degree of competition in our economy.

I. INTRODUCTION

What are the macro-economic effects of dynamic imperfect competition? This question is clearly at the heart of contemporary economic theory. Static micro-economic theory is overwhelmingly favourable to competition. A higher degree of competition is in general associated with higher welfare as can be found in any standard textbook. The typical policy recommendation is then to use government intervention to prevent market failures and to promote fair competition. However, when measured empirically, the static welfare gains of competition often appear very small (see Vickers [1995] for a complete survey on the question, including empirical references). Moreover, casual observation suggests that the relevance of the First Welfare Theorem is probably limited for practical matters given the pervasiveness of increasing returns and thus of imperfect competition. These two facts are consistent with the widespread opinion, outside the world of economists (!), that there can be "too much competition".

Despite these weak empirical micro-validations, economic organisation, based on markets, mitigated by some amount of state intervention (without making precise here how and how much) seems to produce, in the long-run, far better outcomes than alternative forms of economic organisation. The superiority of market economies must then lie in its "dynamics". Recently Nickell [1996], amongst others, has found a positive correlation between competition and productivity growth. This dynamic superiority is also stressed by many popular discussions of the issue. The present chapter seeks primarily to reconcile economic analysis with this popular view.

Our analysis is also motivated by the following theoretical problem. Typically, in endogenous growth models (Romer [1990]), the incentive to produce knowledge is given by the possibility to appropriate a monopoly rent (obtained by the patenting of new blueprints). For the economy to enjoy perpetual growth, knowledge must also have a public good aspect (which cannot be denied) for newcomers to be able to create new knowledge from the existing one. But the necessary assumptions for growth to be consistent in those models with some degree of competition are fairly restrictive since, in the R&D sector, new knowledge is added to the old *without its producer being able to internalise it at all.*

Although it is sensible from a macro perspective, this assumption of "zero dynamic internalisation" seems at odds with the situation in many sectors where only a small number of firms compete. These oligopolistic firms are in many cases likely to internalise, at least partly, the effects of their current R&D effort on their own knowledge and on the knowledge in their sector (see the survey of Tirole [1988, chap 13]). Typically in endogenous growth models, the production of new knowledge $\dot{K}_i$ by Firm i using L_i units of labour for research follows $\dot{K}_i = \gamma K L_i$ and it is not able to internalise the effects of its R&D, $\dot{K}_i$ on the total stock of knowledge (K) since it is atomistic. If this assumption is relaxed, the production function of knowledge might be written $\dot{K}_i = \gamma\left(K_i + K_{-i}\right)L_i$, where K_{-i} is knowledge of competitors. Consequently, the production function then exhibits increasing returns to scale at the firm level. As a consequence of partial dynamic internalisation (however small), the economy would be completely monopolised, at least in the R&D sector. In other words, with partial internalisation of the usual externality-like effects, competition is no longer sustainable in an endogenous growth model.[1]

A possible solution to this problem may be that increasing returns exist only at the level of the technology for a given product (that is we reintroduce some rivalry concerning the use of the accumulative factor). If consumers favour diversity, and if this taste for diversity is sufficiently strong, competition may again be possible despite increasing returns at the firm (technology) level (see Parente [1994]). We consider in our analysis another possibility where the labour force is heterogeneous (i.e., the mirror image of product differentiation in a factor market). Assuming differentiated skills as well as differentiated technologies, firms will find it attractive to hire workers whose skills are "close" to the technology they use, whereas the marginal product of some other "remote" workers can be very low. Despite increasing returns at the firm level, this leaves some room for competition. With this approach, heterogeneity of the labour force provides a natural "brake" to increasing returns so that our model can allow for growth in a fully-fledged imperfectly competitive framework. The justification of our focus on labour is that the labour market is widely

[1] The same type of problem arises in other types of growth models. In "Schumpeterian growth" models (Aghion and Howitt [1992] and [1994] or Grossman and Helpman [1991, chap. 5]), the same assumption of perfect diffusion of knowledge is necessary. Outside those two strands of literature, there are also some perfectly competitive models of growth (Romer [1986] or Lucas [1988]). These theories also require a very specific market imperfection for growth to occur and to be consistent with competition.

acknowledged to present major market imperfections. Moreover, labour is also a crucial part of the growth process, whereas existing models tend to focus on the product markets.

More specifically, we assume two factors of production: labour; and capital, which can be accumulated. Workers, endowed with labour during their youth, save part of their first period income. What is considered as an external effect with a large number of firms can be internalised partly because of the small number of competitors. As a consequence, for growth to be possible the production function must be homogenous of degree one in capital at the firm level. Labour is sufficiently differentiated that no firm has an incentive to hire all the labour force when the market is sufficiently large, despite there being increasing returns at the firm level. The market structure is endogenous. Depending on the size of the market, the equilibrium number of firms can vary between zero and infinity. In short, labour market heterogeneity is the feature we use to make imperfect competition and growth consistent with each other.

Starting from these micro foundations and making suitable assumptions about strategic behaviour, we explore the dynamic macroeconomic implications. Due to increasing returns at the firm level, the First Welfare Theorem is likely not to hold. In our model, free-entry implies too much competition and static efficiency is maximised under monopsony. In fact, to generate growth, increasing returns are needed and the economies of scale they generate are more important than the deadweight loss associated with the monopsony. However this monopsony is dynamically inefficient because it implies a low rate of investment. On the contrary, competition yields a distribution of income more favourable to growth when it increases the income of people with a high propensity to invest. *There is thus a trade-off between static and dynamic efficiency.* This trade-off implies the existence of a growth-maximising degree of competition. As a side result, we also show how labour market institutions can affect the growth rate.

These results can be contrasted with the (small) literature dealing explicitly with competition and growth. In relation to our topic, Smulders and Van de Kuldert [1995] have proposed a model of R&D and growth with imperfect competition, focusing on spillovers but they assume that firms do not behave strategically. Aghion, Harris and Vickers [1995] also focus on R&D and propose a Schumpeterian model of growth where the usual assumption of leapfrogging (or creative destruction) is relaxed in a duopoly setting. They show that when competitors are close to each other, more R&D is performed. From a different perspective Aghion, Dewatripont and Rey [1995] look at the connection between competition, growth and the internal organisation of the firm. They obtain positive dynamic effects of competition that they relate to the incentives given by the market. Those two papers, however, do not derive the market structure endogenously. By contrast, our market structure is endogenous and we consider strategic behaviour explicitly. More fundamentally, we take the view that the distribution of income is important and that competition yields positive dynamic effects when it induces a distribution of income that promotes investment and reduces unproductive rents.

The rest of the chapter is organised as follows. Section II describes the features of the model, while Section III solves the equilibrium under different market structures. Section IV discusses briefly other possible specifications. Finally, Section V contains some concluding remarks.

II. DESCRIPTION OF THE OLG MODEL

Population and preferences

We consider here a model of overlapping generations of worker/consumers (referred to hereafter just as consumers). They live for two periods, leave no bequest and are endowed with one unit of labour during their youth. There is a population of mass N_t for the generation born in t.[2] Our economy starts in period 0, when only youngsters with no endowment of the good are present. There is only one composite good in the economy and its price is normalised to one. This assumption allows us to abstract from all relative prices effects across sectors. This greatly simplifies the analysis and enables us to focus on labour market issues.[3] We denote C^y and C^o, the consumption of respectively young and old

[2] Time subscripts are ignored unless they are necessary.
[3] Besides, this assumption can be justified by a second-best argument since, if all sectors are imperfectly competitive in the economy, the distortions created by an oligopoly within a particular sector are ambiguous.

consumers and we assume the following utility function:

$$U\left(C_t^y, C_t^o\right) = ln\left(C_t^y\right) + b.ln\left(C_{t+1}^o\right), \text{ with } b > 0. \tag{1}$$

Production and consumption

Labour is differentiated and each worker is located on a circle of radius x and indexed by his address $j \in [0,2\pi]$ corresponding to his specialisation. The distribution of skills is uniform over $[0,2\pi x]$ with density one. Then, we observe $2\pi x = N$. At the end of their youth, people receive their first period income (as determined below). This income can be consumed or saved.

To produce the consumption good, two types of technologies are freely available. First, the traditional technology is simply

$$Y_T = \alpha n + k, \tag{2}$$

where n is the quantity of labour and k the quantity of capital. Whatever its characteristics, each unit of labour yields α units of the good, whereas each unit of capital yields one unit of the good. The traditional sector should be thought as a possible reserve sector where the wage is α and the interest rate is unity. There is also a continuum of increasing returns to scale industrial technologies located on the same circle as workers and indexed by $i \in [0,2\pi]$. This creates a matching problem in the labour market since the location of a given worker does not in general coincide with the location of the firm he is working for. A training cost must be incurred by any worker willing to work for a particular firm. Training is time consuming and thus reduces the supply of labour. This training cost increases in the distance between the worker's characteristics and the firm's specialisation (it is a direct analogue of transport cost in spatial economics). The reduction in labour supply is equal to: $c \times \text{distance}$. Then, expressed in terms of numéraire, the training cost for worker j working with firm i is equal to $W_t \times c \times Min\left(|i-j|, |i-j+2\pi|\right)$, whereby W_t is the wage. Without loss of generality we suppose that firm 1 locates at $i = 0$. We also assume that a fraction τ

of this training cost is born by the firm and that the rest is born by the workers.[4] The production function of firm i is assumed to be:

$$Y_i = AK_i L_i^{\beta}, \qquad \beta \leq 1, \tag{3}$$

where K_i is the capital which fully depreciates in one period and L_i is the net labour it uses.[5] Note that endogenous growth models often assume linear returns for all "private" factors at the individual level and linear returns for the reproducible factor (in the "growth-creating" sector) at the aggregate level. The gap between the two usually takes the form of an unspecified externality of the amount of capital on the individual production function (Romer [1986]) or is modelled as a complementarity between inputs to create new knowledge (Romer [1990]). Here, the production function is homogenous of degree 1+β at the firm level. However, due to the imperfect labour market (i.e., all types of labour are not perfect substitutes) and to the timing (described below), when the substitutability in the labour market is sufficiently low (i.e., training costs are sufficiently high), no firm will have an incentive to hire all the labour. Our assumption of increasing returns generates both imperfect competition and a potential for growth since the production function is homogenous of degree one in capital.[6] However, our specification of the production function completely ignores spillovers and purposeful investment which are treated in the next chapter.

We can now solve the model for the saving decision. If W_t^n is the net income of a consumer, his program is:

$$\begin{aligned} &MAX:\ ln\left(C_t^y\right) + b.ln\left(C_t^o\right) \\ &s.t.\ C_t^y = W_t^n - S_t \quad \text{and} \quad C_t^o = S_t r_{t+1}. \end{aligned} \tag{4}$$

[4] We assume τ<1. Otherwise, the existence of a pure-strategy Nash-equilibrium is not guaranteed due to a discontinuity in the best-response correspondence.

[5] Note that our productive factors could be called respectively specific human capital and general human capital (e.g., managerial abilities) instead of labour and capital without changing the set-up.

[6] No explicit R&D process is considered as we consider that in a given sector productivity improvements may be endogenous to the sector (R&D, learning) as well as exogenous to the sector (diffusion from other sectors, adaptation, cheaper inputs). For the sake of simplicity we restrict ourselves to a one-sector model with a production function like (3). See Grossman and Helpman [1994] for an extended discussion on the differences and the similarities between growth models based on research and development and models using learning-by-doing.

We find that: $C_t^y = \dfrac{W_t^n}{1+b}$ and $S_t = \dfrac{b}{1+b} W_t^n$. (5)

Timing of the game

At each period, for each generation the same game is repeated. The timing is as follows:

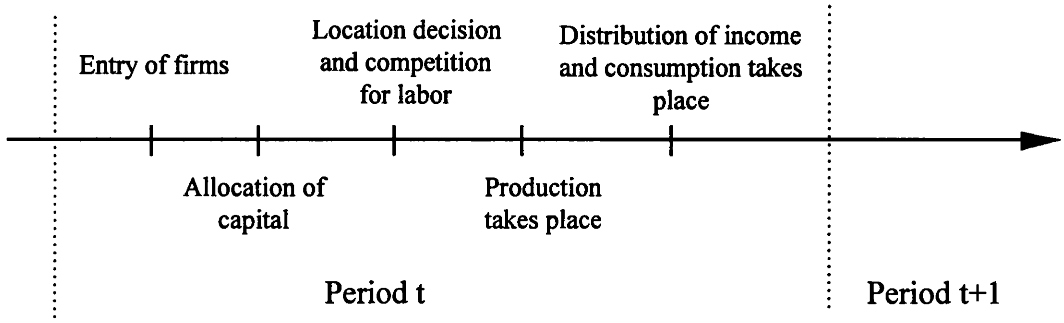


Figure 1

At the beginning of each period, firms willing to operate on the market enter freely. As long as entry is profitable, new firms keep entering. Once expected marginal profits are zero, the number of firms remains stable. Then the savers allocate their capital among firms, assuming that firms cannot commit a given return and that the savings received by firms are not observable until the end of the period. Indeed, firms would always have an incentive to renege *ex-post* on the promised return to capital because there is insufficient income to pay all factors at their marginal productivities (increasing returns). Thus capital is treated as a residual claimant. Because of this impossibility of commitment, savers cannot be given more than residual profits. The assumption of independence is here to prevent any co-ordination of the savers.[7] In a third stage, firms locate and competition for labour takes place. The youngsters make their working decision (i.e., either work with the traditional technology or become an industrial worker and choose a particular firm to work for).

Due to the mobility of capital, the market is contestable at each period so that firms can

[7] However and as usual, a monopolist might be much more profitable than two firms. Savers have then an obvious incentive to co-ordinate. This important question of financial economics is resolved here by our observability assumption.

maximise only current profits. The profit function is thus:

$$\Pi_i \equiv AK_i L_i^{\beta} - W_i L_i - \tau W_i c\, I_i^2/4, \tag{6}$$

where W_i is the wage offered by the firm i and I_i is the measure of the area over which firm i hires workers. Profit is equal to the production minus the wages and the share of the training costs paid by the firm. Of course, due to the training cost, $L_i \leq I_i$. The profit given by equation (6) is distributed to the stockholders of firm i. The return to capital in firm i is $\Omega_i \equiv \Pi_i / K_i$.

III. SOLUTION OF THE OLG MODEL

The monopoly

Before trying to solve the whole game, we focus first on a very simple situation. The monopoly (or properly speaking, monopsony) will be used as a benchmark case. We suppose either that we are in a situation of natural monopoly or that only one firm was given the right to operate in the economy (i.e., the first stage of the game is ignored). This monopolist can be choose to develop more than one technology if she wishes. We also assume in this subsection that the monopolist is unable to discriminate among workers with different skills and that she has all the bargaining power vis-à-vis the workers. This assumption is the analogue of mill-pricing in spatial economics (see Anderson *et al.* [1992]).

Proposition 1 *There is a minimal market size under which no firm can operate.*

Proof Since savers always have the alternative to use their savings with the traditional technology, we need $\Omega_i \geq 1$. From our assumptions, the relation between the market area and the net labour supplied to a monopolist is:

$$L_i = 2\int_0^{I_i/2} (1 - cx)\,dx = I_i - cI_i^2/4. \tag{7}$$

We need $2/c > I^*$ for the productivity of the marginal worker to be positive. The profit function becomes:

$$\Pi_i = AK_i\left(I_i - cI_i^2/4\right)^{\beta} - W_i\left(I_i - cI_i^2/4\right) - \tau W_i c I_i^2/4. \tag{8}$$

From equation (8), we can infer easily the existence of a minimal size of the labour market. If $N \le 1/A$, then $I_i < 1/A$ and $Y_i < K_i$ so that $\Omega_i < 1$ and no industrial firm can be set-up. Similarly, if the training cost c is very high, whatever the market size, very few workers can be hired for any technology, so that no firm can profitably operate. ■

For simplicity, we will suppose below that β=1. The essence of the results remains the same when $\beta < 1$. The next claim states that:

Proposition 2 *There exists a finite profit maximising market area for the technology used by the monopolist.*

Proof Note first that the marginal worker is such that he is indifferent between working for the monopolist or working with the traditional technology. Thus:

$$W_i = \frac{\alpha}{1-(1-\tau)cI_i/2}. \tag{9}$$

Using this last equality, the monopolist maximises her profit with respect to W_i. However, it is technically equivalent to optimising over I_i. Then the monopolist sets her employment area $I^*(K)$ such that $\partial\Pi_i/\partial I^* = 0$. In order for the solution to be a maximum, we need $2/((1-\tau)c) > I^*$, which is satisfied since her marginal worker has a positive productivity. Consequently, the profit maximising employment area $I^*(K)$ must satisfy:

$$\left(AK_i - W^*\right)\left(1 - c.I^*/2\right) - \frac{\tau.c.I^*}{2}W^* - \frac{\partial W^*}{\partial I^*}I^*\left(1 - c(1-\tau)\frac{I^*}{4}\right) = 0, \tag{10}$$

where: $W^* = \dfrac{\alpha}{\left(1 - c(1-\tau)\,I^*/2\right)}$ and $\dfrac{\partial W^*}{\partial I^*} = \dfrac{\alpha(1-\tau)c}{2\left(1 - c(1-\tau)\,I^*/2\right)^2}$.

Since the *lhs* of (10) is monotonic in *I*, whatever *K* and $c>0$, $I^*(K)$ is unique and finite. ■

Moreover we can write:

Proposition 3 *The monopolist introduces only one technology and maximises short-run (or static) productive efficiency.*

Proof Assume the monopolist introduces *n* technologies. Her profit is:

$$\Pi_n = n\left(A\frac{K}{n}L_i - WL_i - \tau c W I_i^2/4\right). \tag{11}$$

Obviously, $\Pi_1 > \Pi_n$, $\forall n > 1$. Moreover, with *n* technologies, production is equal to $Y_n = n(AL_i\,K/n)$. If only one technology is used with the same population, production is equal to $Y_1 = AKL_i + \alpha n I_i$ which implies immediately $Y_1 > Y_n$, $\forall n > 1$. The monopoly consequently maximises production. ■

Proposition 4 *The monopoly leads to a steady-state with no growth.*

Proof in Appendix 1. If we define g_I as the asymptotic growth rate reached with the monopoly, our proposition means that $g_I = 0$. ■

Remark 1 Proposition 3 states that the monopoly is efficient in the short run (or statically). This result can be understood easily. Since firms face increasing returns, then it is no surprise that the First Welfare Theorem does not hold. In our case, the increasing returns are so strong (and they need to be so for the model to be able to generate growth) that static efficiency is achieved by using only one technology. Of course, using less extreme assumptions would give us results less favourable to monopoly, but it would not re-establish the First Welfare Theorem.[8]

Remark 2 Proposition 4 shows that a monopoly induces strong dynamic inefficiencies. The production function is such that endogenous sustained growth is possible, but due to the market power of the monopolist, no growth occurs. The intuition is as follows. The monopolistic behaviour implies that the young are given their reservation wage. Capital accumulation is then limited by the low wages. Of course, if capital becomes abundant for some reason, the monopoly has an incentive to hire more labour. If more labour is hired, higher wages must be given (since discrimination is impossible in the labour market, the intra-marginal worker receives a rent). However, the monopoly wage rises less than equiproportionately with capital (wages increase less than capital). A given increase of capital leads to a smaller increase of wages, which in turn induces an even smaller increase in the capital accumulated. So sustained growth is impossible in a monopoly with our dynamic structure.

Remark 3 What would the results be if we allowed for wage discrimination? They would be reinforced since discrimination enables the monopolist to pay lower wages to the intra-marginal workers. Thus, accumulated capital is bounded again and a lower steady-state is reached. Consequently giving rents to intra-marginal workers implies a higher (but non-increasing) level of production with a monopoly.

The duopoly

Instead of considering only one firm in the market, whatever the population size, we now allow for two firms to compete on the labour market.

Typical situations are:

[8] The partial instead of complete internalisation of increasing returns would just reinforce our argument in favour of the desirability of monopolies who are likely to internalise more than competitive firms.

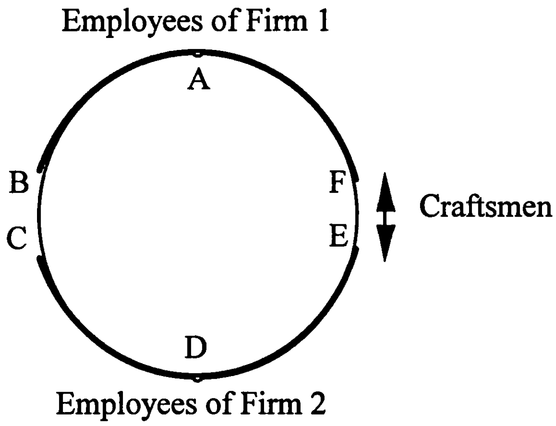


Figure 2a

(abundant labour, high c for instance)

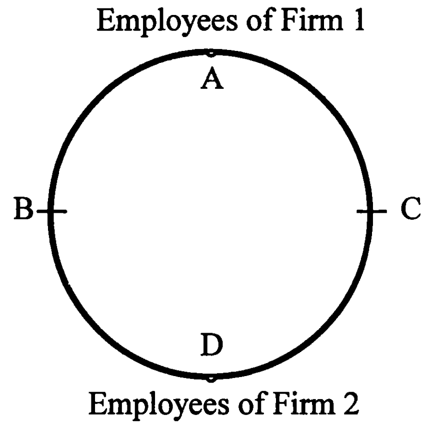


Figure 2b

(scarce labour, low c for instance)

In figure 2a, Firm 1 uses the technology in A and her boundaries are B and F. Firm 2 has boundaries at C and E. The workers with a specialisation between B and C and F and E use the traditional technology. In this case, the two firms are local monopolists for all the workers they use. In figure 2b, on the contrary, the firms have connected employment areas and compete for some workers (at least those located in B and C). Due to our assumption concerning entry, two firms are present, whenever the duopoly is profitable. We assume in the rest of this subsection that the market size is such that only two firms are present.

Proposition 5 *With a duopoly, the economy enjoys perpetual growth as long as firms cannot act as local monopolists for all their workers.*

Proof To demonstrate this proposition and to characterise the situations for which the duopoly occurs, we need to solve the game. The last stage is the labour market equilibrium. When competition is effective between the two firms, marginal workers in equilibrium are indifferent between firms i and j. The indifference condition reads:

$$W_i\left(1-\frac{cl_i(1-\tau)}{2}\right)=W_j\left(1-\frac{cl_j(1-\tau)}{2}\right) \Rightarrow l_i=\frac{2}{(1-\tau)c}\left(1-\frac{W_j}{W_i}\left(1-\frac{cl_j(1-\tau)}{2}\right)\right). \quad (12)$$

Thus the profit function can be written:

$$\Pi_i = \left\{ AK_i \left[1 - \frac{1}{2(1-\tau)} \left(1 - \frac{W_j}{W_i} \left(1 - \frac{cI_j(1-\tau)}{2} \right) \right) \right] - W_i - W_j \left(1 - \frac{cI_j(1-\tau)}{2} \right) \right\}$$
$$\times \frac{1}{(1-\tau)c} \left(1 - \frac{W_j}{W_i} \left(1 - \frac{cI_j(1-\tau)}{2} \right) \right). \tag{13}$$

Using, $I_i + I_j = 2\pi x$, the firm i then sets W_i such that $\partial \Pi_i / \partial W_i = 0$. This condition is necessary and sufficient for maximisation. Appendix 2 shows that the second-order condition is well-behaved. This solves the Nash equilibrium for the wage competition. On a circle, the location stage is trivial: firms locate at each other's opposite. Because of independent allocation by savers, capital is allocated equally among firms ($K_i = K_j = K/2$). At the symmetric equilibrium we find that:

$$W = AK\left(1 - \frac{c\pi x}{2}\right)\left(1 - \frac{c\pi x(1-\tau)}{2}\right) \Big/ \left[1 + \left(1 - \frac{c\pi x(1-\tau)}{2}\right)^2\right]. \tag{14}$$

Thus the economy enjoys a growth rate g_2:

$$g_2 = \frac{b\pi x}{1+b} \frac{A\left(1 - \frac{c\pi x}{2}\right)\left(1 - \frac{c\pi x(1-\tau)}{2}\right)\left(1 - \frac{c\pi x(1-\tau)}{4}\right)}{1 + \left(1 - \frac{c\pi x(1-\tau)}{2}\right)^2} - 1. \tag{15}$$

Then we must check that profits are positive in order to solve step 1. We can plug the equilibrium wage (14) into the profit function. After simplification, we obtain the following existence condition for the duopoly:

$$A\pi x\left(1 + \frac{c\pi x(1-\tau)}{4} - \frac{(c\pi x)^2(1-\tau)(1+3\tau)}{8}\right) \geq \frac{1}{2}\left(1 + \left(1 - \frac{c\pi x(1-\tau)}{2}\right)^2\right). \tag{16}$$

If τ is close to one, we obtain a much simpler expression:

$$A\pi x \geq 1. \tag{16'}$$

The equilibrium wage given by (14) will be observed as long as it is larger than the minimum wage of the local monopolist. In other words, we need the net wage given by (14) to be superior to the wage offered by a monopolist:

$$W > \frac{4\alpha}{4 - c\pi x(1-\tau)}. \tag{17}$$

Otherwise each firm acts as a local monopolist for all her workers and we are trivially driven back to the monopoly case where no growth occurs. ■

Remark 1 Contrary to what happened in the previous sub-section, when there is some competition on the labour market, growth is made possible. This result seems at odds with "casual extrapolation" of standard endogenous growth models. In standard models (e.g., Romer [1986] or [1990]), competition is efficient in the short-run, but induces a sub-optimal rate of growth in the long-run. One might think that introducing imperfect competition with strategic behaviour would reduce static efficiency but increase dynamic efficiency, since imperfectly competitive firms have some market power that enables them to internalise the dynamic effects of their investments. This reasoning is flawed for two reasons. First, to generate imperfect competition and for imperfectly competitive firms to be able to generate growth, increasing returns are needed. In that case, the First Welfare Theorem does not hold anymore and less competition can actually increase static efficiency. Second, income distribution plays an important role in the growth process (see Phelps [1961], Bertola [1993, 1996] or Uhlig and Yanagawa [1996]). In our model, more competition means a distribution of income that favours people who invest in the accumulative factor (a similar effect is present in Bean and Pissarides [1993], but there, it stems from a higher bargaining power of the workers). What is new here is that the income distribution is not given directly by the parameters of the production function, but by the competitive process. So, through a distribution effect, more competition increases the dynamic efficiency of the economy (see below for discussions of the welfare and of alternative dynamic structures).

Remark 2 If the initial level of capital is too small with respect to the productivity of the alternative technology, growth can never occur, although it is feasible. It may also happen that the productivity of the industrial technology A is sufficiently high to be able to generate growth, but too low for the profits to be positive. Conversely, it is possible that with a high initial level of capital, competition is possible but implies negative growth. Then the economy returns to a monopoly situation.

A large number of firms

We now allow for a large number of firms to compete on a large labour market (high x) and derive some conclusions concerning the asymptotic growth rate when the market becomes very large.

Proposition 6 *The growth rate is increasing with τ, decreasing with c, and converges towards g_∞ as defined in equation (20). The asymptotic rate g_∞ does not maximise growth.*

Proof To demonstrate this, note first that the equilibrium wage is determined as previously. We observe then:

$$W_t = \frac{2Ak_t(1-cI_t)(1-c(1-\tau)I_t)}{1+(1-c(1-\tau)I_t)^2}, \tag{18}$$

where k_t is the capital in each firm ($\sum k_{i,t} = K_t$) and I_t is the employment of each firm. Moreover free entry drives profits to their minimum. Consequently $\Pi_{i,t} = k_{i,t}$. After simplification, we get:

$$W_t = k_t\left(AI_t(1-cI_t/2)-1\right)/\left[I_t(1-c(1-\tau)I_t/4)\right]. \tag{19}$$

The system composed of equations (18) and (19) solves for both the equilibrium wage and the equilibrium size of each firm. Note that this system is equivalent to (14) and (16) with equality. No simple closed-form solution can be obtained for the size of the firms.

However we can write the implicit growth rate of the economy and perform some comparative statics on the growth rate. After manipulation, we get:

$$g_{\infty} = \frac{2b}{1+b}\frac{AI_t\left(1-cI_t\right)\left(1-c(1-\tau)I_t\right)}{1+\left(1-c(1-\tau)I_t\right)^2} - 1. \tag{20}$$

The comparative static results are obtained directly by using (18), (19) and (20). The last part of the proposition is straightforward. Note that g_{∞} is also the growth rate obtained with the smallest market size for which a duopoly is present and profitable (the strategic behaviours are the same and the conditions (14) and (18) are equivalent). ■

Remark 1 Now suppose that the labour market is expanding steadily (N is increasing). We can look at the timing of entry for new firms and how the growth rate evolves. If we summarise what has been seen up to now, for a very small economy, only the traditional activity is performed. A monopolist enters for a size I_1 of the labour market. It is such that expression (8) is set to zero with a wage determined according to (9). This monopoly induces bounded capital accumulation. Still, the duopoly cannot exist due to tight competition to attract workers. Competition decreases when the labour market becomes larger so that a duopoly can be sustainable. The second firm enters for a size $2I_2$ such that $\Pi_i = K_i/2$ with a wage determined according to (14). In the duopoly, an increase of the market size N has ambiguous effects on the growth rate g. The first effect is that a larger economy has a bigger labour force, so that productive efficiency is higher. On the contrary, competition is relaxed and it leads to a lower distributive efficiency.[9] Using equation (15), it can be seen that the growth rate within a duopoly increases and then decreases when the market grows larger (if we ignore further entry). For even larger labour markets, two firms, each acting as a local monopolist for all their labour force, may be present. In that case, we are back to the monopoly regime with bounded capital accumulation. With free-entry, due to the equivalence between (14) and (18), the n^{th} competitor enters for an economy whose size is nI_2. The long-run growth rate may vary as follows:

[9] For instance if we set the share of the training cost born by the firm τ arbitrarily close to 1, the growth rate of a duopoly is maximised for $\pi x = 1/c$ and two firms are present for this size of the economy if $\pi x > 1/A$ (see equation (16')).

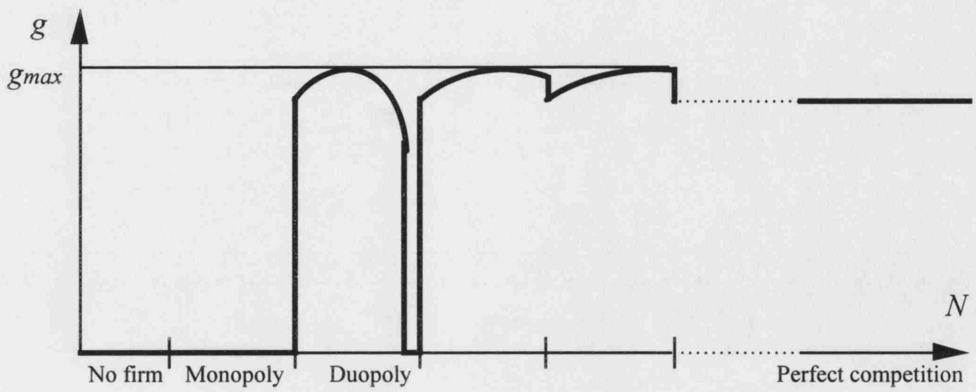


Figure 4

There is then a growth-maximising intensity of competition. Assume a large market, which is initially such that all firms make normal profit. If competition is exogenously relaxed, then the share of production distributed to the young and subsequently invested decreases, but static efficiency (i.e., total production) increases. Overall, investment increases. However, if we relax competition too much, each firm will become a local monopolist for all her labour force and long-run growth will not take place.

Remark 2 Unsurprisingly, the higher the training cost, the lower is the growth rate. More interestingly, as in Thisse and Zénou [1996], the higher the share of the training cost born by the firm, the higher is the growth rate. The intuition for the result is the following. The marginal worker is exerting a pecuniary externality on the wage paid by the firm. If the firm pays a higher share of the training cost, it diminishes the rent paid to the intra-marginal workers. This rent has two different effects. The first effect is intertemporal: lower wages induce lower savings, and therefore a lower rent has an adverse effect on growth. The second effect is an efficiency effect, the lower the rent, the higher the incentive to hire new workers and to raise productive efficiency. In a competitive environment, only the second effect matters (efficiency is increased), since the no-extra-profit condition annihilates the first effect. Then, net wages as well as savings are higher and capital as well as production increase faster.

Remark 3 Despite the specificity of the model, we do believe that this tradeoff (which is opposite to the traditional Schumpeterian tradeoff) has some relevance. Clearly, in a real multi-sector economy, the monopoly can lose some of its short-term appeal because of the usual deadweight loss involved by relative price effects, although this effect is far from being certain if all sectors are imperfectly competitive. It is likely however that free entry will remain inefficient as is usual in differentiation models (see Anderson *et al.* [1992]). The terms of our tradeoff may be altered but, keeping the same dynamic structure, competition is still likely to be desirable because of its dynamic effects and not because of its short-run efficiency. Given that growth is generated without specifying any sophisticated behaviour from either producers or consumers, an extreme interpretation of the results obtained here is that *growth is a by-product of competition in an n^{th}-best world.*

Remark 4 The welfare analysis is ambiguous but straightforward. Clearly, a higher growth rate is harmful for the current old generation since it is achieved with a reduction of the current returns to capital. However if this higher growth rate is obtained by keeping the same market structure, it is beneficial for the all future generations since it increases their income and their welfare unambiguously. If the market structure is modified (e.g., more firms), the higher growth rate will also mean a lower production for the current period and a different structure of wages. Thus, this may hurt the current young generation as a whole as well as its immediate successors (but after some time all the generations will benefit from this). Because of this conflict between generations, the Pareto criterion is not satisfied. When using a utilitarian criterion, as usual the optimal growth rate will be inversely related to the rate of time preference used by the central planner.

Competition between discriminating firms

We suppose now that firms are now able to discriminate among workers depending on their specialisation.

Proposition 7 *The growth rate is higher when firms can discriminate.*

See Appendix 3 for a proof. ■

From the previous results, we can draw the following picture describing the relationship between the asymptotic growth rate and the market structure:

ERRATA

1/ Erratum to page 95. The following paragraph replaces the first paragraph page 95.

Remark 1 This proposition is a bit counter-intuitive since it states that increasing the power of firms over workers induces a higher growth rate, whereas we saw that the growth rate depends on the level of wages. The explanation is that, with discriminatory wages, the total wage bill increases due to an increase in efficiency of the firms. With discriminatory wages, all intra-marginal workers are paid below their marginal productivity and only the marginal worker is paid at his marginal productivity. So the firm expands her market area until her marginal worker is paid at his marginal productivity. By contrast, with non-discriminatory wages, the firm hires workers until the productivity of the marginal worker is equal to the marginal increase of the total wage bill (remember that with non-discriminatory wages, intra-marginal workers exert a negative pecuniary externality and benefit from the wage given to the marginal worker). So competing firms are bigger with discriminatory wages. Thus, in a competitive environment, discriminatory pricing induces higher productive efficiency (unlike in the monopoly case). It is therefore preferable from an efficiency perspective.

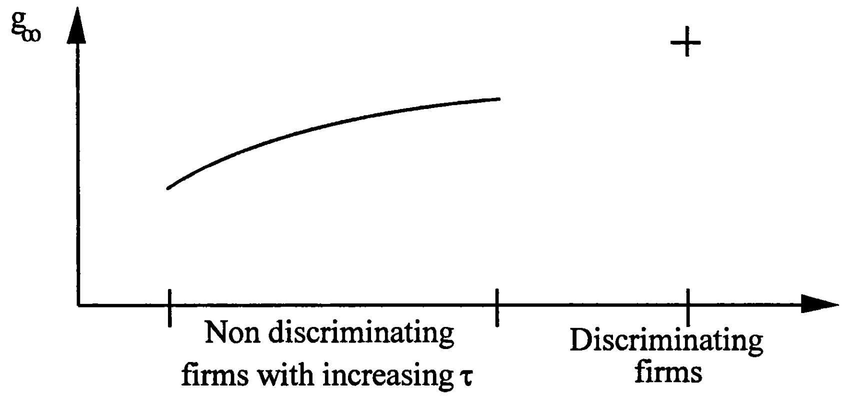


Figure 5

Remark 1 This proposition is a bit counter-intuitive since it says that increasing the power of firms over workers induces a higher growth rate, whereas we saw that the growth rate depends on the level of wages. The explanation is that, with discriminatory wages, we avoid the negative pecuniary externality of intra-marginal workers. All intra-marginal workers are paid below their marginal productivity (and the difference finances the fixed cost) and only the marginal worker is paid at his marginal productivity. So the firm expands her market area until her marginal worker is paid at his marginal productivity. On the contrary, with non-discriminatory wages, the firm hires workers until the productivity of the marginal worker is equal to the marginal increase of wages for *all* her workers. Thus, in a competitive environment, unlike in the monopoly case, discriminatory pricing induces higher productive efficiency. It is therefore preferable from an efficiency perspective.

Remark 2 However, if an economy switches from non-discriminatory to discriminatory wages, inequality may rise considerably with the rise in efficiency. Nonetheless it is possible to argue that some lump-sum transfers can restore a more egalitarian situation, but it would be hard for a social planner to compensate the losers, since the "proximity of the skills to a technology" is hard to measure.

Remark 3 Now, one can think of industrialisation as a switch from the traditional to the industrial technology resulting from a larger market or a smaller training cost. The idea of a larger market can be justified historically by a fall in transport costs. Better transportation may favour mobility in the labour market, the transport of physical goods as well as a better circulation of information. Smaller training costs can result from a change in the technological environment with the possibility of mass production that reduces the role of specific skills.[10] In any case, with our model, industrialisation begins with the formation of a monopoly. This monopoly induces a higher level of production and a slightly higher level of wages without turning all this into sustained growth. This monopoly also implies a rise in inter- and intra-generation inequalities. If the rise in market size or the fall in training costs continues, competition endogenously arises and promotes sustained growth, much higher wages and possibly a reduction in inequalities.[11] Thus, a steady increase in the market size generates a Kuznets curve pattern with first rising inequalities associated with the beginning of industrialisation and then decreasing inequalities generated by competition (see the discussion of these well-known stylised facts in Williamson [1991]). In short, *competition acts as a trickle-down mechanism.*

IV. OTHER SPECIFICATIONS

Our aim in this section is to show that the specification of utility may affect dramatically the dynamic properties of the model. Although the case explored above offers useful insights, it should be seen primarily as a theoretical benchmark. Keeping the same initial model, we now assume that the consumers are infinitely rather than just finitely lived, leaving all static features unchanged (they supply one unit of labour in every period and can accumulate capital over time). We also assume that firms are managed by successive generations of managers or that the rate of time preference is high enough so that, despite the infinite horizon, competition between firms is effective (i.e., no folk theorem applies).

We now get the result:

Proposition 8 *The growth rate is maximised with a discriminatory monopoly.*

Proof The new consumer's program is:

[10] Mokyr [1990] puts the emphasis on this argument.
[11] "After 1750, Anglo-Saxon economies tended to be more competitive than others, cartels and formal barriers to entry were far more common on the continent". Mokyr [1990].

$$U = \sum_{t=0}^{+\infty} \frac{C_t^{1-\sigma}}{1-\sigma} b^t \tag{21}$$

$$s.t. \quad K_{t+1} = \Omega_t . K_t + W_t - C_t .$$

The solution of this program is such that $K_{t+1}/K_t = C_{t+1}/C_t = (\Omega b)^{1/\sigma}$. From this last equation, the rest of the proof is straightforward. The growth rate is positively related to the return of capital. Since, it in turn is equal to the rate of profit, the monopoly is more efficient than competition in the long-run as well as in the short-run. ■

The previous results are thus reversed. The reason is fairly simple. In this extreme case of infinitely-lived agents, savings do not depend on wages, but on capital income; that is the distributive effect disappears. Furthermore, we supposed that collusion among competing firms was impossible. If firms could collude, the degree of collusion could positively influence the growth rate.

Note that the assumption used here is another possible working assumption which can help us understand the previous results. In fact the extreme OLG structure of the benchmark case and the infinitely-lived agent with exponential discounting can be understood as two limit cases of a more general model (see Bertola [1996]).[12] The growth literature has mainly focused on the infinitely-lived agent case. However, this assumption does not seem to receive wide empirical support and has been questioned directly by Wilhelm [1996] or indirectly by Uhlig and Yanagawa [1996]. The distribution effects of our OLG structure could also be derived from an imperfect capital markets argument.

V. CONCLUSION

In this chapter we proposed an analytical framework that combines imperfect competition and growth. We used a standard industrial organisation set-up with increasing returns.

[12] In a more general model where agents live longer and receive income for reproducible as well as non-reproducible factors during all their life, Bertola [1996] shows that qualitatively the dynamic structure remains the same as in our benchmark case provided the intertemporal elasticity of substitution is sufficiently low.

Increasing returns generate both growth and imperfect competition. They also imply that the First Welfare Theorem no longer holds. In our model, we reached the result that in the short-run (i.e., statically), the monopoly maximises productive efficiency. In the long-run, the implications are different. The market structure has an impact on income distribution. In turn, income distribution is crucial since it determines investment. In our model, increased competition on the labour market increases the income received by young agents endowed only with non-renewable specialised labour and decreases rent-seeking. With a higher first period income, they invest more which will increase production in the next period. So in our model, increasing competition (e.g., through a larger market) decreases static efficiency (the volume of production) but increases dynamic efficiency (the share of production devoted to investment), so that there exists a growth-maximising degree of competition.

Then, the (not so original) macroeconomic implication consists in looking for the propensity to enhance growth. In practice, when growth stems from physical capital accumulation, the "growth enhancers" are the savers. This might imply wage moderation when the "capitalists" have a higher propensity to save than the workers. On the contrary, if growth is driven by human capital investment, a distribution favouring young people may be growth enhancing. Although we have to remain cautious in the interpretation, this argument is consistent with the explanations developed by Amsden and Singh [1994] and Eichengreen [1996] for the respective cases of growth in east Asia and Europe after World War II. These situations might have involved a high growth rate, because profits were high (due to weak competition), kept within the firms and invested.

Note also that, in our competitive process, the highly mobile and differentiated factor enjoys a high share of the income because of the competition for it. So if capital becomes more mobile and more differentiated for some reason (e.g., market integration, lower transport costs, higher compatibility between inputs), the competitive process is modified. This also alters factor distribution and thus the growth rate. For instance, it may be that after 1970 the developed economies shifted from a growth process generated by physical capital accumulation to a growth driven by human capital and knowledge accumulation.[13]

[13] And governed by some transitional dynamics after the massive destructions and the scientific discoveries of W.W.II.

At the same time, lower transport costs led to a higher mobility of physical capital and more competition and thus higher returns for physical capital at the expense of other factors. Thus, these changes in the income distribution, favouring physical capital reduced simultaneously the incentive to invest in human capital and knowledge. The net effect was to reduce the growth rate. Of course, more sophisticated models and empirical verifications are needed, but this preliminary conclusion is underlines the possible gains from a further inquiry into the dynamics of market structure and income distribution.

APPENDIX 1 PROOF OF PROPOSITION 4

The idea of the proof is to show that the amount of capital remains bounded with a monopoly. We know from (9) that we need $I^* < 2/c$ and that it implies $W^* < \alpha/\tau$.

• Then, for all $\tau \in (0,1)$, the sum of distributed wages remains bounded. Thus capital accumulated per capita remains bounded and consequently steady growth is impossible.

• If $\tau = 0$, then I^* is such that:

$$\left(AK - \frac{\alpha}{\left(1 - c I^*/2\right)}\right)\left(1 - c.I^*/2\right) - \frac{\alpha c}{2\left(1 - c I^*/2\right)^2} I^*\left(1 - c\frac{I^*}{4}\right) = 0. \tag{A1}$$

This third degree equation has only one real solution which is:

$$I^* = \frac{6AK - \alpha}{3AKc} - \frac{\alpha^2}{c\left[\alpha\left(54A^2K^2 + \alpha^2\right) - 3\alpha AK\sqrt{6\left(27A^2K^2 + \alpha^2\right)}\right]^{1/3}} - \frac{\left[\alpha\left(54A^2K^2 + \alpha^2\right) - 3\alpha AK\sqrt{6\left(27A^2K^2 + \alpha^2\right)}\right]^{1/3}}{3AKc}. \tag{A2}$$

It is then immediate that if K becomes large, α becomes negligible with respect to K. Then we could approximate I^* by:

$$I^* \to \frac{2}{c} - \frac{\left(2 - \sqrt{2}\right)^{1/3}}{(AK)^{1/3} c} - \frac{\alpha^2}{3\left(2 - \sqrt{2}\right)^{1/3} (AK)^{2/3} c}. \tag{A3}$$

Moreover, the capital accumulation equation states that

$$K_{t+1} = \frac{b}{1+b} W_t I^*(K_t) + \left(N - I^*(K_t)\right)\frac{\alpha b}{1+b} \tag{A4}$$

$$\Rightarrow K_{t+1} \to \frac{b}{1+b}\left(\frac{2}{c}-\frac{\left[2-\sqrt{2}\right]^{1/3}}{A^{1/3}K_t^{1/3}c}\right)\times\left(\frac{2A^{1/3}K_t^{1/3}\alpha}{\left[2-\sqrt{2}\right]^{1/3}}\right)+\left(N-I^*(K_t)\right)\frac{\alpha b}{1+b}+o\left(K_t^{-2/3}\right) \quad \text{(A5)}$$

Then we find that $\lim\limits_{K\to+\infty} K_{t+1}/K_t = 0$. Thus capital cannot grow unbounded. ■

APPENDIX 2 SECOND-ORDER CONDITION FOR THE PROFIT FUNCTION

$$\frac{\partial\Pi_i}{\partial W_i}=\left(AK_i-W_i\right)\frac{\partial I_i}{\partial W_i}\left(1-\frac{cI_i}{2}\right)-I_i\left(1-\frac{cI_i}{4}\right)-\frac{\tau c I_i}{2}W_i\frac{\partial I_i}{\partial W_i}-\frac{\tau c I_i^2}{4}, \quad \text{(A6)}$$

with $$\frac{\partial I_i}{\partial W_i}=\frac{2}{\left(W_i+W_j\right)c(1-\tau)}-\left(\frac{2\left(W_i-W_j\right)}{\left(W_i+W_j\right)^2 c(1-\tau)}+\frac{W_j\left(I_i+I_j\right)}{\left(W_i+W_j\right)^2}\right) \quad \text{(A7)}$$

=> $$\frac{\partial I_i}{\partial W_i}=\frac{2W_j}{\left(W_i+W_j\right)^2}\left(\frac{2}{c(1-\tau)}-\left(I_i+I_j\right)\right). \quad \text{(A8)}$$

Then we can write:

$$\begin{aligned}\frac{\partial^2\Pi_i}{\partial W_i^2}&=\left(AK_i-W_i\right)\left[\frac{\partial^2 I_i}{\partial W_i^2}\left(1-\frac{cI_i}{2}\right)-\frac{c}{2}\left(\frac{\partial I_i}{\partial W_i}\right)^2\right]-2\frac{\partial I_i}{\partial W_i}\left(1-\frac{cI_i}{2}\right)\\&-\frac{\tau c}{2}\left[\left(\frac{\partial I_i}{\partial W_i}\right)^2 W_i+I_i\frac{\partial^2 I_i}{\partial W_i^2}W_i+2I_i\frac{\partial I_i}{\partial W_i}\right]\end{aligned} \quad \text{(A9)}$$

The first and the second term of the *rhs* are obviously negative (otherwise the firm acts as a local monopolist and sets its wages as in the previous subsection). It remains to show that

$$\left(\frac{\partial I_i}{\partial W_i}\right)^2 W_i+I_i\frac{\partial^2 I_i}{\partial W_i^2}W_i+2I_i\frac{\partial I_i}{\partial W_i}\geq 0. \quad \text{(A10)}$$

A straightforward injection of the derivatives (A.6) in the previous expression implies that (A.10) is equivalent to:

$$\left(\frac{2W_j}{\left(W_i+W_j\right)^2}\left(\frac{2}{c(1-\tau)}-\frac{\left(I_i+I_j\right)}{2}\right)\right)^2+\left[\frac{2\left(W_i-W_j\right)}{(1-\tau)c\left(W_i+W_j\right)}+W_j\frac{\left(I_i+I_j\right)}{\left(W_i+W_j\right)}\right]\times$$
$$\left[-\frac{4W_iW_j}{\left(W_i+W_j\right)^3}\left(\frac{2}{c(1-\tau)}-\frac{\left(I_i+I_j\right)}{2}\right)+\frac{4W_j}{\left(W_i+W_j\right)^2}\left(\frac{2}{c(1-\tau)}-\frac{\left(I_i+I_j\right)}{2}\right)\right]\geq 0 \qquad \text{(A11)}$$

$$\Leftrightarrow \frac{4}{c(1-\tau)}W_i+\left(W_i-W_j\right)\left[\frac{4}{(1-\tau)c}+\left(I_i+I_j\right)\right]\geq 0. \qquad \text{(A12)}$$

This is always true if $W_i > W_j$. Then there is no incentive to propose a higher wage than the one given by the first-order conditions. A symmetric argument runs for firm *j*. ■

APPENDIX 3 PROOF OF PROPOSITION 7

Bertrand competition implies that firms are ready to bid for workers up to their marginal productivity. It means that the maximum wage offered by firm *i* for a worker located at *x* is:

$$W_{i,max}=AK\left(1-\frac{cx}{2}\right). \qquad \text{(A13)}$$

This wage is only given to the marginal workers. All infra-marginal workers, working for firm *i* are given the wage proposed by the second highest bidder (*i.e.*, the neighbouring competitor). It means that the equilibrium wage for a worker located in *x* is:

$$W_{i,t}(x)=AK_i\left(1-c\left(\frac{I_i+I_j}{2}-x\right)\right), \quad \text{with } W_i(x)\geq B. \qquad \text{(A14)}$$

So we can define $\underline{x}$ such that $\forall x \leq \underline{x}, W(x) = B$. That is, $\underline{x}$ is the maximum distance between the firm's technology and a young worker's specialisation. If the gap between the specialisations is below $\underline{x}$, the firm acts with its workers as a local monopolist. We have:

$$\underline{x} = \frac{I_i + I_j}{2} - \frac{1}{c}\left(1 - \frac{B}{AK_i}\right). \tag{A15}$$

Free-entry induces a no-profit condition. Then:

$$\Pi_i = AK_i I_i \left(1 - \frac{cI_i}{4}\right) - 2\int_0^{I_i/2} W_i(x)dx = K_i. \tag{A16}$$

Depending on $\underline{x}$, two cases must be distinguished:

• $\underline{x} > 0$. At the symmetric equilibrium ($K_i = k$), profit are:

$$\Pi_i = AkI_i\left(1 - \frac{cI_i}{4}\right) - 2\int_0^{\underline{x}} Bdx - 2k\left(1 - c\left(I_i - x\right)\right)dx = k. \tag{A17}$$

After simplifications: $2AkI\left(1 - \frac{c.I^2}{4}\right) - 2BI - k - \frac{1}{Akc}(Ak - B)^2 = 0$. (A18)

Then: $I = \frac{2}{c}\left(1 - \frac{B}{Ak}\right) - \frac{\sqrt{2}}{Akc}\sqrt{B^2 - 2AkB + A^2k^2 - Ack^2}$. (A19)

For k very large, we can approximate (A.19) by:

$$I_\infty = \frac{2}{c} - \frac{\sqrt{2}}{c}\sqrt{1 - \frac{c}{A}}. \tag{A20}$$

We find then: $\underline{x} = \frac{1}{c} - \frac{\sqrt{2}}{c}\sqrt{1 - \frac{c}{A}}$. (A21)

So: $\underline{x} \geq 0 \Leftrightarrow A \leq 2c$. We can now derive an asymptotic growth rate:

$$g_\infty' = \frac{b}{2(1+b)}\left(\frac{A}{c}-1\right)-1. \tag{A22}$$

This latter expression can be compared to (20) with τ arbitrarily close to 1

$$g_\infty' \geq g_\infty \Leftrightarrow \frac{A}{2c}-\frac{1}{2} \geq A\sqrt{\frac{2}{Ac}}\left(1-c\sqrt{\frac{2}{Ac}}\right). \tag{A23}$$

If $2c \geq A$ (i.e., $\underline{x} \geq 0$), straightforward calculation shows that (A.23) is satisfied. It means that discriminatory wages enable a higher growth rate.

• $\underline{x} \leq 0$ (i.e., there are no worker for which the firm can act as a local monopolist). From (A.20), we know that $\underline{x} \leq 0 \Leftrightarrow 2c \leq A$. From (A.16) and after simplification the profit function becomes $\Pi_i = AK_i c I_i^2$. Consequently, the normal-profit condition implies $I = \sqrt{1/(Ac)}$. Thus, we can infer the asymptotic growth rate:

$$g_\infty'' = \frac{b}{(1+b)}\left(2\sqrt{\frac{A}{c}}-1\right)-1. \tag{A24}$$

Then it is immediate to check that $g_\infty'' \geq g_\infty$. ■

CHAPTER 4

CUMULATIVE INVESTMENT AND SPILLOVERS IN THE FORMATION OF TECHNOLOGICAL LANDSCAPES

Abstract: In this chapter, the evolution of industries is modelled as a cumulative cost-reduction (or quality improvement) process subject to spillovers in a differentiated oligopoly. Our results suggest that the long-run outcome is dependent on spillovers. It is found that when they are weak, firms find their niche over time, differentiation remains important and cost-reduction keeps going. On the contrary, if spillovers are strong and if the diffusion function of spillovers is concave, firms tend to agglomerate towards the centre of the market and to use similar technologies. This standardisation process involves less and less investment and may induce instability in the industry. Standardisation leads to a fall in the growth rate of productivity. For spillovers of intermediate strength, complex technological landscapes may arise.

I. INTRODUCTION

The existing theories of technological change tend to offer a uni-dimensional view of the "technological landscape". In a first branch of the literature, models of expanding varieties represent knowledge as a segment whose length increases over time (e.g., Romer [1990]). This segment has very specific properties since location does not matter because of the features of the Dixit and Stiglitz [1977] type of monopolistic competition. In the other stream of literature, models of quality improvements view the growth process as a vertical segment (a ladder) over which firms are moving upwards (e.g. Grossman and Helpman [1991] or Aghion and Howitt [1992]). It seems however that "technological landscapes" are much richer than this simplistic picture.

For instance, the existence of "technological clusters" has long been acknowledged. In this vein, Jaffe [1986] measured empirically technological proximity between firms. This enabled him to provide evidence and stylised facts concerning *technological neighbours* in the product space. For instance, an innovation made by a firm, say in the aeronautics industry, will spillover primarily to its direct competitors using the same type of technologies and then, maybe to a lesser extent, to technologically related industries like the car industry.

When dealing with the knowledge of this landscape over which technologies are located according to their attributes, it seems clear that some areas are still totally unexplored, whereas some others are well known.[1] Among these explored areas, some are "intensively cultivated", whereas some others are deserted and considered as hostile. If we restrict our analysis in terms of costs, there are some hospitable (low costs) "valleys" and some (high costs) "mountains". For instance, when dealing with transport technologies, steam-based or electric technologies can presently be likened to mountains whereas technologies with internal combustion engines are valleys where all car makers work. This technological landscape in the car industry has, however, undergone radical changes in the last century. Before that, the steam-based technologies constituted a valley. This chapter is concerned

[1] The exploration of unknown territories is the main concern of models of expanding varieties. On the same issue, Jovanovic and Rob [1990] provide an interesting alternative approach.

with the evolution and the emergence of relief on the technological landscape. It is a contribution to the understanding of the formation of some market and technological configurations and to the analysis of their implications. In our model, the market structure influences the technological landscape which in turn determines the market structure.

No doubt, the fundamentals of the "technological base" have the potential to explain the technological landscape just as geology is crucial to explaining physical landscapes (e.g., the potential for improvement in electric engines is nowadays acknowledged to be greater than that for steam engines). Some technologies can be naturally more subject to "cost erosion" (or quality improvement) than others. But, as in physical geography, and probably more strongly so, human action is determinant. Besides, there is no *a priori* reason why man's activity should always take place in the technologically most favourable areas. In an influential paper, Arthur [1989] argues that when two products (1 and 2) are imperfect substitutes and "strong" learning-by-doing is present, path-dependency is possible. If customers in the beginning keep buying Product 1, its producers will experience some learning-by-doing for their technology. So the cost of Product 1 decreases, whereas the cost of Product 2 remains very high. Consequently, Product 1 will attract new consumers at the expense of Product 2 and the cost gap between products will widen. This cost advantage may become decisive so that Product 1 can dominate the market after some time. This may be sub-optimal since Product 2 might present a far better potential for further cost reductions. In the comparison between steam and internal combustion engine, Arthur [1989] pretends that steam technologies offered a lot of room for improvements that were never undertaken due to the head-start of internal combustion engines. From our perspective, Arthur [1989] proposes a theory where the landscape results passively from economic activity (like erosion for physical landscapes). It is subject to path-dependency (e.g., erosion due to human activity may not take place in the best land).

Our argument assumes away differences in the potential for improvement and passive learning-by-doing to focus on purposeful investment. In a long-run dynamic framework, firms manufacture differentiated products over a simple Hotelling-type segment (the technological landscape). In the first period, we assume a flat landscape (the cost function is the same whatever variety is manufactured) that is common knowledge for all firms. On the one hand, the market share effect pushes firms towards the centre. On the other hand, the price competition effect favours the dispersion of firms.

Suppose we begin with a duopoly, if transport costs are quadratic, it is well known that firms will tend to adopt the principle of maximum differentiation and locate opposite each other (see d'Aspremont *et al.* [1979]). The distinctive feature of our model is that firms can invest to reduce the cost of their production (or to improve quality). This cost reduction directly benefits one's own technology, but also spills over to neighbouring technologies so that competing firms also benefit from this investment. When horizontal spillovers are weak, the investments made by a firm will only benefit the technologies very close to the ones it uses. If we then recognise that the location of firms in the technological landscape can change over time, the appeal of extreme varieties will be reinforced in the next period because of cost considerations. So, to the standard effects, we have an additional a cost effect since firms prefer to locate where costs are low. In this case of weak spillovers, firms will keep locating on the sides of the market. The cost function over the differentiation space takes an inverse U-shape and firms keep investing over time and thus dig their "niche".

On the contrary if spillovers are strong and if the diffusion function for innovation is concave in technological distance, intermediate technologies can benefit from the cost-decreasing investments made by both firms. This cost reduction in the middle of the segment can exceed those at the extremities. In the next period, firms will then locate closer to one another since the appeal of intermediate technologies is reinforced by this cost advantage of the centre. In this case, the technological landscape takes a U-shape and firms locate closer and closer. As the distance between firms decreases, the spillovers become stronger and the free-rider problem between firms becomes more acute so that they invest less and less. A steady-state can be reached when firms are so close that they do not invest in R&D anymore. This standardisation process can have very adverse welfare implications since firms have a very low degree of differentiation and productivity is stagnant. It is also possible that an equilibrium in pure strategies ceases to exist after a couple of periods. The conjunction of cumulative investment and spillovers can thus generate instability over time.

A possible example of standardisation is provided by the market for personal computers. When Apple launched its first products, the "technological distance" with IBM-compatible computers was important (assuming an axis measuring the ratio of "user-friendliness" over "power"). Then IBM-compatible computers benefited from and copied the innovations brought in by Apple (introduction of the mouse for instance). At the same time, it seems likely that the Apple technology also benefited from investments made by the IBM-compatible computer producers or their suppliers (e.g., improvements in the chip technology). Recently, Apple launched computers that are competing directly with IBM-compatibles (Power PC). This increased proximity is hard to measure but one can note for instance that advertisements in this industry now rely more on emotional appeal than on technical description (see Freeman [1994]). It may seem also that the pace of progress has slowed down in some way.[2] The explanation may be framed in terms of product cycles (all major possible improvements have been exhausted), but it is also possible that the incentive to innovate was reduced by this greater technological proximity. This progressive rise of homogeneity out of heterogeneity seems to be a widespread feature in many industries, from soft drinks to automobiles.

This analysis can be seen as a contribution to the theory of growth. In contrast to previous models, we assume that differentiation is endogenous and that competition is local (each firm competes with only a few firms), whereas the existing literature focuses on global imperfect competition and fixed differentiation. Our main insight is that the addition of a second dimension reveals some new effects generated by the interaction between the two dimensions. This chapter can also be related to the literature on economic agglomeration. In our case, the excessive similarity denounced by Hotelling [1929] is not automatically given and when it occurs, it is only progressively and incompletely. In our model, agglomeration stems from cumulative investment that makes some locations more attractive. In turn subsequent investment reinforces this process. This explanation is new with respect to the existing literature (see Gabszewicz and Thisse [1992] for a survey or Ungern-Sternberg [1988] and Weitzman [1994] for more specific models of endogenous product differentiation).

---

[2] Clearly the performances of computers keep increasing but this may be due to the progress made in the semi-conductor industry.

Moreover, our chapter can also be read in relation to other traditions in the field of industrial organisation. As in Dasgupta and Stiglitz [1980] or Flaherty [1980], both the market structure and the innovative activity are endogenous. What we add here is an endogenous determination of the technological parameters (e.g., costs). As in d'Aspremont and Jacquemin [1988] or Motta [1992], we study a model of investment in an oligopoly model with spillovers. In our case however, differentiation is endogenous and may change over time. Eventually, our model can be viewed as a theory of the life-cycles of industries (see Nelson and Winter [1978], Jovanovic and MacDonald [1994] or Klepper [1996] for alternatives).

The rest of the chapter is devoted to the analysis of the model (Section II). Some extensions are performed in Section III. Section IV contains some concluding remarks.

II. DESCRIPTION OF THE MODEL

Assumptions

DEMAND. As usual in this type of model (Gabszewicz and Thisse [1992]), each consumer buys one unit of product in each period if the price is below his reservation level and we assume that the reservation levels are not binding in what follows. There is a mass one of horizontally differentiated consumers. Consumers are indexed by $x \in [0,1]$ and their distribution is uniform over this segment. They face quadratic transport costs (and mill-pricing). For instance a consumer in x who buys a product located in x_1 has to bear a transport cost of $\tau(x - x_1)^2$. Of course, consumers buy the cheapest available product.

MARKET STRUCTURE AND TIMING. We assume a market which is contestable at each period. Two different firms only can enter at the beginning of the period.[3,4] These two

[3] An interpretation of our framework can be given in terms of patents. When firms enter they have a patent with a length of one period and a limited breadth so that competitors can benefit partly from their investments. At the next period, the present cost function will be available for new innovators.

[4] Our assumption is however supported by David [1975] who states that "... changes are shown as being 'guided' or directed in an ex post manner by previous myopic decisions - decisions having their objective in the minimisation of current (as distinct from future) private cost of production. One can also refer to

assumptions are rather restrictive and are relaxed in section III. First, firms choose a location over $[0,1]$ and a level of cost reducing investment. Price competition takes place in a second stage. The same game is played again at the next period with no additional cost if firms change location. The two firms are indexed by 1 and 2. Without loss of generality, Firm 1 is located on the left of Firm 2.

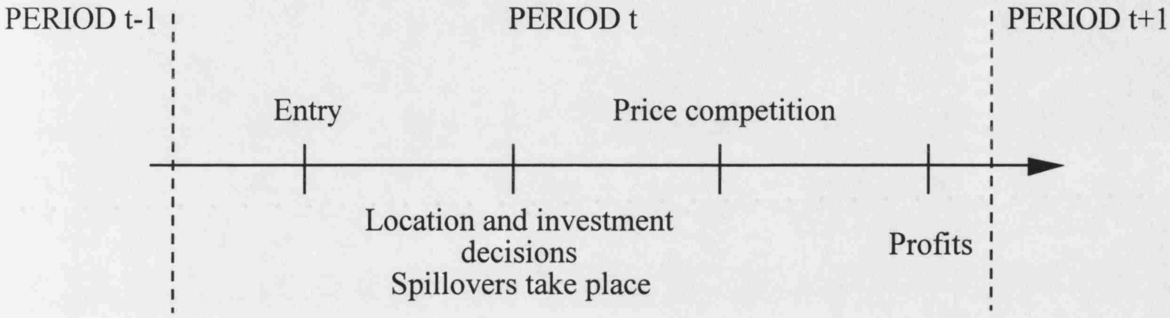


FIGURE 1 TIMING OF EVENTS

PRODUCTION. Firms face a constant marginal cost of production which depends on the variety they manufacture. Each variety on the segment faces a potentially different cost. We denote by $C(x)$ the cost of the variety in x. We assume that $C(x)$ is common for all producers and common knowledge.

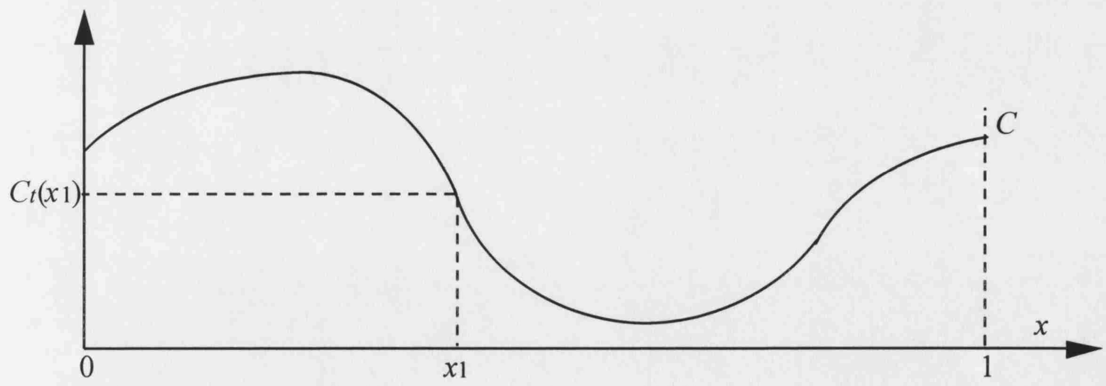


FIGURE 2

INVESTMENT. To reduce its cost in its location, each firm can invest $\underline{C}_{t-1}I_t$ to reduce its costs by $\underline{C}_{t-1}r(I_t)$, where $\underline{C}_{t-1} = Min\{C_{t-1}(x) | x \in [0,1]\}$. This cost-reduction function is such that it is continuous and piecewise twice differentiable with $r' \geq 0$, $r'' \leq 0$, $r(0) = 0$

Rosenberg [1996] to support empirically our assumptions. Theoretically, the assumption is also used by Klepper [1996].

and $\lim_{I \to +\infty} r(I) < 0.5$. The restrictions on the shape of $r(.)$ are required to ensure regularity. The limit condition implies that costs are reduced at most by a half for one firm and that all costs remain positive. The specification of the cost-reduction process also assumes that it is conducted at the beginning of the period by using the product itself manufactured with the cheapest available technology. For instance the cost-reduction function might be of the following shape:

FIGURE 3

SPILLOVERS. The investment performed on a given technology spills over immediately to neighbouring technologies (adding a time-lag would just reinforce our argument). That is the effort undertaken to reduce the cost of production for one product also decreases costs for neighbouring products. If an amount I is invested in x, the reduction of costs in X is equal to $\alpha(|x - X|)r(I)\underline{C}$. These spillovers are additive in the sense that one firm benefits completely from its own cost-reduction and partially from the investment made by the other firm. We assume that the diffusion function, $\alpha(.)$ is continuous and piecewise twice differentiable with $\alpha' \leq 0$, $\alpha(0) = 1$ and $\alpha(.) \geq 0$.

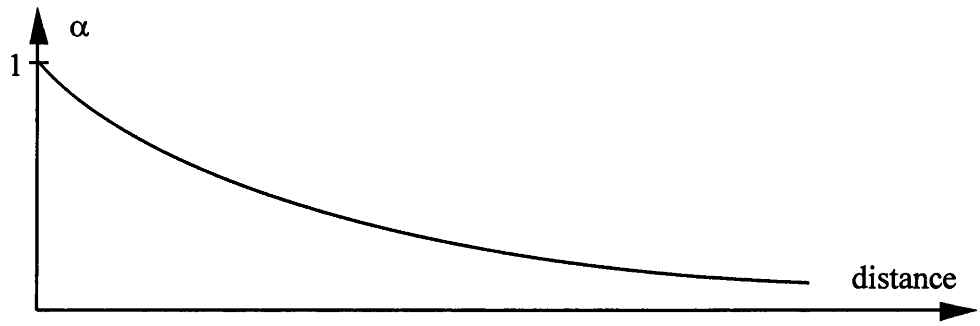


FIGURE 4

For instance, if the cost function is C_{t-1} and if there is only one firm located in x_1 investing I, then the new cost function may be as follows:

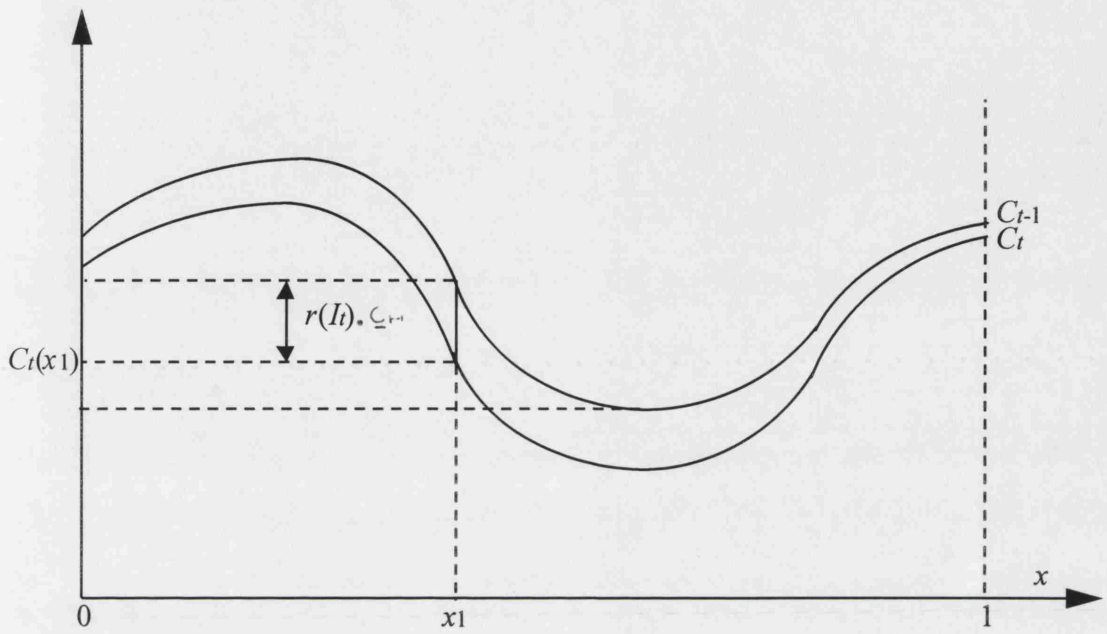


FIGURE 5

Fundamentally, our model enables us to study the interaction between spillovers and oligopoly. Spillovers impact directly on investment through location (as the distance between the two firms changes, the intensity of spillovers varies) and on location through past investments (in each period cost-reduction and spillovers modify the shape of the cost function). The model is setup in terms of cost-reducing investments, but of course our cost index can be interpreted as an inverse quality index since it is well known that horizontal differentiation is a special case of vertical differentiation (Cremer and Thisse [1991]).

The static game

Now we can characterise the equilibrium of the one period game that takes place at date t according to the timing described by Figure 1. We still focus on the case where firms live only one period. The implications of multi-period maximisation are described in Section III. If we summarise our previous assumptions, the cost functions can be expressed by:

$$C_1 = C_t(x_1, I_1) = C_{t-1}(x_1) - \underline{C}_{t-1}\left(r(I_1) + \alpha(x_2 - x_1)r(I_2)\right), \tag{1}$$

and $$C_2 = C_t(x_2, I_2) = C_{t-1}(x_2) - \underline{C}_{t-1}\left(r(I_2) + \alpha(x_2 - x_1)r(I_1)\right). \tag{1'}$$

The cost faced by Firm i in period t is equal to the cost in period t-1 minus the cost reduction given by its own investment and the spillovers from the investment made by Firm j. The profit functions are then:

$$\Pi_1 = (P_1 - C_1)y - \underline{C}_{t-1}I_1, \tag{2}$$

and $$\Pi_2 = (P_2 - C_2)(1-y) - \underline{C}_{t-1}I_2, \tag{2'}$$

where P_i is the price of Firm i and y is the location of the consumer indifferent between the two firms. Note that the cost of investments also depends on the minimum of the cost function in the previous period. This specification enables us to have a level of investment independent of the production costs when we ignore spillovers. That is, if investment is sufficiently productive, the cost-minimising monopolist will keep investing at the same pace over time.[5] Note moreover, that our assumptions are qualitatively the same as the ones made in the R&D models of growth (e.g., Romer 1990 or Grossman and Helpman [1991]).

Using the concept of Perfect-Nash-Equilibrium in pure strategies, we can now state our first result:

Proposition 1 *If the cost function is U-shaped and symmetric, the candidate symmetric equilibrium is unique and such that* $\dfrac{\partial C_{t-1}(x_1^*)}{\partial x_1^*} = -\dfrac{\tau}{2}\left(1 + 4x_1^*\right)$, $\dfrac{\partial r(I_1^*)}{\partial I_1^*} = \dfrac{3}{\left(1 - \alpha\left(1 - 2x_1^*\right)\right)}$

and $P_1^* = P_2^* = C_t(x_1^*) + \tau(1 - 2x_1^*)$.

[5] Abstracting from all location problems, when a quantity Q is produced, the total cost function is $TC_t = QC_t + I_tC_{t-1}$. Following our previous specifications, we get: $TC_t = C_{t-1}\left(Q(1-r(I_t)) + I_t\right)$ and the minimisation of this expression in t is independent of time.

Proof. First, the indifferent consumer on the market is defined by the following arbitrage condition:

$$P_1 + \tau(y - x_1)^2 = P_2 + \tau(x_2 - y)^2. \tag{3}$$

Substituting into the profit function yields:

$$\Pi_1 = (P_1 - C_1)\frac{P_2 - P_1 + \tau(x_2 - x_1)(x_2 + x_1)}{2\tau(x_2 - x_1)} - \underline{C}I_1, \tag{4}$$

and symmetrically for Firm 2. The concavity of $\Pi_1(P_1)$ is immediate so that the first-order condition is necessary and sufficient for a maximum. After simplification, we find the equilibrium prices:

$$P_1 = \frac{1}{3}\left(2C_1 + C_2 + \tau(x_2 - x_1)(2 + x_2 + x_1)\right), \tag{5}$$

and $$P_2 = \frac{1}{3}\left(2C_2 + C_1 + \tau(x_2 - x_1)(4 - x_2 - x_1)\right), \tag{5'}$$

which solves for the second stage of the game. We can inject this result into the profit function. We obtain:

$$\Pi_1 = \frac{\left(\tau(x_2 - x_1)(2 + x_2 + x_1) + C_2 - C_1\right)^2}{18\tau(x_2 - x_1)} - \underline{C}I_1. \tag{6}$$

Now for the first stage, the first-order condition for profit maximisation of Firm 1 with respect to investment reads:

$$\frac{\partial \Pi_1}{\partial I_1} = \frac{\left(\tau(x_2 - x_1)(2 + x_2 + x_1) + C_2 - C_1\right)\left(\frac{\partial C_2}{\partial I_1} - \frac{\partial C_1}{\partial I_1}\right)}{9\tau(x_2 - x_1)} - \underline{C}. \tag{7}$$

First, note that if we had $\tau(x_2 - x_1)(2 + x_2 + x_1) + C_2 - C_1 \leq 0$, using the equilibrium price as defined in equation (5), we would find that $P_1 \leq C_1$ and $\Pi_1 < 0$ which cannot be a maximum. When setting the first-order condition to zero, we find a unique candidate symmetric equilibrium which satisfies:

$$\left(\frac{\partial C_2}{\partial I_1^*} - \frac{\partial C_1}{\partial I_1^*}\right) - 3\underline{C} = 0. \tag{8}$$

Using our specifications concerning the spillovers, we can simplify this further into:

$$\frac{\partial r\left(I_1^*\right)}{\partial I_1^*} = \frac{3}{\left(1 - \alpha\left(1 - 2x_1\right)\right)}. \tag{9}$$

As for location, the first-order condition for Firm 1 is:

$$\frac{\partial \Pi_1}{\partial x_1} = \frac{1}{18}\left(2 + x_2 + x_1 + \frac{C_2 - C_1}{\tau\left(x_2 - x_1\right)}\right)\left(\tau\left(-2 + x_2 - 3x_1\right) + \frac{C_2 - C_1}{x_2 - x_1} + 2\left(\frac{\partial C_2}{\partial x_1} - \frac{\partial C_1}{\partial x_1}\right)\right) \tag{10}$$

When we set this expression equal to zero, two solutions are possible. As previously, $C_1 = \tau(x_2 - x_1)(2 + x_2 + x_1) + C_2$ would imply $\Pi_1 \leq 0$. Consequently, the maximum is such that $C_1 \leq \tau(x_2 - x_1)(2 + x_2 + x_1) + C_2$. Thus the candidate equilibrium is defined by:

$$\tau\left(-2 + x_2 - 3x_1\right) + \frac{C_2 - C_1}{x_2 - x_1} + 2\left(\frac{\partial C_2}{\partial x_1} - \frac{\partial C_1}{\partial x_1}\right) = 0. \tag{11}$$

Using the analogue of (11) for Firm 2, we can prove easily that we have only one candidate symmetric equilibrium location. It is defined by:

$$\frac{\partial C_1}{\partial x_1} = \frac{\tau}{2}\left(-1 - 4x_1\right) + \frac{\partial C_2}{\partial x_1}. \tag{12}$$

We find after further simplifications:

$$\frac{\partial C_{t-1}\left(x_1^*\right)}{\partial x_1^*}=\frac{\tau}{2}\left(-1-4x_1^*\right). \tag{13}$$

Given the convexity of the initial cost function, this candidate symmetric equilibrium is unique. This achieves our proof. ■

Proposition 2 *If the cost function is symmetric with minima in 0 and 1, at the candidate equilibrium, firms locate at each other's opposite and their level of investment is such that* $\frac{\partial r\left(I^*\right)}{\partial I^*}=\frac{3}{(1-\alpha(1))}$.

Proof. The demonstration is essentially the same as in the previous proposition, except that given the shape of $C_{t-1}\left(x_1^*\right)$, we find $x_1^*=0$. Then the investment equation is immediate. Note that this case includes all symmetric functions with an inverse U-shape. ■

Note that these propositions give only necessary conditions for the equilibrium. The issue of existence is tackled later.

We have here five types of effects. As usual, the market share effect tends to favour agglomeration since firms have an incentive to locate at the centre of the market to maximise their sales. This effect is counter-balanced (and even dominated with quadratic transport costs and a flat cost curve) by a standard price effect. When firms locate closer to each other, price competition is stronger and consequently firms face an incentive for increased differentiation.

The other effects stem from the new features of our model. First, there is an R&D effect on location that implies that each firm will tend to locate further away from the centre in order to avoid the competitor's free-riding on its own R&D. However, to free-ride on the competitor's R&D, the firms are encouraged to locate closer to the centre of the market.

With our specifications, at the symmetric equilibrium when it exists, those reciprocal R&D effects cancel each other out.

Second, when the cost curve is convex, the cost effect works as a centrifugal force and it adds to the price effect, so that maximum differentiation is achieved in equilibrium. With a U-shaped cost curve, on the contrary the cost effect pulls firms towards the centre (of course, if the slope of the curve in $x_1 = 0$ is superior to $-\tau/2$, we still obtain a corner equilibrium since the cost effect is not able to overcome the price effect anywhere). This cost effect can be decomposed into two parts. There is first an "absolute cost component". By locating closer to the middle, a firm lowers its cost of production and so raises its profits. Second, there is also a "business stealing" effect, since a lower cost enables firms to be more competitive and to take a bigger market share. The comparative statics of the equilibrium location shows that the closer to the centre, the steeper the cost curve must be at the equilibrium location. This occurs because the cost effect must in equilibrium balance the price effect, which becomes more intense as differentiation decreases. Moreover, the higher the transport costs, the further away firms locate. The reason is that, when firms move towards the centre, the benefits in terms of market share are smaller when transport costs are higher.

Finally, there is an effect of location on R&D. As already noticed by Spence [1984] or d'Aspremont and Jacquemin [1998], in a duopoly, the presence of reciprocal spillovers has detrimental effects on efficiency-enhancing investments. Here, the intensity of spillovers is endogenous and depends on location. The closer the firms, the stronger are the spillovers and thus the smaller the investment. We can now examine the welfare properties of our equilibrium.

Proposition 3 *The symmetric market equilibrium implies too much differentiation and too little investment.*

Proof. The second part of the result stems directly from the existence of spillovers. Note that the total surplus is maximised when total costs are minimised. For the short-run (static) welfare, we define x_1^{opt}, x_2^{opt}, I_1^{opt} and I_2^{opt} which minimise total social cost:

$$TC\left(x_1^{opt}, x_2^{opt}, I_1^{opt}, I_2^{opt}\right) = \underset{x_1, x_2, I_1, I_2, y}{Min} \left\{ \begin{array}{l} \int_0^y \tau(x_1 - z)^2 dz + \int_y^1 \tau(x_2 - z)^2 dz + \underline{C}(I_1 + I_2) \\ yC(x_1, x_2, I_1, I_2) + (1-y)C(x_2, x_1, I_2, I_1) \end{array} \right\} . (14)$$

After solving for the first-order conditions in location and investment, we find after simplification:

$$\frac{\partial r\left(I_1^{opt}\right)}{\partial I_1^{opt}} = \frac{2}{1 + \alpha\left(1 - 2x_1^{opt}\right)}, \qquad I_2^{opt} = I_1^{opt}, \qquad (15)$$

$$\frac{\partial C_{t-1}\left(x_1^{opt}\right)}{\partial x_1^{opt}} = \frac{\tau}{2}\left(1 - 4x_1^{opt}\right) - 2\alpha'\left(1 - 2x_1^{opt}\right) r\left(I_1^{opt}\right)\underline{C} \quad \text{and} \quad x_2^{opt} = 1 - x_1^{opt} . (16)$$

Then a straightforward comparison with equation (13) implies that the equilibrium locations involve too much differentiation. In period 1, we have $x_1^{opt} \geq 1/4$. To determine the optimal location, three effects are at stake, the cost effect (locate where the costs are low), a transport cost effect pushing firms towards the quartiles and a joint-investment effect pushing towards the centre. We can also define a long-run welfare function by taking the net present value of the social cost function over time:

$$\Sigma TC = \underset{x_{1,t}, x_{2,t}, I_{1,t}, I_{2,t}, y_t}{Min} \sum_{t=1}^{+\infty} b^t \left\{ \begin{array}{l} \int_0^{y_t} \tau\left(x_{1,t} - z\right)^2 dz + \int_{y_t}^1 \tau\left(x_{2,t} - z\right)^2 dz + \underline{C}_{t-1}\left(I_{1,t} + I_{2,t}\right) \\ y_t C_t\left(x_{1,t}, x_{2,t}, I_{1,t}, I_{2,t}\right) + \left(1 - y_t\right) C_t\left(x_{2,t}, x_{1,t}, I_{2,t}, I_{1,t}\right) \end{array} \right\} . (17)$$

When solving for the first-order conditions in $x_{1,t}$ and $x_{2,t}$, we can see that the joint investment effect has one more term (the future cost reduction of the joint-investment in *t*). This just reinforces agglomeration for the optimal location since the joint-investment effect is centripetal. Note finally that the equilibrium location can be optimal when the cost curve is sufficiently convex (corner location). ■

We have multiple sources of inefficiency, each corresponding to one decision variable of our game. Firstly, prices are too high (i.e., above the marginal cost as can be seen easily from equations (5) and (5')). In our case, however, the deadweight loss of oligopoly pricing is zero due to our simplified demand function. Secondly, the level of investment is sub-optimal. The inefficiency is twofold. In the short-run, investment is inefficient because of the horizontal spillovers. In the long run, another inefficiency is added due to the short time horizon of firms (see section III). Thirdly, the market equilibrium implies too much dispersion. The location decision generates this inefficiency because of the presence of the (centrifugal) price effect and the absence of the joint-investment effect (centripetal) at the market equilibrium. Now that we have described the outcome of the static game (short-run), we can turn to the dynamic evolution (long-run).

The dynamic behaviour

To explore the dynamics of our model, we assume that the initial situation is such that $C_0(x) = C_0$.[6] The dynamics of the cost curve is driven by the aggregation of the investments over time. The shape of the spillover diffusion function is thus crucial. Potentially complicated dynamics may arise. Before investigating further, we explore two simple dynamic paths for which we can give sufficient conditions under which they arise.

Proposition 4 *If either* $\alpha'' > 0$ *or if* $\alpha'' \leq 0$ *and* $\alpha(1/2) \leq 1/2$, *the cost function takes an inverse U-shape, firms keep locating at each other's opposite and keep investing at the same pace when the symmetric equilibrium exists.*

If $\alpha'' \leq 0$ *and* $\alpha(1) > 0$, *the cost function takes a U-shape, firms locate closer and closer to the centre and investment decreases over time as long as there is a symmetric equilibrium.*

Proof. In the beginning of period 1, firms face a flat cost curve. According to Propositions 1 and 2, the firms will locate at $x_1 = 0$ and $x_2 = 1$. From our assumptions, we have:

[6] We explore here a simple case, but using simulation, a similar analysis can be made for any initial cost curve.

$$C_t(x, x_1, x_2, I_1, I_2) = C_{t-1}(x) - \underline{C}_{t-1}\left(\alpha(|x - x_1|)r(I_1) + \alpha(|x_2 - x|)r(I_2)\right). \qquad (18)$$

At the symmetric equilibrium: $r(I_1) = r(I_2) = r(I^*)$ and $x_2 = 1 - x_1$ so that:

$$C_1(x) = C_0 - C_0(\alpha(x) + \alpha(1-x))r(I^*). \qquad (19)$$

Now we can see that if $\alpha(.)$ is convex, $\alpha(x) + \alpha(1-x)$ is also convex and has two maxima at $x = 0$ and $x = 1$. Then $C_1(x)$ has minima at $x = 0$ and $x = 1$. In period *t*, assuming a cost function with two minima at $x = 0$ and $x = 1$, firms will thus keep locating at $x_1 = 0$ and $x_2 = 1$. The cost function of *t*+1 will also have two minima in 0 and 1 since it is equal to the sum of two functions whose minima are the same, both at $x = 0$ and $x = 1$ (Proposition 2). The same reasoning applies for $\alpha(.)$ concave and $\alpha(1/2) \le 1/2$.

On the contrary, if $\alpha(.)$ is concave with $\alpha(1) > 0$, then $\alpha(x) + \alpha(1-x)$ is concave and has one unique maximum at $x = 0.5$. Then $C_1(x)$ has only one minimum at $x = 0.5$. In period 1, unless $C_1(x)$ is sufficiently steep, the firms will be again at $x_1 = 0$ and $x_2 = 1$. Their investment increases the steepness of the cost function at $x = 0$ and $x = 1$. Eventually, the cost curve can become steep enough for both firms to locate closer. Using the same reasoning as above and Proposition 1, it is immediate that $x_{1,t+1} \ge x_{1,t}$ and that $I_{t+1} \le I_t$. ■

In these two cases, the conjunction of cumulative investment and spillovers gives rise to a simple self-reinforcing mechanism. When spillovers are weak or when the diffusion function $\alpha(.)$ is convex, peripheral technologies benefit from current investments more than the central ones. In turn, this reinforces the appeal of the peripheries in the next period. Firms are "digging their niches" over time. In other words, a high degree of heterogeneity can be sustained by weak spillovers.

On the contrary, when spillovers are strong and when the diffusion function $\alpha(.)$ is concave, cumulative investment reinforces the U-shape of the cost function, thus leading to diminishing differentiation and diminishing investment. A steady-state with no investment can be reached. By contrast with our previous case, the result here states that the "principle of diminishing differentiation holds under sufficient and concave spillovers". Note that our assumption of quadratic transport costs, chosen for tractability reasons, tends to make this case more difficult to exhibit because of the innate tendency of quadratic transport costs towards differentiation (strong price effect).

Moreover, the existence of the equilibrium is not automatically given. When the second-order conditions are not satisfied, this means that firms have an incentive to deviate from the candidate symmetric equilibrium. It is possible that no equilibrium exists in pure strategies. This is a recurrent problem in differentiation models (see Gabszewicz and Thisse [1992] for some development on this issue). A first possible interpretation is that the non-existence result is especially relevant to game theory itself. The second possibility is to relate this possible non-existence to what Porter [1980] calls "industry turmoil". From the simulations that follows it can be shown that the equilibrium becomes more difficult to sustain as the standardisation process takes place. Our interpretation of this situation is the following. The model here is one of incremental innovation (evolutionary). But it is possible that incremental innovations can be large enough to create a decisive cost advantage for one firm, then one must speak of drastic or revolutionary innovations (for a discussion of evolutionary and revolutionary innovations, see Langlois and Robertson [1995]). Hence there is an incentive to relocate in the centre of the market. In this case the competitive process we use is not relevant and another framework, maybe more Schumpeterian, is needed (Reinganum [1985] or chapter 5).

This result provides an alternative explanation of the so-called life cycle of industries. Gort and Klepper [1982], Evans [1987] or Klepper [1996] in his survey, find that the growth rate of an industry falls as it matures. Alternatively, Kamien and Schwartz [1982] and Klepper [1996] survey evidence showing that large firms tend to generate slower productivity growth. All this is usually attributed to a progressive exhaustion of the growth possibilities (see Jovanovic and MacDonald [1994a] for a learning argument on these lines). Here, with sufficient spillovers and a concave diffusion function, a process of rising homogeneity takes place. As stated by Porter [1980], "products have a tendency to become more like commodities over time..." This increased standardisation reduces the incentive to perform independent cost-reducing investments although they are still feasible and socially desirable. In other words, the model we propose here is a model in which technological progress creates the conditions for its own demise.

Our model can also be related to the strategic management literature. For instance, Porter [1980] writes that "every industry begins with an initial structure [...] This structure is usually (though not always) a far cry from the configuration the industry will take later in its development. [...] The evolutionary processes work to push the industry towards its potential structure, which is rarely known completely as an industry evolves. [...] there is a range of structures the industry might possibly achieve, depending on the direction and success of research and development, marketing innovations, and the like. It is important to realise that instrumental in much industry evolution are the investment decisions by both existing firms in the industry and new entrants." He also underlines that overall cost leadership and differentiation are the two most important generic strategies for firms. What our model does is to analyse formally these ideas, explore the welfare of some outcomes, relates some dynamics to some technological fundamentals and finally shows how some generic strategies are optimally chosen depending on the market structure.

Note finally that Proposition 3 states that differentiation is too important with respect to the first-best (short-run and long-run). However, incentives to reduce dispersion alone would have negative dynamic effects since they would imply less investment. However, it is not always true that a market equilibrium with "standardisation" and subsequent dwindling investments is worse than a situation of "niche" formation with maximum differentiation. If we measure roughly the welfare by the asymptotic total cost, it is equal to the total transport costs when niches are formed since production costs tend to zero. Thus we have with maximum differentiation $TC_\infty(0,1) = \tau/12$. With the polar case of standardisation, we find that $TC_\infty\left(x_{1,\infty}, x_{2,\infty}\right) = C_\infty + \tau\left(1 - 6x_{1,\infty} + 12x_{1,\infty}^2\right)/12$. Consequently, standardisation can be asymptotically better than niche formation if we have $TC_\infty\left(x_{1,\infty}, x_{2,\infty}\right) \leq TC_\infty(0,1)$, that is if $\tau x_{1,\infty}\left(1 - 2x_{1,\infty}\right) \geq 2C_\infty$.

A numerical example

For the previous results, we just gave necessary conditions for the equilibrium. However we can derive some examples for which the equilibrium exists. If we set $r(I) = (1 - 1/(1 + 15I))/2$ and $\alpha(|x_1 - x|) = Max\left(0, 1 - \psi(x_1 - x)^2\right)$, one can check that these functional forms satisfy our restrictions concerning the concavity of the investment function. Moreover, we also set $C_0 = 3$ and $\tau = 2$. All our simulations are conducted for different values of ψ, which is an inverse index of the intensity of spillovers (the higher ψ, the weaker the spillovers).

* If $\psi = 4$, we find that $\forall t,\ x_1 = 0$ and $I = 0.0387$. This result conforms to Propositions 2 and 4. The shape of the cost function is as in Figure 6a and the costs of both firms are reduced by more than 20% at each period for the technologies they use.

* If $\psi = 2$, the results are exactly the same except that the cost function is more complex as can be checked on Figure 6b.

* If $\psi = 1.2$, we face an intermediate case for which the cost curve is non-monotonic after one period (see Figure 6c). Of course for $t = 1,\ x_1 = 0$ and $I = 0.0387$. In the beginning of period 2, we have two potential equilibria. The first is the one where both firms locate at each other's opposite. The second is such that Firm 1 locates on the right of the first kink and Firm 2 on the left of the second kink of the cost function. However, due to insufficient steepness, only the first candidate is an equilibrium. However after three periods, the curve becomes sufficiently steep so that maximum differentiation cannot be sustained in equilibrium anymore and both firms locate according to Proposition 1 between the two kinks (Figure 6d). This is a case of a "complex landscape". When cumulative investments and spillovers generate a non-monotonic cost curve, there is room for multiple equilibria or the non-existence of an equilibrium in pure strategies.

* If $\psi = 1$, the first period equilibrium is characterised by $x_1 = 0$ and $I = 0.0387$ (Figure 6e). Then over time, we can observe the process of diminishing differentiation and investments already described.

t	x_1	I
5	0.152829	0.006523
10	0.175594	0.00172415

20	0.182723	0.00022115
30	0.183639	0.00003231
50	0.183769	7.15×10^{-7}
∞	$0.5-\sqrt{0.1} \approx 0.183772$	0

The asymptotic situation is represented graphically by Figure 6f. The same evolution occurs for all values of ψ below 1. For instance, if $\psi = 0.8$, one converges towards a situation where $x_1 = \left(1-\sqrt{0.5}\right)/2$ and $I = 0$. In all those cases, it can be checked numerically that the candidate equilibrium is indeed an equilibrium. However, if we lower the transport costs; or if we set C_0 at a higher initial value (above 10 is sufficient in our simulations), then our candidate equilibrium cannot be sustained any more. The reason is fairly simple: if transport costs are much lower, with a U-shape curve, the centrifugal force is not sufficiently strong to prevent a deviation whereby the firm would locate at the centre of the market and benefit from the lower costs and grab all the market. If the initial level of costs is high enough, then it may be also profitable to deviate and locate at the centre of the market and invest heavily so that a strong cost advantage can be created to expel the other firm.

III. EXTENSIONS

Our basic model suffers from two restrictive assumptions. In what follows, we relax them so that new properties can appear.

Free entry

The first restrictive assumption we made concerned the entry process: only two firms were allowed. To relax this assumption, we assume a fixed cost of entry K. Again, we start with $C_0(x) = C_0$. It seems clear that as K decreases, the number of entrants increases and thus the distance between two adjacent firms decreases. Here, this smaller distance drives investment to zero if r' is bounded from above.

Assume first that the fixed cost is such that initially only two firms can enter. Then, from equation (6), the equilibrium profits are $\Pi = \tau(1-2x_1^*)/2 - \underline{C}I^* - K$. As the two firms get closer, their profits keep decreasing. So if standardisation takes place, then the long-run profit of each duopolist tends towards $\Pi_\infty = \tau(1-2x_\infty)/2 - K$. Then two cases are possible. If $\Pi_\infty \geq 0$, our previous results are unchanged. If $\Pi_\infty < 0$, then the duopoly is asymptotically not sustainable. This means that for *t* sufficiently high, there can be only one firm operating in the market. This firm of course locates at the centre of the market and keeps investing to reduce costs.

Of course, this argument can be extended for a lower fixed cost of entry so that more than two firms can initially enter the market. The first notable feature of this case is that too many firms can enter initially, so that no-one has an incentive to perform any cost-reduction investment. We thus face stagnation in the industry due to too much competition with respect to a second-best situation. The notable feature is that, if spillovers are sufficiently large with a concave diffusion function, the number of firms decreases over time. This implication of the model matches a well-known stylised fact in many sectors where the number of firms increases quickly first and then slowly decreases (see Klepper [1996] for more details on this). The decreasing number of firms in industries is usually explained by a learning-and-selection argument. Firms that fail to keep up with the pace of progress in one sector are driven out of the market because they are not sufficiently competitive (see for instance Jovanovic and MacDonald [1994b] for a presentation of this argument). What we underline here is that the reduction ~~of~~ in the number of firms can arise because of the modification of the "technological conditions" within the market when this modification takes the form of a reduction in the range of competitive technologies. Of course this evolution of the "technological primitives" is itself endogenous and depends on past investments made in the sector. It is also possible that after a few periods of incremental innovations, the technological conditions become such that the market is not "stable" anymore and the equilibrium in pure strategies ceases to exist (industry turmoil). So, with strong spillovers and a concave diffusion function, we observe a tendency towards more standardisation, fewer competitors and more instability of the competitive process over time.

A last implication of our model is that the standardisation process increases the incentive equation (6), the equilibrium profits are $\Pi = \tau\left(1-2x_1^*\right)/2 - \underline{C}I^* - K$. As the two firms get closer, their profits keep decreasing. So if standardisation takes place, then the long-run profit of each duopolist tends towards $\Pi_\infty = \tau(1-2x_\infty)/2 - K$. Then two cases are possible. If $\Pi_\infty \geq 0$, our previous results are unchanged. If $\Pi_\infty < 0$, then the duopoly is asymptotically not sustainable. This means that for t sufficiently high, there can be only one firm operating in the market. This firm of course locates at the centre of the market and keeps investing to reduce costs.

Of course, this argument can be extended for a lower fixed cost of entry so that more than two firms can initially enter the market. The first notable feature of this case is that too many firms can enter initially, so that no-one has an incentive to perform any cost-reduction investment. We thus face stagnation in the industry due to too much competition with respect to a second-best situation. The notable feature is that, if spillovers are sufficiently large with a concave diffusion function, the number of firms decreases over time. This implication of the model matches a well-known stylised fact in many sectors where the number of firms increases quickly first and then slowly decreases (see Klepper [1996] for more details on this). The decreasing number of firms in industries is usually explained by a learning-and-selection argument. Firms that fail to keep up with the pace of progress in one sector are driven out of the market because they are not sufficiently competitive (see for instance Jovanovic and MacDonald [1994b] for a presentation of this argument). What we underline here is that the reduction in the number of firms can arise because of the modification of the "technological conditions" within the market when this modification takes the form of a reduction in the range of competitive technologies. Of course this evolution of the "technological primitives" is itself endogenous and depends on past investments made in the sector. It is also possible that after a few periods of incremental innovations, the technological conditions become such that the market is not "stable" anymore and the equilibrium in pure strategies ceases to exist (industry turmoil). So, with strong spillovers and a concave diffusion function, we observe a tendency towards more standardisation, fewer competitors and more instability of the competitive process over time.

A last implication of our model is that the standardisation process increases the incentive to merge. Yet, anti-trust regulation may prevent mergers. Moreover, in real life, there are large numbers of dimensions for investment decisions and firms. Finding out which activities complement each other is a difficult problem in practice. This slows down the merging process. Finally, mergers induce a monitoring problem since it implies more hierarchical layers (bureaucratic diseconomies of scale). However, a prediction of our model is that mergers should occur more frequently as industries mature.

Non-myopic firms

The second potentially restrictive assumption we made concerns the short-run maximisation objectives of firms. However, it is possible to show that the same kind of results as previously can be derived with a longer time horizon. For this purpose, we assume first that firms maximise over two periods. We write the objective function as:

$$\Omega_t = \Pi_t + b\Pi_{t+1}, \qquad (20)$$

where $b < 1$ is a discount factor. First note that in period t+1, the derivation of the sub-game equilibrium is the same as in section II. Working backwards, the derivation of the equilibrium price in period 1 can be computed as in equations (5) and (5'). As for the equilibrium investment in period t, we find:

$$\frac{\partial \Omega_{1,t}}{\partial I_{1,t}} = \frac{\partial \Pi_{1,t}}{\partial I_{1,t}} + b\frac{\partial \Pi_{t+1}}{\partial I_{1,t}} = 0 \,. \qquad (21)$$

After simplification, we find for the symmetric equilibrium, the necessary condition:

$$\frac{\partial r(I_{1,t})}{\partial I_{1,t}} = \frac{3}{1-\alpha(1-2x_{1,t}) + b\alpha(1-2x_{1,t+1})\underline{C}_{t+1}\left(1-\alpha(1-2x_{1,t+1})\right)/\underline{C}_t} \,. \qquad (22)$$

The FOC with respect to location in the first stage of the game is:

$$\frac{\partial \Omega_{1,t}}{\partial x_{1,t}} = \frac{\partial \Pi_{1,t}}{\partial x_{1,t}} + b\frac{\partial \Pi_{t+1}}{\partial x_{1,t}} = 0\,. \tag{23}$$

After tedious but straightforward simplifications, we find:

$$-\frac{\tau}{2}\left(1+4x_{1,t}\right) - \frac{\partial C_{t-1}\left(x_{1,t}\right)}{\partial x_{1,t}} + \underline{C}_{t-1}r\left(I_{1,t}\right)\left(\alpha'\left(x_{2,t+1}-x_{1,t}\right)-\alpha'\left(x_{1,t+1}-x_{1,t}\right)\right) = 0\,. \tag{24}$$

So if $\alpha(.)$ is concave with $\alpha(1) \geq 0$, a straightforward comparison with equation (13) implies that firms locate further apart in period t than in the case of a one period maximisation. However, one can also check that $x_{1,t+1} > x_{1,t}$. That is, differentiation still decreases over time. So, when we extend the time horizon of the firms, with $\alpha(.)$ concave and $\alpha(1) \geq 0$, firms tend to locate further apart in the beginning and then to get closer over time. Moreover, in equation (22), one can see that investment in the first period is more important than in the case of one-period maximisation (the spillovers are less important and the cost reduction is capitalised over two periods instead of one).

One is then led to wonder what happens if the time horizon becomes infinite. Even though no closed-form solution can be derived for our model, the same phenomenon of standardisation and slow-down of the pace of progress can arise since differentiation may not be sustainable in the long run. The following heuristic argument shows that maximum differentiation can be impossible in the long run if $\alpha(.)$ is concave and if $\alpha(1) \geq 0$. Assuming that firms locate at each other's opposite, the cost difference between the variety in 0 and the one in 0.5 is:

$$C_{t+1}(1/2) - C_{t+1}(0) = C_t(1/2) - C_t(0) + \left(2\alpha(1/2) - 1 - \alpha(1)\right)r\left(I_t\right)C_t(1/2)\,. \tag{25}$$

Then, at the limit we find that:

$$C_\infty(1/2) - C_\infty(0) = C_0\,\frac{2\alpha(1/2) - 1 - \alpha(1)}{2\alpha(1/2)}\,. \tag{26}$$

If the difference of costs between 0 and 0.5 is sufficient, Firm 1 can move towards the centre of the market and enjoy a higher level of profit. This is true if $P_2(x_1 = 0.5, x_2 = 1) \le C_\infty(1)$, that is if $C_\infty(0.5) - C_\infty(0) \ge 5\tau/4$. Consequently for maximum differentiation to be unsustainable in the long run, a sufficient condition is:

$$C_0 \frac{2\alpha(1/2) - 1 - \alpha(1)}{2\alpha(1/2)} \ge \frac{5\tau}{4}. \tag{27}$$

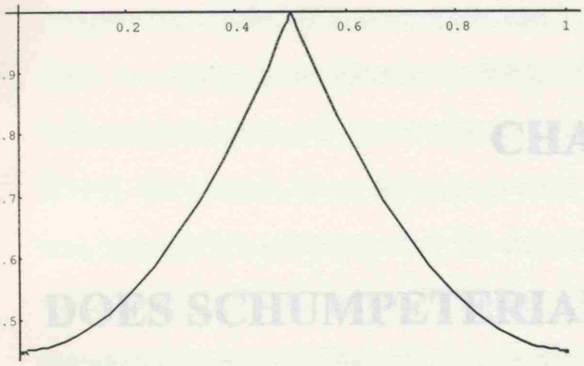
Policy implications

In the face of these persistent inefficiencies, corrective policies can be implemented. As suggested by d'Aspremont and Jacquemin [1988], there are positive effects from co-operative R&D. The main drawback of co-operative R&D in our setting is that by restoring investment it encourages standardisation. Simple policies of R&D subsidies can also be very ineffective since present R&D subsidies favour standardisation which makes further cost reduction more problematic because of stronger spillovers. Effective R&D policies need to control the "location" of investments. This may be very difficult to achieve in practice since it requires the knowledge of the cost curve by the planner (i.e., the knowledge of products which are not manufactured) and the ability to verify the location of firms.

IV. CONCLUDING REMARKS

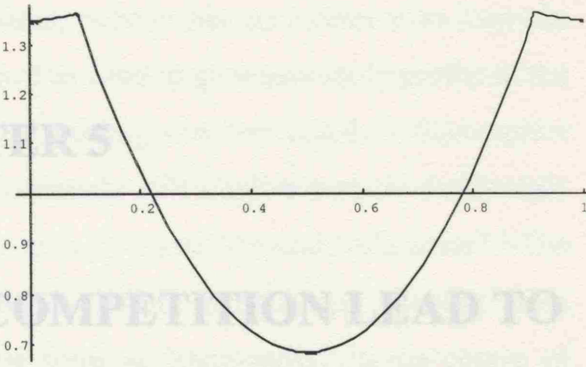
This chapter has proposed a model of the dynamics of industries that is driven primarily by spillovers and cumulative investment. Our model can generate as possible outcomes the formation of niches, a standardisation process or instability. If standardisation occurs, the number of firms in the industry can decrease as well as the intensity of cost-reducing investments. This slowdown in productivity growth does not depend on exhaustion of the potential for growth, but stems rather from the formation of a technological landscape that inhibits growth.

As main determinants of market structure and industry evolution, spillovers offer a quite attractive explanation. However, some problems are in order. Firstly, these explanations based on "invisible technological parameters" may be quite difficult to test. But spillovers can actually leave some paper trails, through patent citations for instance, as demonstrated by Jaffe *et al.* [1993]. In this latter case, the "real industrial landscape" is useful to provide evidence for our concept of technological landscape. Secondly, the vision we propose of industry evolution is opposite to that of management theorists like Porter [1980]. They tend to propose a world in which strategies are the result of free decisions by managers, whereas here the optimal strategy is dictated by the environment. The resolution of this fundamental divergence should be the object of future work. Thirdly, spillovers and their diffusion are taken here as exogenous whereas it seems pretty clear that for instance location decisions influence spillovers (Jaffe *et al.* [1993]), although it remains to understand precisely why and how.



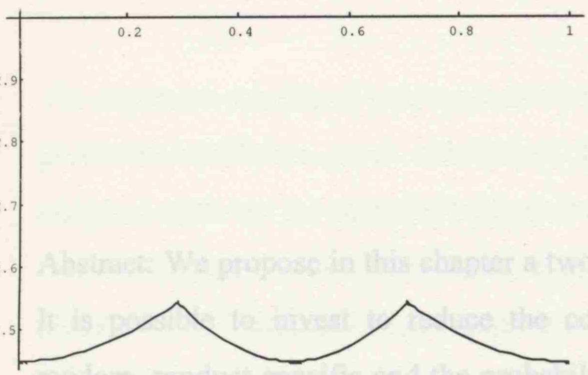
$\psi = 4,\ x_1 = 0,\ I = 0.038746,\ t = 1$

FIGURE 6A



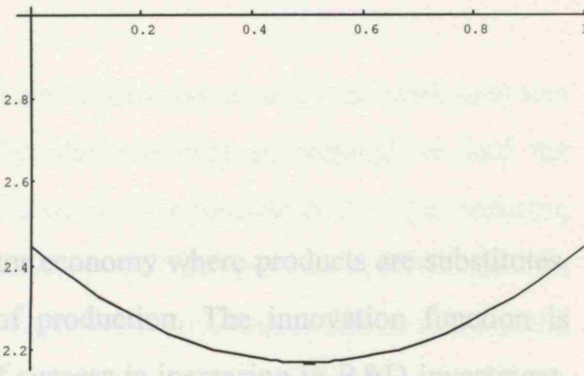
$\psi = 1.2,\ x_1 = 0.1423,\ I = 0.015941,\ t = 3$

FIGURE 6D



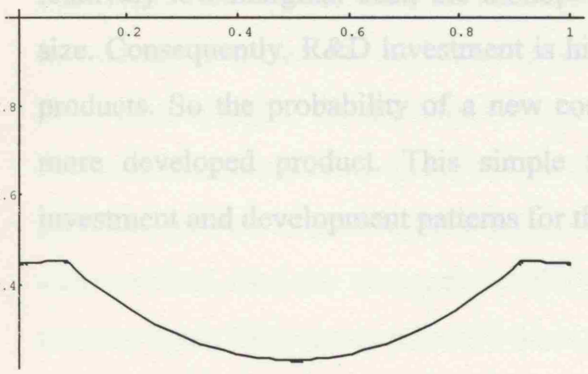
$\psi = 2,\ x_1 = 0,\ I = 0.038746,\ t = 1$

FIGURE 6B



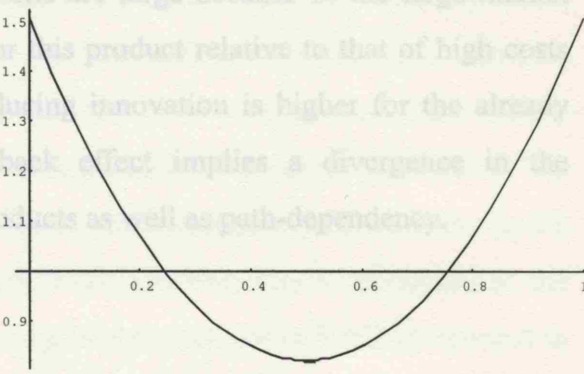
$\psi = 1,\ x_1 = 0,\ I = 0.038746,\ t = 1$

FIGURE 6E



$\psi = 1.2,\ x_1 = 0,\ I = 0.038746,\ t = 1$

FIGURE 6C



$\psi = 1,\ x_1 = 0.1837,\ I = 0,\ t = \infty$

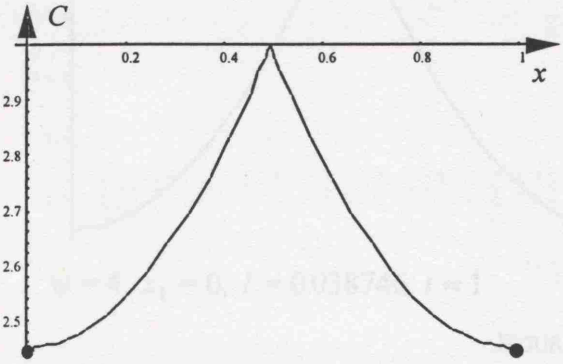
FIGURE 6F

Erratum to page 131. The series of graphs replaces the graphs page 131.

GRAPHICAL REPRESENTATION OF THE NUMERICAL SIMULATIONS OF SECTION II

(Costs are on the vertical axis; the product space is represented by the horizontal axis; the bold dots give the equilibrium locations of the firm on the cost curve; ψ is an inverse measure of the intensity of diffusion for spillovers)

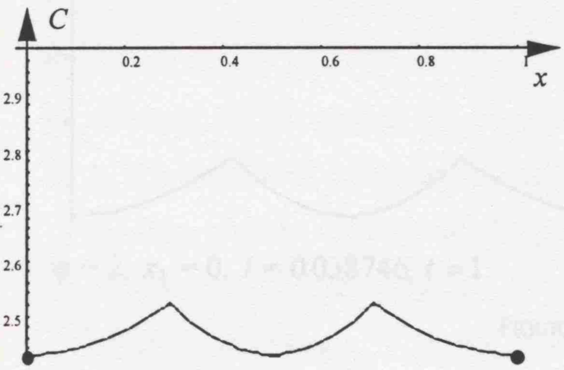
Case 1: $\psi = 4$ -- Cost function for $t = 1$



$x_1 = 0,\ I = 0.038746$

FIGURE 6A

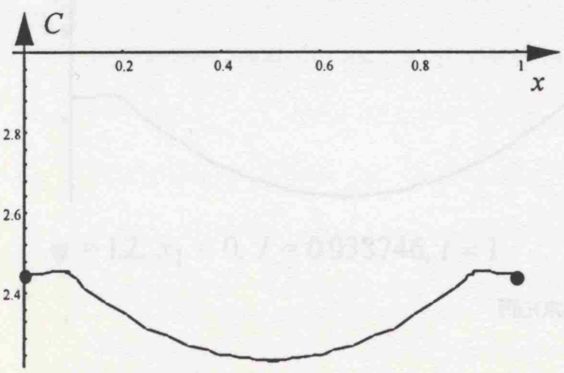
Case 2: $\psi = 2$ -- Cost function for $t = 1$



$x_1 = 0,\ I = 0.038746$

FIGURE 6B

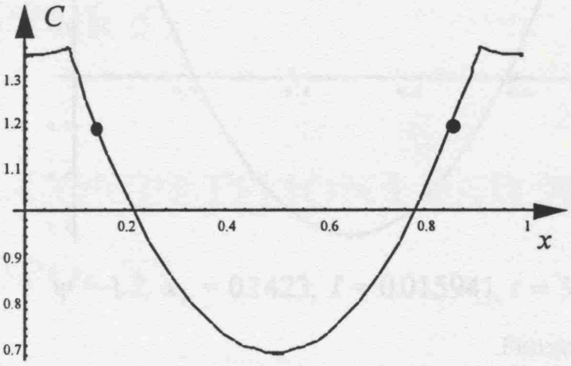
Case 3: $\psi = 1.2$ -- Cost function for $t = 1$



$x_1 = 0,\ I = 0.038746$

FIGURE 6C

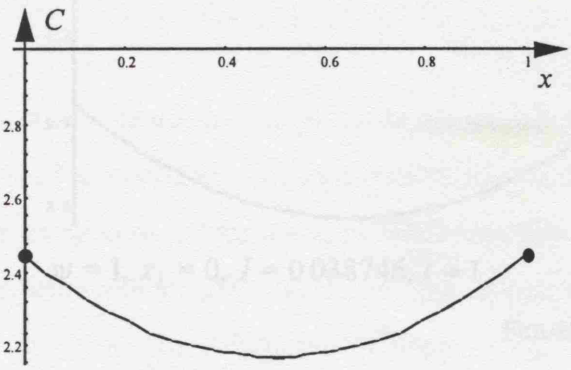
Case 3: $\psi = 1.2$ -- Cost function for $t = 3$



$x_1 = 0.1423,\ I = 0.015941$

FIGURE 6D

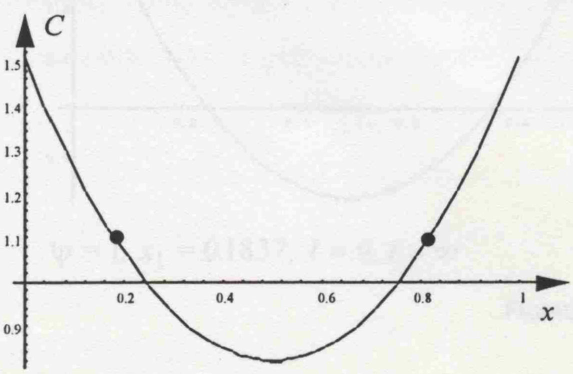
Case 4: $\psi = 1$ -- Cost function for $t = 1$



$x_1 = 0,\ I = 0.038746$

FIGURE 6E

Case 4: $\psi = 1$ -- Cost function for $t = \infty$



$x_1 = 0.1837,\ I = 0$

FIGURE 6F

CHAPTER 5

DOES SCHUMPETERIAN COMPETITION LEAD TO MONOPOLY?

Abstract: We propose in this chapter a two-sector economy where products are substitutes. It is possible to invest to reduce the costs of production. The innovation function is random, product specific and the probability of success is increasing in R&D investment. The successful innovator is a temporary monopolist. As usual for Schumpeterian competition, monopoly profits provide an incentive for R&D. When a product has a relatively low marginal cost, the monopoly profits are large because of the large market size. Consequently, R&D investment is high for this product relative to that of high costs products. So the probability of a new cost-reducing innovation is higher for the already more developed product. This simple feed-back effect implies a divergence in the investment and development patterns for the products as well as path-dependency.

I. INTRODUCTION

The necessity of monopolies for innovations has long been recognised. In a perfectly competitive world, if new designs can be imitated, nobody has an incentive to innovate since the returns to investment in R&D are zero. The need to give monopoly profits to the successful innovator if the economy was to keep growing was first noted by Schumpeter [1911]. Of course, from this perspective the monopoly still implies a static deadweight loss, but this is a necessary evil for the economy to enjoy some "dynamic efficiency".[1] The subsequent Schumpeterian literature has always viewed causality as running from monopoly power to innovation, assuming some form of "stationarity" in the degree of monopoly power.[2] What we ask here, by contrast, is the following: does monopoly power lead to Schumpeterian competition or does Schumpeterian competition lead to monopoly power?

The argument developed below can be sketched easily. Imagine an economy with only two products, which are gross substitutes. Initially, the economy is competitive and the marginal costs are the same for both products. However, it is possible to invest in reducing the costs of production. The innovation function is random, product specific and the probability of a successful innovation is increasing in R&D investment. So, as usual in the Schumpeterian literature, the successful innovator is a monopolist for at least one period. The monopoly profits provide an incentive (and a reward) for R&D. When a product has a relatively low marginal cost, the monopoly profits are large because of the large market size. Consequently, the R&D investment made on this product increases relatively to the investment made on the high cost product. So the probability of a new cost-reducing innovation is higher for the already more developed (i.e., low cost) product. This simple feedback effect implies a divergence in the investment and development patterns for the products, as well as path dependency. Eventually, all the investment will be made on the same product, and the monopoly at the product level will turn into a monopoly at the industry level. This occurs provided there are no diminishing returns to R&D investment in a given product line (or provided the decreasing returns do not appear too early). So our

1 Empirically, the question is hotly debated. See Nickell [1996] for a short review.

2 See Nelson and Winter [1982], Tirole [1988] and Aghion and Howitt [1996a] for surveys on three different strands.

answer is that Schumpeterian competition leads to more monopoly power.

This analysis enriches the existing literature on Schumpeterian competition and growth since we consider here more general demand functions than in previous works. One can refer to Aghion and Howitt [1996a] for a survey of new growth theories based on Schumpeterian incentives. Our findings underline the potentially unbalanced nature of the growth process. In this respect, our work is reminiscent of qualitative considerations that can be found in previous work (see for instance Kamien and Schwartz [1982] or Nelson and Winter [1982]).

The remainder of the chapter is organised as follows. Section 2 provides a real-life example to motivate the argument. The model is set up in Section 3, and the results are presented and discussed in Section 4. Section 5 contains some concluding remarks.

II. AN EXAMPLE

Before undertaking any formal analysis, it is worth having a look at one particular example of the process this chapter focuses on. Of course, real life case studies never exactly match a theoretical argument, especially when this case study is a complex industry. Nonetheless it seems that the evolution of the computer industry since 1975 illustrates our argument.[3] (A detailed description of the rivalry in this sector can be found in Heller [1994]. See also Pugh [1995, 313-136] and Langlois and Robertson [1995] for a short technical chronology.)

The two generic products we identify are, on the one hand, micro-computers and, on the other hand, mainframes and mini-computers. These two products can reasonably be taken as substitutes for each other since the question whether to adopt a micro or mainframe computer strategy arises in many firms. Of course, NASA probably needs more than micro-computers, whereas most households will not buy a mainframe system. But for most firms, the choice between the two is relevant. At the end of the 1970's, the "micro" segment was

[3] Mokyr [1990, pp. 126-135] provides another interesting example with the transport technologies and the rise of the internal combustion engine at the expense of the steam engine.

virtually non-existent and the "mainframe/mini" segment was dominated by IBM. Other competitors on this market were DEC, Wang and Hewlett-Packard. In 1977 Apple Computers, exploiting a technology developed by Xerox, launched the Apple II. This computer contained some major innovations: the mouse, the graphical user interface and the use of floppy disks to store data. This was the first serious attempt to establish a monopoly position on the segment two years after the first micro-computers were introduced. Note, however, that despite (or because of) major improvements with respect to existing products, the product made a slow start. IBM also entered the market in 1981. It tried to establish a dominant position by launching more powerful products and leaving an open access to its standard to other firms (that is, offer a better product and reinforce the lead with a network externality). IBM had 63% of the market in 1984, but the figure had shrunk to 38% by 1987 (Heller [1994]).

Neither Apple, nor IBM, nor any other competitor (such as Compaq) succeeded in establishing a long-run monopoly position. Yet two firms today enjoy very high market power in the micro-computer industry. First, Intel introduced better chips (successively the 80286, 80386, 80486 and Pentium chips) which gave it 2/3 of the market by the beginning of the 1990s', leaving its next challenger, Motorola, far behind (13% of the market). Intel resisted the challenge of the RISC technology, but its monopoly position is still at risk, threatened by a possible drastic technological innovation (be it the SPARC technology or the PowerPC chip). The other dominant firm, Microsoft, is at the other end of the value chain. Its first hit was MS-DOS which became the default operating system for PCs. Then, using the same ingredients as Apple, it brought "user-friendliness" in the PC universe with Windows which became the universal graphical system at the beginning of the 90's. These two Microsoft products (DOS and Windows) successfully withstood the attempts of IBM (OS/2) to capture a major share of the market. However the strong position of Microsoft may be overthrown in the future by innovative ways of supplying computing services (using low-tech terminals and a server on the Internet or in a mainframe system for instance).

Following these major innovations (better chips and user-friendly software), the sales of micro-computers expanded dramatically, whereas the market for mainframe and mini-computers took the opposite direction.[4] This latter market did not really experience any major technological improvement. For instance, the 370 series by IBM had a life cycle of more than 20 years. This lack of innovation does not, however, stem from the absence of R&D since all the major players in the market have invested a lot in their products. Of course, relative to the investments in the micro-computer technology, investments in mainframe systems and mini-computers have been dwindling. (However, the boundary between micro- and mini-computers is blurred because many micro-computers today are much more powerful than most mini-computers 15 years ago.)

After this very brief account, some remarks can be made. First, both markets are characterised by intense rivalry. Firms try to build monopoly positions through a superior technology. Network externalities are pervasive in the industry, but the "standards" are not fixed for as long as they are for keyboards or even for video recorders: these monopoly positions as we saw are temporary. So network externalities must be seen as "multipliers", helpful to reach a dominant situation for a superior product and as barriers to entry when products are good, but without a clear superiority. To summarise, in the computer industry, entrepreneurs, such as Bill Gates and Steve Jobs, invest, create new firms, innovate, launch new products and try to establish temporary monopoly positions. So competition in the computer industry can be held as "Schumpeterian".

Interestingly, the competitive process in this industry has lead to the emergence of "large monopolies". Yet, as might be expected from a Schumpeterian perspective, the competitive process led in each segment to the predominance of one firm, at least on some crucial parts of the value-added chain. But it also led to the domination of nearly all the industry by one type of product. In other words, the monopoly for one segment of the market is now in a dominant position on the whole "industry" since the market for micro-computers is now the dominant market for the computer industry. If one views improvements in micro- versus mini-computers as two different possible directions for technological change, then Schumpeterian competition has led to a unique rather than diversified direction of innovation (the domination of micro-computers over any parallel development of both technologies).

[4] "Have mainframe computers missed the bus?" asked the *Financial Times* (4 September 1996).

Besides, the evolution of the computer industry in the last 20 years has been "unpredictable" or at least "unexpected". It is worth noting that everyone expected a high growth rate, but that most predictions as to its *direction* turned out incorrect. The first justification for this view is that most firms in the mainframe/mini segment continued to invest heavily long after the launch of the first micro-computers. Secondly, IBM, whose knowledge of the computer market cannot be denied, contributed heavily to the improvement of products which eventually took over its own market.[5] Thirdly, many users kept buying those mini-computers or mainframe systems which soon became obsolete. Yet, it is always possible to argue *ex-post* that the rise of PCs was inevitable. But in this case, why didn't everybody emulate Microsoft (or Compaq) at the outset? This leads us to insist on both the inability of people to foresee the systemic implications of innovations and on the role of "historical accidents".[6] The unique direction of progress has been taken following historical accidents. In our case, one can cite poor strategic decisions by IBM, and some successful bets by Microsoft.[7] Of course, here we find again the arguments of the literature on the network externalities (Arthur [1989] or David [1985]). However, they must be amended since it seems that here the competitive and innovative processes must be at the forefront of the story. Despite its late start, the micro-computer has come to dominate the industry and has emerged as a new "general purpose technology" after a succession of purposeful innovations.[8] Indeed, none of these innovations was revolutionary in itself and would have been sufficient to insure the domination of micro-computers, but taken together, they transformed an initially inferior product into a superior one.

The macroeconomic consequences of this are not negligible, since mini-computers and micro-computers cannot be considered as two perfectly substitutable technologies performing the same tasks in the same way. Rather we may interpret them as two different architectures in the sense of Sah and Stiglitz [1986]. In the trade-off between power and

[5] IBM executives, according to a widespread opinion, viewed the micro-computer segment as a niche market, and were unable to forecast its subsequent developments.

[6] Among the many examples given by Rosenberg [1994] and [1996], Marconi saw the usefulness of radio transmission for point to point communication without seeing its implications for the mass entertainment industry. In the 40s, Thomas Watson, chairman of IBM, considered that the market for computers was probably limited to 5 customers and was unable to foresee the potential usefulness of his own products.

[7] Even admitting that without Bill Gates, another businessman would have made them, other decisions than those actually made by IBM could have given a different shape to the industry.

[8] These innovations were "drastic" in the sense that they allowed monopoly positions at the segment level but they remained small at the economy-wide level.

flexibility, mainframe systems maintain their (relative) advantage on the "power" side, whereas micro-computers have an undeniable competitive advantage on the "flexibility" side. Micro-computers are more easily accessible and give more autonomy to their users. Then, concerning R&D activities, one may conjecture that the domination of micro-computers might lead to dispersed research projects of small size.[9] Being conceived and carried on by small-scale teams (often one researcher), these projects may be biased towards the creation of products inducing even more flexibility. The implications of this can be a research process leading to more experiments and the exploration of more research directions, but also smaller innovations (less co-ordination between projects, loss of possible increasing returns to size in research teams, etc). In any case, micro-computer-based research is likely to have a qualitative impact on the content of the growth process.

III. DESCRIPTION OF THE MODEL

TIMING AND DYNAMIC STRUCTURE. Consider an economy populated with overlapping generations of agents. They live for two periods and supply one unit of labour inelastically during their youth. The population of each generation is normalised to one. Labour can be allocated either to the production of (perishable) goods, or can be invested in attempts to improve existing technologies. In case of success in these attempts, the improvement made in period t can be implemented in period $t+1$. A patent is granted to the author of the discovery. The length of the patent is one period only. Following its successful effort made in t, the firm holding this patent will then be able to earn some monopoly profits in $t+1$.

Note that the dynamic structure of our model is standard. The young can save by making a risky investment in more productive technologies (the possibility of other sorts of savings is discussed in Section IV). The expected return on these savings is equal to the monopoly profits in the next period, if any, multiplied by the probability of success. Our dynamic structure is much simplified to ease the analytical derivation of the results. Note also that one-period patents may affect the overall incentive to invest, but this does not play any other role in the argument that deals with the distribution of investments and does not

[9] In comparison to what may occur when research is undertaken using mainframe systems.

change the qualitative features of Schumpeterian competition. An alternative interpretation of our assumption is that patents are not enforceable and that it takes one period for the other producers to learn to imitate.

DEMAND. There are two goods in our economy (our model can be generalised directly to the n-good case). Those goods are horizontally differentiated and are gross substitutes in the benchmark case. For ease of exposition, we assume a CES utility function at each period and risk neutrality.[10] (More general utility specifications are discussed in Section IV.) More formally, the utility function is of the form:

$$U_t = u_t^b u_{t+1}, \tag{1}$$

where $u = \left(q_1^\varepsilon + q_2^\varepsilon\right)^{1/\varepsilon}$, with $0 < \varepsilon < 1$. (2)

Taking the wage as numéraire, the budget constraint is

$$1 - s = P_1 q_1 + P_2 q_2, \tag{3}$$

where s denotes the savings. Then, consumers maximise their utility subject to this budget constraint. The first-order conditions give:

$$s = \frac{1}{1+b}, \tag{4}$$

and $\dfrac{q_1}{q_2} = \left(\dfrac{P_2}{P_1}\right)^{1/1-\varepsilon}$. (5)

Of course, the aggregate demand for good i is the sum of all individual demands:

[10] The assumption of risk neutrality is widespread in the Schumpeterian literature. It follows the claim of Schumpeter [1911] pretending that "the entrepreneur does not take risks". The assumption of risk-aversion is on the contrary at the heart of Knightian theories of entrepreneurship (e.g., Kanbur [1979]). In our model, risk-aversion would reduce overall investment and weaken our results due to the will of diversification. However, in a model with a large number of goods, an application of the law of large numbers would eliminate aggregate fluctuations and would yield similar implications to those of our assumptions here.

$Q_i = \sum_j q_i^j$. Total expenditure is denoted E (and is equal to the consumption expenditures of the young and the wealth of their elders, if any).

SUPPLY. The two goods are produced with constant returns to scale and labour is the only factor of production. The production function of good i by firm k is

$$y_i^k = A_i^k l_i^k \qquad i = 1{,}2\,, \tag{6}$$

where A_i^k is the productivity of labour and l_i^k is the employment of firm k to produce good i. Total supply is defined by $Y_i = \sum_k y_i^k$. We can also define $L_i = \sum_k l_i^k$, the labour demand in sector i. Of course, market clearing conditions imply that $Y_1 = Q_1$, $Y_2 = Q_2$ and $L_1 + L_2 = 1$. Since labour is taken as the numéraire, it is convenient to invert the production function and write the cost function of firm k to produce good i. We have $C_i^k(y) = C_i^k \times y$ where $C_i^k = 1/A_i^k$. The lowest marginal cost in the industry is denoted by $\underline{C}_i = \underset{k}{Min}\left(C_i^k\right)$, the highest productivity is $\overline{A}_i = 1/\underline{C}_i$ and we conventionally define the origin of time such that $\underline{C}_{i,1} \leq 1$.

In all markets, firms compete *a la* Bertrand:

- If no innovation occurred in t-1, $\forall k$, $C_{i,t}^k = \underline{C}_{i,t} = \underline{C}_{i,t-1}$. Then it is immediate that $P_{i,t} = \underline{C}_{i,t}$ and $\pi_{i,t}^k = P_{i,t} A_{i,t}^k l_{i,t}^k - l_{i,t}^k = 0$.
- If an innovation occurred in t-1 and if $\underline{C}_{i,t} \leq \varepsilon \underline{C}_{i,t-1}$ then from the first-order condition for profit maximisation, the monopoly price of good i is

$$P_i = \frac{\underline{C}_i}{\varepsilon}. \tag{7}$$

Assuming the monopolist is unable to internalise the income effect of its own supply, the monopoly profit is equal to

3/ Erratum to pages 140-141. The following replaces the part between equation (7), page 140 and equation (10), page 141.

$$P_i = \frac{\underline{C}_i}{\varepsilon}\left(1+\frac{1-\varepsilon}{x}\right) \quad \text{with} \quad x = \left(\frac{P_i}{P_j}\right)^{\varepsilon/(1-\varepsilon)}. \tag{7}$$

Of course we need the mark-up $(1-\varepsilon+x)/x\varepsilon$ to be below $1/\gamma$ otherwise the limit pricing condition is binding so that

$$P_i = Min\left\{\frac{\underline{C}_i}{\varepsilon}\left(1+\frac{1-\varepsilon}{x}\right), \frac{\underline{C}_i}{\gamma}\right\}. \tag{8}$$

If $(1-\varepsilon+x)/x\varepsilon \leq 1/\gamma$, monopoly profit is equal to

$$\pi_i^M = \frac{(1-\varepsilon)E}{1-\varepsilon+x}, \tag{9}$$

otherwise, when the price is determined by the limit pricing condition, monopoly profit in this case is

$$\pi_i^M = \frac{(1-\gamma)E}{1+\left(\underline{C}_i/\gamma P_j\right)^{\frac{\varepsilon}{1-\varepsilon}}}. \tag{10}$$

$$\pi_i^M = \frac{(1-\varepsilon)E}{1+\left(\underline{C}_i/\varepsilon P_j\right)^{\frac{\varepsilon}{1-\varepsilon}}}. \tag{8}$$

- If an innovation occurred in t-1 and if $\underline{C}_{i,t}/\underline{C}_{i,t-1} = \gamma > \varepsilon$ then the monopoly price is given by the limit pricing condition

$$P_i = \frac{\underline{C}_i}{\gamma}. \tag{9}$$

Thus, the monopoly profit is in this case

$$\pi_i^M = \frac{(1-\gamma)E}{1+\left(\underline{C}_i/\gamma P_j\right)^{\frac{\varepsilon}{1-\varepsilon}}}. \tag{10}$$

RESEARCH. Risky process innovation is the only form of investment we consider. We denote $\Phi(I_i)$, the probability of a process innovation occurring for product i if an amount I_i of labour is invested in the R&D for this product (and of course $I_1 + I_2 = \sum s$). We assume that $\Phi(1) \leq 1$, $\Phi' > 0$ and $\Phi'' < 0$. This hazard function exhibits decreasing returns to scale at the aggregate level. At the individual level, we assume that the probability of one firm getting the patent if an innovation occurs is equal to the ratio of its own investment divided by the total investment in this product. If the innovation occurs, we have $\underline{C}_{i,t} = \gamma \underline{C}_{i,t-1}$ with $\gamma < 1$. The case of a more general improvement function (e.g., inter-industry spillovers and/or improvements of variable size) is discussed later. Note, however, that our assumptions are similar to those used in the recent Schumpeterian competition literature (e.g., Grossman and Helpman [1991, chapter 4] or Aghion and Howitt [1992]) except that they are stated here in a discrete-time framework. Moreover, the model is expressed in terms of cost-reducing investments, but with minor changes it could also be interpreted in terms of quality improvements.

3/ Erratum to page 142-144. The following replaces Proposition 1 and its proof. The first paragraph on page 144 should be ignored.

Proposition 1 *If* $\underline{C}_{i,t} \leq \underline{C}_{j,t}$ *then* $I_{i,t} \geq I_{j,t}$.

Proof From above,

$$P_{i,t+1} = Min\left\{\frac{\underline{C}_{i,t+1}}{\varepsilon}\left(1+\frac{1-\varepsilon}{x_{t+1}}\right), \frac{\underline{C}_{i,t+1}}{\gamma}\right\} \tag{8'}$$

and $$P_{j,t+1} = Min\left\{\frac{\underline{C}_{j,t+1}}{\varepsilon}\left(1+(1-\varepsilon)x_{t+1}\right), \frac{\underline{C}_{j,t+1}}{\gamma}\right\} \tag{8''}$$

when innovations occurs in both sectors at date t. Then we can show that $\partial x_{t+1}/\partial\left(\underline{C}_{i,t+1}/\underline{C}_{j,t+1}\right) \geq 0$ and $x_{t+1} = 1$ if $\underline{C}_{i,t+1}/\underline{C}_{j,t+1} = 1$. If an innovation occurs at date t in sector i only, then $P_{j,t+1} = \underline{C}_{j,t}$. If we write $z_{t+1} \equiv \left(P_{i,t+1}/\underline{C}_{j,t}\right)^{\varepsilon/(1-\varepsilon)}$, we can show easily that $\partial z_{t+1}/\partial\left(\underline{C}_{i,t1}/\underline{C}_{j,t}\right) \geq 0$ and $z_{t+1} \leq 1$ if $\underline{C}_{i,t} \leq \underline{C}_{j,t}$. Then, expected returns on investment in R&D in sector i are

$$\frac{E\left(\pi^M_{i,t+1}\right)}{I_{i,t}} = \frac{\Phi\left(I_{i,t}\right)}{I_{i,t}}\left[\frac{(1-\varepsilon)E_{t+1}}{1-\varepsilon+x_{t+1}}\left(1-\Phi\left(I_{j,t}\right)\right) + \frac{(1-\varepsilon)E_{t+1}}{1-\varepsilon+z_{t+1}}\Phi\left(I_{j,t}\right)\right]. \tag{11}$$

Due to our specification of the utility function (risk neutrality in the second period), the following no-arbitrage condition must hold

$$\frac{E\left(\pi^M_{1,t+1}\right)}{I_{1,t}} = \frac{E\left(\pi^M_{2,t+1}\right)}{I_{2,t}}. \tag{12}$$

After simplification, equation (12) can also be written

$$\begin{aligned}&\frac{\Phi\left(I_{1,t}\right)}{I_{1,t}}\frac{1}{1-\varepsilon+x_{t+1}}\left(1-\Phi\left(I_{2,t}\right)\right) + \frac{\Phi\left(I_{1,t}\right)}{I_{1,t}}\frac{1}{1-\varepsilon+z_{t+1}}\Phi\left(I_{2,t}\right) - \\ &\frac{\Phi\left(I_{2,t}\right)}{I_{2,t}}\frac{1}{1-\varepsilon+1/x_{t+1}}\left(1-\Phi\left(I_{1,t}\right)\right) - \frac{\Phi\left(I_{2,t}\right)}{I_{2,t}}\frac{1}{1-\varepsilon+1/z_{t+1}}\Phi\left(I_{1,t}\right) = 0.\end{aligned} \tag{13}$$

Now suppose, by contradiction, $I_{1,t} < I_{2,t}$ and $\underline{C}_{1,t} \leq \underline{C}_{2,t}$, then, due to the concavity of Φ,

IV. RESULTS

We can now state our first result:

Proposition 1 *If* $\underline{C}_{i,t} \leq \underline{C}_{j,t}\sqrt{\gamma/\varepsilon}$ *and* $\gamma < \varepsilon$ *then* $I_{i,t} \geq I_{j,t}$.

If $\underline{C}_{i,t} \leq \underline{C}_{j,t}$ *and* $\gamma \geq \varepsilon$ *then* $I_{i,t} \geq I_{j,t}$.

In both cases, $I_{i,t}/I_{j,t}$ *increases as* $C_{i,t}/C_{j,t}$ *decreases.*

Proof We consider first the case of $\gamma < \varepsilon$. Expected returns on investment in R&D in sector i are

$$\frac{E\left(\pi_{i,t+1}^{M}\right)}{I_{i,t}} = \frac{\Phi\left(I_{i,t}\right)}{I_{i,t}}\left[\frac{(1-\varepsilon)E_{t+1}}{1+\left(\gamma\underline{C}_{i,t}/\varepsilon\underline{C}_{j,t}\right)^{\frac{\varepsilon}{1-\varepsilon}}}\left(1-\Phi\left(I_{j,t}\right)\right)+\frac{(1-\varepsilon)E_{t+1}}{1+\left(\underline{C}_{i,t}/\underline{C}_{j,t}\right)^{\frac{\varepsilon}{1-\varepsilon}}}\Phi\left(I_{j,t}\right)\right]. \quad (11)$$

Due to our specification of the utility function (risk neutrality in the second period), the following no-arbitrage condition must hold

$$\frac{E\left(\pi_{1,t+1}^{M}\right)}{I_{1,t}} = \frac{E\left(\pi_{2,t+1}^{M}\right)}{I_{2,t}}. \quad (12)$$

Using equation (8) and after simplification, equation (12) can also be written

$$\begin{aligned}&\frac{\Phi\left(I_{1,t}\right)}{I_{1,t}}\frac{1}{1+\left(\gamma\underline{C}_{1,t}/\varepsilon\underline{C}_{2,t}\right)^{\frac{\varepsilon}{1-\varepsilon}}}\left(1-\Phi\left(I_{2,t}\right)\right)+\frac{\Phi\left(I_{1,t}\right)}{I_{1,t}}\frac{1}{1+\left(\underline{C}_{1,t}/\underline{C}_{2,t}\right)^{\frac{\varepsilon}{1-\varepsilon}}}\Phi\left(I_{2,t}\right)-\\&\frac{\Phi\left(I_{2,t}\right)}{I_{2,t}}\frac{1}{1+\left(\gamma\underline{C}_{2,t}/\varepsilon\underline{C}_{1,t}\right)^{\frac{\varepsilon}{1-\varepsilon}}}\left(1-\Phi\left(I_{1,t}\right)\right)-\frac{\Phi\left(I_{2,t}\right)}{I_{2,t}}\frac{1}{1+\left(\underline{C}_{2,t}/\underline{C}_{1,t}\right)^{\frac{\varepsilon}{1-\varepsilon}}}\Phi\left(I_{1,t}\right)=0.\end{aligned} \quad (13)$$

Now suppose, by contradiction, $I_{1,t} < I_{2,t}$ and $\underline{C}_{1,t} \leq \underline{C}_{2,t}\sqrt{\gamma/\varepsilon}$, then, due to the we have $\Phi(I_{1,t})/I_{1,t} > \Phi(I_{2,t})/I_{2,t}$. Consequently, for (13) to hold, we need:

$$\frac{\Phi(I_{2,t})}{1-\varepsilon+z_{t+1}}+\frac{1-\Phi(I_{2,t})}{1-\varepsilon+x_{t+1}}\leq\frac{\Phi(I_{1,t})}{1-\varepsilon+1/z_{t+1}}+\frac{1-\Phi(I_{1,t})}{1-\varepsilon+1/x_{t+1}}. \tag{14}$$

Rearranging yields that a necessary condition for (13) to hold is:

$$\begin{aligned}&\frac{1}{1-\varepsilon+x_{t+1}}+\Phi(I_{2,t})\left(\frac{1}{1-\varepsilon+z_{t+1}}-\frac{1}{1-\varepsilon+x_{t+1}}\right)\leq\\&\frac{1}{1-\varepsilon+1/x_{t+1}}+\Phi(I_{1,t})\left(\frac{1}{1-\varepsilon+1/z_{t+1}}+\frac{1}{1-\varepsilon+1/x_{t+1}}\right).\end{aligned} \tag{15}$$

Since $x_{t+1}\leq 1$, we have

$$\frac{1}{1-\varepsilon+x_{t+1}}\geq\frac{1}{1-\varepsilon+1/x_{t+1}}. \tag{16}$$

So for (13) to hold it thus is necessary that:

$$\Phi(I_{2,t})\left(\frac{1}{1-\varepsilon+z_{t+1}}-\frac{1}{1-\varepsilon+x_{t+1}}\right)\leq\Phi(I_{1,t})\left(\frac{1}{1-\varepsilon+1/z_{t+1}}-\frac{1}{1-\varepsilon+1/x_{t+1}}\right). \tag{16'}$$

Which implies that we need

$$\frac{1}{1-\varepsilon+1/x_{t+1}}+\frac{1}{1-\varepsilon+z_{t+1}}\leq\frac{1}{1-\varepsilon+1/z_{t+1}}+\frac{1}{1-\varepsilon+x_{t+1}} \tag{16''}$$

which cannot be satisfied if $z_{t+1}\leq x_{t+1}$. Of course, $P_{j,t+1}$ is below $\underline{C}_j$ if there is an innovation in t in sector j. Then using (7), one can write

$$P_i-\frac{\underline{C}_i}{\varepsilon}\left(1+\frac{1-\varepsilon}{x}\left(\frac{P_j}{P_i}\right)^{\varepsilon/(1-\varepsilon)}\right)=0. \tag{16'''}$$

Using the implicit function theorem, it is easy to see that $\partial P_i/\partial P_j\leq 1$. This implies then $z_{t+1}\leq x_{t+1}$ and thus a contradiction in equation (16''). As a consequence, if $\underline{C}_{1,t}\leq\underline{C}_{2,t}$, then $I_{1,t}\geq I_{2,t}$. ■

(the numbering of equations remains consistant with the rest of the chapter) concavity of Φ, we have $\Phi(I_{1,t})/I_{1,t} > \Phi(I_{2,t})/I_{2,t}$. Consequently, for (13) to hold, we need:

$$\begin{aligned}&\left(1-\Phi(I_{2,t})\right)\left(1+\left(\gamma\underline{C}_{1,t}/\varepsilon\underline{C}_{2,t}\right)^{\frac{\varepsilon}{1-\varepsilon}}\right)^{-1}+\Phi(I_{2,t})\left(1+\left(\underline{C}_{1,t}/\underline{C}_{2,t}\right)^{\frac{\varepsilon}{1-\varepsilon}}\right)^{-1}<\\&\left(1-\Phi(I_{1,t})\right)\left(1+\left(\gamma\underline{C}_{2,t}/\varepsilon\underline{C}_{1,t}\right)^{\frac{\varepsilon}{1-\varepsilon}}\right)^{-1}+\Phi(I_{1,t})\left(1+\left(\underline{C}_{2,t}/\underline{C}_{1,t}\right)^{\frac{\varepsilon}{1-\varepsilon}}\right)^{-1}.\end{aligned} \quad (14)$$

It is easy to see that this last condition is never satisfied if

$$\frac{1}{1+\left(\underline{C}_{1,t}/\underline{C}_{2,t}\right)^{\frac{\varepsilon}{1-\varepsilon}}} \geq \frac{1}{1+\left(\gamma\underline{C}_{2,t}/\varepsilon\underline{C}_{1,t}\right)^{\frac{\varepsilon}{1-\varepsilon}}}. \quad (15)$$

That is, after simplification,

$$\gamma\underline{C}_{2,t}^2 \geq \varepsilon\underline{C}_{1,t}^2. \quad (16)$$

Hence we have a contradiction. The demonstration is exactly the same for $\gamma \geq \varepsilon$ except that in equation (12) one needs to use equation (10) instead of equation (8). We find

$$\frac{\Phi(I_{1,t})}{I_{1,t}}\frac{1}{1+\left(\underline{C}_{1,t}/\underline{C}_{2,t}\right)^{\frac{\varepsilon}{1-\varepsilon}}}-\frac{\Phi(I_{2,t})}{I_{2,t}}\frac{1}{1+\left(\underline{C}_{2,t}/\underline{C}_{1,t}\right)^{\frac{\varepsilon}{1-\varepsilon}}}=0, \quad (13')$$

and the contradiction is immediate. The last part of the proposition is given by an application of the implicit function theorem on equations (13) and (13'). For equation (13), note that the first and second term can simplify into a first-order term with only $\Phi(I_1)$ and a second order term in $\Phi(I_2)$ and $\Phi(I_1)$. A symmetric argument applies for the third and fourth terms. For equation (13'), the application of the implicit function theorem is straightforward. ■

Ignore - see Errata sheet.

~~First note that our restriction for the cases when $\gamma < \varepsilon$ is not very strong since it is satisfied when $\underline{C}_{1,t} \leq \gamma \underline{C}_{2,t}$. This means that, if starting from the same initial cost, product 1 has benefited from one more innovation, the condition is satisfied. This condition is sufficient, whereas for $\gamma \geq \varepsilon$, the condition given in the proposition is of course necessary and sufficient.~~

The intuition for the result is very simple. If one product enjoys a cost advantage, then its market size is bigger. Concerning R&D, on the one hand, the cost reduction that will be achieved on this product is going to be smaller in absolute terms (it is proportional to the current cost). On the other hand, it will take place over a greater quantity because the substitution effect dominates the income effect. This can be seen from equation (5), where marginal cost pricing implies that $Q_1/Q_2 = (C_2/C_1)^{1/(1-\varepsilon)}$. Since the monopoly profit margin is equal to a markup over marginal cost, the expected total profit is thus proportional to the ratio of marginal costs to the power of $\varepsilon/(1-\varepsilon)$. Thus when products are substitutes, the incentive to invest is stronger for the product with a lower marginal cost. In case of a symmetric Cobb-Douglas utility function (the limiting case with $\varepsilon = 0$), the incentive to invest is independent of marginal cost. Finally, when the products are complements ($\varepsilon < 0$) the incentive to invest is on the contrary higher in the product with the highest marginal cost.

How robust is this result? Note first that this proposition deals only with I_i/I_j. It is thus independent of our dynamic structure which determines only the total amount of savings. Second, note that this result is valid not only for CES utility functions, but also for any demand function for which the substitution effect is stronger than the income effect. Finally, this result offers some similarities with those obtained in the patent race literature (using different assumptions, see for instance Harris and Vickers [1987]). One possible way to weaken the result is to assume that the cost reduction may become more important for the laggard as the gap between marginal costs increases. However, even if $\gamma\left(\underline{C}_j/\underline{C}_i\right)$ becomes very small (i.e., important cost reduction), the profit is determined by the markup $1/\varepsilon$ as in equation (8). If the gap is big enough, this implies that our property is still valid unless $\gamma\left(\underline{C}_j/\underline{C}_i\right)\underline{C}_j < \gamma\left(\underline{C}_i/\underline{C}_j\right)\underline{C}_i$ (which seems unlikely and would create cycles). We can now explore the dynamic properties of our model.

Proposition 2 *Asymptotically, we observe that the ratio of marginal costs converges to either* $\lim_{t\to+\infty} \underline{C}_{1,t}/\underline{C}_{2,t} = 0$ *or* $\lim_{t\to+\infty} \underline{C}_{1,t}/\underline{C}_{2,t} = +\infty$.

Proof See the Appendix. ■

The Appendix shows that technically our dynamic process is equivalent to an urn process (a non-linear Polya process). This associated urn process is such that at every period, one event occurs. This event is randomly drawn from the following set: {no innovation, innovation for product 1, innovation for product 2, innovations for both products}. The probability of each event depends on the history. In other words, each event is equivalent to adding a ball of a specified colour in an urn and the probability of each colour depends on the composition of the urn. Indeed, knowing the number of past innovations for each product, it is possible to compute their levels of productivity.[11] In turn, the productivity gap determines the probability of each event at the next period through the investment functions. After having established that our process is equivalent to an urn process, it can be checked that the deterministic process associated with our stochastic process converges so that one can apply the theorems of Arthur, Ermoliev and Kaniovski [1987]. Those theorems state that our process converges towards a stable fixed point of the probability function. The fixed point for which $\overline{A}_{1,t} = \overline{A}_{2,t}$ is unstable given the result of Proposition 1 above. Thus our system converges towards a fixed point for which $log\,\overline{A}_{1,t} \neq log\,\overline{A}_{2,t}$, that is for which $\lim_{t\to+\infty} \underline{C}_{1,t}/\underline{C}_{2,t} = 0$ or $\lim_{t\to+\infty} \underline{C}_{1,t}/\underline{C}_{2,t} = +\infty$.

The analysis of urn processes has a long history in mathematics. This type of process was first explored by the Russian mathematician Polya in 1931 (Polya [1931]). His problem was the following. Consider an urn, starting with one red ball and one white ball. At each period, a ball is chosen randomly and replaced. If it is red (resp. white), another red (resp.

[11] Take the initial level of productivity and multiply by the exponent of the number of innovations by the size of the innovations.

white) ball is added. The outcome of this simple process is quite surprising since the proportion of red balls converges towards a random variable uniformly distributed between 0 and 1. Later, the analysis was generalised to more general (i.e., non-linear) urn processes (Hill, Lane and Sudderth [1980]). For instance, if the probability of adding a red ball is 40% if their proportion is less than one half and 60% if it is more than one half, the composition of the urn converges towards either 40% or 60% of red balls. On the contrary if the probability of adding a red ball is 40% if their proportion is more than one half and 60% if it is less than one half, the composition of the urn converges towards 50% of red balls. In the latter (simplistic) case, the process is predictable and ergodic whereas in the former (equally simplistic) case it is not predictable, nonergodic and not necessarily path-efficient (i.e., the best asymptotic outcome may not be picked-up). Using a different approach, Arthur, Ermoliev and Kaniovski [1987] generalised the theorems to the n-dimension case with time variable probability functions.

Returning to the model, we can now understand the intuition of our result. Starting from equal levels of productivity, due to the randomness of the innovative process, some asymmetries are bound to occur over time.[12] These asymmetries do not wash out in the long-run, but on the contrary are self-reinforcing. Due to the mechanism of Proposition 1, when a product enjoys a cost advantage over the other, its price is lower and thus its market size is higher. More specifically, the profit is equal to $\pi^m = \text{Quantity} \times \text{Markup}$, where the markup is equal to $\varepsilon/(1-\varepsilon) \times \underline{C}$. Moreover, Quantity = Expenditures ÷ (Cost + Markup). Consequently, in the expression of the monopoly profit, the marginal cost vanishes and the monopoly profit is just a function of the market size. The bigger the market size, the higher is the monopoly profit and thus the more attractive is R&D for this product. So, a cost advantage implies a positive feed-back effect. This feed-back effect induces divergence between products in the long-run (but still the possibility of leapfrogging when the gap is not too big). So the "principle of vanishing diversity" holds in the long-run under Schumpeterian competition when varieties are gross substitutes. Path-dependency is prominent here since a simple historical accident can trigger divergence.

Graphically, the process can be represented in a two-dimensioned space. On the horizontal

[12] The probability of only innovations for both products to occur is zero.

axis we put the productivity of product 1 divided by the sum of log productivity levels. On the vertical axis, we represent the conditional probability of an innovation for product 1 against product 2.

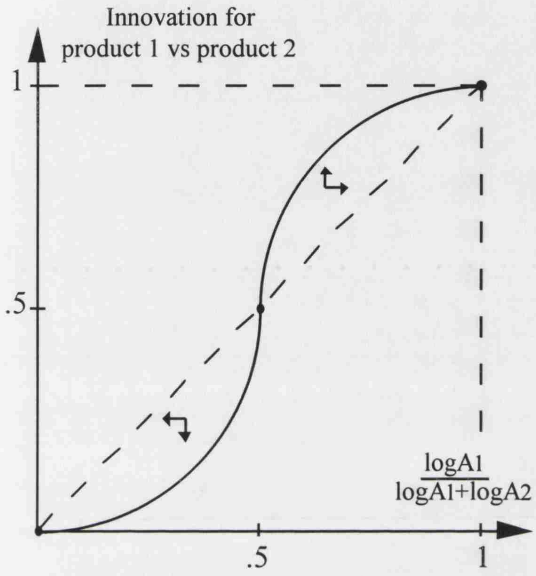


Figure 1a

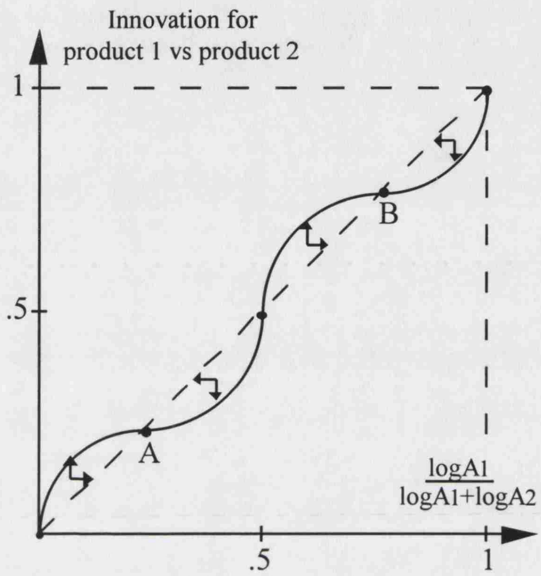


Figure 1b

Two cases can arise. First, if there are some interior fixed points other than {0.5,0.5}, the process converges towards one of the stable fixed point of the set (e.g., A or B in figure 1b). There, the ratio of productivity levels diverges, but it converges in log. In the other case, if {0,0} and {1,1} are the only stable fixed points, productivity levels are going to diverge "strongly" since even the ratio of their log is going either to zero or infinity.

The result of Proposition 2 is reversed when both products are complements ($\varepsilon < 0$). In this case, when a product has a cost advantage, it is more profitable to invest in the other product.[13] As a consequence, productivity gaps vanish in the long-run. In case of inelastic preferences (i.e., Cobb-Douglas), the logs of the productivity levels are draws of a normal distribution due to straightforward application of the Central Limit Theorem (the nature of the result on this limit case is similar to Grossman and Helpman [1991, chapter 4]).

[13] When products are gross complements, the price is determined by a limit-pricing behaviour. The intuition is that the cost reduction for one product will imply a price reduction and thus will boost the demand for the other product making R&D for this second product more attractive.

In relation to existing literature, note that path-dependency is present here as in Arthur [1989]. However, the mechanism that generates it is different. In Arthur [1989], path-dependency stems from increasing returns to adoption and it is viewed essentially as a passive process. Here path-dependency is generated by the interaction between market demand and R&D-based competition to dominate a segment of the industry.

Our model also bears a strong Schumpeterian flavour. Firstly, the competitive process we assume is similar to the one described by Schumpeter [1911]. Competing entrepreneurs perform some R&D for which innovations are random and depend on the effort made. When successful, they reap some temporary monopoly profits. This type of competitive process has been used in the industrial organisation literature where it enjoys quite a long tradition (Tirole [1988]) as well as in the more recent growth literature (Aghion and Howitt [1996a]). Secondly, our model emphasises the development and the emergence of "large monopolies" as a result of the competitive process. Thirdly, it has an "evolutionary" aspect where by evolutionary we mean its non-ergodic and non-predictable dynamics.

Welfare analysis is difficult since we use an overlapping generation framework. The problem is that an increase in investment by a central planner would eventually raise the average growth rate and thus benefit future generations, but at the cost of a decrease in welfare for the current generation. Thus, the Pareto criterion is not really interesting since any level of saving can potentially lead to a different Pareto-optimal path. If now turn to the expected growth rate, it is useful to define the following index:

$$A = \overline{A}_{1,t} \frac{Y_{1,t}}{Y_{1,t} + Y_{2,t}} + \overline{A}_{2,t} \frac{Y_{2,t}}{Y_{1,t} + Y_{2,t}}. \tag{17}$$

Then we can state:

Proposition 3 *The expected growth rate $E(\Delta A/A)$ increases as the ratio of the highest productivity divided by the lowest increases.*

Proof Using Jensen's inequality, the result stems directly from the concavity of the hazard function. For instance, if $\overline{A}_{1,t} = \overline{A}_{2,t}$, we find that $E(\Delta A/A_t) = \Phi(1/2(1+b))$. On the contrary, when $\overline{A}_{1,t} >> \overline{A}_{2,t}$, we find that $E(\Delta A/A_t) = \Phi(1/(1+b))$. ∎

So the expected growth rate of the economy is higher when one product is much more advanced than the other. The reason is the following. The overall amount invested in R&D remains the same whatever the productivity gap. When one product dominates the market, there is no "dilution effect" of R&D investment. So more asymmetry between products has a clear positive effect on the growth rate. However, it would be desirable to enrich the model with the possibility of other sorts of savings (e.g., a storage technology with a sure return). In this case, the total investment in R&D would not be fixed anymore but determined by arbitrage between the different possible investments. This would create a threshold of minimum profitability for the investment in the lagging product with the possibility of a corner solution (all investments in this product stop). Moreover, the divergence between products would imply a negative effect on total R&D investment because of the concavity of the hazard function (the reduction of investment on the lagging product will not transfer completely to the leader but will benefit partly the storage technology).

V. IMPLICATIONS AND CONCLUSIONS

Our analysis has focused on the long-run consequences of Schumpeterian competition. Unlike previous analysis which concentrated either on a one-sector economy (Aghion and Howitt [1992]) or a multi-sector economy with iso-elastic utility functions (Grossman and Helpman [1991, chap. 4]), we consider more general demand functions (even if a simplified dynamic structure is needed for that purpose). When considering an industry where the products are substitutes, our main findings are the following:

Drastic innovations lead to monopolies in each product market (monopolies at the segment level) as usual in this type of model. An innovation for Product 1 may also give it a cost advantage over Product 2. This cost advantage enlarges its market. Then, due to this larger market, R&D in Product 1 offers higher expected returns. Thus in equilibrium, the R&D

investment is higher for Product 1. This increases the chances of the next innovation incurring in the same segment of the industry. This positive feedback effect leads eventually to the domination of one product over the entire industry and the emergence of monopolies at the industry level. Those monopolies however have a shorter life-expectancy since all the investment concentrates on a single product and thus increase the chances of a future innovation on this product.

One can interpret this mechanism as the gradual rise of a new "General Purpose Technology" (GPT). So here, the emergence of the GPT is progressive and occurs without the assumption of exogenous very large innovations as in Helpman and Trajtenberg [1994] or Aghion and Howitt [1996b]. Small innovations complement each other and gradually build-up to form a GPT. Subsequent product development is going to take place around the GPT (see Mokyr [1990, pp. 292-297] for a lengthy discussion of the concepts of micro-invention and macro-invention). For producers, a trade-off between compatibility on the factor side (i.e., necessity to use the GPT) and differentiation on the product side (to avoid competition) is likely to arise.

Our analysis also has implications concerning the direction of technological change. Our competitive process determines not only the rate of growth, but also the "structural change" in the economy. Interestingly, in the model proposed above the average growth rate is predictable, whereas the direction of change depends on historical accidents and cannot be predicted *ex-ante*. If according to Schumpeter [1942, p. 84], "The problem that is usually visualised is how capitalism administers existing structures, whereas the relevant problem is how it creates and destroys them", the message of our analysis is that this destruction and creation of new structures is not deterministic but evolves myopically and randomly despite purposeful investment. Of course, the issue of structural change goes far beyond the analysis proposed here and should be the object of future work.

APPENDIX PROOF OF PROPOSITION 2

The behaviour of our economy can be represented by a stochastic dynamic process with four dimensions. At each period, the following four events can take place {no innovation, innovation for product 1, innovation for product 2, innovations for both products}. The variables $X_{i,t}$, $i = 1,2,3,4$, denote the proportion of each event in the history at date *t*. We can define the following stochastic variables:

$$\beta_1\left(X_{1,t}\right) = \begin{cases} 1 \text{ with probability } \Psi_{1,t} = \Phi\left(I_{1,t}\right)\left(1 - \Phi\left(I_{2,t}\right)\right) \\ 0 \text{ with probability } 1 - \Psi_{1,t}, \end{cases} \tag{A1}$$

$$\beta_2\left(X_{2,t}\right) = \begin{cases} 1 \text{ with probability } \Psi_{2,t} = \Phi\left(I_{2,t}\right)\left(1 - \Phi\left(I_{1,t}\right)\right) \\ 0 \text{ with probability } 1 - \Psi_{2,t}, \end{cases} \tag{A2}$$

$$\beta_3\left(X_{3,t}\right) = \begin{cases} 1 \text{ with probability } \Psi_{3,t} = \Phi\left(I_{1,t}\right)\Phi\left(I_{2,t}\right) \\ 0 \text{ with probability } 1 - \Psi_{3,t}, \end{cases} \tag{A3}$$

$$\beta_4\left(X_{4,t}\right) = \begin{cases} 1 \text{ with probability } \Psi_{4,t} = \left(1 - \Phi\left(I_{1,t}\right)\right)\left(1 - \Phi\left(I_{2,t}\right)\right) \\ 0 \text{ with probability } 1 - \Psi_{4,t}, \end{cases} \tag{A4}$$

with $I_{1,t}$ and $I_{2,t}$ determined by equation (12) from $\overline{A}_{1,t}$ and $\overline{A}_{2,t}$ and by the market clearing condition $I_{1,t} + I_{2,t} = 1/(1+b)$ (which follows from equation (4)). To study the dynamics, let us introduce the following variables

$$b_{i,t+1} = b_{i,t} + \beta_{i,t}\left(X_{i,t}\right), \ i = 1,2,3,4. \tag{A5}$$

The proportions are equal to

$$X_{i,t} = b_{i,t} \Big/ \left(b_{1,1} + b_{2,1} + b_{3,1} + b_{4,1} + t\right). \tag{A6}$$

One can also notice that, following our assumptions, we have

$$log\,\overline{A}_{i,t+1} = log\,\overline{A}_{i,t} + \left(\beta_i\left(X_{i,t}\right) + \beta_3\left(X_{3,t}\right)\right).log\left(1/\gamma\right),\ i = 1,2. \tag{A7}$$

We can also set the initial values at $b_{1,1} = -log\,\overline{A}_{1,1}/log\gamma$, $b_{2,1} = -log\,\overline{A}_{2,1}/log\gamma$, $b_{3,1} = 0$ and $b_{4,1} = 0$. Thus, the probability of each event is given by equations (A1)-(A4), whereas the system (A5) rules the dynamic behaviour of the model. This system is well-defined. The vector of proportions $\{X_i\}$ is obtained from the vector $\{b_i\}$ through equation (A6) which also allows us to calculate $\overline{A}_{1,t}$ and $\overline{A}_{2,t}$. Then using equations (12) and (4), it is possible to know $I_{1,t}$ and $I_{2,t}$. Those investments along with the levels of productivity in turn are sufficient to calculate the vector of probabilities $\{\beta_{i,t}\}$.

To apply Theorem 1 of Arthur *et al.* [1987], one can see immediately that the probability functions $\beta_i(.)$ are time-invariant so that their convergence is trivially given. It remains to check that the deterministic dynamic system associated with equation (A1)-(A5) also converges. To see this, note from proposition 1 that if $\overline{A}_{1,t} > \overline{A}_{2,t}$, then $I_{1,t} > I_{2,t}$ and $E\left(\beta_{1,t}\right) > E\left(\beta_{2,t}\right)$. The process is self-reinforcing until we reach $E\left(\beta_1\right) = \Phi\left(1/1+b\right)$, $E\left(\beta_2\right) = 0$, $E\left(\beta_3\right) = 0$, and $E\left(\beta_4\right) = 1 - \Phi\left(1/1+b\right)$. If $\overline{A}_{1,t} < \overline{A}_{2,t}$, the convergence goes along the same mechanism in the opposite direction. Eventually for if $\overline{A}_{1,t} = \overline{A}_{2,t}$, the vector $\{E\left(\beta_i\right)\}$ remains constant at

$$z_0 = \left\{\Phi\left(1/2(1+b)\right), \Phi\left(1/2(1+b)\right), \Phi\left(1/2(1+b)\right)^2, \left(1 - \Phi\left(1/2(1+b)\right)\right)^2\right\}. \tag{A8}$$

We also define Z the set of fixed points of Ψ: $Z = \{x, \Psi(x) = x\}$. Note that $z_0 \in Z$ and that Z contains a finite number of connected components due to the continuity of β and the monotonicity of the investment functions. As a consequence, the vector of proportions converges with probability one to a point z of Z (Arthur *et al.* [1987]).

Next, we can see that z_0 is the only fixed point for which $X_{1,t} = X_{2,t}$. We can also prove easily that this point z_0 is a nonvertex unstable point of β. It stems directly from the last part of Proposition 1. If after a small perturbation $X_{1,t} > X_{2,t}$, then $I_{1,t} > I_{2,t}$ and $E(\beta_{1,t}) > E(\beta_{2,t})$ (and conversely if $X_{2,t} > X_{1,t}$). We can now apply theorem 3 of Arthur *et al.* [1987] and prove that the process cannot converge towards z_0. Then the process must converge towards a fixed point such that $X_{1,t} \neq X_{2,t}$, that is such that $I_{1,t} \neq I_{2,t}$. Therefore, $\overline{A}_{1,t}$ and $\overline{A}_{2,t}$ will grow asymptotically at different rates. This proves our result. ■

REFERENCES

AGHION P., DEWATRIPONT M., AND REY P., 1995, "Competition, Financial Discipline and Growth", mimeo.

AGHION P., HARRIS C., AND VICKERS J., 1995, "Competition and Growth with Step-by-Step Technological Progress", mimeo.

AGHION P., AND HOWITT P., 1992, "A Model of Growth Through Creative Destruction", *Econometrica* 60, 323-351.

AGHION P., AND HOWITT P., 1996a, "A Schumpeterian Perspective on Growth and Competition", *Advances in Economic Theory*, edited by Kreps and Wallis, Cambridge University Press, forthcoming.

AGHION P., AND HOWITT P., 1996b, "On the Macroeconomic Effects of Major Technological Change", mimeo.

AMSDEN A., AND SINGH A., 1994, "The optimal degree of competition and dynamic efficiency in Japan and Korea", *European Economic Review* 38, Papers and Proceedings, 941-951.

ANDERSON S., DE PALMA A., AND THISSE J.-F., 1992, *Discrete Choice Theory of Product Differentiation*, MIT Press, Cambridge (Mass.).

ARROW K., 1962, "The Economic Implications of Learning by Doing", *Review of Economic Studies* 29, 155-173.

ARNOTT R., AND STIGLITZ J., 1979, "Aggregate Land Rents, Expenditure on Public Goods, and Optimal City Size", *Quarterly Journal of Economics* 92, 471-500.

ARTHUR W.B., 1989, "Competing Technologies, Increasing Returns and Lock-in by Historical Events", *Economic Journal* 99, 116-131.

ARTHUR W. B., ERMOLIEV Y., AND KANIOVSKI Y., 1987, "Path-Dependent Processes and the Emergence of Macrostructure", *European Journal of Operational Research* 30, 294-303.

ASCHAUER D., 1989, "Is public expenditure productive?", *Journal of Monetary Economics* 23, 177-200.

AZARIADIS C., AND DRAZEN A., 1990, "Threshold Externalities in Economic Development", *Quarterly Journal of Economics* 105, 501-526.

BARRO R., 1990, "Government Spendings in a Simple Model of Endogenous Growth", *Journal of Political Economy* 98, S103-S125.

BARRO R., AND BECKER G., 1989, "Fertility Choice in a Model of Economic Growth", *Econometrica* 57, 481-501.

BARRO R., AND SALA-I-MARTIN X., 1992, "Public Finance in Models of Economic Growth", *Review of Economic Studies* 59, 645-662.

BARRO R., AND SALA-I-MARTIN X, 1995, *Economic Growth,* McGraw-Hill, New-York.

BEAN C., AND PISSARIDES C., 1993, "Unemployment, Consumption and Growth", *European Economic Review* 37, 837-54.

BECKER G., 1965, "A Theory of the Allocation of Time", *Economic Journal* 75, 493-517.

BECKER G., MURPHY K., AND TAMURA R., 1990, "Human Capital, Fertility and Economic Growth", *Journal of Political Economy* 98, S12-S37.

BÉNABOU R., 1993, "Workings of a city: Location Education and Production", *Quarterly Journal of Economics* 108, 619-652.

BENHABIB J., AND PERLI R., 1994, "Uniqueness and Indeterminacy: On the Dynamics of Endogenous Growth", *Journal of Economic Theory* 63, 113-142.

BERTOLA G., 1993, "Factor Shares and Savings in Endogenous Growth", *American Economic Review* 83, 1184-1198.

BERTOLA G., 1993, "Factor Shares and Savings in Endogenous Growth", *American Economic Review* 83, 1184-1198.

BLANCHARD I. (ED.), 1994, *Labour and Leisure in Historical Perspective, Thirteenth to Twentieth Centuries*, Franz Steiner Verlag, Stuttgart.

CHARI V.-V., AND HOPENHAYN H., 1991, "Vintage Human Capital, Growth and the Diffusion of New Technology", *Journal of Political Economy* 99, 1142-1165.

COLE H., MAILATH G., AND POSTLEWAITE A., 1992, "Social Norms, Savings Behavior, and Growth", *Journal of Political Economy* 100, 1126-1152.

CORNEO G., AND JEANNE O., 1994, "The Macroeconomic Consequences of Conspicuous Consumption", mimeo.

CRAFTS N., 1995, "Endogenous Growth: Lessons for and from History", mimeo.

CREMER H., AND THISSE J.-F., 1991, "Horizontal Models of Differentiation: a Special Case of Vertical Differentiation Models", *The Journal of Industrial Economics* 39, 1092-1125.

D'ASPREMONT C., GABSZEWICZ J., AND THISSE J.-F., 1979, "On Hotelling's Stability in

Competition", *Econometrica* 47, 1145-1150.

D'ASPREMONT C., AND JACQUEMIN A., 1988, "Cooperative and Non-cooperative R&D in Duopoly with Spillovers ", *American Economic Review* 82, 1133-1137.

DASGUPTA P., AND STIGLITZ J., 1980, "Industrial Structure and the Nature of Innovative Activity", *Economic Journal* 90, 266-293.

DAVID P., 1975, *Technical Choice, Innovation and Economic Growth*, Cambridge University Press.

DAVID P., 1985, "Clio and the Economics of QWERTY", *American Economic Review* 75, Papers and Proceedings, 332-337.

DE LONG B., AND SUMMERS L., 1991, "Equipment Investment and Economic Growth", *Quarterly Journal of Economics* 106, 445-502.

DE LONG B., AND SUMMERS L., 1993, "How strongly do developing economies benefit from equipment in investment ?", *Journal of Monetary Economics* 32, 395-415.

DIXIT A, AND STIGLITZ J., 1977, "Monopolistic Competition and Optimal Product Diversity", *American Economic Review* 67, 297-308.

DOMAR E., 1947, "Expension and Employment", *American Economic Review* 37, 343-355.

DURANTON G., 1997, " Urban Structure, Urbanization and Growth", forthcoming in *Economics of Cities*, edited by J.-M. Huriot and J. Thisse.

EDEL M., AND SCLAR E., 1974, "Taxes, Spending, and Property Values: Supply Adjustment in a Tiebout-Oates Model", *Journal of Political Economy* 82, 941-954.

EICHENGREEN B., 1996, "Institutions and Economic Growth: Europe After World War II", in Crafts N., and Toniolo G. (eds), *Economic Growth in Europe since 1945*, CEPR and Cambridge University Press.

ERIKSSON K., 1996, "Economic Growth with Endogenous Labor Supply", *European Journal of Political Economy* 12, 533-544.

EVANS D., 1987, "Tests of Alternative Theories of Firm Growth", *Journal of Political Economy* 95, 657-675.

FERSHTMAN C., MURPHY K., AND WEISS Y., 1996, "Social Status, Education and Growth", *Journal of Political Economy* 104, 108-132.

FLAHERTY T., 1980, "Industry Structure and Cost-Reducing Investment", *Econometrica* 48, 1187-1210.

FLATTERS F., HENDERSON J.-V., AND MIESZKOWSKI P., 1974, "Public goods, efficiency and regional fiscal equalization", *Journal of Public Economics* 3, 99-112.

FOGEL R., 1994, "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy", *American Economic Review* 84, 369-395.

FREEMAN L., 1994, "Computer Selling Moves into 'Warm, Fuzzy' Phase ", *Advertising Age* 65 (48), S4-S11.

FUJITA M., AND THISSE J.-F., 1996, "Economics of Agglomeration", CEPR Discussion Paper 1344.

FUJITA M., 1989, *Urban Economic Theory, Land Use and City Size*, Cambridge University Press.

GABSZEWICZ J.J., AND THISSE J.-F., 1992, "Location", *Handbook of Game Theory with Economic Application* vol I, dir Auman et Hart, Elsevier Science Publishers (North Holland).

GEORGE H., 1879, *Progress and Poverty*, NY.

GRAMLICH E., 1994, "Infrastructure Investment: A Review Essay", *Journal of Economic Literature* 32, 1176-1196.

GORT M., AND KLEPPER S., 1982, "Time Paths in the Diffusion of Product Innovations", *Economic Journal* 92, 630-653.

GROSSMAN G., AND HELPMAN E., 1991, *Innovation and Growth in the Global Economy*, MIT Press, Cambridge MA.

GROSSMAN G., AND HELPMAN E., 1994, "Technology and Growth", NBER Working Paper 4926.

HAHN F., 1989, "Solovian Models of Growth", University of Cambridge Economic Theory DP 137.

HAMILTON B., 1976, "The Effects of Property Taxes and Local Public Spending, on Property Values: A Theoretical Comment", *Journal of Political Economy* 84, 647-650.

HARRIS C., AND VICKERS J., 1987, "Racing with Uncertainty", *Review of Economic Studies* 54, 1-21.

HARROD R. F., 1948, *Towards a Dynamic Economics*, McMillan, London.

HELLER R., 1994, *The Fate of IBM*, Little, Brown and Company, London.

HELPMAN E., AND TRAJTENBERG M., 1994, "A Time to Sow and a Time to Reap: Growth Based on General Purpose Technologies", CEPR Discussion Paper 1080.

HENDERSON J.-V., 1995, "Will homeowners impose property taxes?", *Regional Science and Urban Economics* 25, 153-181.

HENDERSON J.V., AND SLADE E., 1993, "Development Games in Non-Monocentric Cities", *Journal of Urban Economics* 34, 207-229.

HILL B., LANE D., AND SUDDERTH W., 1980, "A Strong Law for some Generalized Urn Process", *Annals Probabilistic* 8, 214-226.

HOCHMAN O., PINES D., AND THISSE J.-F., 1995, "On the Optimal Structure of Local Governments", *American Economic Review* 85, 1224-1240.

HOTELLING H., 1929, "Stability in Competition", *Economic Journal* 39, 41-57.

HOYT W., 1991, "Competitive Jurisdictions, congestion, and the Henry George Theorem: When should property be taxed instead of land", *Regional Science and Urban Economics* 21, 351-370.

JAFFE A., 1986, "Technological Opportunity and Spillovers of R&D: Evidence from Firm's Patents, Profits and Market Values", *American Economic Review* 76, 984-1001.

JAFFE A., TRAJTENBERG M., AND HENDERSON R., 1993, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations", *Quarterly Journal of Economics* 108, 577-598.

JONES C., 1995a, "R&D-Based Models of Economic Growth", *Journal of Political Economy* 103, 759-784.

JONES C., 1995b, "Time Series Tests of Endogenous Growth Models", *Quarterly Journal of Economics* 110, 495-525.

JONES L., MANUELLI R., AND ROSSI P., 1993, "Optimal Taxation in Models of Endogenous Growth", *Journal of Political Economy* 101, 485-517.

JOVANOVIC B., 1995, "Research, Schooling, Training and Learning-by-Doing in the Theory of Growth", mimeo.

JOVANOVIC B., AND ROB R., 1990, "Long Waves and Short Waves: Growth Through Intensive and Extensive Search", *Econometrica* 58, 1391-1409.

JOVANOVIC B., AND MACDONALD G., 1994a, "Competitive Diffusion", *Journal of Political Economy* 102, 24-52.

JOVANOVIC B., AND MACDONALD G., 1994b, "The Life Cycle of a Competitive Industry", *Journal of Political Economy* 102, 322-347.

KAMIEN M., AND SCHWARTZ N., 1982, *Market Structure and Innovation*, Cambridge

University Press.

KANBUR S. M., 1979, "Of Risk Taking and the Personal Distribution of Income", *Journal of Political Economy* 87, 769-797.

KANEMOTO Y., 1984, "Pricing and Investment Policies in a System of Competitive Commuter Railways", *Review of Economic Studies* 51, 665-682.

KANEMOTO Y., AND KIYONO K., 1993, "Regulation of Commuter Railways and Spatial Development", mimeo.

KING R., PLOSSER C., AND REBELO S., 1988, "Production, growth and business cycles. I. The basic neo-classical model", *Journal of Monetary Economics* 21, 195-232.

KING R., AND REBELO S., 1990, "Public Policy and Economic Growth: Developing Neoclassical Implications", *Journal of Political Economy* 98, S126-S150.

KLEPPER S., 1996, "Entry, Exit, Growth and Innovation over the Product Cycle", *American Economic Review* 86, 562-583.

KRUGMAN P., 1995, *Development, Geography and Economic Theory*, MIT Press, Cambridge (Mass.).

KUZNETS S., 1959, *Six Lectures on Economic Growth*, The Free Press, NY.

KUZNETS S., 1973, "Modern Economic Growth: Findings and Reflections", *American Economic Review* 63, 247-258.

LAFFONT J.-J., AND TIROLE J., 1993, *A Theory of Incentives in Procurement and Regulations*, MIT Press (Cambridge, Mass.).

LANGLOIS R., AND ROBERTSON P., 1995, *Firms, Markets and Economic Change*, Routledge, London.

LUCAS R., 1988, "On The Mechanics of Economic Development", *Journal of Monetary Economics* 22, 3-42.

MADDISON A., 1991, *Dynamic Forces in Capitalist Development*, Oxford University Press.

MANKIW G., 1995, "The Growth of Nations", *Brooking Papers on Economic Activity* 1, 275-326.

MANKIW G., ROMER D., AND WEIL D., 1992, "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics* 107, 407-37.

MARSHALL A., 1890, *The Principles of Economics*, McMillan (N.Y.), 1925.

MARX K., AND ENGELS F., 1967, *The Communist Manifesto*, Harmondsworth, Penguin Books, London.

MEADOWS G.-R., 1976, "Taxes, Spending, and Property Values: A Comment and Further Results", *Journal of Political Economy* 84, 869-880.

MIDGLEY P., 1994, "Urban Transport in Asia", World Bank Technical Paper 224.

MIESZKOWSKI P., 1972, "The property tax: an excise tax or a profit tax", *Journal of Public Economics* 51, 415-435.

MIESZKOWSKI P., AND ZODROW G., 1989, "Taxation and the Tiebout Model: The Differential Effects of Head Taxes, Taxes on Land Rent, and Property Taxes", *Journal of Economic Literature* 27, 1098-1146.

MIRRLEES J., 1972, "The Optimum Town", *Swedish Journal of Economics* 74, 114-135.

MOKYR J., 1990, *Lever of Riches*, Oxford University Press.

MOTTA M., 1992, "Cooperative R&D and Vertical Product Differentiation", *International Journal of Industrial Organization* 10, 643-661.

NELSON R., AND WINTER S., 1978, "Forces Generating and Limiting Concentration under Schumpeterian Competition", *Bell Journal of Economics* 9, 524-48.

NELSON R., AND WINTER S., 1982, *An Evolutionary Theory of Economic Change*, Harvard University Press, Cambridge (Mass.).

NICKELL S., 1996, "Competition and Corporate Performance", *Journal of Political Economy* 104, 724-746.

OATES W.E., 1969, "The Effects of Property Taxes and Local Spending on Property Values: An Empirical Study of Tax Capitalization and Tiebout Hypothesis", *Journal of Political Economy* 77, 957-971.

OLSEN E., 1987, "The demand and supply of housing service: a critical survey of the empirical literature", *Handbook of Regional and Urban Economics* vol II, 989-1022, dir. Mills, Elsevier Science Publishers (North-Holland).

PARENTE S., 1994, "Technology Adoption, Learning by Doing and Economic Growth ", *Journal of Economic Theory* 63, 346-369.

PENCAVEL J., 1986, "Labor Supply of Men", *Handbook of Labor Economics* vol I, 1-103, Ashenfelter and Layard (dir.), Elsevier Science Publishers (North-Holland).

PHELPS E., 1961, "The Golden Rule of Capital Accumulation: A Fable for Growth-men", *American Economic Review* 51, 638-643.

PINES D., 1991, "Tiebout without Politics", *Regional Science and Urban Economics* 21, 469-489.

POLYA R., 1931, "Sur Quelques Points de la Théorie des Probabilités", *Annales de l'Institut Henri Poincaré* 1, 117-161.

PORTER M., 1980, *Competitive Strategy*, The Free Press, New York.

PRUD'HOMME R., 1987, "Financing Urban Public Services", *Handbook of Regional and Urban Economics* vol II, 1179-1206, dir. Mills, Elsevier Science Publishers (North-Holland).

PUGH E., 1995, *Shaping an Industry and its Technology*, MIT Press, Cambridge (Mass.).

REINGANUM J., 1985, "Innovation and Industry Evolution", *Quarterly Journal of Economics* 100, 81-99.

ROMER P., 1986, "Increasing Returns and Long-Run Growth", *Journal of Political Economy* 94, 1002-1037.

ROMER P., 1990, "Endogenous Technical Change", *Journal of Political Economy* 98, S71-S102.

ROSENBERG N., 1994, *Exploring the black box: Technology, economics, and history*, Cambridge University Press.

ROSENBERG N., 1996, "Uncertainty and Technological Change", mimeo.

SAH R.-K., AND STIGLITZ J., 1986, "The Architecture of Economic Systems: Hierarchies and Polyarchies", *American Economic Review* 76, 716-727.

SAMUELSON P., 1954, "The Pure Theory of Public Expenditures", *Review of Economy and Statistics* 36, 387-389.

SCHUMPETER J., 1911, *The Theory of Economic Development*, Harvard University Press, Cambridge MA.

SCHUMPETER J., 1942, *Capitalism, Socialism and Democracy*, Harper and Row, New York.

SCOTCHMER S., 1986, "Local Public Good in an Equilibrium", *Regional Science and Urban Economics* 16, 463-481.

SHELL K., 1966, "Toward a Theory of Inventive Activity and Capital Accumulation", *American Economic Review* 56, 62-68.

SMITH A., 1776, *An Inquiry into the Nature and the Causes of the Wealth of Nations,* Modern Library, New York, edition 1937.

SMITH-MORRIS M. (EDITOR), 1990, *The Economist Book of Vital World Statistics*, The Economist Books Ltd, London.

SMULDERS S., AND VAN DE KULDERT T., 1995, "Imperfect Competition, Concentration and Growth with Firm Specific R&D", *European Economic Review* 39, 139-160.

SOLOW R., 1956, "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics* 70, 65-94.

SOLOW R., 1957, "Technical Change and the Aggregate Production Function", *Review of Economics and Statistics* 39, 312-320.

SOLOW R., 1994, "Perspectives on Growth Theory ", *Journal of Economic Perspectives* 8, 45-54.

SPENCE M., 1984, "Cost Reduction, Competition and Industry Performance", *Econometrica* 52, 101-121.

STARRETT D., 1978, "Market Allocations of Location Choice in a Model with Free Mobility", *Journal of Economic Theory* 17, 21-37.

STIGLITZ J., 1994, *Whither Socialism?*, MIT Press, Cambridge (Mass.).

THISSE J.-F., AND ZENOU Y., 1996, "How to Finance Education When the Labor Force is Heterogenous ", mimeo.

TIEBOUT C., 1956, "A Pure Theory of Public Expenditures", *Journal of Political Economy* 64, 416-424.

TIROLE J., 1988, *The Theory of Industrial Organization*, MIT Press, Cambridge (Mass.).

UHLIG H., AND YANAGAWA N., 1996, "Increasing the Capital Income Tax May Lead to Faster Growth", *European Economic Review* 40, 1521-1540.

VON UNGERN-STERNBERG T., 1988, "Monopolistic Competition and General Purpose Products", *Review of Economic Studies* 55, 231-246.

UZAWA H., 1965, "Optimal Technical Change in an Aggregative Model of Economic Growth", *International Economic Review* 6, 18-31.

VICKERS J., 1995, "Concepts of Competition", *Oxford Economic Papers* 47, 1-23.

VICKREY W., 1977, "The city as a firm", in Feldstein M., et Inman R., *The Economics of Public Services*, MacMillan, London.

WASSMER R., 1993, "Property taxation, property base, and property values: An Empirical test of the 'new view'", *National Tax Journal* 46, 135-159.

WEITZMAN., 1994, "Monopolistic Competition with Endogenous Specialization", *Review of Economic Studies* 61, 45-56.

WILDASIN D., 1987, "Theoretical Analysis of Local Public Economics", *Handbook of*

Regional and Urban Economics vol II, 1131-1178, dir. Mills, Elsevier Science Publishers (North-Holland).

WILDASIN D., 1989, "Interjuridictional Capital Mobility: Fiscal Externality and Corrective Subsidiy", *Journal of Urban Economics* 25, 193-212.

WILHELM M., 1996, "Bequest Behavior and the Effect of Heirs' Earnings: Testing the Altruistic Models of Bequests", *American Economic Review* 86, 874-899.

WILLIAMSON J., 1991, *Inequality, Poverty , and History*, Basil Blackwell, Cambridge (Mass.).

WORLD BANK, 1994, *World Development Report.*

YOUNG A., 1928, "Increasing Returns and Economic Progress", *Economic Journal* 38, 527-542.