London School of Economics and Political Science

Essays in Applied Microeconomics

Patrick Blanchenay

Thesis submitted to the Department of Economics of the London School of Economics and Political Science for the degree of Doctor of Philosophy
London, July 2013

To my father Pierre Blanchenay (1943–2007), who passed away at the beginning of this project, and to my nephew Romain Blanchenay (2013–) who was born at its end.

Declaration

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

I declare that my thesis consists of 27505 words.

Statement of conjoint work

I confirm that Chapter I was jointly co-authored in equal proportions with Emily Farchy.

Statement of inclusion of previous work

I can confirm that Chapter II is a generalization of a previous study for the Masters by Research in Economics at the LSE, awarded in 2008.

Abstract

This thesis addresses three questions using the same tool of microeconomic modelling. In the first chapter (joint with Emily Farchy), I examine the role of individual's decision to acquire broad versus specialist knowledge. I show that a worker can afford to become more specialized on a narrower set of skills by relying on other workers for missing skills. This yields a new explanation of the urban wage premium, and in particular of why workers tend to be more productive in bigger cities, where the existence of better networks of workers provides more incentives to acquire specialized skills. This conclusion matches well established empirical findings on workers' productivity in the literature.

In the second chapter, I look at the dynamics of human capital acquisition over time and show the possibility of what I term a social poverty trap. Namely, parents who do not instil in their offspring the culture of social cooperation (modeled as a higher discount rate) deny them the possibility of future good outcomes; in turn, this new generation will be unable to invest resources in the socialization of their offspring, and so on. This creates a poverty trap where some dynasties are stuck in a bad equilibrium.

In the last chapter, I model political parties campaigning on different issues to voters with limited attention. I assume that the relative salience of the different issues depend on how much time parties devote to each issue. In this setting, I show that campaigning might result in excessive focus on divisive issues (for political differentiation) to the detriment of Pareto-improving ones.

Acknowledgements

It is a trite cliché to say that it takes a village to raise a child, but I am enclined to think that it took a megalopolis to write this PhD thesis.

First and foremost, I would like to acknowledge the help of my supervisor Gilat Levy, for her generous help and constant availability. Her ability to immediately grasp the subtleties of a question and provide insightful comments on them still amazes me to this day. I would also like to thank all the faculty of the Theory Group—in particular Erik Eyster, Philipp Kircher, Francesco Nava, Ronny Razin, Balázs Szentes—for their help, encouragements and comments during and oustide seminars, and the intellectual open-mindedness and honesty that prevailed in their corridors.

I would also like to thank the two examiners for my PhD: Prof. Socorro Puy from Universidad de Málaga, and Prof. Cecilia Testa from Royal Holloway, University London. They have both provided me with excellent suggestions for improvement that have been incorporated in the present volume, and interesting research avenues in order to develop the papers further in the future.

I have also had the privilege to be surrounded by incredible students and friends here. First and foremost my companions in arms, Can Çeliktemur, Kara Contreary, Tara Mitchell and Sebastian Kodritsch, truly amazing people whith whom I have had many lively debates and passionate discussions. Such discussions also abunded in EOPP Happy Hour, where the participation of Madhav Aney, Konrad Buchardi and Stefan Penczynski was so enjoyable. Everyone else in STICERD was also fantastic, with a special thought for Sue Coles who kindly took at heart to help students overcome moments of doubts. A lot of other students also helped with their suggestions, squash challenges and general joviality, in particular Sarah Auster, Giuseppe Berlingieri, Guilhem Cassan, Jason R. Donaldson, Ana McDowall, Oliver Pardo and Rosa Sanchis-Guarner.

In particular, my flatmates Dimitri Szerman, Nathan Converse, Niklas Buehren and Jaime Millan should also be commended for tolerating the smoke in the kitchen and the guitar playing. And of course Laura Derksen for putting up with me during the last and most stressful year of the project, even though she was herself facing challenges.

Support also came from abroad, with little concern for distance. Emily Farchy, whom I have known for a quarter of my life, who co-authored the first chapter, and who has changed me for the better in more ways than she will ever realize. Le soutien espiègle et fidèle de La Compagnie royale des Slacqueurs. Florian Fabre, parce que notre amitié à toute épreuve a bravé les années et les absences. Jean Rives, malheureusement disparu, dont les sages conseils m'ont personnellement touché. Et évidemment, ma famille, aux quatre coins de France et du monde, et en particulier ma mère Anne-Marie, et mes deux frères Nicolas et Bertrand, qui m'ont inconditionnellement offert le réconfort dont j'avais besoin quand c'était nécessaire. J'ai dit au revoir à mon père en commençant ce projet, je dis bonjour à mon premier neveu, Romain, en le terminant.

Contents

Introduction

The present thesis offers three essays that make use of a common method, microeconomic modelling, to provide insight into three distinct settings in which the interaction of self-interested agents can generate important social inefficiencies.

In the first chapter, co-authored with Emily Farchy, I put forth a new explanation for the urban wage premium. Previous study of the geographic distribution of economic activity has identified a persistent stylized fact: individuals living in denser metropolitan areas are more productive and earn more than individuals living in rural areas, even when controlling for higher local prices. Explanations for this regularity broadly fall into two categories (Duranton and Puga 2004). The first relies on sorting: more skilled individuals and firms choose to locate with other skilled agents, while the less skilled are priced out of metropolitan areas. There is indeed conclusive evidence that part of the skill differential between urban and rural areas can be explained by differences in individual characteristics (e.g. Overman and Puga 2010). The other explanation is that of agglomeration externalities, namely that part of the urban wage premium is location-specific, and that there is something intrinsically beneficial from living in denser areas (e.g. learning, sharing factors of production, and improved matching in the labour market).

The model presented in Chapter I combines these two explanations. We suggest that individuals living in metropolitan areas enjoy higher social capital through the existence of networks. Networks generate positive externalities by allowing workers to compensate for their missing skills by relying on other members of the labour force. Workers then have a greater incentive to become more specialized, i.e. to acquire deeper knowledge on a narrower set of skills, and hence become more productive.

Chapter II also posits that differences in social capital can explain individual-level heterogeneity of economic outcomes. However, I approach the question from a very different angle, by looking at the persistence of this heterogeneity from generation to generation, a phenomenon often referred to as poverty trap. I examine individuals' decisions to expend resources in order to socialize their offspring toward cooperative behaviour. In order to do that, I model the propensity of an individual to take advantage of mutually beneficial cooperative opportunities as his or her discount factor. A large literature on Folk Theorems shows indeed that in repeated Prisoner's Dilemma games, patient individuals find it easier to sustain cooperation, and therefore to secure higher payoffs. This is consistent with empirical evidence that more patient individuals tend to achieve better economic outcomes (G. S. Becker and Mulligan 1997; Heckman and Rubinstein 2001).

In such a setting, individuals who have benefited from greater socialization efforts by their parents are able to secure more resources and, in turn, expend more on the socialization of their own offspring. This generates what I call a social poverty trap, in which poor dynasties can never secure enough resources to socialize their offspring to cooperative behaviour, in turn condemning them to lower economic outcomes. The non-convexity, generated by the discontinuity of payoffs between cooperators and defectors, creates the conditions under which heterogeneity in invidivuals' outcomes may perpetuate itself from generation to generation. This suggests another channel through which social mobility can be hampered, based on the intergenerational transmission—or lack thereof—of an element of social capital that I identify as the culture of cooperation.

The inefficiency highlighted in Chapter III is political rather than economic. I attempt to address the following puzzle: why have consensual issues, such as the defense of the environment, taken so long to emerge in the political debate (Poguntke 2002), despite the fact that they gather widespread non-partisan agreement (at least in Europe)? It may indeed seem strange that political parties and candidates sometimes ignore such issues, and take the risk of letting a single-issue party capitalize on such opportunities.

To answer this question, I set up a political economy model of electoral competition in which parties maximize their vote shares by choosing the issues on which to compaign. There are two issues: one, horizontal, which is divisive, and the other, vertical, which is consensual. Each party decides which issue should be part of the platform and which position to take on each issue.

The crucial difference with the standard literature on electoral competition is that the relative salience of the two issues is endogenous. More specifically, I assume that the more parties feature an issue as part of their platform, the more that issue becomes important to voters. The endogenous salience makes it difficult for a single party to "bowl alone" on a given issue, since it might not make the issue important enough to voters. This generates an incentive for parties to emphasize the same issue as their competitors; in such cases, parties prefer to compete on the horizontal, divisive, issue because it offers more scope for political differentiation.

I therefore suggest that consensual issues may fail to surface in the political arena because they do not offer enough scope for parties to differentiate themselves, and because a single-issue party may be unable to put such issues on the (endogenous) agenda of the electoral campaign.

CHAPTER I

Explaining the Urban Wage Premium: The Role of Network Capital

Lord Finchley tried to mend the Electric Light
Himself. It struck him dead: And serve him right!
It is the business of the wealthy man
To give employment to the artisan.

—

I.1 Introduction

The existence of cities is, prima facie, counterintuitive. The agglomeration of people and economic activity into areas where the cost of living and de cost of living and doing business is the highest appears to counter basic logic. Yet while the agglomeration externalities that attract companies to cities, and prompt them to pay higher wages to city workers have by now been well documented, they have not yet been well understood.

This chapter provides an explanation for the productivity benefit that accrues from urban density through the mechanism of social capital and networks. To the extent that networks are denser in denser agglomerations, this represents an alternative explanation of the urban wage premium. [1]

However, where the literature on social capital as a mechanisms for agglomeration externalities has previously focussed on its ability to improve matching (see for example Calvó-Armengol and Jackson 2004), our model sees social capital feeding directly into the production function in the form of the quality of the human capital embodied in the agents network. Thus agglomeration externalities result, not from enhanced learning in denser areas, but from the enhanced ability to rely on the learning of others.

In our model skills are both horizontally and vertically differentiated. Each individual decides on what skills to acquire, but in doing so faces a trade-off between breadth of skills and depth: individuals may become very specialized in a narrow set of skills, or sacrifice depth of skill in favour of breadth becoming

Contrary to most of the literature on specialization as a source of agglomeration externalities (Chaney and Ossa 2013; R. A. Becker and Henderson 2000), we do not rely on increasing returns to specialization to drive the productivity premium associated with density. Conversely we assume decreasing returns to skill depth so that, other things equal, workers do not have an incentive to specialize but prefer to maintain a broad range of skills. In the model developed in this paper we allow specialized workers to *insure* against the risk inherent in specialization by relying, when necessary, on the human capital embodied in the members of their network. As such the larger networks associated with urban density provide complementary incentives for workers to specialize in their skill sets making big cities more productive.

Our model is able to predict, not only the agglomeration externality associated with urban density, but that the magnitude of this externality depends on the aggregate skill level of the city. Furthermore, contrary to much of the literature

[1] As long ago as 1950, Festinger, Schachter, and Back the spatial dependance of social connections has been widely accepted. Since then Glaeser and Sacerdote (2000) have found that in big cities, individuals living in apartment buildings are more likely to engage in social activities with their physical neighbours. From a theoretical perspective, Zenou (2011) examines the interaction between spatial distances and network distances endogenising the location decision of individual agents based upon their position in a social network.[2] Zenou assumes complementarity in the effort invested across all members in a network so that aggregate interactions increase with the density of the network (i.e. cities).

focusing on specialization as a mechanism through which the agglomeration wage premium accrues, our model does not imply perfect segregation according to skills across cities.

Finally this is, to our knowledge, the first attempt to explain why the urban wage premium is maintained but depreciates when a worker leaves a city. This depreciation with distance (De La Roca and Puga 2013) suggests a location-specific component of human capital that is not easily accounted for by a learning mechanism. The model outlined in this chapter is consistent with this phenomenon to the extent that networks are dependent on density and depreciate with distance.

I.2 Related literature

There are two mechanisms that have been put forward to explain the higher productivity of workers observed in cities. The first explanation is that city workers may be systematically more skilled than their counterparts in less dense areas. Such systematic skill disparities may arise either through positive assortative matching among workers—such that more skilled workers sort into cities[3]—or they may simply be the result of differing initial distributions. That is the incentives to acquire skills—and hence the skills observed in equilibrium—may vary systematically with density.[4]

However, Combes, Duranton, Gobillon, and Roux (2012) posit that the relatively minor differences between in- and out-migrants suggests "a modest role for migration in accounting for sorting". Instead they suggest "a role for differences in skill formation prior to the entry on the labour market or for stronger worker

[3] In support of selective sorting much work (see for example Overman and Puga 2010) finds that a large part of the urban wage premium can be explained by observable worker characteristics or, when restricting attention to migrants, to worker fixed effects (as unobservable skills).

[4] Unpacking those types of skill favoured by agglomeration, Bacolod, Blum, and Strange (2009) find that large cities attract at the same time workers with cognitive and interpersonal skills, and unskilled workers, so that larger cities display more heterogeneous distribution of skills. Indeed, in addition to a higher average skill level of urban workers, Combes, Duranton, Gobillon, Puga, et al. (2012) find that the effect of density on the distribution of skills is not well described by first moment. That is, the skill distribution of workers in denser areas is at the same time dilated—workers with very high and low skills being over-represented—and truncated at the lower end. Furthermore, using employment descriptions matched with micro data on occupational outcomes Michaels, Rauch, and Redding (2013) document how, since the end of the 19th Century, urban employment has made a distinct move towards interactive occupations.

learning in cities." Differences in skill formation prior to labour market entry reflect incentive structures that vary systematically with employment density. It is these differing incentive structures and the concomitant optimal skill aquisition decisions that our model hopes to capture.

Furthermore, despite systematic skill differences, both Glaeser and Maré (2001) and much work that has since followed (Combes, Duranton, Gobillon, and Roux 2010; Gibbons, Overman, and Pelkonen 2010) have found evidence that, above and beyond differences in worker characteristics, the higher productivity observed in city workers arises from some form of externality that accrues from the spatial agglomeration of workers. Interestingly these two effects—worker characteristics and agglomeration externalities—appear to interact. That is, the magnitude of agglomeration externalities has been found to differ with skill level of the city.[5]

Puga (2010) seperates these agglomeration externalities into those resulting from sharing,[6], from improved matching[7] from learning, and from specialization. It is on those last two elements that our model focuses. Most closely related to the model of this chapter is the literature—with its roots in the work of Adam Smith—on the role of the market depth that results from employment density on facilitating

---

5 Recent work has found evidence of complementarity between skills and density (Glaeser and Resseger 2010, for example, find agglomeration benefits are present only in relatively skilled cities), while other work suggests low skill workers may benefit disproportionately from agglomeration (Glaeser, Kahn, and Rappaport 2008; Eeckhout, Pinheiro, and Schmidheiny 2010).

6 Literature documenting agglomeration externalities accruing from the ability of co-located firms to share inputs documents the sharing of facilities—see Scotchmer (2002) for a review of the literature on the role of sharing facilities. Burchfield et al. (2006), for example, finds that residences tend to be more clustered in cities where water is provided through shared public facilities; sharing suppliers—regressing, by sector, geographic concentration on proxies for different agglomeration motives, S. S. Rosenthal and Strange (2001) find limited evidence sharing suppliers relative to other motives for agglomeration.Overman and Puga (2010) argue that the importance of input sharing will depend also on the spatial concentration of input suppliers; and sharing a common labour pool—they find that sectors experiencing more idiosyncratic volatility are more spatially concentrated. Ellison, Glaeser, and Kerr (2010) find evidence of co-agglomeration among industries that share labour mix requirements.

7 The increased efficiency of matching in agglomerations has been explored in models such as that of Helsley and Strange (1990) who model the density of urban labour markets, such that: (i) employers skill requirements and (ii) workers skills, are distributed along the circumference of a circle (à la Salop 1979). A thicker skill distribution reduces the mismatch and hence cost of retraining to meet requirements of employer. In addition, the improved chance of finding a match in denser markets may also enable workers to be choosier in accepting offers (Coles and E. Smith 1998; Berliant, Reed, and Wang 2006).

specialization. Along these lines two papers in particular, are closely to related to our own. The first, R. A. Becker and Henderson (2000), models a firm who chooses how to spread one unit of time on tasks where the fulfillment of tasks is subject to *increasing returns* to investment but where, at the same time, the output of each firm is positive only if all tasks are completed. That is, if there is a task to which no firm devotes any time, then the output for all firms is zero. The second related paper by Chaney and Ossa (2013) models a mechanism through which increased market size induces a deeper division of labour, and hence higher productivity, among teams whose costs are subject to both a fixed and marginal component. Here the *increasing returns* to specialization result from the fixed cost component of establishing a team. In our model we try to avoid such an assumption. Indeed rather than assuming increasing returns to specialization (as both Becker and Henderson and Chaney and Ossa) we assume the opposite. We show that decreasing returns to specialization at the individual level still generate increasing returns to specialization on aggregate.

Also related to our model is the literature modelling the benefits from agglomeration in terms of its ability to enhance learning (see Duranton and Puga 2004, for a review of microfoundations). Our model is related to the literature on learning in the sense that density increases the human capital an agent has at her disposal. However, rather than increasing her own human capital the externality in our model arises from the ability of an agent to use the human capital of those in her network. Thus, depending whether one subscribes to the learning explanation for the agglomeration wage premium, or the network explanation outlined in this chapter, the finding that the agglomeration benefit decays with time spent away from a city (De La Roca and Puga 2013) must be explained, either as a result from the erosion of knowledge that accompanies relocation in a less dense area, or by the erosion of networks that results from distance (see Belot and Ermisch 2009, for an empircal investigation of the importance of location in the maintenance of social capital).

This chapter sets out a model to explain these interactions. It provides both a mechanism for the systematic skill differences in urban areas and for the skill dependent agglomeration externalities.

I.3 Model

I.3.1 Workers

The environment is characterized by two sectors: a tradeable sector and non-tradeable sector. Workers may work in either of these sectors and bring to their work a skill level θ.

The skills embodied in a worker are differentiated along two tangents. They are vertically differentiated such that worker i is more deeply skilled than worker j if $\theta^i > \theta^j$ and, in addition, they are horizontally differentiated along a continuous circle of length 1 (à la Salop 1979)—these horizontally differentiated skills represent the different areas of specialization upon which a worker may focus. We model the cost of acquiring skill depth as an opportunity cost—that of foregone breadth. Thus a worker with skill depth θ has a skill support of $\frac{1}{\theta}$. Denote $\eta \in [0, 1]$ the proportion of workers for whom $\theta^i > 1$.

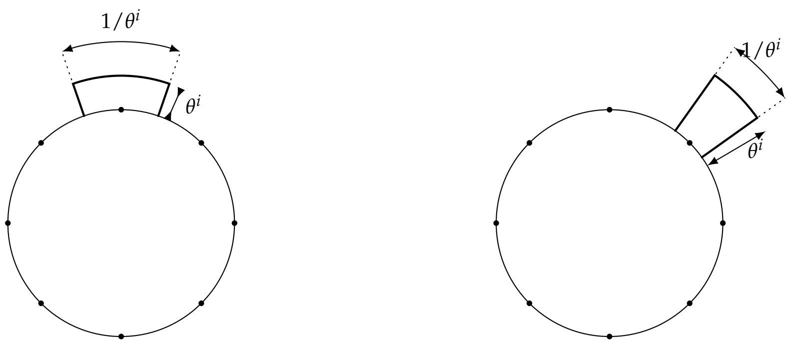


Figure I.1: Two workers with different levels of skill. The one on the left is less specialized and has more breadth.

Each of N workers maximizes a quasi-linear utility taking the form

$$U = 2\sqrt{c_{NT}} + c_T \tag{I.1}$$

where c_T and c_{NT} denote the consumption levels of tradeable and non-tradeables respectively.[8] This choice of utility function is designed to capture a salient consump-

[8] We obtain the same qualitative results when using a more general quasi-linear utility function of the form $U = kc_{NT}^{\beta} + c_T$.

tion pattern in that, as income increases, non-tradeable needs are more easily satiated where consumption of tradeables remains unbounded. As such richer individuals tend to expand their consumption towards tradeable goods—e.g. faster cars.

We normalize the (exogenous) price of tradeable goods to $p = 1$ and denote w the wage (effective price) in the non-tradeable sector.

Conditional on an interior solution, a worker receiving income m, will therefore achieve utility:

$$U = m + \frac{1}{w} \tag{I.2}$$

Each worker is endowed with one unit of labour that she inelastically supplies to the sector—tradeable or non-tradeable—in which she chooses to work. This labour translates into an effective output that is dependent on her skill depth (level of specialization) θ, such that each worker can produce θ^{δ}.

The parameter $\delta \in (0, 1)$ captures the decreasing returns to specialization. Thus contrary to models of the the gains of specialization enabled by market depth—such as that of Chaney and Ossa (2013) outlined above—we do not rely on increasing returns to specialization.

I.3.2 Projects in the Tradeable Sector

Projects in the tradeable sector are indivisible. Such projects can be thought of, for example, as consultancy projects or international business services. These projects are subject to international competition and, as such are characterized by perfectly elastic demand at a price normalized to $p = 1$ exogenous to local supply. Each worker in the tradeable sector receives one project such that the number of tradeable projects in the economy is equal to the number of workers working in the tradeable sector, ηN.

Each project in the tradeable sector requires a specific skill. The required skill is distributed uniformly along the skill circle, independently of the skill location of the individual that receives the project, and independently of the skill location of other projects. Moreover we assume that the completion of projects in the tradeable sector requires a degree of specialization $\theta > 1$. As a result each worker in the tradeable sector requires some degree of specialization and is therefore obliged to leave some range of the skill set uncovered. If this worker's project falls in this uncovered part of the skill set (outside her support), she will be unable to complete the project herself.

To insure against this eventuality tradeable workers can network with other workers in the sector. If networked, a worker whose project requires skills outside her horizontal skill set can rely on the expertise of a member of her network. In this eventuality, transmission costs will erode the efficiency of the skill by $1 - \alpha$ where $\alpha \in (0, 1)$. Thus a worker relying on her network for the specialization to perform her task will have an effective skill level of only $\alpha\theta^{-i}$ where θ^{-i} is the level of skill of the worker on whom she relies. Her effective output will then be $(\alpha\theta^{-i})^{\delta}$.

In Appendix A.2 we identify the sufficient condition under which it will be optimal for any worker working in the the tradeable sector to get networked. Under this condition given by Equation (A.11), the decision to work in the tradeable sector will imply the decision to get networked. In the remainder of the chapter we assume this condition holds.

I.3.3 Non-Tradeable Sector

The non-tradeable sector does not require specialized skills but only labour. As a result the unit price of the output will be equal to the price of one unit of labour—the wage. The non-tradeable wage/price in this sector is determined in equilibrium by local demand and supply.

Aggregate supply of non-tradeables results from each non-tradeable worker inelastically supplying their one unit of labour. Total output in the non-tradeable sector is therefore $(1 - \eta)N$.

Aggregate demand for non-tradeable goods is simply the aggregation of the local individual demands that result from the utility maximization of workers—both in the tradeable and non-tradeable sectors (Equation (I.1)).

I.3.4 Timing

Workers take two decisions. Each decides (i) in which sector they will work and, if they work in the tradeable sector, (ii) their level of specialization—the depth of skill, θ, to acquire.

The timing is as follows:

1. Workers choose their sector and, if tradeable, their level of specialization θ^i.

2. Tradeable workers are each assigned a project.

3. Each worker receives their revenue and makes their consumption choices to maximize their utility.

To identify a subgame perfect equilibrium in which workers are indifferent between working in the tradeable sector and working in the non-tradeable sector we proceed by backward induction examining the payoffs of workers in each sector.

I.4 Equilibrium payoffs

I.4.1 Payoffs in the Tradeable sector

Payoffs in the tradeable sector will depend upon the specialization choice of the worker and the location of their project on the skills circle. To find the indirect utility, therefore, we first calculate the optimal allocation of consumption for a worker working in the tradeable sector.

A worker, producing tradeable goods using a given skill level θ will earn: θ^δ such that Equation (I.2) becomes:

$$U_T(\theta) = \theta^\delta + \frac{1}{w} \tag{I.3}$$

where, recall, w is the price of the non-tradeable good. Therefore, the indirect utility of completing a project in the tradeable sector depends positively on the skill depth θ being used, but at a rate decreasing according to the magnitude of $\delta \in (0, 1)$.

If the worker possesses the appropriate (horizontally differentiated) skill, her indirect utility of working on a project in the tradeable sector will be $U_T(\theta^i) = (\theta^i)^\delta + \frac{1}{w}$.

However, it is by no means guaranteed that a worker possesses the relevant skills in her support. Recall that a worker of specialization θ^i has a skill breadth $1/\theta^i$. Projects arriving according to a uniform distribution will, therefore, only fall in the agents support with probability $1/\theta^i$. When, on the other hand, the project falls outside the worker's skills support and on the skill set of another worker in her network, worker i must rely on the skill of this other worker, θ^{-i}, to perform the project. Recall that, under this scenario, the productivity of skills will be eroded in transmission to $\alpha\theta^{-i}$, where $\alpha \in (0,1)$ captures the strength of the skill network such that individual i receives a utility of only $U_T(\alpha\theta^{-i})$.[9]

Since we will later focus on a symmetric equilibrium, where all individuals in the tradeable sector choose the same level of (vertical) skill specialization, we assume now that all other workers in the tradeable sector (aside from worker i) choose the same level of specialization θ^{-i}.

The decision of each tradeable worker regarding how much to specialize will dictate how much of the skill circle will be covered and the probability $\tilde{p}$ that a project assigned to worker i will fall outside her skill set but inside another worker's skill set. Appendix A.1 identifies conditions under which no worker in the tradeable sector leaves skill gaps with her neighbours. Under these conditions non-tradeable workers thus cover the whole skill circle. Given that worker i skill set covers a fraction $1/\theta^i$ of the skill circle, the probability that her project will fall into another worker's skill set is:

$$\tilde{p} = 1 - \frac{1}{\theta^i} \tag{I.4}$$

and worker i's optimal choice of θ^i, given θ^{-i} is that which maximizes her expected utility:

$$\mathbb{E}U_T^i = \underbrace{\frac{1}{\theta^i} \cdot U_T(\theta^i)}_{\text{Own skill}} + \underbrace{\left(1 - \frac{1}{\theta^i}\right) \cdot U_T(\alpha\theta^{-i})}_{\text{Other worker's skill}} \tag{I.5}$$

From Equation (I.5) it is clear that worker i's expected utility depends on her level of skill θ^i and on that of the other workers in her network θ^{-i}. In choosing θ^i to maximize expected indirect utility a worker faces two opposing incentives. In the first place, by raising θ^i the worker raises her payoff in the event that the tradeable project she draws requires skills in which she specializes. However, at the same time, by becoming more specialized the agent also decreases the likelihood that she has the relevant skills to fulfill her project, $(1 - \frac{1}{\theta^i})$. Increased specialization therefore increases the probability that she must rely on the skills of her network—reducing the efficiency with which she fulfill's her project, and hence her utility.

[9] When the project falls in an area covered by both worker i and someone's else skill set, worker i will perform the project.

This tradeoff is clarified in the differential of the expected utility function:

$$\frac{\partial \mathbb{E} U_T^i}{\partial \theta^i} = \frac{1}{\theta^i} \left\{ \underbrace{U_T'(\theta^i)}_{\substack{\text{Direct effect} \\ \text{on payoff}}} - \frac{1}{\theta^i} \underbrace{\left[U_T\left(\theta^i\right) - U_T\left(\alpha \theta^{-i}\right) \right]}_{\text{Cost of relying on network}} \right\} \tag{I.6}$$

As is clear from the above, the higher is α the smaller is the opportunity cost of relying on another worker's skill set; the incentive to specialize is larger when the network is of higher quality. Optimization of Equation (I.6) under the assumption that all individuals in the tradeable sector are evenly spaced on the skill circle[10], and therefore at a distance $\frac{1}{\eta N}$ from each other, yields the following Lemma.

Lemma I.1. *When δ and α are sufficiently high, worker i will choose level of specialization θ^{i*} that satisfies:*

$$\frac{1}{\theta^{i*}} = \frac{2}{\eta N} - \frac{1}{\theta^{-i}} \tag{I.7}$$

such that they neither overlap nor leave gaps with their neighbours.

Proof. See Appendix A.1 □

In other words, provided that the returns to specialization are sufficiently high (high δ) and that the opportunity cost of specialization is sufficiently low (high α), workers will choose to be as specialized as possible. They will choose a skill set such that they neither overlap nor leave gaps with their neighbours.

The condition identified for the optimal choice of specialization highlights the trade-off that workers are facing when choosing how much to specialize. On one hand, given the decreasing personal returns to specialization (captured by $\delta < 1$), the expected utility of an isolated worker is higher when the worker is not specialized, i.e. when the worker's skill set is broad. This effect is stronger when δ is low, as it decreases the incentives to become specialized. On the other hand, the opportunity cost of specialization (by being unable to complete a project) is overcome by the ability to rely on one's network. This is even more so the case when the quality

[10] This could be endogenized through a location choice stage in the game. Because of the opportunity cost of relying on one's network, there are incentives for each worker to be as far as possible from each other, thereby yielding to a uniform spacing on the skill circle. One could further induce such spacing by arguing as in Salop (1979) that this relaxes price competition.

of the network, captured by α, is high. As a result, when both δ and α are high, specialization is beneficial, on one hand because personal productivity remains high, and because relying on one's network for missing skill is not too penalizing. This induces workers to specialize as much as possible (without leaving gaps with their neighbours).

Note that Lemma I.1 points to the strategic complementarity that exists in the choice of specialization. The more the other workers are specialized, the smaller is the opportunity cost to worker i of relying on another worker's skill, and the larger the incentive to be specialized.

At the optimal level of specialization, a skilled worker in the tradeable sector faces an expected payoff of:

$$\mathbb{E}U_T^i = \frac{1}{w} + \left[\left(\frac{2}{\eta N} - \frac{1}{\theta^{-i}}\right)^{1-\delta} + \left(1 - \frac{2}{\eta N} + \frac{1}{\theta^{-i}}\right)(\alpha\theta^{-i})^{\delta}\right] \tag{I.8}$$

When workers' choice of specialization is neither so broad that it overlaps with the skill support of other workers, nor so narrow that parts of the skill circle are left uncovered[11], the symmetric equilibrium in specialization choice is

$$\frac{1}{\theta^*} = \frac{1}{\eta N} \qquad \text{or} \qquad \theta^* = \eta N \tag{I.9}$$

Proposition I.1. *The higher the* number *of networked workers the more each specializes (see Figure I.2).*

Proof. Follows immediately from Equation (I.9) □

In the symmetric equilibrium[12], the utility of a worker in a tradeable sector is therefore:

$$\mathbb{E}U_T^*(w) = \frac{1}{w} + (\eta N)^{\delta}\left[\frac{1}{\eta N} + \alpha^{\delta}\left(1 - \frac{1}{\eta N}\right)\right] \tag{I.10}$$

---

[11] This will be the case when α and δ are sufficiently large, see Appendix A.2 for the precise condition.

[12] When all $\theta = \eta N$, the condition (A.13) becomes $\alpha^{\delta}(\eta N)^{2-\delta} \geq (1-\delta)$, which will be satisfied for any reasonable parametric assumptions.

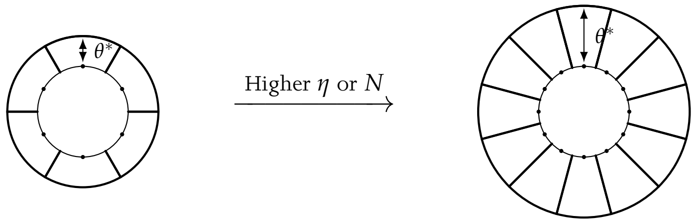


Figure I.2: Effect of an increase in number of tradeable sector workers on the level of specialization.

I.4.2 Equilibrium in the non-tradeable labour market

Effective labour in the non-tradeable sector does not depend upon specialized knowledge. Each worker supplies one unit of labour, accruing revenue w, and producing one unit of non-tradeable good.

After optimization of consumption choices, the indirect utility of a worker in the non-tradeable sector is therefore

$$U^*_{NT}(w) = \frac{1}{w} + w \tag{I.11}$$

Recall that workers in the non-tradeable sector supply their labour inelastically. Thus aggregate supply of non-tradeables is locally defined by the number of individuals working in that sector:

$$L_{NT} = (1 - \eta)N \tag{I.12}$$

Since individual demand for non-tradeables does not depend on income (due to the quasi-linear utility), it is identical for all workers, independently of the sector they work in. Aggregate demand for non-tradeables is therefore given by:

$$C_{NT} = \frac{N}{w^2} \tag{I.13}$$

The equilibrium wage in the non-tradeable sector is therefore that which equates supply and demand such that $L_{NT} = C^{NT}_{NT} + C^{T}_{NT}$:

$$w^*_{NT}(\eta) = \frac{1}{\sqrt{1-\eta}} \tag{I.14}$$

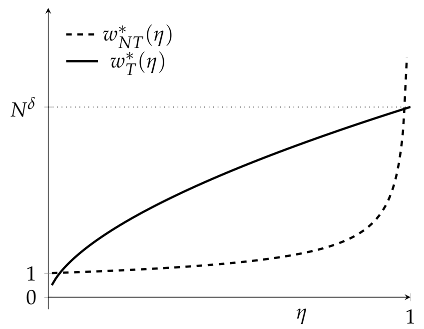


Figure I.3: Equilibrium individual revenues w^*_{NT} and w^*_T as a function of η the proportion of workers in the tradeable sector.

Equilibrium wage, and hence individual revenue, in the non-tradeable sector is increasing and convex in η, and goes from 1 to $+\infty$ as η goes from 0 to 1.

Given the optimal choice of specialization in Equation (I.9), individual revenue in the tradeable sector is: $w^*_T(\eta) = (\eta N)^\delta$ and goes from 1 to N^δ as the number of workers in the tradeable sector goes from 1 to N. We plot individual revenues for workers in the tradeable and non-tradeable sectors in Figure I.3.

Final indirect utility is utility estimated at equilibrium wage $w^*_{NT}(\eta)$ defined in Equation (I.14). In the non-tradeable sector:

$$V_{NT}(\eta) \equiv U^*_{NT}(w^*_{NT}(\eta)) = \sqrt{1-\eta} + \frac{1}{\sqrt{1-\eta}} \tag{I.15}$$

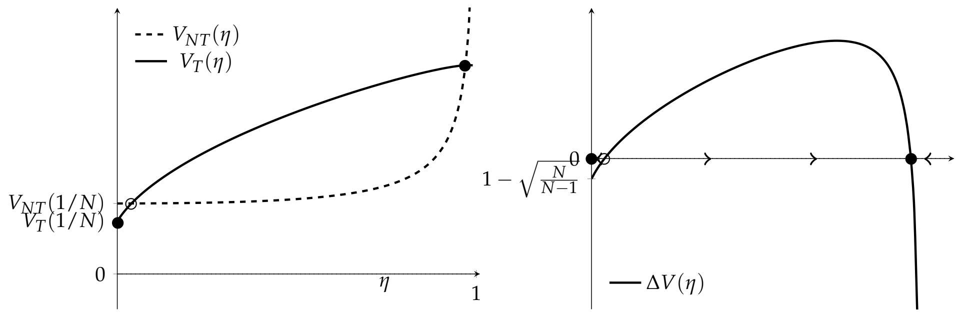
which is increasing and convex in η.

Similarly in the tradeable sector:

$$V_T(\eta) \equiv \mathbb{E}U^*_T(w^*_{NT}(\eta)) = \sqrt{1-\eta} + (\eta N)^\delta \left[\frac{1}{\eta N} + \alpha^\delta \left(1 - \frac{1}{\eta N}\right)\right] \tag{I.16}$$

We plot both of those in Figure I.4.

Note that Equation (I.16) only holds for $\eta > 1/N$, that is when at least one other worker is working in the tradeable sector. The decision facing the first worker to decide to enter the tradeable sector, when there are no workers yet working in that sector is analysed in Section I.5.1 below.



(a) Payoffs $V_T(\eta)$ and $V_{NT}(\eta)$.

(b) Difference $\Delta V(\eta)$ and dynamics.

Figure I.4: Plotting payoffs as a function of η, using parameters $N = 100$, $\alpha = 0.6$, $\delta = 0.6$. Black dots denote stable equilibria; the white dot is an unstable equilibrium.

I.5 Equilibria analysis

All else equal workers will choose to work in the sector yielding the highest payoff. They will make their specialization (and network) choices dependent on the sector in which they work as outlined above.

That is, as long as the payoff V_T is higher than V_{NT}—the payoff they receive if they do not acquire a network—marginal agents will continue to specialize and network such that η will increase. Denote $\Delta V(\eta) \equiv V_T(\eta) - V_{NT}(\eta)$ the difference between the payoffs achievable in the tradeable and non-tradeable sector. Our variable of interest is η^* the proportion of agents who choose to network in equilibrium. We depict the payoffs as a function of η in Figure I.4, with the corresponding dynamics and equilibria.

I.5.1 Equilibrium with only unskilled agents

A corner equilibrium $\eta^* = 0$ (no one acquires skill) exists if the payoff of acquiring skills *alone* is lower than remaining unskilled in the non-tradeable sector. [13]

$$V_T\left(\frac{1}{N}\right) \tag{I.17}$$

If all workers work in the non-tradeable sector, all agents produce and consume one unit of non-tradeable good, therefore getting indirect utility $V_{NT}(0) = 2$.

If all but one workers work in the non-tradeable sector, equilibrium wage in the non-tradeable sector would be $w^*_{NT} = \sqrt{1 - \frac{1}{N}}$. A worker who would be alone in the tradeable sector would not be able to rely on her network. If the level of skill required for projects in the tradeable sector is $\bar{\theta}$, she would choose level of specialization $\theta^i = \bar{\theta}$, therefore earning payoff $V_T(1/N) = \bar{\theta}^{\delta-1} + \sqrt{\frac{N}{N-1}}$, which is decreasing in $\bar{\theta}$.

Comparing the payoffs by both the above scenario's gives us the following proposition on a non-skilled equilibrium:

Proposition I.2. *If the skill requirement in the tradeable sector, $\bar{\theta}$, is higher than a threshold $\left(\frac{\sqrt{N-1}}{2\sqrt{N-1}-\sqrt{N}}\right)^{\frac{1}{1-\delta}}$, there exists an equilibrium in which all individuals work in the non-tradeable sector ($\eta^* = 0$) and no-one acquires skill.*

I.5.2 Interior equilibria

Since there is no restriction on who can get networked (and specialized), an interior equilibrium requires a non-arbitrage condition.

More precisely, an interior equilibrium exists when there exists $\eta^* \in (0, 1)$ such that:

$$V^i_T(\eta^*) = V^i_{NT}(\eta^*) \qquad \Longleftrightarrow \qquad \Delta V(\eta^*) = 0 \tag{I.18}$$

Since the function $V_T(\eta)$ is weakly concave in η, and since the function $V_{NT}(\eta)$ is strictly convex in η, their difference $\Delta V(\eta)$will be strictly concave in η.

[13] While η is continuous when more than one worker is working in the tradeable sector, the payoff of the first worker in the tradeable sector cannot be obtained by taking the limit of V_T as η tends to 0. This is because the first individual in the tradeable sector cannot rely on any network.

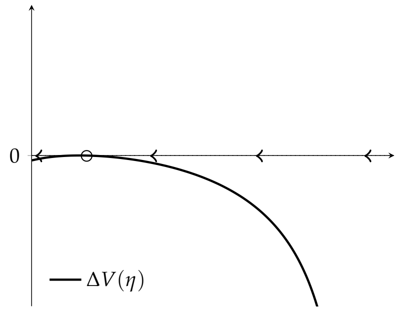


Figure I.5: Case of a unique interior equilibrium: the equilibrium is unstable because of the dynamics for $\eta < \eta^*$.

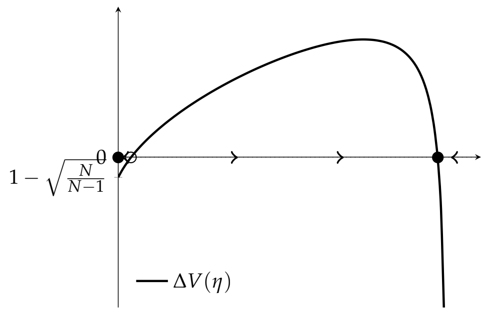


Figure I.6: Case of two interior equilibria (and one stable corner). Only the right interior equilibrium is stable.

If the difference function $\Delta V(\eta)$ crosses the zero-axis (i.e. if there exists an interior equilibrium), it must either:

1. Cross it in a unique point then fall (Case 1, illustrated in Figure I.5), or;

2. Cross it twice: once upward sloping, once downward-sloping (Case 2, illustrated in Figure I.6).

We consider the *stability* of these potential interior equilibria.

Case 1: There is a unique η^* such that $\Delta V(\eta^*) = 0$.
If there is a unique η^*, as in Figure I.5, then it will be unstable; any small deviation to a smaller $\eta < \eta^*$ leads to $\Delta V(\eta) < 0$. This means that agents can get a higher payoff by being unspecialized, which in turns would lead to a lower η, and so on until η^* degenerates to zero.

Case 2: There are two distinct levels of networks $0 < \eta_1^* < \eta_2^* < 1$ such that $\Delta V(\eta^*) = 0$.

When there are two distinct equilibria, Figure I.6, ΔV must be increasing at η_1^* and decreasing at η_2^*. This is because, since $\lim_{\eta \to 0} \Delta V(\eta) < 0$, and $\lim_{\eta \to 1} \Delta V(\eta) < 0$, and since $\Delta V(\cdot)$ is continuous and strictly concave, it must be that $\Delta V'(\eta_1^*) > 0$ and $\Delta V'(\eta_2^*) < 0$.

It therefore follows that:

- The interior equilibrium at η_1^* is *unstable*. This is because: A small deviation to the left would lead to $\Delta U(\eta) < 0$ and hence take the system to the stable equilibrium, outlined above, at $\eta^* = 0$. Conversely, a small deviation to the right yields $\Delta U(\eta) > 0$ leading to a convergence on the stable equilibrium, outlined below, at η_2^*.
- The interior equilibrium η_2^* is *stable* since the dynamics work in the opposite direction.

Let us now focus on the interior stable equilibrium.

I.5.3 Comparative statics on the stable interior equilibrium

Let us focus on the case where there is a *stable interior* equilibrium. Our main result is to show that the equilibrium proportion η^* of workers in the tradeable sector is increasing in the quality of the networks α and in the city size N. In other words, bigger cities, and better connected ones, yield a bigger share of the population working in the tradeable sector, where they can network, which in turns yield higher specialization and higher productivity.

We start by analyzing the effect of network quality α. The population share working in the tradeable sector η^* , but it is pinned down by the no-arbitrage condition that payoffs should be equalized in the two sectors: $\Delta V(\eta^*) = 0$. We perform the comparative static analysis of the effect of α on η^* using the Implicit Function Theorem:

$$\frac{d\eta^*}{d\alpha} = -\left.\frac{\partial \Delta V(\eta)/\partial \alpha}{\partial \Delta V(\eta)/\partial \eta}\right|_{\eta=\eta^*} \tag{I.19}$$

We know that, at $\eta = \eta^*$, the denominator is negative because, as the previous section showed, by definition, ΔV is downward sloping at the stable equilibrium. So we can look at the sign of the numerator. A higher α affects how beneficial the network is for workers in the tradeable sector. More precisely:

$$\frac{\partial V_T(\eta)}{\partial \alpha} = \alpha^{\delta-1}\delta\left(1 - \frac{1}{\eta N}\right)(\eta N)^{\delta} \qquad > 0 \tag{I.20}$$

The positive sign of which establishes that $\frac{\mathrm{d}\,\eta^*}{\mathrm{d}\,\alpha} > 0$ yielding our next proposition.

Proposition I.3. *A city with a higher quality of networks α is more productive.*

Proof. The number of workers in the tradeable sector η^* is increasing in α, the quality of networks in that sector. This induces each worker in the tradeable sector to specialize more, as $\theta^* = \eta^* N$, thereby becoming more productive and inducing positive externalities on other workers' in the sector. □

Note that the effect of better networks is only indirect. *Ceteris paribus*, it does not induce a higher choice of specialization (see Equation (I.9)). However by increasing the attraction of the tradeable sector, where workers can benefit from networks, it makes it possible for each worker in the tradeable sector to become more specialized.

We now examine the effect of city size on the productivity of the city. Again we use the Implicit Function Theorem to establish the sign of:

$$\frac{\mathrm{d}\,\eta^*}{\mathrm{d}\,N} = -\left.\frac{\partial \Delta V(\eta)/\partial N}{\partial \Delta V(\eta)/\partial \eta}\right|_{\eta=\eta^*} \tag{I.21}$$

And again, we need only look at the sign of the numerator:

$$\frac{\partial \Delta V}{\partial N} = \frac{\delta\left(\alpha^{\delta}\left(1 - \frac{1}{\eta N}\right) + \frac{1}{\eta N}\right)(\eta N)^{\delta}}{N} - \left(1 - \alpha^{\delta}\right)\frac{(\eta N)^{\delta}}{\eta N^2} \tag{I.22}$$

When N is large, the second term becomes negligeable, so that $\partial \Delta V/\partial N > 0$ will be positive for any sufficiently large values of N, the city size. Yielding our main result:

$$\frac{\mathrm{d}\,\eta^*}{\mathrm{d}\,N} \geq 0 \tag{I.23}$$

That is, above a minimum size threshold, bigger cities yield a higher proportion of networked workers, supplying the tradeable good. Moreover, we can also look at the effect of city size on the level of specialization chosen by the networked agents. As illustrated in Equation (I.9), in equilibrium, networked workers will choose a level of specialization $\theta^* = \eta^* N$. Using the product rule:

$$\frac{\mathrm{d}\,\theta^*}{\mathrm{d}\,N} = \eta^* + N\frac{\mathrm{d}\,\eta^*}{\mathrm{d}\,N} \quad \geq 0 \tag{I.24}$$

That is, bigger cities lead to a higher level of specialization among networked workers. This result stems from two effects:

- A direct effect as workers become more specialized in order to avoid overlapping with the skill set of other workers as the skill circle becomes more crowded.
- An indirect effect resulting from the improvement in the quality of the skill insurance. That is, as other workers become more specialized, the expected cost of being obliged to rely on the skills of a member of ones network—in the event that the project falls outside ones skill support—reduces. This is because, being more specialized, each member of the network is able to provide knowledge to complete the project to a higher level.

We summarize the last two equations in our main result:

Proposition I.4. *In the stable interior equilibrium, the numbers of networked workers is increasing in city size, and the networked workers are more specialized, thereby yielding a higher average productivity.*

The intuition behind this result is that, as N increases, and the number of workers in the tradeable sector increases, so each worker becomes increasingly specialized in order to avoid the situation in which her skills overlap with those of her skill *neighbours*. In the absence of networks this increased specialization would reduce productivity—due to the decreasing returns to specialization. In our model, however, this effect is counteracted by the skill dependent externality that accrues from the human capital embodied in other workers in the network. When α and δ are sufficiently large this externality effect is large enough to outweigh the decreasing returns to specialization.

In terms of the urban wage premium there is a sorting effect as the proportion of specialized workers is increasing in the city size and there is an externality effect arising through the the impact of each networked individuals specialization level on the productivity of each of the members of her network.

I.6 Empirical Predictions of Chapter I

The empirical literature in this field is both highly developed and rapidly moving. As such, this chapter will not attempt to add to this but will, instead, highlight some of the stylized facts that have emerged from recent work that are consistent with the implications from our model.

The conclusion of the model developed above, that workers earn more in denser urban areas, has been widely supported by the empirical literature.

Our model avoids the pitfalls often associated with relying on the specialization enabled by market depth and endogenously yields a heterogeneous skill distribution without assuming *ex ante* individuals of different types. Thus the model anticipates skills both at the high and low end of the distribution will co-locate in cities and that the disparities between high and low skills will be more significant in larger cities. As far as we are aware, this is the first model able to match this result that has recently emerged from the empirical literature (see Combes, Duranton, Gobillon, Puga, et al. 2012) and the result does not rely on selective migration—which, as Combes, Duranton, Gobillon, and Roux (2012) note, appears to explain relatively little of the skill difference in denser cities.

Yet sorting does not account for the entirety of the urban wage premium, neither in the empirical literature (Glaeser and Maré 2001; Overman and Puga 2010), nor in our model. Many recent empirical works have attempted to quantify the relative magnitude of the sorting component and the externality component of the urban wage premium. What appears to have emerged, is that both play an interdependent role. In addition to providing an alternative explanation for this urban wage premium, the mechanism outlined in the model above demonstrates the empirical finding that this urban externality is skill dependent; that the externality is larger in cities in which the number of skilled workers is highest and that there exists no such externality in the least skilled cities (Glaeser and Resseger 2010).

In this manner we have shown how our model is able to predict, not only the sorting effect and the externality effect that characterize enhanced productivity of urban workers, but also the interaction between these two effects as outlined in Section I.1.

The reliance of the externality outlined in our model on the diversity of horizontally differentiated skills is hard to identify empirically at the city level. However, a recent paper by Bosquet and Combes (2013) examining the productivity of French academics (according to the quantity and quality of their annual publications) finds that those departments containing a broad range of specializations within a field are associated with larger externalities resulting from departmental size.[14]

Finally the findings of two pieces of recent work, while not directly predicted by the model outlined in this chapter, do provide some support for the assumptions underlying it. In the first place, work by Bacolod et al (2009) finds cognitive and people skills are particularly rewarded in cities. This is consistent with the model to the extent that, because benefits of urban location accrue through networks, the payoff from locating in an urban area will be higher for those for whom network maintenance is less costly—those with strong people skills. In the second place the notion that the urban wage premiun is driven by some form of embodied capital is reinforced by recent findings of De La Roca and Puga (2013) that suggest the premium takes time to accrue. And, while this delayed accrual is consistent both with the literature that posits enhanced learning as the primary mechanism through which the urban wage premium accrues and networks mechanism—as outlined in this chapter—in our opinion the depreciation of the premium upon leaving a city, identified in the same paper, is less consistent with the learning hypothesis than with our own. Our model outlines a mechanism that can account for the location-specific component of human capital, that we posit to be networks.

[14] The same paper finds that women and older academics endow a larger externality on their peers despite having, on average, a poorer publication performance themselves. In the context of a productivity cost associated with offering ones skills for use in the production function of ones peers, this finding could be taken as support for the mechanism at the basis of our model that the externality is conferred through the ability to utilize the human capital embodied in those in ones network.

CHAPTER II

The Culture of Cooperation: A Social Poverty Trap

II.1 Introduction

The model presented in this chapter explains how the ability or inability to transmit a culture of cooperation to future generations can lead to a time-stable distribution of wealth with multiple equilibria, which I call a *social poverty trap*. I use this term to refer to a situation in which the process of cultural transmission from parents to children may lead to different long-run levels of wealth due to the inability to instill in future generations a culture of cooperation.

Although there is a consensus in the population about the best trait to be transmitted, resource constraints will lead to very different outcomes for different dynasties. And, as is common in poverty trap models, the outcome of a given dynasty in this model will be governed solely by the initial characteristics of that dynasty.

Following the literature on cooperation in repeated games, I model the ability of an individual to cooperate as her discount factor. However, a discount factor will here be interpreted more broadly than strict time-preference. In repeated games, cooperation can be sustained as an equilibrium provided players are sufficiently patient, even though it does not represent a Nash equilibrium of the stage game. Therefore, the discount factor represents some type of social ability to engage in cooperative behaviour despite the possibility of non-cooperative behaviour by other players. In any situation involving complementarities, cooperation might be beneficial and will therefore help a relationship to reach a higher economic outcome. Therefore the discount factor can be viewed as some type of economic asset, or social capital, that is (imperfectly) transmitted from one generation to the next. Parents can invest effort into building the social capital of their children: they can, for instance, achieve this by teaching their children an attitude that favours the long-term over the short term; this is beneficial in my setup, in which a long-lived cooperative relationship yields better payoffs than short-lived defections. Parents can also devote effort to building the social network and connections of their children, thereby enhancing their likelihood of living in a cooperative environment and being able to seize profitable opportunities that would have otherwise not arisen, which is in a way analogous to the mechanism at work in Chapter I.[1]

The main mechanism at work in my model is the following. Individuals face an infinitely repeated Prisoner's Dilemma (henceforth PD) game. Patient individuals will be able to secure a high level of wealth from a cooperative interaction, and will therefore be able to invest more resources into transmitting this social ability to their children, who in turn will also be able to secure a high level of wealth. On the other hand, impatient individuals will instead choose to defect in repeated relationships, and choose instead to enjoy profit rather than sustained cooperation. They will end up in a defecting equilibrium, in which they get a lower level of wealth. Because they will be resource-constrained, they will fail to transmit the culture of cooperation to their children, who will end up in the low-wealth equilibrium. The result is driven by the non-convexity of the cultural transmission problem, because of the different strategies (trigger strategy versus systematic defection) adopted in the game by individuals with different traits.

Section II.2 surveys related literature. Section II.3 presents the set-up of this repeated-game model; Sections II.4 and II.5 develop the possibility of the time-stable distribution of wealth and social capital with multiple equilibria. Section II.7 concludes.

[1] Combes, Duranton, Gobillon, and Roux (2012) suggest for example that parental efforts prior entry on the labour market may affect individuals economic outcomes.

II.2 Related literature

This article draws from three main fields in the literature: models of poverty traps and multiple equilibria, literature on cultural tranmission, and literature on cooperation in repeated interactions.

In spirit, the model presented here is similar to that of Galor and Zeira (1993) and Banerjee and Newman (1993). In their article, Galor and Zeira develop a model in which several equilibria of wealth can arise, as the result of the decision to invest or not in human capital (education) in order to work in the skilled sector. The investment in education can help reach a higher level of wealth, and therefore transmit more assets to the next generation. However, because of imperfect borrowing, and because the cost of education is indivisible, poor individuals cannot invest in human capital, and therefore leave bequests too low for their children to invest in education. Poor dynasties are therefore stuck in the low equilibrium of unskilled job. Banerjee and Newman develop a similar model where occupational choice between self-employment, factory job, and entrepreneurship, is the main vector of poverty transmission. They highlight the possibility of an equilibrium in which the country largely populated by self-employed artisans and peasants depending on the initial distribution of wealth.[2] This model looks at a different type of human capital, which I call the culture of cooperation, namely the ability to sustain cooperative behaviour in repeated Prisoner's Dilemma-like situations.

A key feature of poverty trap models is the lack of intergenerational mobility. This conclusion applies in this model as well. A common point of these models, which is also present in my model, is that the long-run distribution of wealth, or occupational choice, is largely dependent on the initial distribution of wealth in the population.

The model presented here is also linked to the literature on socialization and intergenerational transmission of a trait. Seminal ork by Bisin and Verdier (2001) shows how cultural traits can remain heterogeneous in the population, even if the cultural transmission is stochastic. The main result is driven by the desire for parents to socialize their children to their own cultural trait (they term this desire "imperfect empathy"), which leads to the perpetuation of heterogeneous groups, more specif-

[2] See also the simplified version of Ghatak and Nien-Huei Jiang (2002).

ically when the cultural transmission exhibits "cultural substitutability" between the vertical and horizontal socialization processes. Sáez-Martí and Sjögren (2008) extend their result by showing how an heterogeneous equilibrium distribution of a trait can arise, even when there is consensus about the best trait to transmit.

In the present model, the population can indeed converge to a heteregenous distribution of the discount factor, although for simplicity I have assumed a deterministic and completely vertical transmission process. Transmission of social capital seems to be consistent with evidence by Dohmen et al. (2012) that trust and risk attitudes are transmitted from parents to children.

Finally, this article is related to various articles analyzing the benefit of patience and the link between patience and the ability to cooperate. Several articles highlight how patience can lead to improved economic outcomes. Krusell and A. A. Smith (1998) explain how small difference in discount factor can lead to different saving behaviours in the presence of idiosyncratic earning risk, therefore leading to important differences in long-run accumulated wealth. Doepke and Zilibotti (2008) focus on the choice between different occupations, with different earning profiles. They explain how more patient individuals will choose activities with a steeper earning curve, while impatient individuals will choose occupations flatter earning curves, therefore failing to invest enough resources into teaching patience to their children. G. S. Becker and Mulligan (1997) and Heckman and Rubinstein (2001) present empirical evidence that more patient individuals tend to perform economically better.

The importance of time preferences has also been studied in articles examining the possibility of sustaining cooperation in repeated games (these results are usually referred to as Folk Theorems). The seminal paper of Fudenberg and Maskin (1986) showed how a non-Nash payoff can be sustained in an infinitely repeated game as payoffs in a subgame perfect equilibrium. My model heavily relies on such equilibria, by allowing some individuals to profitably cooperate in a Prisoner's Dilemma game.

Although generally derived with homogeneous time preferences in the populations, a direct interpretation of Folk Theorem results is that patient players have the potential to sustain higher equilibrium payoffs than impatient players. This is the line for instance adopted by Poulsen and Svendsen (2005), who model social capital as the ability to cooperate in a one-shot PD game. They focus on the possible emergence of a unified norm of cooperation. Instead, I endogenize time preferences and show the possibility of equilibria where hetereogenous time preferences persist, which then generate persistent difference in payoffs. Therefore, contrary to Krusell and A. A. Smith (1998) and Doepke and Zilibotti (2008), this model focusses on the ability to cooperate as the main channel through which patience affects economic outcomes.

II.3 Model

The model outlined below is designed to explain how the lack of socialization towards cooperative attitude can generate a *social poverty trap* in which some dynasties remain in a low-payoff equilibrium of non-cooperation, while others enjoy a high-payoff due to cooperation. Individuals (and, in equilibrium, dynasties) only differ by their discount factor. Let's denote $\beta_i \in [0, 1)$ the discount factor of individual i.

Every generation is divided into two periods:

1. In the *cooperation period*, each individuals plays an infinitely repeated Prisoner's Dilemma game with random matching and perfect public monitoring.

2. Then in the *transmission period*, each individuals has one child, decides how much to invest in socialization her child, and dies.

II.3.1 Cooperation period

In the cooperation period, each individual in the economy plays an infinitely repeated game, where at each stage, she is matched against a different player. For simplicity, the matching is i.i.d. uniform; in particular, the probability of two individuals matching does not depend on their characteristic discount factors. I also assume that the population is sufficiently large, so that the probability of meeting the same player twice can be neglected.

Note that in many applications, the assumption of random matching is not innocuous, in particular, there are many instances in which socio-economic characteristics may affect the matching.[3] Geographical or skill-based sorting, as discussed

[3] Such an example can be found in opportunities on the labour market, where one is more likely to hear from job opportunities from their social networks, see Calvó-Armengol and

in Chapter I, school choice, occupational choice, etc. are all instances of mechanisms that tend to increase the probability of interaction between similar individuals, such as in Dixit (2003) or Tabellini (2008) circular matching models. Such skewed matching would reinforce incentives for cooperators to cooperate, and would increase their average payoffs by increasing the likelihood of a "good match". I expect higher feedback effects in the building of social capital, thereby increasing the range of parameters over which multiple equilibria can arise.

Players' discount factor are private information.

	C	D
C	c, c	a, b
D	b, a	d, d

Figure II.1: Payoffs in the Prisoner's Dilemma game.

The game played at each stage is a Prisoner's Dilemma game: the payoffs, presented in Figure II.1, obey the standard assumption that $b > c > d > a$. All the payoffs are positive, without loss of generality. As in all Prisoner's Dilemma games strategy D is strictly dominant in every stage game, and unless there is some sufficiently high subsequent reward for cooperation, individuals will choose to defect.

For cooperation to be sustainable in a setting with random matching, it is necessary to provide some way for players to coordinate in order to provide long-lived intertemporal incentives. Since the objective of the model is not *per se* to obtain a Folk Theorem but rather to show the existence of equilibria where heterogeneous payoffs can persist, I assume a strong information structure, namely that every past move of every player is known to all other players. I show in Section II.4 how to sustain cooperation in such a setting. As numerous papers have shown[4], it is possible to relax the informational structure and still obtain a Folk Theorem, which could then lead to similar results as the ones obtained in this model.

Jackson (2004) and Bayer, Ross, and Topa (2008). This can be mostly explained by the fact that geographic proximity and network proximity can coincide, as Zenou (2011) demonstrated.

4 Although Kandori (1992) already varies the degree of observability, a stronger result is shown by Ellison (1994) in the case of anonymous random matching. See also Kandori (2002) for a review of private monitoring.

Let's denote by π_{it} the payoff that individual i gets in the t^{th} repetition of the PD game; the total discounted payoff of individual i for this infinitely repeated game is:

$$\pi_i = \sum_{t=0}^{\infty} \beta_i^t \pi_{it} \tag{II.1}$$

and the corresponding average payoff is:

$$\overline{\pi_i} = \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T} \beta_i^t \pi_{it} \tag{II.2}$$

II.3.2 Intergenerational transmission period

Each individual has deterministically one parent and one child, and I identify each dynasty with the subscript i. This set-up ignores potential replicator dynamics in the line of Boyd and Richerson (1985), where the reproduction rate of an individual could depend on the payoff that she gets in the PD game, or any type of heterogeneous, possibly endogenous fertility rate. This would make the dynamics of the population more complex; however Bowles (1998) shows that in the presence of replicator dynamics, heteregenous distribution of a trait can be evolutionary stable even if equilibrium payoffs are distinct, as it is the case in my model. In the case of chosen fertility rates, several authors have identified a correlation, albeit negative, about wage and fertility rate. This negative relationship could account for poverty traps, as Kremer and Chen (1999) argue. Such an addition may lead to a starker result if poorer families have to spread their investment in social capital among more children; but if social capital built by parents can be transmitted fully to each children, as is the case with social networks, potential increasing returns in social capital could overcome the fertility trap.

Since I will focus on time invariant distribution, I abstract from time indices.

Each individual values both consumption and her child's welfare, and can therefore invest in effort to socialize their child to some specific discount factor, according to some production function of patience $f(\cdot)$. The technology that captures how parents can transform socialization effort into the actual discount factor of their child is captured by the production function$f(\cdot)$, on which I make the following assumption:

Assumption II.1. *The production function of patience $f(\cdot)$ is a continuous and strictly increasing mapping from $[0, +\infty)$ to $[0, 1]$ such that $f(0) = 0$.*

Hidden in this specification is the fact that an individual's discount factor solely depends on their parent's effort investment towards them; in particular, it does *not* depend on their parent's own discount factor. This only makes the results stronger. A world where individuals partially inherit this trait from their parents (who could then invest additional effort to raise it further) would generate an even stronger correlation between parent's and offspring's discount factor. It would be even easier for individuals with a high discount factor to socialize their children to a high discount factor, compared to individuals with low discount factors. This only increases the likely of a poverty trap, compared to a model where all individuals start from a zero discount factor, as specified in Assumption II.1.

For tractability, I assume that individual i cares about her own consumption c_i and about the level of effort e_i she puts in educating her child to cooperation. Therefore I do not model the parent' optimal choice of a discount factor for their offspring, but assume that the parent cares about instilling some pro-social behaviour (in the form of a higher discount rate) to their offspring. Adriani and Sonderegger (2009) offer a justification as to how this can be optimal from the parent's point of view, even though such pro-social behaviour may not maximize the offspring's payoff. Their justification relies on the positive externality of pro-social behaviour, which also arises in my model (the more cooperators there are in a given generation, the more beneficial it is to cooperate because the more likely a player is likely to be matched with a cooperator.)

Effort is resource-consuming. It can be thought, for instance, as the opportunity cost of devoting some time to the child's socialization to cooperative behaviour, i.e. teaching him trust, cooperative attitude, or the ability to inspire trust, etc. This views the transmission process as pure cultural education. An alternative view of the parents' effort would be to consider it as an "investment" in favours: parents do others a favour, in the hope that they will be able to call in favours for their children in return. Effort is then devoted to build the quality of the social network of their children, which will enhance the likelihood of cooperation for the children. Whether this effort is spent at every stage of the PD game or in a one-time expense at the end of one's life makes no different here, and for simplicity, can be summed up into one expense at the end of the PD game.

Therefore each individual solves:

$$\max_{c_i, e_i} U_i(c_i, e_i) \qquad \text{s.t.} \qquad c_i + e_i \leq \overline{\pi_i} \tag{II.3}$$

where c_i, e_i and $\overline{\pi_i}$ respectively denote individual i's consumption, socialization effort and PD game average payoff[5], with the usual assumption that U_i is a continuous everywhere-differentiable quasi-concave function such that $\frac{\partial U_i}{\partial c_i} > 0$ and $\frac{\partial U_i}{\partial e_i} > 0$.

The assumption that individuals care about the socialization effort e_i *per se*, as opposed to the utility of their child, is mainly done for simplicity, to reflect concerns for future generation while avoiding to solve a non-convex recursive problem in the intergenerational transmission process. However, it is worth pointing such an assumption implies that individuals have no interest in transmitting some wealth directly to their offspring. Socialization effort is the only way that a parent can affect the outcome of their child, and there is therefore no monetary transfer from one generation to the next.

I further assume that every individual's concern about their offspring is embodied by the same altruism factor $\gamma > 0$ such that $\frac{\partial^2 U_i}{\partial e_i \partial \gamma} > 0$. In other words, a higher γ increases the marginal utility of effort spent on one's child socialization.[6] The altruism factor γ may or may not depend on individual's discount factor β_i. Traditional models models of bequests assume that the altruism factor is correlated with the discount factor β_i. Mulligan (1997) also made a case that parental priorities may depend on their socio-economic characteristics. Here it need not be the case. In Section II.6, I analyze two cases, one in which altruism depends on β_i and one where it does not.

Denote $e_i^*(\overline{\pi_i}, \beta_i)$ the optimal level of child socialization chosen by a parent with resources $\overline{\pi_i}$ and discount factor β_i. Given the assumptions on U_i, it will be continous with respect to $(\overline{\pi_i}, \beta_i)$. I make a further assumption:

5 It is possible to introduce relative prices in the constraint, but this would just amount to some rescaling.

6 Note here that contrary to Bisin and Verdier (2001), there is a general consensus in the population about the socialization optimum: absent any restrictions, all parents want to socialize their child to the highest possible social capital. However, they might (and will) be constrained by their own resources (payoffs in the PD game).

Assumption II.2. *The optimal level of socialization effort e_i^* chosen by individual i satisfies:*

$$\frac{\partial e_i^*(\overline{\pi_i}, \beta_i)}{\partial \overline{\pi_i}} > 0 \qquad\qquad \frac{\partial e_i^*(\overline{\pi_i}, \beta_i)}{\partial \beta_i} \geq 0 \tag{II.4}$$

This assumption states that when the resource constraint is relaxed, more resources will be devoted to the child's socialization. It will play a crucial role in the poverty trap, because it means that a parent who fares better in the cooperation period will be able to socialize their child better, hence enabling them, in turn, to do better in their cooperation period. The assumption does rule out some utilities such as perfect complements or quasilinear utility in e_i. However, as is common in poverty trap models, the objective of the this model is to capture situations where resources are going to be a constraint on the socialization process, usually implied by limited borrowing ability or imperfect capital markets.

The assumption is more lenient about the effect of a parent's own characteristic β_i on the provision of socialization effort, only requiring it to be weakly positive. In Section II.6, I examine a case where it is zero and a case where it is strictly positive.

II.4 Sustaining cooperation in the PD game with random matching

In this section I now turn to show that in the setting of this model, there can exist a *social poverty trap*. I define a social poverty trap as a time-stable distribution of social capital and wealth in which some dynasties repeatedly enjoy higher payoffs while others are stuck with low payoffs. The difference between dynasties is due to the fact that some will be able to transmit higher discount factor, therefore sustaining cooperative behaviour in the PD game, while the others fail to accumulate enough resources to socialize their offspring towards pro-social behaviour (a high discount factor).

Let's first look at the infinitely-repeated PD game and how it can generate different payoffs for different individuals in the population. In such a setting, it is possible to support an infinite number of payoffs as subgame perfect. In particular there always exists a rather uninteresting equilibrium in which everyone always defects in the PD game, which would yield a completely homogeneous population. I instead focus on the cooperation that will yield the maximum payoff for the maximum share of the population. Namely, I show how it is possible to sustain a subgame perfect equilibrium in which all players with a discount factor higher than a threshold coordinate on cooperating with each other (the "cooperators"), while players with discount factors below the threshold always defect ("defectors").

Although cooperating is a dominated strategy in the PD stage game, it is a well-known result that cooperation can be sustained as a Subgame Perfect equilibrium in the iterated PD, provided players are patient enough. Fudenberg and Maskin (1986) have shown that any convex combinations of payoffs above the minimax payoffs can be sustained as a subgame perfect equilibrium. The incentive to cooperation resides in the fact that minimaxing a player can be used as a credible punishment against deviation from cooperation. Provided that players are patient enough, the punishment can be inflicted sufficiently long to make sustained cooperation preferrable to a one-time deviation followed by several periods of punishment.

The crucial element in Folk Theorems such as Fudenberg and Maskin's result is the players' discount factor. The threat of punishment can only be effective if the players value future enough that they prefer to enjoy the payoff of cooperation repeatedly, rather than cheat and get a high payoff in one period, and then being minimaxed for a (possibly infinite) number of periods. The inability of punishment to discipline players with a low discount factor will be at the heart of the poverty trap that I illustrate in this article.

An implicit assumptions of Fudenberg and Maskin's result is that the set of players is constant from one stage game to the next, i.e. that the same players interact with each other, and that players have full memory of the history of the game, so that they can punish past deviation of the other players. In my setting, the fact that players are randomly matched against each other prevent this type of punishment, the set of players being different from one stage game to the next.

Here, I have assumed that the population is "sufficiently large", i.e. that the probability of meeting the same opponent again can be taken as equal to zero. Of course, this is very unlikely to be the case if the population is finite, because the PD game is infinitely repeated; then, a player will meet a past opponent almost surely, and would theoretically be able to retaliate. However, I can assume that the population is large enough that this will occur with zero probability. This means that in case of a deviation, the player being cheated cannot retaliate very soon. The punishment is then discounted so much that it would only be effective on players with a very high discount factor.

However, in the presence of public information, i.e. when players past actions are observable, I can use a strategy profile similar to that of Kandori (1992) under public observability, to show how cooperation could be sustained even though the set of players is not fixed from one period to the next. Under random pairing, the observed actions stated above can be supported as a perfect equilibrium with the following strategy profile. A player with a low[7] discount factor always defects. A player with a high discount factor applies the following cooperative strategy:

1. If an opponent defects, then in all subsequent matches[8], defector and opponents minimax each other (i.e. they both play D), while all other continue to cooperate.
2. The punishment does not apply to an opponent who has previously played D as part of a retaliation process.

The idea behind this strategy profile is that random pairing does not allow repeated punishment by the same set of players anymore, therefore preventing punishments in iterated Prisoner's Dilemmas à la Fudenberg and Maskin. However, in order to sustain cooperation by the high types, it is now possible to spread the burden of punishment across all subsequent matches with different individuals, who are successively going to incur the (D,D) payoff for one period only. The publicity of past actions now makes it possible to coordinate this punishment over time.

Kandori's result, much like that of R. W. Rosenthal (1979), applies to a setting of random matching in which the population is divided into two subsets, and a member of one subset always paired (randomly) against a member of the other subset. In this model, I do not make such an assumption: although in equilibrium there are two types of players, each type of player can be matched against a player of the same type or of the other type. The counterpart to this is that I need to assume more public information available.

[7] "High" and "low" are defined more precisely later in this section, in Equation (II.8).

[8] It is possible to sustain cooperation using punishment limited to T periods only, but infinite punishment (used in "grim trigger strategy") provide the most incentives for cooperation.

Let's call "defectors" agents who adopt the strategy of always defecting, and "cooperators" those who adopt the two-part strategy described above. As is common in Folk Theorem results, there exists a threshold discount factor, that I denote β^*, such that it is a subgame perfect equilibrium for players with a discount factor lower than β^* to be defectors, and for those above to be cooperators. This is what I now show.

Let's denote α_t the proportion of the cooperators in the population for a given generation t. Then in equilibrium, a share $(1-\alpha_t)$ of the generation t population is perpetually defecting. For that generation, in each stage game, the probability that a given player will be matched against a "defector" is $(1-\alpha_t)$. Because there is public information about the past actions of the opponent, a cooperator can identify a defector if matched together, and the cooperator will play D against a defector.[9] Because the defector is always defecting, she is always being punished. Therefore a cooperator playing D against a defector will not incur punishment in subsequent matches.

Therefore, the expected payoff of the cooperating strategy for an individual with discount rate β and identified as a cooperator is, over the cooperative phase periods:

$$\pi_C = \sum_{s=0}^{\infty} \beta^s(\alpha_t c + (1-\alpha_t)a) = \frac{1}{1-\beta}(\alpha_t c + (1-\alpha_t)d) \tag{II.5}$$

while that of defecting against a cooperator is:

$$\pi_D = b + \sum_{s=1}^{\infty} \beta^s d = b + \frac{\beta}{1-\beta}d \tag{II.6}$$

Therefore, an individual identified as a cooperator, and with discount rate β, will never defect against a cooperator as long as:

$$(b-d)\beta \geq b - d - \alpha_t(c-d) \tag{II.7}$$

The condition for cooperation can be rewritten as

$$\beta \geq \beta^* \qquad \text{where} \qquad \beta^*(\alpha_t) \equiv 1 - \alpha_t \cdot \frac{c-d}{b-d} \tag{II.8}$$

[9] Note that in this game, the minimax payoff of each player is d.

Equation (II.8) precisely identifies the cut-off discount rate that determines whether a given individual will adopt the cooperating or defecting strategy. This proves the following proposition.

Proposition II.1 (Existence of an intragenerational cooperation equilibrium). *In any given generation, there exists a cut-off discount rate $\beta^*(\alpha_t) \equiv 1 - \alpha_t \cdot \frac{c-d}{b-d}$ such that a separating equilibrium with eternal punishment exists where:*

- *individuals with a discount rate higher than β^* will adopt the cooperative strategy,*
- *individuals with a lower discount rate will defect whenever matched against a cooperator, and a* fortiori *against a defector (and hence always defect).*

Note that the proportion of cooperators in the population in this equilibrium will depend on the distribution of the discount factors in the population. Because I have assumed that the socialization of children occur at the end of the PD game, the distribution of discount factors does not change for the iterated PD game in a given generation. Denoting $H_t(\cdot)$ the cumulative distribution of the discount rate in the population at a given generation t, α_t solves :

$$\alpha_t = 1 - H_t(\beta^*) = 1 - H_t\left(1 - \alpha_t \cdot \frac{c-d}{b-d}\right) \tag{II.9}$$

subject to $0 \leq \alpha_t \leq 1$. As I will show, provided the distribution H of discount factors does not change from one generation to the next, this share α will too be constant.

It might also be noteworthy to examine the stability of this equilibrium to trembling errors. In Kandori (1992), when a cooperator accidently defects, this creates a cascade of defection that eventually contaminates the whole population and destroys the coooperation pattern. Here, when a cooperator accidently defects, she will be punished forever, and therefore becomes a defector. As the population is assumed to be sufficiently large, this does not affect the equilibrium α; strategies that were optimal for the other players before the accidental defection still remains optimal. This result is of course very sensitive to the assumption of a large population. If α was affected (reduced) by the accidental defection, this would increase the threshold β^*; this might induce former cooperators to adopt the defecting strategy, thereby increasing the α further, and so on. The possibility of such a cascade, and its stopping point, depend on the distribution of discount factors, specially around the initial threshold. It can potentially entail the breakage of the cooperation pattern. However, with the assumption of a "sufficiently large" population, where no single individual weighs enough, the possibility of such cascades can be ignored.

Long-run average payoffs Because of the discrete change in strategies implied by the threshold β^*, a share of a population enjoys a discretely higher average payoff in the cooperation game. Given the equilibrium specified in Section II.4, the average payoff for a cooperator and for a defector respectively are

$$\overline{\pi_C} \equiv \alpha_t c + (1 - \alpha_t) d \qquad\qquad \overline{\pi_D} \equiv d \tag{II.10}$$

As I have assumed that $c > d$, cooperating always yields a higher average payoff in this equilibrium, regardless of the share α_t of cooperators in the population. These payoffs make it apparent that the cooperators will be able to secure more wealth from the cooperation stage, and therefore be able to transmit more to their children, following Assumption II.2. It is this discontinuity in payoffs that generate the possibility of a povery trap.

II.5 Cultural transmission equilibrium

As explained in Section II.3.2, individual i will devote a certain amount e_i^* of her resources, to the transmission of cooperative behaviour to her child. Given that e_i^* is continuous in the payoff π_i, and that there is a discontinuity in payoffs between cooperators and defectors (see Equation (II.10)), there will be a discontinuity in the provision of socialization effort between cooperators and defectors. More specifically:

The investment in cultural transmission of individual i, depending on whether she has cooperated or defected in the PD game, is therefore :

$$e_i = \begin{cases} e_i^{*D}(\beta_i) \equiv e_i^*(\overline{\pi_C}, \beta_i) & \text{if } \beta_i \geq \beta^* \\ e_i^{*C}(\beta_i) \equiv e_i^*(\overline{\pi_D}, \beta_i) & \text{if } \beta_i < \beta^* \end{cases} \tag{II.11}$$

Given that the production of patience $f(\cdot)$ is continuous, the discontinuity in effort provision between cooperators and defectors will translate into a discontinuity in the discount factor of the future generation depending on their parents' type:

$$g(\beta_i) = \begin{cases} \beta'_{iC} \equiv f(e_i^*(\overline{\pi_C}, \beta_i)) & \text{if } \beta_i \geq \beta^* \\ \beta'_{iD} \equiv f(e_i^*(\overline{\pi_D}, \beta_i)) & \text{if } \beta_i < \beta^* \end{cases} \tag{II.12}$$

where $g(\cdot)$ denotes the final transmission function from generation to the next. Because of that discontinuity, we now define β^{*-} (resp. β^{*+}) as the limit of β approaching β^* by below (resp. above).

It will be useful for later analysis to notice that the higher is individual i's discount factor, the higher (weakly) will be the discount factor effectively transmitted to individual i's child.

Lemma II.1. *The final transmission of discount factor to their offspring as a function of an individual's discount factor is continuous and (weakly) increasing on* $[0, \beta^*) \cup (\beta^*, 1]$.

Proof. Direct consequence of Assumption II.2 that socialization effort $e_i^*(\pi_i, \beta_i)$ is continuous and increasing in resources π_i, continuous and weakly increasing in discount factor β_i Assumption II.1 that $f(\cdot)$ is continuous and increasing. The discontinuity at β^* comes from the discrete jump in the payoff of the cooperation period from $\overline{\pi_D}$ to $\overline{\pi_C}$. □

II.5.1 Multiple equilibria in social capital

The concept of poverty trap relies on the existence of multiple levels of cooperation, captured by the discount factor, that are absorbing. This means that an individual with such a discount factor will transmit exactly the same to his child, who in turn will transmit it to the next generation. The existence of such absorbing states on each side of β^* means that a cooperator will socialize his child into a cooperator, and a defector into a defector. This means that some dynasties will be stuck in low payoff, generation after generation, while other enjoy higher payoffs through cooperation. It is the perpetuation of this differential in payoffs through cooperation in a social situation (or lack thereof) that I call *social poverty trap*. I define these concepts more precisely below and turn to explore conditions for the existence of such a poverty trap.

Let's define $m : \beta \mapsto g(\beta) - \beta$ such that the roots of $m(\cdot)$ are the fixed points of $g(\cdot)$.

Definition II.1. *A* dynastic equilibrium *is a fixed point of the transmission function* $g(\cdot)$, *i.e. a level of social capital* $\beta^E \in [0, 1]$ *such that* $m(\beta^E)$.

When a dynasty is in a dynastic equilibrium, it means that every generation receives the same discount factor as the previous one. Since I restrict attention to stable equilibrium, it is useful to precise my definition of stability:

Definition II.2 (Stable dynastic equilibrium)**.** *A dynastic equilibrium* $\beta^E > 0$ *is* stable *if there exists an interval interval* $\mathcal{I}$ *around* β^E *such that* $\forall \beta < \beta^E \in \mathcal{I}, m(\beta) > 0$ *and* $\forall \beta > \beta^E \in \mathcal{I}, m(\beta) < 0$. *If* $\beta^E = 0$, *it is stable if for all positive* β *in the neighbourhood of 0,* $m(\beta) < 0$.

The definition of stability above relies on the dynamics of patience transmission from one generation to the next. Whenever $m(\beta) > 0$, the following generation in a dynasty receives a higher discount factor than its parent, and conversely when $m(\beta) < 0$. In such a situation, a small deviation from an equilibrium β^E to a slightly small β means that $m(\beta) > 0$, which means that social capital will accumulate again and converge back towards β^E (and vice versa for a deviation to the right of β^E).

Dynastic equilibria apply to individual dynasties. Our definition of social equilibrium relies on every dynasty being in equilibrium[10]:

Definition II.3 (Social equilibrium)**.** *The economy is in* social equilibrium *if every individual dynasty is in a dynastic equilibrium.*

My definition of social equilibrium implies that the share of cooperators α^* in the economy is constant in a social equilibrium, since the level of discount factor in each dynasty does not change over time, which means that cooperators remain cooperators and defectors remain defectors. I can then focus on a constant threshold discount factor $\beta^* \equiv \hat{\beta}(\alpha^*)$.

Since the focus of this chapter is the persistence of heterogeneity, it is necessary to define the subset of social equilibria where such persistent inequality exists, which I call social poverty traps:

[10] A weaker definition could be to only require that the distribution of discount factor in the population is constant over time.

Definition II.4 (Social poverty trap)**.** *A* social poverty trap *is a social equilibrium such that there exists at least one stable dynastic equilibrium on each side of β^*.*

The existence of stable dynastic equilibria on each side of discount threshold β^* mean that some dynasties will be persistently earning the higher payoffs accrued to cooperators, while other dynasties will be stuck in the lower payoff equilibrium of defection. I identify conditions under which such a poverty trap can exist:

Proposition II.2. *A social poverty trap with a stable proportion of cooperators exists if the following two conditions hold:*

$$\lim_{\beta \to \beta^{*-}} g(\beta) < \beta^* \tag{C1}$$

$$g(\beta^*) > \beta^* \tag{C2}$$

Proof. By Lemma II.1 the function $m(\cdot)$ is continuous on $[0, \beta^*)$. By definition of $f(\cdot)$, it must be that $m(0) = g(0) \geq 0$. Condition C1 implies that $\lim_{\beta \to \beta^*} m(\beta) < 0$. By the Theorem of Intermediate Values, there exists at least one β where $m(\cdot)$ changes sign from positive to negative, i.e. a stable equilibrium, on $[0, \beta^*)$.

The reasonining is analogous on $(\beta^*, 1]$. Given Assumption II.1, $m(1) \leq 0$. Condition C2 implies that $m(\beta^*) > 0$, and I can use the Theorem of Intermediate Values again, to imply the existence of at least one stable equilibrium on $[\beta^*, 1]$. □

Conditions C1 and C2 mean that optimal choice of socialization efforts and the production function are such that defectors even close to the left β^* do not socialize their child to become cooperators; and conversely, cooperators on the right of β^* increase the discount factor transmitted to their children and therefore converge towards the stable equilibrium β^H. Taken together, the conditions imply *a fortiori* that *any* dynasty with a discount factor smaller (resp. higher) than β^* will socialize their child to a discount factor smaller (resp. higher) than β^*. In other words, the share of cooperators and defectors stays constant over time.

I illustrate a generic social poverty trap in Figure II.2. Dynastic equilibria β_0^E, β_2^E and β_3^E are stable, while dynastic equilibrium β_1^E is unstable. Given the dynamics of the system all dynasties that started in $[0, \beta_1^E)$ will converge to β_0^E, all dynasties that started in (β_1^E, β^*) to β_2^E, all dynasties that started in $[\beta^*, 1]$ to β_3^E.

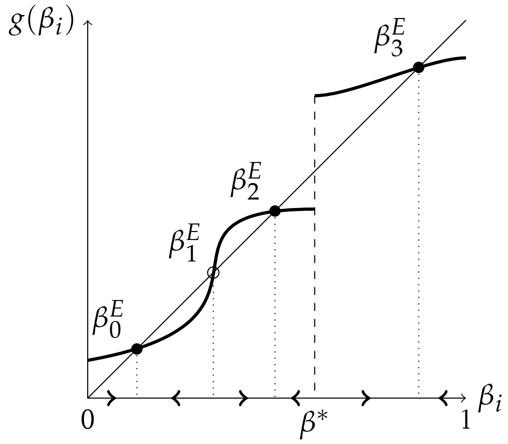


Figure II.2: Generic case of Social poverty trap when Conditions C1 and C2 are satisfied. Black dots denote stable dynastic equilibria, white dots unstable ones. Arrows indicate dynamics.

Therefore in equilibrium, there is persistence of heterogeneous payoffs over time. All dynasties that started in $[0, \beta^*)$ fall in the defection range generation after generation and therefore fail to accumulate enough resources to socialize their child into the cooperative range of discount factors.

Therefore, Proposition II.2 shows that under some conditions, there can exist a poverty trap, i.e. the possibility that a share of the population will converge and remain to a high level of wealth, whereas the other will remain in a low level of wealth. Some dynasties are trapped in a low equilibrium because of their inability to educate their children to cooperative behaviour, which in term prevent them from securing enough resources to make investment in cultural transmission of cooperation worthy. They are therefore caught in a vicious circle that prevent future generations from reaching higher levels of wealth.

Two main mechanisms are leading to the poverty trap. The first one is the fact that more impatient individuals can only reach a lower wealth when interacting with other agents, due to their inability to seize cooperative opportunities, and fail to accumulate enough resources. The second mechanism is directly due to the assumption that more impatient people put (weakly) less value on the education of their children, and invest less resources in the transmission process.

The poverty trap comes from the non-convexity of the investment problem, which can be generated by the jump in payoffs at β^*, making gradual investment to a higher equilibrium impossible. This is different from Galor and Zeira (1993), in which the non-convexity is due to the indivisibility of the investment in education, or from that of Banerjee and Newman (1993) or Doepke and Zilibotti (2008), in which the non-convexity is essentially due to different occupational choices.

II.6 Examples

In this section I show two examples of social poverty trap to illustrate that Conditions C1 and C2 can be obtained readily in standard setups. In the first example, the level of altruism towards next generation is exogenous, while in the second one it depends on the current discount factor of the individual.

II.6.1 Example with exogenous altruism towards next generation

As in Equation (II.3), let individuals maximize $U_i(c_i, e_i)$ by choosing their consumption c_i and the level of effort dedicated to their child, e_i, subject to the resource constraint $c_i + e_i \leq \pi_i$. I now refine Assumption II.2 by assuming the utility function is such that the optimal level of socialization effort by individual i satisfies:

$$\frac{\partial e_i^*(\overline{\pi_i}, \beta_i)}{\partial \beta_i} = 0 \tag{II.13}$$

This means that the level of effort that an individual will devote to socializing their offspring does not depend on their own discount factor, but will instead only depend on their available resources $\overline{\pi_i}$. For convenience, I denote it $e_i^*(\overline{\pi_i})$. In other words, defectors will devote socialization effort $e_i^*(\overline{\pi_D})$ which will yield discount factor $\beta^L \equiv f(e_i^*(\overline{\pi_D}))$, while cooperators will devote $e_i^*(\overline{\pi_C})$ and socialize their child to discount factor $\beta^H \equiv f(e_i^*(\overline{\pi_C}))$. In this setup, Conditions C1 and C2 become:

$$f(e_i^*(\overline{\pi_D})) < \beta^* \tag{C1$'$}$$

$$f(e_i^*(\overline{\pi_C})) > \beta^* \tag{C2$'$}$$

All dynasties that started with $\beta_i \in [0, \beta^*)$ converge in one generation to β^L, while all dynasties that started with $\beta_i \in [\beta^*, 1]$ will converge in one generation to β^H. I illustrate the social poverty trap of this setting in Figure II.3.

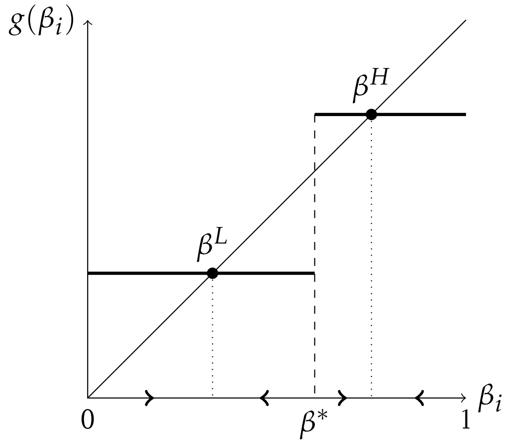


Figure II.3: Social poverty trap when the level of altruism towards next generation γ does not depend on the discount factor. β^H and β^L are both stable equilibria.

II.6.2 Example where altruism depends on the discount factor

In this example, individuals maximize a Cobb-Douglas utility function where the weight put on the socialization effort is their discount factor β_i. Namely:

$$U_i(\pi_i, e_i) = \ln c_i + \beta_i \ln e_i \tag{II.14}$$

Optimally, each individual will devote a share $\frac{\beta_i}{1+\beta_i}$ of resources towards socialization of their children:

$$e_i^*(\overline{\pi_i}, \beta_i) = \frac{\beta_i}{1+\beta_i}\overline{\pi_i} \tag{II.15}$$

which, for a given level of resources $\overline{\pi_i}$, is increasing and concave in β_i (hence satisfying Assumption II.2). and the corresponding socialization effort is given by:

$$g(\beta_i) = \begin{cases} f\left(\frac{\beta_i}{1+\beta_i}\overline{\pi_C}\right) & \text{if } \beta_i \geq \beta^* \\ f\left(\frac{\beta_i}{1+\beta_i}\overline{\pi_D}\right) & \text{if } \beta_i < \beta^* \end{cases} \tag{II.16}$$

In such a setting, Conditions C1 and C2 become:

$$f\left(\frac{\beta^*}{1+\beta^*}\overline{\pi_D}\right) < \beta^* \tag{C1$''$}$$

$$f\left(\frac{\beta^*}{1+\beta^*}\overline{\pi_C}\right) > \beta^* \tag{C2$''$}$$

In Figure II.4, I depict an example where Conditions C1$''$ and C2$''$ are satisfied, with the additional assumptions that $f(\cdot)$ is concave and that $f(0) = 0$. All dynasties that

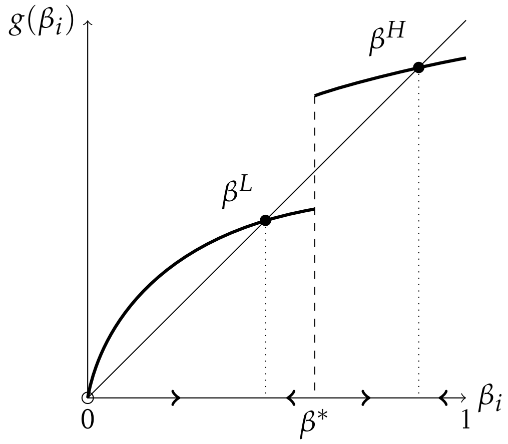


Figure II.4: Social poverty trap when the level of altruism depends on the discount factor β_i, using a Cobb-Douglas utility and a concave transmission technology $f(\cdot)$. Both β^H and β^L are stable equilibria. The origin is an unstable equlibrium.

started in $[0, \beta^*)$ will converge to β^L and enjoy defection payoffs $\overline{\pi_D}$; dynasties that started in $[\beta^*, 1]$ will converge to β^H and enjoy the cooperation payoffs $\overline{\pi_C}$. Unlike the example in Section II.6.1, convergence here is progressive.

II.7 Conclusion of Chapter II

The model presented in this article highlights the role of social capital and cooperative ability as economic assets, that can be transmitted—culturally—from one generation to the next. Because of the feedback effects between cultural transmission and economic success, the system can potentially exhibits several equilibrium levels of social capital and wealth, a situation that I have called *social poverty trap*. The model therefore offers another channel through which patience can affect economic outcome. By considering the propensity to cooperative behaviour as an asset that is only imperfectly transmitted from one generation to the next, this work offers an additional explanation for the lack of social mobility. Assumptions of exogenous fertility and random matching were done for simplicity, but relaxing them would only strengthen the conclusions of the model, by increasing the benefits cooperation among dynasties with high discount factor.

This social capital can be interpreted as the ability to consider cooperative relationships in the long-run rather than in the short-run, or the building of social connections that will enable individuals to take advantage of cooperative opportunities in their professional life. By failing to transmit those, some families fail to provide their children with a potential to economic success, therefore entering a vicious circle of relative poverty.

The model would suggest the possibility of breaking vicious cycles of social poverty traps, by promoting the interaction of low social capital individuals with individuals with higher social capital. Such a possibility would provide better economic prospects for the former, which in turn would incentivize cooperation and enable transmission of social capital to future generations. This justification is behind programs such as Moving To Opportunity, which has demonstrably generated better behaviours among youths of participating families that relocated (Katz, Kling, and Liebman 2001), although there is debate about the long-term success of the programme (Gennetian et al. 2012). The model also suggests that social poverty traps can be broken by improving transmission of social capital from parents to children; the implementation of early childhood intervention such as Head Start, designed to compensate potentially deficient parenting, as proven to be effective (Currie 2001).

CHAPTER III

Issue Selection in Electoral Campaigns with Endogenous Salience

III.1 Introduction

Many political scientists have been intrigued by the slow speed at which some consensual issues have emerged in the political arena. A canonical illustration is the slow emergence of green parties in Europe, which only started to be politically representated from the late seventies, and which have had only recent and limited success at putting the environmental questions on the political agenda (Poguntke 2002). Since environmental protection is largely a consensual issue in Europe, it is legitimate to wonder why the established parties haven't used this popular issue as part of their electoral campaigns. In other words, why has this issue until very recently been, in Europe, the province of specialized parties? And, why did it take so long for green parties to bring salience to the environment question in the political debate? Other examples of non-partisan issues may include security of the national territory, as examplified in the USA after 9/11 attack, educational attainment, the fight against corruption, etc. And why, instead, does a legislative body such as the US Congress seem to focus its activiy on partisan issues (Baumgartner et al. 2009)? It may seem that when electoral pressures are involved, parties and politicians may focus their effort on divisive issues, on which underlying disagreement exist, instead of consensual issues, on which preferences are essentially a matter of degree.

The model presented here analyzes the question of issue choice in an electoral campaign, by looking at parties that can choose their platforms over a divisive (partisan) issue and a consensual (non-partisan) issue. I highlight the fact that when parties are constrained in the number of issues that they can address in a campaign, they prefer to (a) restrict the number of issues they campaign on, and (b) focus on divisive rather than consensual issues. Furthermore, it explains why a party may not be able to single-handedly make an issue salient enough that established parties want to take a position on it, even if (and especially if) it is a consensual one.

The main departure from the standard literature on multidimensional political competition is twofold. First, I analyze a game with three parties instead of two, in order to analyze the possibility that a new party challenges established parties that competed along a traditional issue. The main justification is that when there are three parties, each party has less influence on the political agenda of the electoral campaign. In my analysis I compare the two- and the three-party cases as a way of varying the influence of each single party on the campaign. In addition, multi-dimensional spatial models of electoral competition are typically plagued by problems of non-existence of equilibrium; I address this problem by using discrete strategies.

Secondly and most importantly, the relative salience of the dimensions is the result of the parties strategies. Most models of multidimensional electoral competition use Euclidean preference with exogenous weights in the voters' preferences; the relative salience of the different dimensions is fixed (see for instance the seminal work of Riker and Ordeshook 1973). In that vein, one paper similar to the present chapter is Colomer and Llavador (2012) model of issue choice, in which parties choose how to allocate effort, trading-off consensual issues versus exogenously salient ones. In my model however, I allow the salience of issues to depend on the parties' strategies in the campaign. Specifically, the more emphasis parties give to one of the issues, the more this issue will become salient to voters. There is an implicit time-constraint here; parties therefore have to decide how to allocate their public time among the two issues, and which position to take on those issues.

With such assumption of endogenous salience, my model is strongly related to the literature on agenda formation (See seminal works of Austen-Smith 1987; Baron and Ferejohn 1989), but also to a rich literature on political influence (Most notably G. S. Becker 1983; Grossman and Helpman 1994), and more recent papers analyzing competition for attention (Levy and Razin 2008; Eliaz and Spiegler 2011), where players can spend money to influence the outcome of an election or influence the set of proposals from which a decision must be made. In that strand, a related model is that of Amoros and Puy (2007) in which parties can spend resources to make one issue more salient than another. In the same spirit, recent papers (Aragonés, Castanheira, and Giani 2012; Denter 2013) have looked at the decision by parties to invest resources to improve their perceived competence on specific issues. In my model however, I abstain from vertical (quality) differentation and do not allow for financial transfers or lobbying; instead, parties affect voters' perception directly through their political message and their allocation of time between issues.

This model is also linked to the literature on issue ownership. The classical Downsian model of electoral competition assumes that candidates or parties take a position in the policy space. Instead, the literature on issue ownership argues that parties have more or less fixed policy positions but compete by trying to make salient issues on which they are perceived to be competent (Stokes 1963; Riker 1993; Petrocik 1996), a strategy often referred to as "priming". There is still a debate as to whether this is empirically verified. Some argue that parties do campaign on their idiosyncratic issues[1]; others find evidence against this hypothesis[2]. A related strand of literature analyzes niche parties and how "third parties" may enter the political arena on issues that are not represented by established parties (Rosenstone, Behr, and Lazarus 1984; Hug 2001; Cantillon 2001; Meguid 2005).

My model combines the Downsian with the issue ownership approach by making the assumption that parties' platform choices affect the relative salience to the voters of one issue versus the other. The conclusion suggests that the attention or media constraint that parties are facing—namely how much they suffer from having a wide as opposed to a focused platform—affects the type of issues being debated. In particular, I show that when the attention constraint is more severe, parties will tend to focus on partisan issues, on which they can differentiate politically.

[1] Budge and Farlie 1983; Budge and Laver 1992; Bélanger and Meguid 2008; Green and Hobolt 2008.

[2] Sigelman and Buell 2004; Kaplan, Park, and Ridout 2006.

Although the question is quite different, the setting is inspired by Cantillon (2001). The main difference is that in my model, the relative salience of the issues is not exogenous but is the result of the parties' campaign strategies.

III.2 Model

Three parties are competing in an electoral campaign, in which they try to maximize their vote share. Parties can take a platform position on up to two binary issues: a horizontal issue $\{L, R\}$ and a vertical issue $\{U, D\}$. To illustrate, LU and R are two examples of platform choice, the former where the party takes position on both dimensions, the latter on the horizontal dimension only.

There is continuum of voters whose individual preferences are summarized by their coordinates (x, y) in the $[0, 1] \times [0, 1]$ space. Voters are uniformly distributed over $[0, 1] \times [0, 1]$. The utility for a voter (x, y) of a given policy by party i is a weighted combination of utilities derived on each dimensions (horizontal and vertical) dimensions:

$$\mathcal{U}^i_{x,y} = (1 - s) \cdot f(1 - e_i) \cdot v_1 + s \cdot f(e_i) \cdot v_2 \tag{III.1}$$

This utility has three components. First, the terms v_1 and v_2 capture some Euclidean preference along the horizontal and vertical dimensions respectively:

$$\begin{cases} v_1(L) = 1 - x \\ v_1(R) = x \end{cases} \qquad \begin{cases} v_2(U) = y \\ v_2(D) = -\infty \end{cases} \tag{III.2}$$

On the horizontal dimension, since $v_1(L) - v_1(R) = 1 - 2x$, a voter with a low x will strongly favour a party offering L rather than R in its platform.

On the vertical dimension, the fact that $v_2(D) = -\infty$ captures the crucial assumption that *the vertical issue is non-partisan*. A voter's y coordinate simply indicates how strongly they care about policy U. The assumption that $v_2(D) = -\infty$ implies that parties will never choose to offer a policy that contains position D on the vertical issue: they could do strictly better by replacing it by U. Hence, parties are constrained to the following action space: $\{L, LU, U, RU, R\}$.

Secondly, in the utility function given in Equation (III.1), each dimension is weighted by a component specific to the party's platform under consideration, respectively $f(1-e_i)$ for the horizontal dimension and $f(e_i)$ for the vertical dimension. I define e_i as the emphasis that party $i \in \{1,2,3\}$ puts on the vertical issue in its campaign, depending on the electoral platform that it offers:

$$e_i = \begin{cases} 0 & \text{for platforms } L \text{ or } R \\ \frac{1}{2} & \text{for platforms } LU \text{ or } RU \\ 1 & \text{for platforms } U \end{cases} \tag{III.3}$$

And I define the function $f(\cdot)$ accordingly:

$$\begin{cases} f(0) = 0 \\ f(1) = 1 \\ f\left(\frac{1}{2}\right) = \delta \quad \text{with } \delta \in [0,1] \end{cases} \tag{III.4}$$

The function $f(\cdot)$ captures the effect of party's choice to either focus on one issue only ($f(1) = 1$) or to split its platform on both issues ($f(\frac{1}{2}) = \delta$). When δ is close to 1, each dimension receives a high weight in the utility of voters: parties lose very little by splitting their platforms over the two issues. On the contrary when δ is close to zero, parties that cover both dimensions in their platforms are heavily penalized. The special case of $\delta = \frac{1}{2}$ can be thought of as constant returns to scale.

One interpretation of δ is an attention constraint. During electoral campaigns, parties have limited media exposure and limited resources to publicise their political message. They may therefore have to focus their message on narrow set of issues, rather than talking about very varied issues with the risk of appearing shallow. How binding this constraint might be is captured by δ.

The limitation could also be coming from the voters themselves, who might favour political legibility. In this alternative but related interpretation, a party campaigning on too many issues may be hard to pin down on a traditional political spectrum, or it might make it unclear which of the issues will be prioritised if the party is elected. In this sense, δ captures by how much voters shy away from ambiguous parties. When δ is high, voters accept well that a party takes a stance of a broad set of issues. A low δ makes political competition stronger, since spreading one's agenda means voters are more likely to turn to another party: voters are more volatile.

Finally, each dimension in the utility function is weighted by an aggregate component, respectively s for the vertical issue and $1 - s$ for the horizontal issue. More specifically, I define the salience of the vertical issue:

$$s = \frac{1}{3} \sum_i e_i \tag{III.5}$$

which is an average of the emphasis each party gives to the vertical dimension in their platform.

The interesting feature of this model is that the tastes of the voters are not fixed but depend on the campaign strategies of the parties. The more parties emphasize one of the two issues, the more this issue will become salient to voters. Parties therefore not only choose policy positions, but also how much emphasis (if any) to put on each issue. How much total emphasis is put by parties on each issue will affect the weights of the weighted Euclidian preferences of the voters. The more parties offer U as part of their platform (possibly as the unique issue), the more voters will care about the vertical issue. Conversely, $(1 - s)$ captures the salience of the horizontal issue, which will be more important the more parties' platform feature the horizontal issue. The relative salience of the vertical and the horizontal issue, captured by s and $(1 - s)$ affects how much weight voters put on each dimension.

An interpretation is that the amount of time spent by the parties on a given issue affects the likelihood that a given voter will consider the issue relevant when making a decision, in a way similar to the consideration sets in Eliaz and Spiegler (2011). An alternative explanation is that voters have better information on the issues that are more discussed (Demange and Van Der Straeten 2009; Egorov 2012) and hence are more likely to use such issues to judge a party's platform.

Table III.1 summarizes the utility derived by a voter with characteristics (x, y) from each possible party platform.

Finally in order to avoid further strategic considerations, we make the following two assumptions:

Platform	Utility of voter (x, y)
L	$(1-s)(1-x)$
LU	$(1-s)\delta(1-x)+s\delta y$
U	sy
RU	$(1-s)\delta x+s\delta y$
R	$(1-s)x$

Table III.1: Utility derived by voter with characteristics (x, y) from each possible platform.

- voters vote sincerely for the party whose policy will give them the higher utility[3],
- parties can commit to the policies they offer during the campaign.

Since it is cumbersome to compute scores, and consequently parties' best-responses, with a generic δ, I analyze two "extreme" cases: $\delta = 1/2$ and $\delta = 1$. As will become apparent later, it is unnecessary to consider values of δ smaller than $\frac{1}{2}$, since it yields the same equilibria as when $\delta = \frac{1}{2}$. I compare the results in Section III.5.

III.3 Analysis of the case where $\delta = \frac{1}{2}$

The case where $\delta = 1/2$ can be thought of as constant returns to scale in the political communication process. Splitting their effort (or emphasis) on both the horizontal and the vertical issues will simply halve the weight that voters give to each.

III.3.1 Payoffs

To compute the payoffs of the games, I determine, for every combination of two platforms, what share of voters would prefer one platform to the other.

In order to do this, it is useful to first compute, *for a given salience s*, the coordinates of all voters indifferent between any two stategies. Given the Euclidean structure of the utility function, this determines a straight line locus of indifferent voters in the policy space. All voters to the left of that line will prefer one platform while

[3] Since there is a continuum of voter, the event that a given voter is decisive has measure zero.

voters to the right will prefer the other. For any $\delta \in (0,1)$, we get the following indifference loci:

$$L/LU \qquad y = \frac{1-s}{s} \cdot \frac{1-\delta}{\delta}(1-x) \qquad \text{(III.6)}$$

$$L/U \qquad y = \frac{1-s}{s}(1-x) \qquad \text{(III.7)}$$

$$L/RU \qquad y = \frac{1-s}{s} \cdot \frac{1}{\delta} - \frac{1-s}{s} \cdot \frac{1+\delta}{\delta}x \qquad \text{(III.8)}$$

$$L/R, LU/RU \qquad x = \frac{1}{2} \qquad \text{(III.9)}$$

$$LU/U \qquad y = \frac{1-s}{s} \cdot \frac{\delta}{1-\delta}(1-x) \qquad \text{(III.10)}$$

$$LU/R \qquad y = \frac{1-s}{s}\left[\frac{1+\delta}{\delta}x - 1\right] \qquad \text{(III.11)}$$

$$U/RU \qquad y = \frac{1-s}{s} \cdot \frac{\delta}{1-\delta}x \qquad \text{(III.12)}$$

$$U/R \qquad y = \frac{1-s}{s}x \qquad \text{(III.13)}$$

$$RU/R \qquad y = \frac{1-s}{s} \cdot \frac{1-\delta}{\delta}x \qquad \text{(III.14)}$$

Using Equations (III.6) to (III.14), we can compute the indifference loci in the special case where $\delta = \frac{1}{2}$.

$$L/U, L/LU, LU/U: \qquad y = \frac{1-s}{s}[1-x] \qquad \text{(III.15)}$$

$$U/RU, U/R, RU/R: \qquad y = \frac{1-s}{s}x \qquad \text{(III.16)}$$

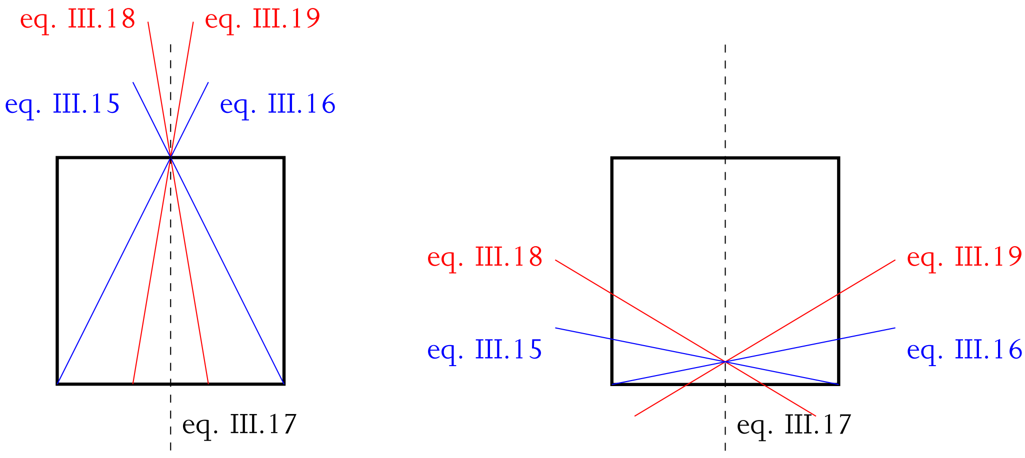
$$L/R, LU/RU: \qquad x = \frac{1}{2} \qquad \text{(III.17)}$$

$$L/RU: \qquad y = \frac{1-s}{s}[2-3x] \qquad \text{(III.18)}$$

$$LU/R: \qquad y = \frac{1-s}{s}[3x-1] \qquad \text{(III.19)}$$

As an illustration, these indifference lines are plotted in Figure III.1 for two given levels of salience, $s = \frac{1}{3}$ in Figure III.1a and $s = \frac{5}{6}$ in Figure III.1b.

More generally, each configuration of platform choices offered by the parties gives rise to a specific level of salience s, as defined in Equation (III.5). When there are three parties, the salience s of the vertical issue must belong to the set



(a) $s = \frac{1}{3}$ (b) $s = \frac{5}{6}$

Figure III.1: Loci of indifferent voters for two different level of salience, $\delta = \frac{1}{2}$.

$\{0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}, 1\}$. When there are two parties, the salience s of the vertical issue now must be belong to the set $\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$.

Given the salience and the parties' choices, it is possible to calculate the corresponding score for each party, in terms of the share of the population voting for them. I calculate these scores geometrically, by computing the areas delimited by the indifference loci. For any combinations of platforms, a given section of the electoral space will have a preferred option among the platforms offered by the parties. Doing so when there are three platforms offered is not a problem since voter's Euclidian preferences are transitive. It is assumed that when several parties choose the same platforms, they split their votes evenly.

All payoffs are written in Appendix B, in Table B.1 for the three-platform competition, and in Table B.2 when only two players compete.

III.3.2 Equilibria with three parties.

I now analyze the game with three parties, contrasting the simultaneous move game with the case where one party moves before or after the other two. In the sequential game, I will interpret that party moving alone as a new party or a small party, possibly trying to put an issue on the campaign agenda. I then compare the results to that of the two-party version (in Section III.3.3) to understand the importance of the weight a single party can have on the campaign.

Nash equilibria in the simultaneous game

Using payoffs from Table B.1 (Appendix B.1.1) we get the following result in the three-player simultaneous platform:

Result III.1. *When $\delta = \frac{1}{2}$, the Nash equilibria of the simultaneous game are:*

- *all permutations of* (L, L, R) *or* (L, R, R),
- *the strategy profile* (U, U, U)

Result III.1 shows that in the case where $\delta = \frac{1}{2}$, the campaign will display limited political differentiation: parties may take difference position but along the same issue. In that situation, it is too costly for any party to "dilute" its platform by campaigning on both issues, and this results in the parties focusing on one issue only.

More precisely, the choice to spread the party's platform over the two issues induces three effects:

- a *predatory effect* of covering a wider political space,
- a *dilution effect* of putting less emphasis on the initial issue,
- a *salience effect* of increasing the salience of the other issue.

By covering both issues, a party displays a more predatory attitude by competing with the other parties on their grounds, therefore stealing some voters from them by choosing a position close to them. That is what I call the *predatory effect*.

But by shifting from a single-issue platform to a wide platform, the party will put less emphasis on the issues, meaning that they will be less relevant for the voters. So there is also an indirect *salience effect* taking place: the salience of the new issue increases relative to that of the old one, simply because the new issue is more present in the political debate preceding the vote.

For a given party, whichever effect is stronger depends on the strength of the *dilution effect*, captured by δ, as well as on the configuration chosen by the other two parties, which determines how successful the predation can be, and how salient each issue is relative to the other.

In a setting where $\delta = \frac{1}{2}$, the direct dilution effect is always stronger. This implies, as can be seen from Table B.1, that LU and RU are dominated strategies: no matter what the other parties do, it is always better to choose a single-issue political platform, that is: either L, U or D.

Moreover, the predatory effect always dominates of the salience effect. That is to say, if the two other parties are fighting along one of the issue, say if they have chosen L and R, the best-response then is for the third party to play along the same issue, i.e. play L or R, instead of trying to emphasize the other issue. By itself, a party is not important enough to move the emphasis of the campaign to another issue.

When the other parties are emphasising the horizontal issue, the salience s of the vertical issue will be low. Therefore the benefit of incorporating the vertical issue on the one's agenda will be small. However, since s is small, the salience of the horizontal issue $(1 - s)$ will be high; reducing the party's emphasis on the horizontal issue will therefore be very detrimental.

Similarly, when the two other parties are campaigning on U, the best-response is for the third player to campaign on U as well, therefore giving rise to the (U, U, U) equilibrium.

The bottom line is that when $\delta = \frac{1}{2}$, spreading one's platform is too costly. Parties will campaign on the same single issue, i.e. the electoral campaign will be limited in scope.

Equilibria in the sequential game

Result III.1 is not entirely surprising given the entire symmetry of the game, but one may wonder less direct confrontation could obtain if one of the player was different. We now turn to a slightly modified situation in which a "new" party competes with two established parties. By relaxing the assumption that all parties move simultaneously, we can look at a situation where two established parties are already in place, maybe fighting along the horizontal dimension. The rise of a new vertical dimension may prompt a third party to (exogenously) enter the arena.

The question is now to analyze what happens when we break the symmetry between parties by assuming either that the two established parties move fist, followed by the new party; or that this new party moves first, and then the two established parties move simultaneously. We will look at subgame perfect equilibria in both games, once again using the payoffs calcultated in Table B.1.

Let's first start by the case where two parties, maybe because of historical reasons, are established in the political spectrum, followed by the arrival of a new party. The new party threatens to enter the campaign after both mainstream parties choose their policies. A party willing to have the issue U discussed in the campaign could enter if mainstream parties do not offer it as part of their policies. However, Result III.2 shows that this threat is not credible:

Result III.2. *When the two mainstream parties move first, followed by the third party, the subgame perfect equilibria of the game are:* (L, R) *or* (R, L) *followed by either* L *or* R*. The equilibrium salience of the non-partisan issue is* 0.

In other words, the only credible entry that the new party can make is to enter on the same issue as the other two parties. What happens is that the horizontal issue is already made salient by the other two parties, and it is more rewarding for the new party to enter on that issue rather than try to take the campaign on the horizontal one.

This may explain why the (endogenous) political clivages exhibit some inertia. A new party willing to compete with established party will not be able to take advantage of a non-partisan in order to enter the political arena. By itself it cannot affect the salience of that issue enough: established parties anticipate this failure and therefore have no incentive to pre-empt entry by covering the vertical issue.

The situation where the new party moves first can be considered as an instance where the party "discovers" the vertical issue and may be able to use this in order to create surprise.[4]

Result III.3. *When a party moves first, and the two other parties then move simultaneously, the outcome of the subgame perfect equilibria of the game with* $\delta = \frac{1}{2}$ *is the first party playing* L *or* R*, followed by the two other parties playing* (L, R)*. The equilibrium salience of the non-partisan issue is* 0.

[4] Let's however remember that the new party cares only about its vote share, not about which issue is represented in the campaign. In this model, parties are purely office-motivated.

(The proof, using elimination of Pareto-dominated Nash equilibria in the second stage, is provided in (The proof is provided in Appendix B.1.2.) p. 87.)

Note that, in the second stage subgame, the Nash equilibria where the same actions is played by the "main" parties, i.e. (L, L) or i.e., are Pareto-dominated by (L, R), when they differentiate. By playing the (L, R) Nash equilibrium the main parties get a payoff of $(\frac{1}{4}, \frac{1}{2})$ instead of $(\frac{1}{4}, \frac{1}{4})$. We may therefore imagine that, maybe for historical or ideological reasons, the mainstream parties are able to coordinate on this equilibrium.

Once again, the fact that the vertical issue is non-partisan does not offer enough scope for differentiation. In other words, the new party is not able, even by moving first, to put the new non-partisan issue on the table, since the mainstream parties prefer to fight along the horizontal issue only.

The new party cannot credibly choose U in the first stage, because this would be followed in the second stage by a (L, R) Nash equilibrium that would yield a payoff of zero to the first-mover. Therefore the subgame perfect equilibria are very similar in Results III.2 and III.3, once we remove the Pareto-dominated ones in the Result III.3.

The result is again that only the partisan horizontal issue is represented in the debate. That is, a third party cannot by itself, move the campaign away from the divisive issue. When $\delta = \frac{1}{2}$, the salience effect is never strong enough for the third party to put the non-partisan issue on the platforms of its competitors.

III.3.3 Two-player game

It might be useful as a benchmark, to analyze the same game when there are only two parties competing. We compute the payoffs in a very similar way to Section III.3.2, using Equations (III.15) to (III.19) for salience levels $s \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$.

The corresponding payoffs are calculated in Table B.2 in the appendix.

Result III.4. *When $\delta = \frac{1}{2}$, a strategy profile is a Nash equilibria in the two-player simultaneous game of platform choice if and only if is is a combination of single-issue platforms, i.e.*

- (L, L), (R, R), (U, U)
- (L, U), (L, R), (U, R)

In all equilibria, the payoffs are $(\frac{1}{2}, \frac{1}{2})$.

If we allow player 1 to move first, then we can analyze the sequential version of the two-player platform game, which yields the same type of equilibria (with subgame perfection).

Result III.5. *When* $\delta = \frac{1}{2}$*, the subgame perfect equilibria in the two-player sequential game of platform choice are* L*,* R *or* U*, followed indifferently by* L*,* R *or* U*. In all equilibria, the payoffs are* $(\frac{1}{2}, \frac{1}{2})$.

With two players only, Results III.4 and III.5 are less strong than Results III.1 to III.3 with three players, in the sense that one party alone cannot force the competing party on the same issue, and the model becomes less predictive on the type of issue that would be part of the campaign. When there are only two parties, each player has a greater incidence of the relative salience of each issue. For instance, by moving from a platform L to LU will increase the salience of the vertical issue by $\frac{1}{4}$ instead of $\frac{1}{6}$ when there are three parties. This makes the salience effect stronger.

As a result, parties do not necessarily fight along the same issues[5], because each party can now offset by itself the salience that the other is giving to the other issue. The predatory and salience effect exactly cancel each other, which may lead to a less confrontational campaign. It is therefore possible to conjecture that campaigns featuring a larger number of parties, where each is less able to steer the focus of the campaign on its own, is more likely to exhibit the features discussed above, i.e. parties focussing on a limited set of issues, and focussing on the divisive issue as a means to differentiate politically.

However, although less predictive about the type of issues that would become part of the campaign, the results with two parties are partially similar to Results III.1 to III.3 with three parties: parties will focus on a single issue rather than splitting their platforms.

Proposition III.1. *When* $\delta = \frac{1}{2}$*, it is too costly for parties to split their platforms, and parties only campaign on single-issue platforms. When are there three parties, unless they move simultaneously, the non-partisan issue is never part of the campaign.*

[5] Very much like in Amoros and Puy (2007).

III.4 Analysis of the $\delta = 1$ case

In the extreme case of $\delta = 1$, parties face no *direct* penalty from splitting their platforms over both issues. A party taking position on both issue will be preferred by any voter (x, y) over a party only taking position over one of those issues, *for any given level of salience*.

For instance, for any given level of salience, all voters prefer a party offering LU to a party only offering L. Therefore, any party offering either L or U, against a party offering LU, would get no vote (and similarly for R and U against RU).[6]

For the other combinations of platforms, when $\delta = 1$, we get the following indifference loci:

$$L/U \qquad y = \frac{1-s}{s}(1-x) \tag{III.20}$$

$$L/RU \qquad y = \frac{1-s}{s} \cdot (1-2x) \tag{III.21}$$

$$L/R, LU/RU \qquad x = \frac{1}{2} \tag{III.22}$$

$$LU/R \qquad y = \frac{1-s}{s} \cdot (2x-1) \tag{III.23}$$

$$U/R \qquad y = \frac{1-s}{s}x \tag{III.24}$$

Once again, one should keep in mind that changing from a single-issue to a split issue is also going to affect the total salience of the respective issue. So holding the other strategies constant, although there is no direct "penalty" from splitting one's platform over both issues (because $\delta = 1$), there could be an effect of the overall salience. Therefore, a single-issue platform, say R, could potentially do better than the mixed issue platform RU, if the focusing only on the horizontal issue could improve its salience (thereby justifying this choice).

Using Equations (III.20) to (III.24), we calculate the payoffs for all combinations of platforms (see Table B.3 in Appendix B Appendix B.2.1). From this we can analyze simultaneous and sequential versions of this platform game. Since the procedure

6 Some extreme voters (more precisely, voters with $y = 0$) could actually be indifferent and split their votes between, say, for instance L and LU . However, those voters have measure zero and can be ignored in the parties' payoffs.

is similar to that of Section III.3, I do not detail the calculations but only state the results, before discussing them in Section III.5

Result III.6. *When $\delta = 1$, a Nash equilibrium in the simultaneous game is any strategy profile where a party plays LU and the two other RU, or a party playing RU and the two others playing LU.*

(Once again, the proof follows directly from the payoffs in Table B.3, in the appendix.)

Result III.7. *In the case where $\delta = 1$, when a party moves first, and the two other parties then move simultaneously, the outcome of the subgame perfect equilibria of the game are:*

- *LU followed by (LU, RU) (or permutation), or (RU, RU)*
- *RU followed by (LU, RU) (or permutation), or (LU, LU)*

As in Result III.3, when the "main" parties (second-movers) play the same strategies, the Nash equilibrium in the second stage is Pareto-dominated by the (LU, RU) one. Contrary to the case where $\delta = 1/2$, the order of the move in the sequential game now matters when $\delta = 1$.

Result III.8. *When $\delta = 1$ and the two mainstream parties move first, the subgame perfect equilibria of the game are: (LU, RU) followed by either LU or RU.*

And finally, when the game is played only with two players:

Result III.9. *When $\delta = 1$, all equilibria in the simultaneous and sequential platform game with two players have each party playing either LU or RU.*

Results III.6 to III.9 follow from the payoffs detailed in Tables B.3 and B.4 in Appendix B.2. We can summarize them in the following proposition:

Proposition III.2. *When $\delta = 1$, political parties prefer to offer wide platforms and the electoral campaign does not feature single-issue parties. The non-partisan issue is always offered by all parties.*

III.5 Interpretation of the results

Proposition III.2 captures the fact that when the attention constraint is not severe ($\delta = 1$), all equilibria involve parties campaigning on both issues. This is because $\delta = 1$ essentially shuts down the dilution effect. Parties, then, do not suffer from splitting their platforms and have more incentives to include U in their platform. When doing so, a party will make the vertical issue more salient (salience effect) *and* taking a stance on that issue (predatory effect), without suffering from what could be seen as a less focused platform.

On the other hand, Proposition III.1 highlight the fact that when $\delta = \frac{1}{2}$, the dilution effect provides too strong a disincentive for parties to split their platforms. The joint benefit of the predatory and salience effects, i.e. taking a stance on an issue and thereby making such a move more rewarded by voters through increased salience, are outweighed by the deep penalty associated with non-focused platforms. This, combined with the positive externalities that arise from campaigning on the same issue as the other party, gives rise to Proposition III.1.

Taken together, Propositions III.1 and III.2 capture the impact of δ on the type of campaign. When the electoral competition is strong or the media space very constrained (low δ), parties choose to focus their platform on one issue, while they choose to "spread" when δ is high. Since δ represents how much a party suffers from splitting their manifesto over both issues, the effect of δ on the incentive to focus one's platform is clearly monotically decreasing.

It is therefore possible to conjecture that there exists a $\delta^* \in \left(\frac{1}{2}, 1\right)$ such that, when $\delta < \delta^*$, the campaign will feature only single-issue parties and when $\delta > \delta^*$, the campaign will feature multi-issue parties. Such a conjecture suggests that even when an issue is non-partisan, such issue may not be part of the parties' platforms because spreading one's platform is too detrimental ($\delta < \delta^*$), i.e. it would overly reduce the benefit of campaigning over the divisive issue. Moreover, adopting the non-partisan issue U as a unique policy would not provide enough political space if the other two parties focus on the horizontal L-R issue. As a result, U may not be part of any platform at all.

Assuming that there are mechanisms incentivizing the winning party to stick to its campaigning platform (e.g. reputation or mid-term elections), the non-partisan policy U would then not be implemented at all, even after the elections have taken place and its electoral stakes are gone. It is therefore possible that non-partisan issues, which would be beneficial to all voters, are altogether left out of the political arena.

Moreover, a comparison of results between two- and three-party games in our model suggests that the smaller the parties are, the more likely they are to bunch on similar issues. This is an interesting addendum to Duverger's Law, in which majoritarian systems, which tend to favour a smaller number of parties, should also result in more diverse election issues, while parliamentary systems should result in more focused campaigns, a prediction that could potentially be tested.

III.6 Conclusion of Chapter III

Our model of endogenous political agenda allows party not only to take positions on issues but also to influence how salient voters perceive each issue to be. The more parties devote their public time to talk about an issue, the more this issue will be salient to voters. In turn, this makes the issue more appealing from the point of view of the parties. The positive externalities associated with campaigning on the same issue as the other parties are however mitigated by the need to defend one's own political space, which pushes parties to campaign on the divisive issue rather than the non-partisan issue, in order to differentiate politically. But our model further suggests that when there are numerous parties, the incidence of each party on the issue salience is limited; this makes it impossible for parties to "bowl alone", even on a non-divisive issue, while the other parties remain focused on the divisive one. A party cannot, on its own, change the face of a political campaign.

The positive externalities associated with issue salience may then explain the inertia of the (endogenous) political agenda: a new party, even purely ideologically-motivated, may not be able on its own to shift the focus away from the divisive issue towards the non-partisan one. This would even be more true if traditional parties have more power over issue salience (for instance through established connections with the media): they could then maintain the political spectrum along traditional cleavages, and discourage emergence of parties fighting for new issues.

Moreover, when the air-time and public messages by parties are constrained, or when ambivalence is severely disliked by voters, parties will focus their platforms on the divisive issue, and not even include the consensual one as part of their public campaign. This explains why non-divisive issues like the ecology or consumer rights may emerge only slowly in the political space: they offer little scope for differentiation and because of resource constraints on the public communication, parties may not prioritise them.

One could check the robustness of the results by introducing a dogmatic party that only cares about making the vertical issue as salient as possible in the campaign.[7] In such a configuration, the equilibrium of the sequential game still features mainstream parties offering (L, R) for a low δ. The conclusion is similar as before: the party that will be elected will be one of the mainstream parties, and one that does not have U as part of its electoral platform.

When the attention constraint is very binding, as captured by a low δ in our model, a new party entering the race does not have much influence over the issues being debated during the campaign. Even if such a party is ideologically motivated by the issue U, it does not have enough weight on its own to make the non-partisan issue salient enough to force the other two parties on the issue. As a consequence, if entry in the political arena were costly, such new party may well refrain to enter at all.

This may help explain why, in some European countries, green parties haven taken so long to materialize the issue politically. The fact that environmental issues start to be taken up by mainstream—although only half-heartedly—may suggest a change in the constraint of the communication space (higher δ): more media availability, increased Internet communication, etc. which may have induced more competition over the issue. It may also be that the debate relating environmental issues, while becoming more focused and subtle, has also become more divisive, with credible argumented positions on both sides of specific questions (e.g. on nuclear power, the use of GMOs, or the adaptation vs. mitigation debate in climate change). The model suggests that even in countries that try to level the playing field—e.g. through strict airtime regulations—smaller parties may have difficulties making a difference on the agenda. This can be particularly handicapping for small parties in majoritarian systems, where their electoral prospects are usually limited

[7] This could be the case of a small activist party that cannot get elected because of a majoritarian electoral system.

to forcing some issues on the agenda competitively discussed by the main parties. When the attention constraint is particularly acute, this might not even be possible for them to achieve.

Our model therefore suggests a conclusion different from that of Page (1976), who predicted that politicians would emphasize the non-partisan issue and remain ambiguous on the divisive ones. Further, only when political communication is not too constrained will parties include the consensual issues as part of their public campaign. This suggests a link between the structure of the media, the legal requirements about public campaigning, and the dimensionality of the political ideological space.

APPENDIX A

Appendix to Chapter I

A.1 Proof of Lemma I.1

A.1.1 Workers do not want to leave gaps with their neighbours

Denote ν the number of individuals networked in the tradeable sector. By assumption, all these individuals are evenly spaced on the skill circle, therefore at a distance $1/\nu$ of each other.

Remember that $\tilde{p}(\theta^i, \theta^{-i}, \nu)$ denotes the probability that the skill worker i requires is covered by another worker in her network.

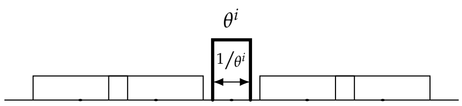


Figure A.1: Situation where worker i's choice of specialization does not lead to skill overlap with the neighbours. This happens when $\frac{1}{\theta^i} \leq \frac{2}{\nu} - \frac{1}{\theta^{-i}}$.

When $\frac{1}{\theta^i} \leq \frac{2}{\nu} - \frac{1}{\theta^{-i}}$, worker i's skill set does not overlap with that of its immediate neighbours, as pictured in Figure A.1.[1] In that case, this probability, $\tilde{p}$, will be independent of the breadth of worker i's skills. Rather, it will depend on the breadth

[1] All networked workers are evenly spaced on the skill circles at a distance $\frac{1}{\nu}$ from each other. If all *other* workers overlap between each other, the skill space they cover has width $\tilde{p} = 1 - \frac{2}{\nu} + \frac{1}{\theta^{-i}}$, where θ^{-i} is the level of specialization chosen by worker's i two immediate neighbours (and possibly the others on the circle).
If, on the other hand the other workers do not overlap between each others, then $\tilde{p} = \sum_{k \neq i} \frac{1}{\theta^k} = (\nu - 1)\theta^{-i}$, where the last eaquality comes from assuming that all these other

of the other workers' skill sets. Therefore, when $\frac{1}{\theta^i} \leq \frac{2}{\nu} - \frac{1}{\theta^{-i}}$ such that there is indeed no overlap in the skill support of networked agents, worker i's expected utility, from Equation (I.5), can be rewritten as:

$$\mathbb{E}U_T^i = \frac{1}{\theta^i} \cdot U_T(\theta^i) + cstt \tag{A.1}$$

And since $\frac{1}{\theta^i} U_T(\theta^i) = (\theta^i)^{\delta-1} + \frac{k^2}{\theta^i w}$ is decreasing in θ^i, worker i would never find it optimal to leave parts of the skill distribution uncovered—that is to leave gaps with her skill neighbours.

(a) No gap, no overlap. (b) No gap, but overlap.

Figure A.2: A worker will not leave gaps with their workers.

She will instead choose θ^i such that the breadth of her skill support is sufficient either to exactly meet the skill support of her neighbours (Figure A.2a) or, broader still, to overlap with the support of her neighbours (Figure A.2b). That is to say agent i will choose a skill breadth

$$\frac{1}{\theta^i} \geq \frac{2}{\nu} - \frac{1}{\theta^{-i}} \tag{A.2}$$

where Equation (A.2) holds with equality when the skill sets are just touching. In either case, we can make the same reasoning for *all* workers on the circle. Hence, there will be no 'gaps' along the skill circle—no horizontal skills left uncovered—and the probability that worker i's project falls in the support of a worker with whom she is networked is simply $1 - \frac{1}{\theta^i}$. This, then, is $\tilde{p}$, the probability that worker i is forced to rely on the skills of her network θ^{-i} and gains utility $U_T(\alpha\theta^{-i})$—eroded by α due to the cost associated with the transmission of specialized knowledge through networks.

workers had chosen the same level of specialization θ^{-i}. In both cases, it does not depend on θ^i.

A.1.2 Workers do not want to overlap with their neighbours

When Equation (A.2) holds, expected utility for networked workers is therefore:

$$\mathbb{E}U_T^i = \underbrace{\frac{1}{\theta^i} \cdot U_T(\theta^i)}_{\text{Own skill}} + \underbrace{\left(1 - \frac{1}{\theta^i}\right) \cdot U_T(\alpha\theta^{-i})}_{\text{Other worker's skill}} \tag{A.3}$$

$$\mathbb{E}U_T^i = \frac{k^2}{w} + \left[(\theta^i)^{\delta-1} + \left(1 - \frac{1}{\theta^i}\right)(\alpha\theta^{-i})^{\delta}\right] \tag{A.4}$$

To pin down the optimal choice of θ^i, let's denote $d(\nu, \theta^{-i}) \equiv \frac{2}{\nu} - \frac{1}{\theta^{-i}}$ the width between a worker's neighbours on either side of her position on the skill spectrum. We showed in Appendix A.1.1 that the expected utility is decreasing once the skill set does not overlap with that of one's neighbours, i.e. when the width of the worker's skill set is smaller than the width to its neighbours: $\frac{1}{\theta^i} < d(\nu, \theta^{-i})$.

We now a necessary and sufficient condition for the expected utility to peak at that particular width, namely that the expected utility would decrease when the width of the skill set becomes bigger than the criticial width $d(\nu, \theta^{-i})$. As long as the change in expected utility as a function of skill width, $\frac{\partial \mathbb{E}U_T^i}{\partial(1/\theta^i)}$, evaluated at that precise width $\frac{1}{\theta^i} = d(\nu, \theta^{-i})$ is negative, workers will find it optimal to be as specialized as is possible (whilst ensuring that there remains no gaps, see Appendix A.1.1). This is because the worker's expected utility $\mathbb{E}U_T^i$ will be decreasing on both sides of criticial width $d(\nu, \theta^{-i})$. More precisely:

$$\frac{\partial \mathbb{E}U_T^i}{\partial(1/\theta^i)} = (1-\delta)\left(\frac{1}{\theta^i}\right)^{-\delta} - (\alpha\theta^{-i})^{\delta} \tag{A.5}$$

is negative as long as $(\alpha\theta^{-i})^{\delta} \geq (1-\delta)\theta^{i^{\delta}}$. This, evaluated at $\frac{1}{\theta^i} = \frac{2}{\nu} - \frac{1}{\theta^{-i}}$, gives the condition:

$$(\alpha\theta^{-i})^{\delta} \geq (1-\delta)\left[d(\nu, \theta^{-i})\right]^{-\delta} \tag{A.6}$$

which requires that α and δ are sufficiently high.

When Equation (A.6) is satisfied, the expected utility of worker i peaks when $\frac{1}{\theta^i} = \frac{2}{\nu} - \frac{1}{\theta^{-i}}$: the individual will choose to neither overlap nor leave gaps with neighbours. Individual optimal choice then is θ^{i*} such that:

$$\frac{1}{\theta^{i*}} = \frac{2}{\nu} - \frac{1}{\theta^{-i}} \tag{A.7}$$

As shown in Appendix A.2, all workers in the tradeable sector will decide to get networked, such that $\nu = \eta N$, yielding Equation (I.7) in Lemma I.1.

Note that Equation (A.6) is necessary and sufficient. For sufficiency, first note that Appendix A.1.1 shows the worker's expected utility $\mathbb{E}U_T^i$ is decreasing in $1/\theta^i$ when $1/\theta^i \in (0, d(\nu, \theta^{-i})]$. Moreover, $\mathbb{E}U_T^i$ is concave over $(d(\nu, \theta^{-i}), 1]$ since on that interval $\frac{\partial^2 \mathbb{E}U_T^i}{\partial 1/\theta^{i2}} = -\delta(1-\delta)\left(\frac{1}{\theta^i}\right)^{-(1+\delta)} < 0$. By concavity, if the function is downward-sloping as $1/\theta^i$ approaches $\rightarrow d(\nu, \theta^{-i})$ from above, it is downward-sloping everywhere else on $(d(\nu, \theta^{-i}), 1]$. This implies that the expected utility is maximized at the kink. Necessity is implied by concavity again, since if the function were not downward-sloping at the kink, it would be maximized somewhere else than the kink on the interval $(d(\nu, \theta^{-i}), 1]$.

A.2 Decision to get networked

A.2.1 Condition for getting networked

An individual in the tradeable sector who decides not to get networked is unable to perform the project if she does not have the required skill. Her expected payoff therefore is simply:

$$\mathbb{E}U_T^0 = \frac{1}{\theta^i} U_T(\theta^i) = (\theta^i)^{\delta - 1} + \frac{1}{\theta^i w} \tag{A.8}$$

Given the decreasing returns to skill, a non-networked agent has incentives not to specialize and choose $\theta^{i*} = 1$, yielding expected utility:

$$\mathbb{E}U_T^0 = 1 + \frac{1}{w} \tag{A.9}$$

Denote ν the number of networked workers in the tradeable sectors (including worker i). The expected payoff of an individual who chooses optimal choice of θ^i in Equation (A.7) is given by:

$$\mathbb{E}U_T^i = \frac{1}{w} + \left[\left(\frac{2}{\nu} - \frac{1}{\theta^{-i}}\right)^{1-\delta} + \left(1 - \frac{2}{\nu} + \frac{1}{\theta^{-i}}\right)(\alpha\theta^{-i})^{\delta}\right] \tag{A.10}$$

Comparing Equations (A.9) and (A.10), a worker will choose to get networked if the payoff difference satisfies doing so satisfies:

$$\left(\frac{2}{\nu} - \frac{1}{\theta^{-i}}\right)^{1-\delta} + \left(1 - \frac{2}{\nu} + \frac{1}{\theta^{-i}}\right)(\alpha\theta^{-i})^{\delta} - 1 \geq 0 \tag{A.11}$$

Here a higher number of networked individuals has two effects:

- it increases the optimal choice of skill level among networked individuals. (see Equation (I.7))
- it increases the incentive for tradeable workers to get networked (see Equation (A.11) in Appendix A.1)

The two effects combined mean that the incentives to network are larger when others also do so. However, while this strategic complementarity is the mechanism at the heart of our paper, the focus is not to deal with the coordination problem induced by it, and we therefore assume that tradeable workers coordinate on getting networked.[2]

In that case, every tradeable worker gets networked:

$$\nu = \eta N \tag{A.12}$$

and we can then rewrite the "no overlap" condition (A.6) as

$$(\alpha\theta^{-i})^{\delta} \geq (1-\delta)\left[\frac{2}{\eta N} - \frac{1}{\theta^{-i}}\right]^{-\delta} \tag{A.13}$$

A.2.2 Adding a cost of networking

If there is a cost $q > 0$ of getting networked, the condition to get networked becomes:

$$q \leq \left(\frac{2}{\nu} - \frac{1}{\theta^{-i}}\right)^{1-\delta} + \left(1 - \frac{2}{\nu} + \frac{1}{\theta^{-i}}\right)(\alpha\theta^{-i})^{\delta} - 1 \tag{A.14}$$

which only raises marginally the constraints on δ and α, and makes the coordination problem marginally more acute.

However, we feel that the discounted efficiency of networks (through $\alpha < 1$) captures the cost of relying on one's network, and an additional cost q would only replicate this effect without adding much to the model.

[2] See Appendix A.2 for details on the decision to get networked. To rationalize the coordination problem, one could think that, in reality, this choice is done in an overlapping manner over time, and hence any new worker faces a set of already networked individuals.

APPENDIX B

Appendix to Chapter III

B.1 Payoffs and proofs for the case $\delta = \frac{1}{2}$.

B.1.1 Three-player game payoffs with $\delta = \frac{1}{2}$.

Player 1 is the table-mover, player 2 is the row-mover, player 3 is the column-mover.

Player 3

Player 2	L	LU	U	RU	R
L	$\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$	$\frac{9}{20}, \frac{9}{20}, \frac{1}{10}$	$\frac{3}{8}, \frac{3}{8}, \frac{1}{4}$	$\frac{19}{60}, \frac{19}{60}, \frac{11}{30}$	$\frac{1}{4}, \frac{1}{4}, \frac{1}{2}$
LU	$\frac{9}{20}, \frac{1}{10}, \frac{9}{20}$	$\frac{3}{4}, \frac{1}{8}, \frac{1}{8}$	$\frac{1}{2}, 0, \frac{1}{2}$	$\frac{7}{12}, 0, \frac{5}{12}$	$\frac{1}{2}, 0, \frac{1}{2}$
U	$\frac{3}{8}, \frac{1}{4}, \frac{3}{8}$	$\frac{1}{2}, \frac{1}{2}, 0$	$\frac{1}{4}, \frac{3}{8}, \frac{3}{8}$	$\frac{5}{12}, \frac{1}{4}, \frac{1}{3}$	$\frac{1}{2}, 0, \frac{1}{2}$
RU	$\frac{19}{60}, \frac{11}{30}, \frac{19}{60}$	$\frac{7}{12}, 0, \frac{5}{12}$	$\frac{5}{12}, \frac{1}{3}, \frac{1}{4}$	$\frac{7}{12}, \frac{5}{24}, \frac{5}{24}$	$\frac{1}{2}, 0, \frac{1}{2}$
R	$\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$	$\frac{1}{2}, \frac{1}{2}, 0$	$\frac{1}{2}, \frac{1}{2}, 0$	$\frac{1}{2}, \frac{1}{2}, 0$	$\frac{1}{2}, \frac{1}{4}, \frac{1}{4}$

Player 1 plays L

Player 3

Player 2	L	LU	U	RU	R
L	$\frac{1}{10}, \frac{9}{20}, \frac{9}{20}$	$\frac{1}{8}, \frac{3}{4}, \frac{1}{8}$	$0, \frac{1}{2}, \frac{1}{2}$	$0, \frac{7}{12}, \frac{5}{12}$	$0, \frac{1}{2}, \frac{1}{2}$
LU	$\frac{1}{8}, \frac{1}{8}, \frac{3}{4}$	$\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$	$\frac{1}{8}, \frac{1}{8}, \frac{3}{4}$	$\frac{1}{4}, \frac{1}{4}, \frac{1}{2}$	$\frac{5}{24}, \frac{5}{24}, \frac{7}{12}$
U	$0, \frac{1}{2}, \frac{1}{2}$	$\frac{1}{8}, \frac{3}{4}, \frac{1}{8}$	$\frac{1}{10}, \frac{9}{20}, \frac{9}{20}$	$\frac{3}{16}, \frac{5}{8}, \frac{3}{16}$	$\frac{1}{3}, \frac{1}{4}, \frac{5}{12}$
RU	$0, \frac{5}{12}, \frac{7}{12}$	$\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$	$\frac{3}{16}, \frac{3}{16}, \frac{5}{8}$	$\frac{1}{2}, \frac{1}{4}, \frac{1}{4}$	$\frac{5}{12}, 0, \frac{7}{12}$
R	$0, \frac{1}{2}, \frac{1}{2}$	$\frac{5}{24}, \frac{7}{12}, \frac{5}{24}$	$\frac{1}{3}, \frac{1}{4}, \frac{5}{12}$	$\frac{5}{12}, \frac{7}{12}, 0$	$\frac{11}{30}, \frac{19}{60}, \frac{19}{60}$

Player 1 plays LU

Player 3

Player 2	L	LU	U	RU	R
L	$\frac{1}{4}, \frac{3}{8}, \frac{3}{8}$	$\frac{1}{2}, \frac{1}{2}, 0$	$\frac{3}{8}, \frac{1}{4}, \frac{3}{8}$	$\frac{1}{4}, \frac{5}{12}, \frac{1}{3}$	$0, \frac{1}{2}, \frac{1}{2}$
LU	$\frac{1}{2}, 0, \frac{1}{2}$	$\frac{3}{4}, \frac{1}{8}, \frac{1}{8}$	$\frac{9}{20}, \frac{1}{10}, \frac{9}{20}$	$\frac{5}{8}, \frac{3}{16}, \frac{3}{16}$	$\frac{1}{4}, \frac{1}{3}, \frac{5}{12}$
U	$\frac{3}{8}, \frac{3}{8}, \frac{1}{4}$	$\frac{9}{20}, \frac{9}{20}, \frac{1}{10}$	$\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$	$\frac{9}{20}, \frac{9}{20}, \frac{1}{10}$	$\frac{3}{8}, \frac{3}{8}, \frac{1}{4}$
RU	$\frac{1}{4}, \frac{1}{3}, \frac{5}{12}$	$\frac{5}{8}, \frac{3}{16}, \frac{3}{16}$	$\frac{9}{20}, \frac{1}{10}, \frac{9}{20}$	$\frac{3}{4}, \frac{1}{8}, \frac{1}{8}$	$\frac{1}{2}, 0, \frac{1}{2}$
R	$0, \frac{1}{2}, \frac{1}{2}$	$\frac{1}{4}, \frac{5}{12}, \frac{1}{3}$	$\frac{3}{8}, \frac{1}{4}, \frac{3}{8}$	$\frac{1}{2}, \frac{1}{2}, 0$	$\frac{1}{4}, \frac{3}{8}, \frac{3}{8}$

Player 1 plays U

Table B.1: Payoffs in the three-player simultaneous game, with $\delta = \frac{1}{2}$. In each cell, the first number is the payoff for the table-player, the second for the row-player, the third for the column-player.

Player 3

Player 2		L	LU	U	RU	R
	L	$\frac{11}{30}, \frac{19}{60}, \frac{19}{60}$	$\frac{5}{12}, \frac{7}{12}, 0$	$\frac{1}{3}, \frac{5}{12}, \frac{1}{4}$	$\frac{5}{24}, \frac{7}{12}, \frac{5}{24}$	$0, \frac{1}{2}, \frac{1}{2}$
	LU	$\frac{5}{12}, 0, \frac{7}{12}$	$\frac{1}{2}, \frac{1}{4}, \frac{1}{4}$	$\frac{3}{16}, \frac{3}{16}, \frac{5}{8}$	$\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$	$0, \frac{5}{12}, \frac{7}{12}$
Player 2	U	$\frac{1}{3}, \frac{1}{4}, \frac{5}{12}$	$\frac{3}{16}, \frac{5}{8}, \frac{3}{16}$	$\frac{1}{10}, \frac{9}{20}, \frac{9}{20}$	$\frac{1}{8}, \frac{3}{4}, \frac{1}{8}$	$0, \frac{1}{2}, \frac{1}{2}$
	RU	$\frac{5}{24}, \frac{5}{24}, \frac{7}{12}$	$\frac{1}{4}, \frac{1}{4}, \frac{1}{2}$	$\frac{1}{8}, \frac{1}{8}, \frac{3}{4}$	$\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$	$\frac{1}{8}, \frac{1}{8}, \frac{3}{4}$
	R	$0, \frac{1}{2}, \frac{1}{2}$	$0, \frac{7}{12}, \frac{5}{12}$	$0, \frac{1}{2}, \frac{1}{2}$	$\frac{1}{8}, \frac{3}{4}, \frac{1}{8}$	$\frac{1}{10}, \frac{9}{20}, \frac{9}{20}$

Player 1 plays RU

Player 3

Player 2		L	LU	U	RU	R
	L	$\frac{1}{2}, \frac{1}{4}, \frac{1}{4}$	$\frac{1}{2}, \frac{1}{2}, 0$	$\frac{1}{2}, \frac{1}{2}, 0$	$\frac{1}{2}, \frac{1}{2}, 0$	$\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$
	LU	$\frac{1}{2}, 0, \frac{1}{2}$	$\frac{7}{12}, \frac{5}{24}, \frac{5}{24}$	$\frac{5}{12}, \frac{1}{3}, \frac{1}{4}$	$\frac{7}{12}, \frac{5}{12}, 0$	$\frac{19}{60}, \frac{11}{30}, \frac{11}{30}$
Player 2	U	$\frac{1}{2}, 0, \frac{1}{2}$	$\frac{5}{12}, \frac{1}{4}, \frac{1}{3}$	$\frac{1}{4}, \frac{3}{8}, \frac{3}{8}$	$\frac{1}{2}, \frac{1}{2}, 0$	$\frac{3}{8}, \frac{1}{4}, \frac{3}{8}$
	RU	$\frac{1}{2}, 0, \frac{1}{2}$	$\frac{7}{12}, 0, \frac{5}{12}$	$\frac{1}{2}, 0, \frac{1}{2}$	$\frac{3}{4}, \frac{1}{8}, \frac{1}{8}$	$\frac{9}{20}, \frac{1}{10}, \frac{9}{20}$
	R	$\frac{1}{4}, \frac{1}{4}, \frac{1}{2}$	$\frac{19}{60}, \frac{19}{60}, \frac{11}{30}$	$\frac{3}{8}, \frac{3}{8}, \frac{1}{4}$	$\frac{9}{20}, \frac{9}{20}, \frac{1}{10}$	$\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$

Player 1 plays R

Table B.1: (Continued from p. 85) Payoffs in the three-player simultaneous game, with $\delta = \frac{1}{2}$. In each cell, the first number is the payoff for the table-player, the second for the row-player, the third for the column-player.

B.1.2 Proof of Result III.3 (on Page 70)

When $\delta = \frac{1}{2}$, for each platform choice from Player 1 in the first stage, we look at Nash equilibria in the 2x2 simultaneous game of the second stage. In case of multiple Nash equilibria in the second stage, we discard the ones that are Pareto-dominated for players 2 and 3.

First mover	NE in 2$^{\text{nd}}$ stage	Non-dominated NE	Player's 1 payoff
L	(L,R) and (R,R)	(L,R)	1/4
LU	(L,U) and (L,R)	(L,U) and (L,R)	0
U	(L,R) and (U,U)	(L,R)	0
RU	(R,U) and (L,R)	(R,U) and (L,R)	0
R	(L,R) and (L,L)	(L,R)	1/4

This makes L or R followed by (L, R) to be the subgame perfect outcome of this sequential game.

B.1.3 Two player game payoffs with $\delta = \frac{1}{2}$.

Payoffs are indicated in Table B.2.

		Player 2				
		L	LU	U	RU	R
	L	$\frac{1}{2}, \frac{1}{2}$	$\frac{5}{6}, \frac{1}{6}$	$\frac{1}{2}, \frac{1}{2}$	$\frac{11}{18}, \frac{7}{18}$	$\frac{1}{2}, \frac{1}{2}$
	LU	$\frac{1}{6}, \frac{5}{6}$	$\frac{1}{2}, \frac{1}{2}$	$\frac{1}{6}, \frac{5}{6}$	$\frac{1}{2}, \frac{1}{2}$	$\frac{7}{18}, \frac{11}{18}$
Player 1	U	$\frac{1}{2}, \frac{1}{2}$	$\frac{5}{6}, \frac{1}{6}$	$\frac{1}{2}, \frac{1}{2}$	$\frac{5}{6}, \frac{1}{6}$	$\frac{1}{2}, \frac{1}{2}$
	RU	$\frac{7}{18}, \frac{11}{18}$	$\frac{1}{2}, \frac{1}{2}$	$\frac{1}{6}, \frac{5}{6}$	$\frac{1}{2}, \frac{1}{2}$	$\frac{1}{6}, \frac{5}{6}$
	R	$\frac{1}{2}, \frac{1}{2}$	$\frac{11}{18}, \frac{7}{18}$	$\frac{1}{2}, \frac{1}{2}$	$\frac{5}{6}, \frac{1}{6}$	$\frac{1}{2}, \frac{1}{2}$

Table B.2: Payoffs in the two-player simultaneous game, with $\delta = \frac{1}{2}$.

B.2 Payoffs for the case $\delta = 1$.

B.2.1 Three-player game payoffs with $\delta = 1$.

Player 1 is the table-mover, player 2 is the row-mover, player 3 is the column-mover.

Player 3

Player 2		L	LU	U	RU	R
	L	$\frac{1}{3},\frac{1}{3},\frac{1}{3}$	$0,0,1$	$\frac{3}{8},\frac{3}{8},\frac{1}{4}$	$\frac{9}{40},\frac{9}{40},\frac{11}{20}$	$\frac{1}{4},\frac{1}{4},\frac{1}{2}$
	LU	$0,1,0$	$0,\frac{1}{2},\frac{1}{2}$	$0,1,0$	$0,\frac{1}{2},\frac{1}{2}$	$0,\frac{11}{20},\frac{9}{20}$
Player 2	U	$\frac{3}{8},\frac{1}{4},\frac{3}{8}$	$0,0,1$	$\frac{1}{4},\frac{3}{8},\frac{3}{8}$	$\frac{1}{4},0,\frac{3}{4}$	$\frac{1}{2},0,\frac{1}{2}$
	RU	$\frac{9}{40},\frac{11}{20},\frac{9}{40}$	$0,\frac{1}{2},\frac{1}{2}$	$\frac{1}{4},\frac{3}{4},0$	$\frac{3}{8},\frac{5}{16},\frac{5}{16}$	$\frac{9}{20},\frac{11}{20},0$
	R	$\frac{1}{4},\frac{1}{2},\frac{1}{4}$	$0,\frac{9}{20},\frac{11}{20}$	$\frac{1}{4},\frac{1}{4},\frac{1}{2}$	$\frac{9}{20},0,\frac{11}{20}$	$\frac{1}{2},\frac{1}{4},\frac{1}{4}$

Player 1 plays L

Player 3

Player 2		L	LU	U	RU	R
	L	$1,0,0$	$\frac{1}{2},0,\frac{1}{2}$	$1,0,0$	$\frac{1}{2},0,\frac{1}{2}$	$\frac{11}{20},0,\frac{9}{20}$
	LU	$\frac{1}{2},\frac{1}{2},0$	$\frac{1}{3},\frac{1}{3},\frac{1}{3}$	$\frac{1}{2},\frac{1}{2},0$	$\frac{1}{4},\frac{1}{4},\frac{1}{2}$	$\frac{5}{16},\frac{5}{16},\frac{3}{8}$
Player 2	U	$1,0,0$	$\frac{1}{8},\frac{3}{4},\frac{1}{8}$	$1,0,0$	$\frac{1}{2},0,\frac{1}{2}$	$\frac{3}{4},0,\frac{1}{4}$
	RU	$\frac{1}{2},\frac{1}{2},0$	$\frac{1}{4},\frac{1}{2},\frac{1}{4}$	$\frac{1}{2},\frac{1}{2},0$	$\frac{1}{2},\frac{1}{4},\frac{1}{4}$	$\frac{1}{2},\frac{1}{2},0$
	R	$\frac{11}{20},\frac{9}{20},0$	$\frac{5}{16},\frac{3}{8},\frac{5}{16}$	$\frac{3}{4},\frac{1}{4},0$	$\frac{1}{2},0,\frac{1}{2}$	$\frac{11}{20},\frac{9}{40},\frac{9}{40}$

Player 1 plays LU

Player 3

Player 2		L	LU	U	RU	R
	L	$\frac{1}{4},\frac{3}{8},\frac{3}{8}$	$0,0,1$	$\frac{3}{8},\frac{1}{4},\frac{3}{8}$	$0,\frac{1}{4},\frac{3}{4}$	$\frac{1}{2},\frac{1}{4},\frac{1}{4}$
	LU	$0,1,0$	$0,\frac{1}{2},\frac{1}{2}$	$0,1,0$	$0,\frac{1}{2},\frac{1}{2}$	$0,\frac{3}{4},\frac{1}{4}$
Player 2	U	$\frac{3}{8},\frac{3}{8},\frac{1}{4}$	$0,0,1$	$\frac{1}{3},\frac{1}{3},\frac{1}{3}$	$0,0,1$	$\frac{3}{8},\frac{3}{8},\frac{1}{4}$
	RU	$0,\frac{3}{4},\frac{1}{4}$	$0,\frac{1}{2},\frac{1}{2}$	$0,1,0$	$0,\frac{1}{2},\frac{1}{2}$	$0,1,0$
	R	$\frac{1}{2},\frac{1}{4},\frac{1}{4}$	$0,\frac{1}{4},\frac{3}{4}$	$\frac{3}{8},\frac{1}{4},\frac{3}{8}$	$0,0,1$	$\frac{1}{4},\frac{3}{8},\frac{3}{8}$

Player 1 plays U

Table B.3: Payoffs in the three-player simultaneous game, with $\delta = 1$. In each cell, the first number is the payoff for the table-player, the second for the row-player, the third for the column-player.

Player 3

Player 2		L	LU	U	RU	R
	L	$\frac{11}{20}, \frac{9}{40}, \frac{9}{40}$	$\frac{1}{2}, 0, \frac{1}{2}$	$\frac{3}{4}, \frac{1}{4}, 0$	$\frac{5}{16}, \frac{3}{8}, \frac{5}{16}$	$\frac{11}{20}, \frac{9}{20}, 0$
	LU	$\frac{1}{2}, \frac{1}{2}, 0$	$\frac{1}{2}, \frac{1}{4}, \frac{1}{4}$	$\frac{1}{2}, \frac{1}{2}, 0$	$\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$	$\frac{1}{2}, \frac{1}{2}, 0$
Player 2	U	$\frac{3}{4}, 0, \frac{1}{4}$	$\frac{1}{2}, 0, \frac{1}{2}$	$1, 0, 0$	$\frac{1}{8}, \frac{3}{4}, \frac{1}{8}$	$1, 0, 0$
	RU	$\frac{5}{16}, \frac{5}{16}, \frac{3}{8}$	$\frac{1}{4}, \frac{1}{4}, \frac{1}{2}$	$\frac{1}{2}, \frac{1}{2}, 0$	$\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$	$\frac{1}{2}, \frac{1}{2}, 0$
	R	$\frac{11}{20}, 0, \frac{9}{20}$	$\frac{1}{2}, 0, \frac{1}{2}$	$1, 0, 0$	$\frac{1}{2}, 0, \frac{1}{2}$	$1, 0, 0$

Player 1 plays RU

Player 3

Player 2		L	LU	U	RU	R
	L	$\frac{1}{2}, \frac{1}{4}, \frac{1}{4}$	$\frac{11}{20}, 0, \frac{9}{20}$	$\frac{1}{4}, \frac{1}{4}, \frac{1}{2}$	$0, \frac{9}{20}, \frac{11}{20}$	$\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$
	LU	$\frac{11}{20}, \frac{9}{20}, 0$	$\frac{3}{8}, \frac{5}{16}, \frac{5}{16},$	$\frac{1}{4}, \frac{3}{4}, 0$	$0, \frac{1}{2}, \frac{1}{2}$	$\frac{9}{40}, \frac{11}{20}, \frac{9}{40}$
Player 2	U	$\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$	$\frac{1}{4}, 0, \frac{3}{4}$	$\frac{1}{4}, \frac{3}{8}, \frac{3}{8}$	$0, 0, 1$	$\frac{3}{8}, \frac{1}{4}, \frac{3}{8}$
	RU	$0, \frac{11}{20}, \frac{9}{20}$	$0, \frac{1}{2}, \frac{1}{2}$	$0, 1, 0$	$0, \frac{1}{2}, \frac{1}{2}$	$0, 1, 0$
	R	$\frac{1}{4}, \frac{1}{4}, \frac{1}{2}$	$\frac{9}{40}, \frac{9}{40}, \frac{11}{20}$	$\frac{3}{8}, \frac{3}{8}, \frac{1}{4}$	$0, 0, 1$	$\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$

Player 1 plays R

Table B.3: (Continued from p. 88) Payoffs in the three-player simultaneous game, with $\delta = 1$. In each cell, the first number is the payoff for the table-player, the second for the row-player, the third for the column-player.

B.2.2 Two player game payoffs with $\delta = 1$.

Player 2

Player 1		L	LU	U	RU	R
	L	$\frac{1}{2}, \frac{1}{2}$	$0, 1$	$\frac{1}{2}, \frac{1}{2}$	$\frac{5}{12}, \frac{7}{12}$	$\frac{1}{2}, \frac{1}{2}$
	LU	$1, 0$	$\frac{1}{2}, \frac{1}{2}$	$1, 0$	$\frac{1}{2}, \frac{1}{2}$	$\frac{7}{12}, \frac{5}{12}$
	U	$\frac{1}{2}, \frac{1}{2}$	$0, 1$	$\frac{1}{2}, \frac{1}{2}$	$0, 1$	$\frac{1}{2}, \frac{1}{2}$
	RU	$\frac{7}{12}, \frac{5}{12}$	$\frac{1}{2}, \frac{1}{2}$	$1, 0$	$\frac{1}{2}, \frac{1}{2}$	$1, 0$
	R	$\frac{1}{2}, \frac{1}{2}$	$\frac{5}{12}, \frac{7}{12}$	$\frac{1}{2}, \frac{1}{2}$	$0, 1$	$\frac{1}{2}, \frac{1}{2}$

Table B.4: Payoffs in the two-player simultaneous game, with $\delta = 1$.

Bibliography

Adriani, Fabrizio and Silvia Sonderegger (Dec. 2009). "Why do parents socialize their children to behave pro-socially? An information-based theory". In: *Journal of Public Economics* 93.11-12, pp. 1119–1124.

Amoros, Pablo and M. Socorro Puy (Oct. 2007). *Dialogue or issue divergence in the political campaign?* CORE Discussion Papers 2007084. Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).

Aragonés, Enriqueta, Micael Castanheira, and Marco Giani (June 2012). *Electoral Competition through Issue Selection*. CEPR Discussion Papers 9012. C.E.P.R. Discussion Papers.

Austen-Smith, David A. (1987). "Sophisticated Sincerity: Voting Over Endogenous Agendas". In: *The American Political Science Review* 81.4, pp. 1323–1330. DOI: 10.2307/1962591.

Bacolod, Marigee, Bernardo S. Blum, and William C. Strange (Mar. 2009). "Skills in the city". In: *Journal of Urban Economics* 65.2, pp. 136–153.

Banerjee, Abhijit V. and Andrew F. Newman (1993). "Occupational Choice and the Process of Development". In: *The Journal of Political Economy* 101.2, pp. 274–298. DOI: 10.2307/2138820.

Baron, David P. and John A. Ferejohn (1989). "Bargaining in Legislatures". In: *The American Political Science Review* 83.4, pp. 1181–1206. DOI: 10.2307/1961664.

Baumgartner, Frank R., Jeffrey M. Berry, Marie Hojnacki, Beth L. Leech, and David C. Kimball (2009). *Lobbying and Policy Change: Who Wins, Who Loses, and Why*. University of Chicago Press. Chap. 5.

Bayer, Patrick, Stephen L. Ross, and Giorgio Topa (Dec. 2008). "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes". In: *Journal of Political Economy* 116.6, pp. 1150–1196.

Becker, Gary S. (1983). "A Theory of Competition Among Pressure Groups for Political Influence". In: *The Quarterly Journal of Economics* 98.3, pp. 371–400.

Becker, Gary S. and Casey B. Mulligan (1997). "The Endogenous Determination of Time Preference". In: *The Quarterly Journal of Economics* 112.3, pp. 729–758. DOI: 10.2307/2951254.

Becker, Randy A. and Vernon Henderson (2000). "Intra-industry Specialization and Urban Development." In: *Economics of Cities: Theoretical Perspectives*. Ed. by Huriot Jean-Marie and Jean-FranÃ§ois Thisse. Cambridge: Cambridge University Press, pp. 138–166.

Bélanger, Éric and Bonnie M. Meguid (Sept. 2008). "Issue salience, issue ownership, and issue-based vote choice". In: *Electoral Studies* 27.3, pp. 477–491. DOI: 10.1016/j.electstud.2008.01.001.

Belot, Michèle and John Ermisch (2009). "Friendship ties and geographical mobility: evidence from Great Britain". In: *Journal of the Royal Statistical Society Series A* 172.2, pp. 427–442.

Berliant, Marcus, Robert R. III Reed, and Ping Wang (July 2006). "Knowledge exchange, matching, and agglomeration". In: *Journal of Urban Economics* 60.1, pp. 69–95.

Bisin, Alberto and Thierry Verdier (Apr. 2001). "The Economics of Cultural Transmission and the Dynamics of Preferences". In: *Journal of Economic Theory* 97.2, pp. 298–319.

Bosquet, Clément and Pierre-Philippe Combes (Apr. 2013). *Do Large Departments Make Academics More Productive? Agglomeration and Peer Effects in Research*. SERC Discussion Papers 0133. Spatial Economics Research Centre, LSE.

Bowles, Samuel (Mar. 1998). "Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions". In: *Journal of Economic Literature* 36.1, pp. 75–111.

Boyd, Robert and Peter J. Richerson (1985). *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Budge, Ian and Dennis J. Farlie (1983). *Explaining and Predicting Elections: Issue Effects and Party Strategies in Twenty-Three Democracies*. London: Allen and Unwin.

Budge, Ian and Michael J. Laver (1992). *Party Policy and Government Coalitions*. Palgrave Macmillan.

Burchfield, Marcy, Henry G. Overman, Diego Puga, and Matthew A. Turner (May 2006). "Causes of Sprawl: A Portrait from Space". In: *The Quarterly Journal of Economics* 121.2, pp. 587–633.

Calvó-Armengol, Antoni and Matthew O. Jackson (Sept. 2004). "The Effects of Social Networks on Employment and Inequality". In: *American Economic Review* 94.3, pp. 426–454. DOI: 10.1257/0002828041464542.

Calvó-Armengol, Antoni and Matthew O. Jackson (June 2004). "The Effects of Social Networks on Employment and Inequality". In: *The American Economic Review* 94.3, pp. 426–454. DOI: doi:10.1257/0002828041464542.

Cantillon, Estelle (Apr. 2001). *Electoral Rules and the Emergence of New Issue Dimensions*. CEPR Discussion Paper 2769. Center for Economic Policy Research.

Chaney, Thomas and Ralph Ossa (May 2013). "Market Size, Division of Labor, and Firm Productivity". In: *Journal of International Economics*.

Coles, Melvyn G and Eric Smith (Feb. 1998). “Marketplaces and Matching”. In: *International Economic Review* 39.1, pp. 239–54.
Colomer, Josep and Humberto Llavador (Mar. 2012). “An agenda-setting model of electoral competition”. In: *SERIEs* 3.1, pp. 73–93.
Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux (Nov. 2012). “The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection”. In: *Econometrica* 80.6, pp. 2543–2594.
Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux (June 2010). “Estimating Agglomeration Economies with History, Geology, and Worker Effects”. In: *Agglomeration Economics*. Ed. by Edward L. Glaeser. NBER Books. National Bureau of Economic Research, Inc, pp. 15–66.
— (2012). “Sorting and local wage and skill distributions in France”. In: *Regional Science and Urban Economics* 42.6, pp. 913–930.
Currie, Janet (2001). “Early childhood education programs”. In: *The Journal of Economic Perspectives* 15.2, pp. 213–238.
De La Roca, Jorge and Diego Puga (Jan. 2013). *Learning By Working In Big Cities*. CEMFI Working Papers 2013-1301. CEMFI.
Demange, Gabrielle and Karine Van Der Straeten (Nov. 2009). *A communication game on electoral platforms*. IDEI Working Papers 589. Institut d’Économie Industrielle (IDEI), Toulouse.
Denter, Philipp (Feb. 2013). *A theory of communication in political campaigns*. Economics Working Paper Series 1302. University of St. Gallen, School of Economics and Political Science.
Dixit, Avinash (Dec. 2003). “Trade Expansion and Contract Enforcement”. In: *Journal of Political Economy* 111.6, pp. 1293–1317.
Doepke, Matthias and Fabrizio Zilibotti (2008). “Occupational Choice and the Spirit of Capitalism”. In: *Quarterly Journal of Economics* 123.2, pp. 747–793. DOI: 10.1162/qjec.2008.123.2.747.
Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde (2012). “The Intergenerational Transmission of Risk and Trust Attitudes”. In: *The Review of Economic Studies* 79.2, pp. 645–677. DOI: 10.1093/restud/rdr027.
Duranton, Gilles and Diego Puga (2004). “Micro-foundations of urban agglomeration economies”. In: *Handbook of Regional and Urban Economics*. Ed. by J. V. Henderson and J. F. Thisse. Vol. 4. Handbook of Regional and Urban Economics 4. Elsevier. Chap. 48, pp. 2063–2117.
Eeckhout, Jan, Roberto Pinheiro, and Kurt Schmidheiny (Dec. 2010). *Spatial Sorting: Why New York, Los Angeles and Detroit attract the greatest minds as well as the unskilled*. CEPR Discussion Papers 8151. C.E.P.R. Discussion Papers.
Egorov, Georgy (Dec. 2012). *Single-Issue Campaigns and Multidimensional Politics*. Working paper. Northwestern University.

Eliaz, Kfir and Ran Spiegler (2011). "Consideration Sets and Competitive Marketing". In: *Review of Economic Studies* 78.1, pp. 235–262.

Ellison, Glenn (July 1994). "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching". In: *The Review of Economic Studies* 61.3, pp. 567–88.

Ellison, Glenn, Edward L. Glaeser, and William R. Kerr (Sept. 2010). "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns". In: *American Economic Review* 100.3, pp. 1195–1213. DOI: 10.1257/aer.100.3.1195.

Festinger, L., S. Schachter, and K. Back (1950). "The Spatial Ecology of Group Formation". In: *Social Pressure in Informal Groups: A Study of Human Factors in Housing*. Ed. by L. Festinger, S. Schachter, and K. Back. Stanford: Stanford University Press. Chap. 4.

Fudenberg, Drew and Eric Maskin (1986). "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information". In: *Econometrica* 54.3, pp. 533–554. DOI: 10.2307/1911307.

Galor, Oded and Joseph Zeira (1993). "Income Distribution and Macroeconomics". In: *The Review of Economic Studies* 60.1, pp. 35–52. DOI: 10.2307/2297811.

Gennetian, Lisa A, Lisa Sanbonmatsu, Lawrence F Katz, Jeffrey R Kling, Matthew Sciandra, Jens Ludwig, Greg J Duncan, and Ronald C Kessler (2012). "The long-term effects of Moving to Opportunity on youth outcomes". In: *Cityscape*, pp. 137–167.

Ghatak, Maitreesh and Neville Nien-Huei Jiang (Oct. 2002). "A simple model of inequality, occupational choice, and development". In: *Journal of Development Economics* 69.1, pp. 205–226. DOI: 10.1016/S0304-3878(02)00059-7.

Gibbons, Steve, Henry G. Overman, and Panu Pelkonen (Oct. 2010). *Wage Disparities in Britain: People or Place?* SERC Discussion Papers 0060. Spatial Economics Research Centre, LSE.

Glaeser, Edward L., ed. (July 2010). *Agglomeration Economics*. NBER Books. National Bureau of Economic Research, Inc.

Glaeser, Edward L., Matthew E. Kahn, and Jordan Rappaport (Jan. 2008). "Why do the poor live in cities The role of public transportation". In: *Journal of Urban Economics* 63.1, pp. 1–24.

Glaeser, Edward L. and David C. Maré (Apr. 2001). "Cities and Skills". In: *Journal of Labor Economics* 19.2, pp. 316–42.

Glaeser, Edward L. and Matthew G. Resseger (2010). "The Complementarity between Cities and Skills". In: *Journal of Regional Science* 50.1, pp. 221–244.

Glaeser, Edward L. and Bruce Sacerdote (Mar. 2000). "The Social Consequences of Housing". In: *Journal of Housing Economics* 9.1-2, pp. 1–23.

Green, Jane and Sara B. Hobolt (Sept. 2008). "Owning the issue agenda: Party strategies and vote choices in British elections". In: *Electoral Studies* 27.3, pp. 460–476. DOI: 10.1016/j.electstud.2008.02.003.

Grossman, Gene M. and Elhanan Helpman (1994). “Protection for Sale”. In: *The American Economic Review* 84.4, pp. 833–850. DOI: 10.2307/2118033.

Heckman, James J. and Yona Rubinstein (2001). “The Importance of Noncognitive Skills: Lessons from the GED Testing Program”. In: *The American Economic Review* 91.2, pp. 145–149. DOI: 10.2307/2677749.

Helsley, Robert W. and William C. Strange (Sept. 1990). “Matching and agglomeration economies in a system of cities”. In: *Regional Science and Urban Economics* 20.2, pp. 189–212.

Hug, Simon (Sept. 2001). *Altering Party Systems: Strategic Behavior and the Emergence of New Political Parties in Western Democracies (Interests, Identities, and Institutions in Comparative Politics)*. University of Michigan Press.

Kandori, Michihiro (1992). “Social Norms and Community Enforcement”. In: *The Review of Economic Studies* 59.1, pp. 63–80. DOI: 10.2307/2297925.

— (Jan. 2002). “Introduction to Repeated Games with Private Monitoring”. In: *Journal of Economic Theory* 102.1, pp. 1–15.

Kaplan, Noah, David K. Park, and Travis N. Ridout (2006). “Dialogue in American Political Campaigns? An Examination of Issue Convergence in Candidate Television Advertising”. In: *American Journal of Political Science* 50.3, pp. 724–736.

Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman (2001). “Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment”. In: *The Quarterly Journal of Economics* 116.2, pp. 607–654. DOI: 10.1162/00335530151144113.

Kremer, Michael and Daniel Chen (1999). “Income-Distribution Dynamics with Endogenous Fertility”. In: *The American Economic Review* 89.2, pp. 155–160. DOI: 10.2307/117098.

Krusell, Per and Anthony A. Smith (1998). “Income and Wealth Heterogeneity in the Macroeconomy”. In: *The Journal of Political Economy* 106.5, pp. 867–896.

Levy, Gilat and Ronny Razin (July 2008). “Attention for Sale: Endogenous Agenda Formation in Sequential vs. Simultaneous Debates”. mimeo.

Meguid, Bonnie M. (2005). “Competition Between Unequals: The Role of Mainstream Party Strategy in Niche Party Success”. In: *American Political Science Review* 99.03, pp. 347–359. DOI: 10.1017/S0003055405051701.

Michaels, Guy, Ferdinand Rauch, and Stephen J. Redding (Jan. 2013). *Task Specialization in U.S. Cities from 1880-2000*. CEPR Discussion Papers 9308. C.E.P.R. Discussion Papers.

Mulligan, Casey B. (1997). *Parental Priorities and Economic Inequality*. Chicago Series in Law and Society. University of Chicago Press.

Overman, Henry G. and Diego Puga (June 2010). “Labor Pooling as a Source of Agglomeration: An Empirical Investigation”. In: *Agglomeration Economics*. Ed. by Edward L. Glaeser. NBER Chapters. National Bureau of Economic Research, Inc, pp. 133–150.

Page, Benjamin I. (1976). "The Theory of Political Ambiguity". In: *The American Political Science Review* 70.3, pp. 742–752. DOI: 10.2307/1959865.

Petrocik, John R. (1996). "Issue Ownership in Presidential Elections, with a 1980 Case Study". In: *American Journal of Political Science* 40.3, pp. 825–850. DOI: 10.2307/2111797.

Poguntke, Thomass (2002). "Green parties in National Governements: From Protest to Aquiescence?" In: *Green Parties in National Governments*. Ed. by Ferdinand Müller-Rommel and Thomas Poguntke. Routledge. Chap. 7.

Poulsen, Anders and Gert Svendsen (Apr. 2005). "Social Capital and Endogenous Preferences". In: *Public Choice* 123.1-2, pp. 171–196. DOI: 10.1007/s11127-005-0266-0.

Puga, Diego (2010). "The Magnitude And Causes Of Agglomeration Economies". In: *Journal of Regional Science* 50.1, pp. 203–219.

Riker, William H. (1993). *Agenda Formation*. Ann Arbor: University of Michigan Press.

Riker, William H. and Peter C. Ordeshook (1973). *An Introduction to Positive Political Theory*. Prentice-Hall.

Rosenstone, Steven J., Roy L. Behr, and Edward H. Lazarus (1984). *Third Parties in America: Citizen Response to Major Party Failure*. Princeton Univ Pr.

Rosenthal, R. W. (1979). "Sequences of Games with Varying Opponents". In: *Econometrica* 47.6, pp. 1353–1366. DOI: 10.2307/1914005.

Rosenthal, Stuart S. and William C. Strange (Sept. 2001). "The Determinants of Agglomeration". In: *Journal of Urban Economics* 50.2, pp. 191–229.

Sáez-Martí, María and Anna Sjögren (Mar. 2008). "Peers and Culture". In: *Scandinavian Journal of Economics* 110.1, pp. 73–92.

Salop, Steven C. (1979). "Monopolistic Competition with Outside Goods". In: *Bell Journal of Economics* 10.1, pp. 141–156.

Scotchmer, Suzanne (2002). "Local public goods and clubs". In: *Handbook of Public Economics*. Ed. by A. J. Auerbach and M. Feldstein. Vol. 4. Handbook of Public Economics. Elsevier. Chap. 29, pp. 1997–2042.

Sigelman, Lee and Emmett H. Buell (2004). "Avoidance or Engagement? Issue Convergence in U.S. Presidential Campaigns, 1960-2000". In: *American Journal of Political Science* 48.4, pp. 650–661. DOI: 10.1111/j.0092-5853.2004.00093.x.

Stokes, Donald E. (1963). "Spatial Models of Party Competition". In: *The American Political Science Review* 57.2, pp. 368–377. DOI: 10.2307/1952828.

Tabellini, Guido (2008). "The Scope of Cooperation: Norms and Incentives". In: *Quaterly Journal of Economics*.

Zenou, Yves (Feb. 2011). *Spatial versus Social Mismatch: The Strength of Weak Ties*. Research Papers in Economics 2011:5. Stockholm University, Department of Economics.