POLICY COMPETITION AND THE SPATIAL ECONOMY

PLATFORM31	This dissertation is part of the Platform31 research project "Stedelijke netwerken". Financial support from Platform31 is gratefully acknowledged.

ISBN 978 90 361 0376 3

Cover design: Crasborn Designers, Valkenburg a.d. Geul. Cover photo: Museo Guggenheim, Bilbao (Spain) by nito/Shutterstock.com.

This is book no. 573 of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

VRIJE UNIVERSITEIT

POLICY COMPETITION AND THE SPATIAL ECONOMY

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. F.A. van der Duyn Schouten,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Economische Wetenschappen en Bedrijfskunde
op maandag 20 januari 2014 om 15.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Michiel Jan-Arnold Gerritse

geboren te Wageningen

promotoren: prof.dr. H.L.F. de Groot
prof.dr. E.T. Verhoef

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

Symbol	Description
A	aggregate productivity (may depend on labor supply)
a	individual worker productivity
B_r	expenditure share of global expenditure on region r
$B(t)$	financial assets bought at time t
$b_{r,q}$	share of (manufacturing) expenditure from region r to region q
$c(i)$	consumption of manufacturing variety i
C	consumption index/final good consumption
E	expenditure
F	fixed labor requirement in production
G	consumption of government-provided good
H	aggregate land/housing supply
h	consumption of housing
i	index to denote individual firms
I	funds spent on public inputs
K	high-skilled workers
L	low-skilled workers
l_f	leisure
l_s	labor supply (individual)
m	consumption of government-provided amenity
MP	market potential
N	population
n	number of firms
$p(i)$	price of manufacturing variety produced by firm i
P	(CES) harmonized price index/final good price
Q	regional aggregate production
$q(z)$	product of workers of skill z
r	land rent (inside a city)
r_a	agricultural rent (outside a city)
r_f	financial interest rate
S	subsidy

s	share of budget devoted to subsidies
s_n	share of global number of firms in region 1
T	tax rate
t	time index
TC	firm's total costs
U	utility function
V	indirect utility function
W	welfare function
w	wage rate
$y(i)$	quantity of an input produced
Y	final product assembled from inputs $y(i)$
z	skill index
α	preference for (tradable) manufacturing goods
β	time discount factor
γ	preference for government-provided goods
δ	governs scale externalities
ε	governs elasticity of substitution between goods from different regions
ϕ	the freeness of trade, $\tau^{1-\sigma}$
η	elasticity of housing supply
Θ	(personal) time endowment
θ	commuting costs
κ	share of workers supplying labor in region 1 (implies commuting if $\kappa > \lambda$)
λ	share of workers that live in region 1
μ	preference for non-tradable (local) goods
ξ	disutility parameter for commutes
Π	firm's profit function
$\pi(t)$	the holding costs of a house for one period
ρ	governs elasticity of substitution between low and high-skilled workers in production of a final good
σ	the constant elasticity of substitution between different manufacturing goods (ch. 2,3), intermediates (ch. 5,6) or workers of different skill (ch. 7)
τ	iceberg transport cost (τ units need to be shipped for one unit to arrive)
ν	employment share in the non-tradables industry
ψ	governs elasticity of substitution between low and high-skilled workers in externalities when producing a final good

PREFACE

The set of interesting, maybe naive ideas that I worked on has somehow, almost unnoticed, changed into this thesis. Obviously, I invested considerable time and effort writing this dissertation, but I was enjoying the effort so much that I would not be disappointed if I were only halfway now – my advisors wisely told me to stop writing the next chapter.

Erik en Henri, ik ben jullie inhoudelijk en persoonlijk veel verschuldigd. Erik, ik heb me keer op keer verbaasd hoe snel je mijn hersenkronkels doorzag en er nuttige dingen over zei. Het is een genot om als slecht formulerende, eigenwijze AIO zo snel op snelheid te zijn - ik hoop dat ik niet te eigenwijs ben geweest, of te veel advocaat van de duivel heb gespeeld. Henri, we hebben veel gepraat over het proefschrift maar eigenlijk delen we veel bredere interesses. Ik heb genoten van discussies over allerlei economische hobby's en uitstapjes naast mijn proefschrift. Dank je voor veel wijze inzichten over de grotere lijnen: de grote beleidsvragen, de stand van het vakgebied, en hoe we precies met elkaar om moeten gaan in de economie. Ik wil jullie allebei ook bedanken voor de grote vrijheid die ik voelde bij het schrijven van dit proefschrift. Ik kon altijd mijn eigen ideeën onderzoeken en hoewel ze regelmatig niet werkten, was het erg de moeite waard. Mijn nog diepere respect moet ik betuigen voor het lezen van alle tussenversies - nu ik het spelletje iets meer begin te doorzien realiseer ik me pas dat jullie eerste jaar met mij zwaar moet zijn geweest. Desalniettemin was de sfeer uitstekend met zijn drieën om tafel.

Reading a thesis like mine - and even worse, judging it - is undoubtedly hard work. I sincerely want to thank Andrés Rodríguez-Pose, Philip McCann, Lex Meijdam, Harry Garretsen and Peter Nijkamp for all their efforts and patience with this text. I highly appreciate your efforts, and I am flattered to have you on the reading committee. I want to mention Harry Garretsen specifically. Harry, you were, in part, the architect of me working with Erik and Henri, and on and off-stage helped me in various ways. I look forward to running into each other every now and then.

Dit proefschrift was niet geschreven zonder de financiële steun van Platform 31 (voormalig NICIS) onder de projectnaam Stedelijke Netwerken. Daar ben ik ze erg dankbaar voor. Ik heb ook profijt gehad van interacties met partners binnen het project: discussies en samenwerking met de de afdeling economie in Groningen (waaronder Gerard Marlet en Abdella Oumer), veel nuttige observaties uit de praktijk van de stedelijke partners die een belangrijke rol speelden in het forensenmodel in twee hoofdstukken in dit proefschift, en een samenwerking met het Centraal Planbureau (met name Wouter Vermeulen) om de effecten van infrastructuur-investeringen te bestuderen.

Genetisch gezien hebben mijn ouders dit boekje volledig mogelijk gemaakt. Maar daar hield het niet op - jullie interesse in mijn werk, en die van Doranne natuurlijk, zijn belangrijk voor me. Mam, dank je voor alle steun rond het schrijven van dit boekje. Pap, is het niet frappant dat ik in jouw interessegebied terecht ben gekomen, terwijl je me juist hebt aangemoedigd veel om me heen te kijken.

Colleagues and friends of a PhD student cannot escape listening to boring discussions on equilibrium and simulation issues. I appreciate the people that tolerated my bother – at least all my office mates, and fellow teachers Mark and Chris (despite his limited knowledge of pre-70s music). Sergej, we entered and leave the VU more or less at the same time. We got along straight away and I have enjoyed the many discussions over beers between bankers since. Dani and Masha, thanks for the increasingly off campus hangouts especially if they involved jamón, and of course for showing me that few things have more sex-appeal than a bowtie.

Mobile as academics are, I regret not seeing a lot of people more often. I want to thank many people for showing me Milan, like Andrea, Idil and Olaf. It is great to know that people close and far away are thinking similar thoughts, like Ana, David, Elvan, Heike, Lena, Ricardo, Stephan to name a few; and I think Gunther Maier's efforts to bring young researchers together cannot be over-appreciated.

Work with my co-authors never entered this thesis – "hobby" is a pejorative term but it does describe the enjoyment with which I worked on co-authored projects. Ana, you found me at the right time (for me) indeed, thanks for asking all the relevant questions and introducing me to Groningen. Dani, what you do is fascinating and I believe we are good complements. Gerrit, het is ontzettend plezierig met je te werken, en ik kijk op naar hoe groot jouw wereld buiten de economie is.

Ik kan me soms verliezen in een paar vergelijkingen, en gelukkig zijn er mensen die me doen inzien wat voor onzin die vergelijkingen zijn. Anne, Bregtje, Coty, ML en Wouter, goed dat we nog biertjes (en icetea) doen. Anne, Beek, Fleur, Marieke en Thierry, jullie zijn de leukste mensen om

Macchiavelli van te verliezen. Joris, ik denk niet dat ik snel weer op honderd meter van je woon, maar ik heb mijn best gedaan. Lieve familie, inclusief kleine neefjes en hele kleine neefjes, etentjes zijn altijd een warm bad. Tom, Resi, Ceciel, Jasper en Tjerk, dank jullie voor de wijn, bij mijn promotiediner en elders.

Bart, Koen, Lucas, Pieter en Sander, jullie brengen relativeren op een heel ander niveau. Steevast aan de flep en mijn favoriete vriendjes om poëzie mee te schrijven. Ik wilde graag één paranimf uit de Trekschuit, maar dan blijven er vier over. Aannemende dat ik nog een keer trouw en/of overlijd (beide één spreker) zal ik dan maar proberen in ieder geval namen op drie kinderen te plakken.

Lieve Lotte, we hebben heel dicht bij elkaar en heel ver van elkaar gewoond (waaronder Milaan-Barcelona). We hebben samen best wat gereisd, maar los van elkaar ook veel kilometers afgelegd. Ik kan dus goed vergelijken: als jij in de buurt bent, is eten lekkerder en zijn steden bruisender. Ik vind het moeilijk om onder woorden te brengen, maar met jou erbij is alles leuker.

Herfst 2013, Groningen, Utrecht en Amsterdam,

Michiel

CHAPTER 1

INTRODUCTION

Can we expect governments to choose the right policies? Governments and politicians easily spend up to half of a nation's income (World Bank, 2011), so the mistakes that they make can be expensive. Moreover, as a government's budget consists of painfully raised tax revenue, its citizens often care deeply whether their taxes are put to good use. Unfortunately, there are reasons to believe governments will fail to develop the best policies in the interest of their citizens – not because they are malevolent, but because there are obstacles between intent and actual policy. For economies that are connected to other economies in various ways, spatial relations form a hurdle to perfect government. If firms move around freely, government may adapt their tax rates or industrial policies to attract firms and their employment opportunities from other locations, thus competing with other governments.

Competition among governments may take many forms. Much less so than with firms, however, is it always obvious how they compete, and why they compete. An idealized view of governments holds that those in power will use any instrument that improves the fate of their own inhabitants, if only to be (re-)elected. Even so, if all governments act in the best interest of their citizens, significant problems can arise if economies are interrelated. For instance, policymakers can adjust their taxation of firms to encourage firm location, realizing that lower taxes on many firms still yield higher budgets than high taxes on few firms. Clearly, such practices are good news to firms, and in its most extreme form for mail box companies and holding firms that are set up for the sole purpose of avoiding taxation. The Rolling Stones have set up headquarters in Amsterdam and paid most of their taxes there,[1] even though they neither live in Amsterdam nor perform there disproportionately often. Similarly, excise taxes on gas in Luxemburg are (much) lower than in surrounding areas. Individually, this tax level is rational for the government of Luxemburg: tax revenues are higher with low taxes on many sales than with high taxes on few sales. Collectively, however, the aggregate tax revenue from fuel may

[1] http://www.independent.co.uk/news/uk/this-britain/stones-paid-just-16-tax-on-163240m-royalties-410232.html

well fall, which results in fewer resources for governments to finance their public tasks. From the individual perspective, low taxes are understandable, but in a broader perspective, they may reduce welfare. However, tax rates need not be the only instrument at hand. Especially when looking at a subnational level, where tax rates on labor and capital are not for the government's choosing, governments find other ways to attract firms or people. These may range from local tax holidays (Klemm, 2010) and explicit subsidies on the supply of public infrastructure (Keen and Marchand, 1997) like roads or office space and facilities that attract specific (types of) persons or firms (Falck et al., 2011), but which take up precious government budget. Therefore, this competition between governments is sometimes referred to as a "race to the bottom": governments gear their policies to the attraction of firms, entering a wasteful process that reduces the funds for their other public responsibilities.

A different view on these worries, however, arises when examining the geography of economies more closely. Large spatial inequalities indicate that firms will not choose any location just for its appealing policies: a look into international income differences or into the urban landscape suggests that firms prefer to locate in large markets. For that reason, agglomeration effects play a role in policy outcomes. A simple intuition behind how agglomeration forces affect governments' competition is the following: if the government of a large city or country sets higher taxes, firms will not unavoidibly move. Surely, they would gain from paying lower taxes if they left, but they would also lose from leaving the access to consumers, or to their suppliers, and benefit less from externalities of knowledge and on the labor market. Therefore, lowered tax rates or other efforts to encourage firm establishment may not work for small regions' governments, and large regions' governments do not feel pressured to forego public services to retain local firms.

It is exactly the role of these agglomeration forces in policy competition that this thesis further examines. The thesis provides a theoretical view on policy competition that puts economic geography under further scrutiny. In doing so, the thesis advances several insights that come about upon closer examination of the agglomeration argument in policy competition. For instance, it shows that the existence of agglomeration forces does not necessarily reduce the need for policy harmonization (as was argued by earlier literature in this context); that people's mobility provides a plausible but imperfect counterweight against poor policy resulting from tax competition; and that comparable policy biases can emerge even if cities are isolated and do not compete with other governments.

Most conclusions are based on a common theoretical framework in the tradition of the New Economic Geography (e.g., Fujita et al., 2001). The New Economic Geography was the framework that initially started the discussion of scale effects in tax competition (Baldwin and Krugman, 2004). To provide a context of the contributions of this thesis, the next section provides a discussion on policy competition and the role of scale effects that are central in the New Economic Geography literature.

1.1 Policy competition and location

A central (decades-old) case to interfere in local governments' administration relies on the belief that footloose firms can use their freedom to move to other jurisdictions, thus enforcing the policies that they like (Wilson and Wildasin, 2004; Fischel, 2001). If firms and capital are able to move across regions and countries, governments are forced to take into account that firms might leave if they dislike the local policies. If policies can be designed to attract firms, government can bias their taxation and expenditure in the favor of local industry, hence attracting firms that provide a tax base, employment, and other benefits to the region. Many instruments lend themselves to attract economic activity. Some governments supply outright subsidies or provide firms with tax breaks. Governments can also provide services and goods to their local industries: physical infrastructure, credit provision, industrial sites and zoning, or facilitating a talented labour pool. Likewise, countries and cities can present themselves as good business sites, by facilitating headquarter services, city marketing and "urban icons": landmarks that distinguish their location compared to their peers. Clearly, as economic activity transcends intra- as well as international borders, these mechanisms apply on several layers of government: urban, regional, national and even supranational governments face competition for firms and have instruments to manipulate firms' location choices.

If all jurisdictions attempt to attract firms, however, the incentive for firms to relocate is small: firms are likely to be indifferent between two locations if both set high tax rates, but, equally, they will be indifferent if both locations set low tax rates. However, if other governments set low tax rates, high local taxes may lead to severe losses of economic activity, and therefore regions are forced to set lower taxes to retain economic activity. The low tax rates translate into an erosion of the social and public services that governments provide to their citizens. Competition in policies can therefore be wasteful, and is referred to as a “race to the bottom”. Efforts to improve cities’, regions’ or countries’ “competitiveness” are understandable from the perspective of citizens and policymakers, but they may eventually fail to produce desirable outcomes when considered in the system of cities or countries to which they belong.

If local policies have effects outside the own region, efficient policy formation is likely to be problematic. If policymakers use policies that have disparate consequences for local inhabitants’ welfare and for global welfare, policies are not likely to be optimal. In other words, better policies could be available if governments took into account all the effects that their policies generate. Unfortunately, this is often impossible in dichotomized policy-formation; governments care most about the well-being of citizens inside their borders out of benevolence, or because those are the voters that need to (re-)elect them.

An argument based on the spatial nature of the economy, however, claims that a careful study of firms’ location motives paints a less gloomy picture. If firms prefer to settle in locations where they have easy access to large markets or where they can easily obtain their inputs, there is an agglomerative force. Should this happen, then governments can exploit the firm’s preference to locate in large markets by making them pay higher taxes. Under such agglomerative forces, a policy bias towards firms is not or less likely to be required to retain firms and jobs.

The argument that agglomeration effects can eliminate the race to the bottom rests on insights in the spatial nature of economies developed in the New Economic Geography (NEG) literature. The crucial advance can be attributed to the regional scale effects that NEG portrays (see Krugman, 1991, or Baldwin et al., 2003 for an introduction). Such scale or agglomeration effects are not novel, however. At least since the Industrial Revolution, agglomeration forces have formed a centerpiece of economic and geographical inquiry with, among others, central contributions from Alfred Marshall in 1890. Various sources of agglomeration can be identified, such as the benefits of others in learning and knowledge spillovers, sharing inputs and improved matching quality and fewer search frictions between employer and employees or buyers and sellers in large markets (Duranton and Puga, 2004). Indeed, the results from the seminal tax-competition-with-scale-effects result from Baldwin and Krugman (2004), based on NEG, can be motivated by a wide variety of scale effects (Krogstrup, 2008).

Yet, with its full-fledged trade model, and the ability to explain the firms' internal returns to scale, the New Economic Geography literature has brought important new aspects of space into mainstream economics according to many economists, including the Nobel prize committee.[2] The Dixit and Stiglitz (1977) monopolistic competition model, that allows for firm indivisibilities and transport cost permits a central focus on trade, along with the ability to explain agglomeration through firms' or people's migration. The unified explanation of trade and location sets the NEG apart from many earlier models of scale effects. In the context of this thesis, the NEG model is an advancement on earlier models of agglomeration (that use, for instance, productive spillovers or other localization and urbanization economies), which do not necessarily explain trade and other interaction between regions (Fujita and Thisse, 2002, provide an overview). Similarly, other trade models, based on factor endowments or Ricardian comparative advantages, are less suited to study migration and the growth and decline of locations, which is a decisive process in tax competition. The ability to explain agglomeration processes as well as trade patterns in a single framework therefore makes the NEG an obvious choice to study spatial interactions.

Despite a wide recognition of the NEG results, and a surge in research into its theoretical and empirical implications, its foundations are not undebated. To many economists, the exact behavior of economies under trade costs is controversial, while many geographers view the role of space and distance frictions in New Economic Geography as an analytical oversimplification. Although improving upon spaceless economies, the introduction of an idealized, abstract geometric space (Garretsen and Martin, 2010) remains insufficient to call new economic geography a part of geography (Duranton and Rodríguez-Pose, 2005). Yet, despite their very plain treatment of space, as said, NEG models gave rise to novel views on tax competition. Given the intimate relation between tax competition and space, it is not surprising that further scrutinizing the role of space alters the analysis and its conclusions. Tax and policy competition, by their nature, deal with interactions between different but connected economies, and therefore geography is likely to feature prominently in the analysis.

[2] Paul Krugman won the Nobel Prize in economics in 2008. See Fujita and Thisse (2009) or Behrens and Robert-Nicoud (2009) for an appraisal.

The multiplicity of locations, and their scale- and agglomeration effects, inherently deal with localization effects and the role of distance in trade, labor mobility and migration. It is therefore exactly the role of space and geography in economic policy competition that this thesis aims to explore.

1.2 Central questions

In this thesis, the consequences of taxation as well as the corresponding policies that are used to steer economic size and composition are given a spatial context. Not only do taxes change location decisions; the infrastructure that they finance could improve local productivity, they could affect firm entry and exit, and they could change a region's economic composition and change the local quality of life. Therefore, studying the broader spatial picture requires that interregional interactions be examined: even if not in the own region, increased consumption options due to additional firm entry in another region may benefit local inhabitants, and employment opportunities outside the own city still favor commuters. Therefore, the complete set of government tasks (taxation, public services, industrial policies, etc.) involves interactions well beyond the location-distorting impacts of taxation.

The central research questions adressed in this thesis therefore all revolve around how spatial interactions form obstacles to achieving policies that maximize aggregate welfare. Although some economists like to believe differently, the broader economy is replete with barriers and frictions that make the economies of some locations look very different from other locations. It is costly to ship goods over distance, residents have a limited spatial radius to travel to work, and even information does not flow freely across space. As a result, representative economies do not exist; rather, citizens, firms and policymakers face a collection of locations that are not identical. How do spatial interactions between regions and their governments impede the formation of optimal policies?

The framework of spatial interaction in policy gives rise to a wide variety of questions that can be studied. As economies are not islands, there are many ways in which regions and cities are connected, and equally, there are many types of questions that can be adressed that deal with space. The nature of interplay between governments themselves has implications on how policy is formed: does it matter whether governments choose their policies simultaneously or in sequence? The labor market is also a source of interaction: workers can commute, and the type of workers that governments try to attract can be very specific (e.g., highly skilled or educated, rich individuals). Many countries substantially subsidize commuting. Should central governments promote or discourage commuting to allow local governments to reach better decisions? Should the amenities that highly skilled workers like, such as museums, be concentrated in a few cities, or be scattered across space? This ties in with the discussion of the effectiveness of using cultural amenities as an (urban) development strategy (e.g., the “Bilbao effect”). And if infrastructure is constructed between cities, how will this affect jobs and location patterns? The answer to this question can not only help to predict, for instance, how the Chinese urban landscape will change after massive infrastructure investments are completed, but also whether infrastructure investments effectively promote employment in lagging regions. Similarly, if people can move freely, citizens leaving one housing market will enter another. If workers can migrate away from cities that spend all their budget on firms instead of on citizens, will the movers put sufficient pressure on policymakers to design good policies? Should a central government intervene, by promoting locating in certain cities? Does it matter whether houses are rented or owned for optimal policies come about in a democracy?

Some broader themes that this thesis addresses deserve to be discussed. A tension that surfaces in several chapters of the thesis (notably chapters 3 and 6) is that while firms are attracted by policies geared toward business, inhabitants may be driven away from them. If individuals can migrate to places where they would like to live, then cities and regions that provide citizen-friendlier policies are more likely to grow large. This phenomenon is also popularly known as voting by feet. As a result, there is a tension between two types of competition: welfare-reducing competition for firms or capital (that distorts taxes and policy), and welfare-improving competition for citizens (that forces policies that optimize local welfare). While the joint presence of such mechanisms complicates the analysis (see chapter 4), it also ties together different strands of literature. On the one hand, there is a strand of literature focusing on (international) capital taxation (see, e.g., the overview in Baldwin et al., 2003, chapter 15). On the other hand, there is a tradition in regional economics and local public finance that focuses on efforts to attract workers (e.g., Oates and Schwab, 1988). Chapter 6 fits more into this latter tradition because it allows people to freely migrate to their location of choice, but the incentive to attract economic activity from the tax competition literature can also be clearly identified. The possible tension between attracting firms and voting by feet is not often found in the literature, and it probably requires empirical investigation to establish which of the two mechanisms is dominant.

Another theme that is increasingly gaining attention (and rightfully so) is that little is known about the actual motives of government. This thesis follows an established tradition by (usually) assuming that governments maximize the average welfare in their jurisdiction. As chapter 3 shows, however, the predicted behavior of governments can vary according to the government's objectives. Chapter 2 shows that a minor simplification of the assumed government's objectives (that is often chosen for analytical convenience) can have large consequences for the policy recommendations that follow from the analysis. In that sense, the chapter makes the argument that despite governments' best intentions, the best policies do not ensue. However, another plausible source of poor policy is that those in government face incentives to select other than welfare-maximizing taxes or industrial policies. Chapter 4 of this thesis moves away from the benevolent view of government in favor of a political view; this clearly has implications for the choices that governments make. By choosing a democratic representation, chapter 4 fits in a tradition of political economics (Drazen, 2000; Persson and Tabellini, 2000; Acemoglu and Robinson, 2006). An argument for pursuing this road is that at times, the political constraint on governments may be far more pressing or relevant than the identification of optimal policy; and this may apply to local governments too.

While these broader themes run through the chapters of this thesis, the conclusions differ per individual chapter as they study different building blocks of spatial economies. These building blocks are discussed in detail in the next section, but Table 1.1 towards the end of this introduction summarizes the issues that the thesis treats chapter by chapter. It can also be used as a reading guide to select chapters of particular interest.

1.3 An overview of the thesis

Given the size of the potential losses involved, and the various possibilities to coordinate or control the behavior of governments, a clearer insight into policy competition is desirable. This thesis contributes to a deeper understanding of the economics of space in policy competition, by studying various manifestations of space in such frameworks. In addition to affecting location choices, cities and regions can manipulate relations between their goods markets, labor markets, and housing markets. This section discusses the contribution of each chapter of the thesis in this respect.

The second chapter investigates the effects of government competition for firms through subsidies in the presence of agglomeration forces. This captures the idea that inhabitants of a city benefit from the presence of businesses, but that it is costly (in terms of public goods provision to households) to attract business. It shows that agglomeration effects will lead larger regions' governments to provide larger subsidies, and that the race to the bottom (or to the top, if firms are attracted to high subsidies instead of low taxes) need not occur. More importantly, it shows that once spillovers on the goods market are taken into account, and governments face a level playing field, policy harmonization can improve welfare. This is in sharp contrast to earlier insights, that advocate against a uniform policy. The chapter shows that the case for no harmonization in earlier literature rests on two assumptions: i) larger governments have timing advantages, and ii) governments are oblivious to the real effects in good market spillovers that their subsidy schemes have.

The third chapter builds on the analysis of the second chapter, but explicitly considers the internal organization of cities. It acknowledges that cities are not a (very small) point in space, but that they require physical space. Therefore, citizens of larger cities will have longer commutes on average. This captures the notion that the increased size of a city, next to all sorts of positive effects, also brings about negative consequences, and the city size is determined by the balance of those opposing forces. While this feature is accomodated in a relatively standard framework, it generates a striking change in how policies come about. Specifically, it suggests that the urban structure leads to coordination problems between governments that are trying to set optimal tax rates. Because both the optimal policy and many other policies can be sustained as a choice that maximizes local welfare in either city, a "lock-in" effect occurs. The intuition is that neither government wants to move the agglomeration. Not moving the agglomeration requires that the tax rates of both regions are not further apart than a certain critical distance: the large region loses the agglomeration if its tax rate exceeds a ceiling, the small region receives the agglomeration if its tax rate is below a tax floor. Effectively, the optimal tax rates are therefore defined as the critical distance apart from the other region's tax rate. Because they are defined in relative terms, there are many tax rate pairs that could classify as optimal (satisfying the tax floor/ceiling), so many tax rate pairs can form part of an agglomeration-preserving policy. It then depends on the initial conditions whether tax rates maximize global welfare, but they could be so gravely off target, that tax-revenue (i.e., budget, not welfare) maximizing local governments can perform better than benevolent governments from a welfare perspective.

The fourth chapter is similar to the third in that it also considers the effects of urban structure on policy formation. However, in contrast to chapter 3, chapter 4 focuses on the politics inside cities. It questions whether the policy inefficiencies observed in the tax competition model can also arise within cities that do not (necessarily) interact with other cities. The chapter introduces an urban model with durable housing and a democratic government. In this framework, homeowners are most likely to vote for productivity- and wage-increasing industrial policies if that increases the price of their house: if wages go up, locations near the labor market are more desirable to live in. The resulting increase in their wealth induces proprietors of a house to vote for extensive industrial policies. As a result, the political candidate that is democratically elected has a bias towards excessive industrial policies: in a closed economy, democratic outcomes are not optimal from a welfare perspective. If migration is allowed in and out of the city, these conclusions are reinforced. Voting for productivity-enhancing infrastructure attracts new inhabitants. Their increased demand for housing causes the future house value improvements, which, in turn, provide a similar political inclination towards expenditure on industrial policies.

The fifth and the sixth chapter are also related; they both focus the option of commuting between regions. The fifth chapter investigates the spatial organization of the economy if firms as well as residents are footloose, and if workers can choose to commute to areas other than their home region. In particular, the model can be used to predict the ramifications of improving interregional infrastructure. In general, increasing connectivity leads residential spreading, and decentralization of the jobs per head. Whether absolute employment spreads or concentrates, depends on preferences and the current stock of infrastructure. This is consistent with recent empirical results on the consequences of infrastructure investment. Moreover, the chapter finishes by demonstrating that the NEG model, which has no closed-form solution, can be derived as a limiting case of the chapter's model. In that specific case, like in the NEG, results can no longer be found using pencil and paper; in fact, numerical simulation suggests that a host of different equilibria can be found. In the setup of the model, however, arriving at that limiting case requires very specific assumptions. Therefore, the chapter's model has the advantage it produces a unique, closed-form equilibrium for all other plausible parameter ranges. The predictions derived from closed-form equilibria therefore seem more general, as they can be found both within and outside the NEG parameter sets. One of those predictions is that jobs will only decentralize following infrastructure investment under very stringent conditions.

The sixth chapter builds on the fifth by introducing governments that can supply public inputs for local production. In contrast to the commonly found policy bias towards industries (including in various chapters in this thesis), the analysis suggests that the public support of firms can also be too low, rather than too high. Because commuters do not take all of the benefits of commuting into account (in particular, their productive effects on co-workers), the commuting flow is suboptimally small. In the long run, migration of residents forces local governments to converge to the best available policies, although the commuting flow is too small. Therefore, citizens' mobility exerts a corrective force onto governments, but it works imperfectly because migration only partially remedies the inefficient size of the commuting flow. A central or national government could improve the situation by instating (differentiated) subsidies for housing (such as mortgage interest rate deductions), but its optimal design crucially depends on housing market properties.

Finally, the seventh chapter deals with the geography of skill, rather than of firms. It thus extends the labor market perspective taken in the fifth and sixth chapter, but abandons the commuting possibilities, to study effects of worker heterogeneity. It departs from the view that many local governments aim to attract specific types of people and industries. In particular, highly educated and skilled workers, and the firms that employ them, are likely to have beneficial effects locally (Moretti, 2004): they improve productivity (also of others), improve technological progress and advance the quality of life. Coupled with the good empirical track record of high-skilled cities (Glaeser and Saiz, 2004), it is no surprise that cities aim to attract the highly skilled. Certain amenities, among which culture, educational institutions and health infrastructure, particularly appeal to the more highly skilled, and can therefore be used to attract highly skilled workers. However, every high-skilled worker won in one region is lost in another region. This chapter examines whether equilibrium policies are optimal when cities interact in attracting highly skilled workers. Given the desirability of highly skilled workers, symmetric cities will bias their policies towards them, thus overproviding the relevant amenities like culture, green areas, education, transport and such. However, as such policy-tailoring makes workers sort according to skill, a likely alternative is that specialized cities emerge. The main benefit of specialization is that policy is more efficient (policy preferences become more homogenous so there is more agreement on the local policy), but there are costs in terms of production: specialization may come at the cost of productivity. Since city governments internalize the specialization benefits of efficient policy fully, but internalize the productive cost of specialization only partially (attracting high skills to the high-skilled regions also renders the low skilled city more lowly skilled), specialized cities overspecialize in their skill. Lowly skilled cities become too lowly skilled, highly skilled cities become too higly skilled. The trade of goods can partially alleviate the policy problems, but as long as some cities are of mixed skill, inefficiencies remain within cities as well as in the patterns of specialization.

Finally, a summary of the topics treated in the various chapters is listed in Table 1.1.

Table 1.1: An overview of the themes by chapter

Chapter	2	3	4	5	6	7
Government interaction	✓	✓				✓
Commuting		✓	✓	✓	✓	
Goods market spillovers	✓					✓
Labor market				✓	✓	✓
Housing market		✓	✓	✓	✓	
Tax competition/voting by feet		✓			✓	✓
Politics			✓			

1.4 Reading guide

While all chapters in this thesis question the efficiency of policy in a spatial context, the spatial settings differ throughout the thesis. Table 1.1 organizes the themes and focuses of different chapters and can be used as a reading guide to select chapters of interest. Readers interested in labor market issues, for instance, may be particularly interested in the chapters that deal with commuting and the chapter that studies skill differences (chapters 5, 6 and 7). Similarly, the housing market features in many chapters, but especially in chapter 4. The ability of citizens to relocate to jurisdictions that they like, which is a central topic in local public finance, is studied in chapters 3, 6, and 7.

The theoretical nature of the models leads to extensive use of letters and Greeks symbols to label variables and parameters in the models. Many chapters share common building blocks, however, and where possible, the chapters employ a uniform notation. A summary of that notation is given in the list of symbols at the start of this thesis.

The eighth and final chapter will return to the questions raised in this introduction. It discusses the findings of different chapters and highlights central conclusions and policy implications to conclude the thesis.

CHAPTER 2

COMPETITION FOR FIRMS UNDER AGGLOMERATION: POLICY INTERACTIONS AND WELFARE

2.1 Introduction

For decades, economists have advocated policy harmonizations based on a fear of "races to the bottom". If mobile firms can be attracted by setting lower taxes, it is rational for governments to expand their tax bases with lower taxes. The fiscal externalities thus generated, however, eventually cause suboptimally low tax rates and foregone provision of public goods (see Wilson and Wildasin, 2004, or Zodrow and Mieszkowski, 1986). Indeed, the possibility that governments dress down policies to attract economic activity is a large concern with the increased globalization and integration of regions and countries.

However, with the advent of New Economic Geography, these conclusions were revised (Baldwin and Krugman, 2004). Introducing trade costs and imperfect competition gives rise to agglomeration rents for firms (i.e., higher operating returns in larger markets), which make firms willing to pay higher taxes to locate in larger regions. In that case, governments of smaller markets seeking to attract firms both need to undercut the large region's taxes and compensate the foregone benefits of clustering with other firms. Larger governments can exploit agglomeration rents by setting their taxes such that governments of smaller markets do not find it worthwhile to attract firms: this leads to higher taxes in large markets (e.g. Jofre-Monseny, 2013). If such agglomerations occur, harmonization "always harms at least one nation" (Baldwin and Krugman, 2004), because optimal tax rates are not equal under agglomeration and harmonization forces at least one country away from its ideal tax rate. Rather, the policy prescription is to set a tax floor: this prevents small regions from undercutting tax rates and so allows large regions to set still higher taxes.

Along with agglomeration effects, however, New Economic Geography models of tax competition invariably introduced other changes to the original tax competition model. In this chapter, we show that these modifications introduced alongside are central in arriving at the welfare conclusions. Firstly, along with the agglomeration rent, the large region typically has a first-mover advantage: it selects and commits to a tax rate before the small region does. Earlier tax competition models, by contrast, study simultaneous policy-making. Secondly, in the models with increasing returns to scale, policies are assumed to have no effect on the real economy: taxes are levied on firms and disappear from the accounts. This is in contrast to earlier models in which governments transform tax revenue into public services. Both of these adaptions have a clear role in simplifying the analysis of tax competition. However, they also obscure whether it is agglomeration effects or other changes with respect to the original tax competition model that account for the novel policy conclusions.

As a main contribution, this chapter shows that the agglomeration effects from the New Economic Geography do not make a case against harmonization – policy harmonization can also improve welfare if such agglomeration effects are present. The desirability of policy harmonization crucially depends on the dynamics of the policy competition game and the presence of policy spillovers. The intuition for our results is that if governments set taxes simultaneously, small regions have an incentive to compete for the agglomeration. This happens if there are goods market spillovers of policy. Small regions' threats to take over the agglomeration leads large regions to subsidize firm variety, which improves consumption options and hence welfare in peripheral regions. The small region's threat to compete is worthless in a sequential game (because the large region commits to a limit tax), but credible in a simultaneous game. Harmonized policies can form Pareto-improvements only if the peripheral region acts competitively, which only occurs if governments set policy simultaneously. Our results thus suggest that part of the changes in strategic incentives in NEG-based policy competition is not due to agglomeration effects, but to the timing assumptions.

The policy conclusions of agglomeration in tax competition therefore depend on the whether large regions have timing advantages in policy formation. Whether simultaneous or sequential policy-setting is more realistic, probably depends on numerous factors, including the institutional arrangements and political cycles. If governments could choose when to change their policies, a leadership for large economies is unlikely (Kempf and Rota-Graziosi, 2010). The objective of this chapter is not to show which timing structure is more realistic; rather, it shows that like in the regular tax competition model, timing, and not agglomeration, is responsible for a large share of the policy implications (Gordon, 1992; Wang, 1999).

The chapter is set up as follows. In section 2.2, we briefly position the chapter. We lay out the model of the economy in section 2.3, where we treat government policy as given. This sketches how the economy responds to changes in government policy, and in terms of game theory, it shows how strategy pairs pay off. In section 2.4, we endogenize government policy, and study the solutions to the strategic situation that arises in this economy. Section 2.5 concludes.

2.2 Local policy competition: policy effects and interaction

The seminal article of Baldwin and Krugman (2004) examines international competition in tax rates when there are agglomeration rents. The crucial insight taken from the new economic geography models is that real factor rewards are higher in the core, and therefore the mobile factor can be taxed to some extent without the risk of it leaving the region. Baldwin and Krugman assume a first-mover (Stackelberg) advantage for the large, core region and argue that the core's government exploits its first-mover advantage and the real reward difference between the regions to set a higher tax rate than the periphery's government does. The tax rate is so low, however, that it deters the peripheral government from effectively undercutting the core's tax rate and "stealing" the core – i.e. the core sets a limit tax. Baldwin and Krugman "conjecture that [their] results hold in a broad range of models" next to the footloose entrepreneur model they use. A number of modifications has indeed confirmed their intuition. For instance, models of incomplete agglomeration (Borck and Pflüger, 2006), two-factor models (Kind et al., 2000) and models of welfare-maximizing instead of Leviathan governments (Ludema and Wooton, 2000) show that economically large regions exploit the benefits associated with co-location.

The results of the model of Baldwin and Krugman change when governments do not consume tax revenue, but spend their budget in the economy. When taking into account the goals of levying taxes, taxation does not necessarily distort the mobile factors' decisions adversely. For instance, taxes may finance local public goods that increase the attractiveness of a location, or local public inputs that improve the local productivity (Keen and Marchand, 1997). In particular, Brakman et al. (2002) show that increasing returns in public goods production or productivity-enhancing public investment may foster agglomeration. Likewise, Commendatore et al. (2008) show that productive public expenditure may attract more firms, but a tax on local labor reduces demand, making the outcome of government policy ambiguous on balance. The observation that government expenditure has spatial implications is in line with the general equilibrium nature of new economic geography models. In this chapter's model, given the high level of goods market integration, the impact of government policy on local industry is felt in nearby regions as the number of imported varieties grows when nearby governments subsidize firms. As we show later in this chapter, such policy spillovers may affect policy formation.

More importantly for the present purpose, apart from market interactions, regions interact in policy-setting. In the model of Baldwin and Krugman, the larger region has a first-mover advantage in addition to the advantage stemming from its size. The timing allows the larger region to credibly select a limit tax, and to discourage the smaller region from setting a competitive tax. Ludema and Wooton (2000) use a parallel requirement, namely that in a stable equilibrium, the core re-emerges as the core after the policies have been set. The sequential structure of the game avoids the problem that there are no simultaneous best responses in pure strategies when firms tend to agglomerate and the world is "lumpy". Often, when smaller regions compete for the agglomeration, the larger region's optimal response is to compete as well. In that case, the smaller region's best response is not to compete for the majority of firms using low tax rates. The dynamic advantage of the large region simplifies the model considerably, because it generates the limit tax as a subgame perfect equilibrium.

When comparing the literature on policy competition with and without agglomeration externalities, it is worthwhile to isolate the agglomeration effect from the Stackelberg effect, because the first-mover advantage also changes conclusions if there are no agglomeration effects (Gordon, 1992; Wang, 1999). Kempf and Rota-Graziosi (2010) show that the sequential tax-setting results in lower taxes and reduces the inefficiency associated with tax competition. In their model of endogenous Stackelberg leadership, they also demonstrate that a less productive country is likely to be the leader in tax-setting, arguing that smaller countries, or those endowed with less capital fit the leading position. This is opposite to the timing advantages that agglomeration models of tax competition ascribe to large regions.

Sequential and simultaneous policy setting describe different institutional contexts, but they also have different game theoretical characteristics. A less appealing property of simultaneous mixed strategies is that their interpretation is less straightforward, compared to pure strategies. The mixed-strategy profile is a description of the likelihood of strategies being played, not an observable single pure strategy. On the other hand, the mixed-strategy equilibrium is robust to some criticisms of the Stackelberg form of the policy game. If policy setting is a repeated process, the Stackelberg follower may adopt rational punishing strategies that compromise the one-shot Stackelberg equilibrium (Cruz, 1975; Aoyagi, 1996). Since the mixed strategy equilibrium is a simultaneous subgame-perfect Nash equilibrium, its solution is robust to repetition. Similarly, the mixed-strategy equilibrium does not suffer from the sequential game's sensitivity to multiple players (Fudenberg and Tirole, 1991, p. 98). Finally, and most important in the context of the model, the mixed-strategy equilibrium avoids one conceptual step. In the mixed-strategy equilibrium, the asymmetric position of smaller and larger regions ensures that the larger region has an incentive to set higher subsidies. A region's propensity to select competitive strategies is a function of the number of firms in the region at the start of the game. This is in contrast to the sequential structure, where the number of firms entitles the largest region to a first-mover advantage, which in turn drives the outcome of the game. Put differently, if the first-mover advantage was given to the smaller region, the outcomes of the sequential game would no longer be clear-cut.

2.3 Model setup

The economy underlying the policy competition game in this chapter relies on agglomeration forces in the New Economic Geography tradition. In particular, we use a "footloose entrepreneur" model (Forslid and Ottaviano, 2003) with vertical linkages. It assumes that the benefits of proximity to suppliers of inputs can form a centripetal force for firms to cluster. Although earlier tax competition models were based on a marginally different model (e.g. footloose entrepreneurs without vertical linkages in Baldwin and Krugman), its agglomerative forces and welfare conclusions are qualitatively the same (Ottaviano and Robert-Nicoud, 2006). However, an essential difference is that under vertical linkages, taxes and subsidies have an effect on firm entry and exit, so the size of the firm population is endogenous. This is much in line with the Dixit-Stiglitz model, which was originally designed to study whether firm diversity is optimal. Endogenous entry plays an important role as it implies that policies have goods market spillovers: benefits of policy that increase firm entry are felt in other regions. These spillovers are absent in regular "footloose entrepreneur" models, where the number of firms is fixed (equal to the number of entrepreneurs). The other alternative model, the basic core-periphery (CP) model (Fujita et al., 2001) relies on labor mobility, which runs counter to the idea that workers try to attract mobile firms, and that undertaxation is a concern. Moreover, the footloose entrepreneneur with vertical linkages model shares the analytical closed-form solutions that enabled the original footloose entrepreneneur (without vertical linkages) to yield a more comprehensive discussion of tax competition than the CP model, which would rely on numerical simulations.

The role of the government is to provide subsidies and provide public services. As argued before, it may be desirable to incorporate the effects of policies on the real economy into the model. To stress that role of the government, we shall assume that the government makes an expenditure decision: they spend their budget either on public services, or on subsidies that can attract firms. This expenditure decision is a more accurate description for local governments that have less discretion in taxation but do have substantial industrial policies. Due to our assumption on public goods, public goods are always in positive demand and positive government budgets are always required for welfare-maximizing governments. If there was no central source of financing, public goods would be financed by negative subsidies, or effectively, a firm-level tax. For governments, the opportunity costs of lowering taxes or raising subsidies is, equally, determined by the marginal utility derived from public services. Therefore, whether governments finance public services from a business tax, or have a fixed budget to allocate on business subsidies or public services should not change the intuition of our results.

We first describe the basic structure of the economy. There are three types of actors in the economy: consumers/workers, firms, and governments. There are two sectors in the economy. The agricultural sector produces under constant returns to scale, employing only labor; this will be the numéraire sector. The manufacturing sector acquires a fixed factor that is constructed from the output of other firms. The price of the fixed factor poses a fixed cost, ensuring increasing returns to scale in production. Workers are mobile between sectors, so wage rates equalize between the sectors. The agricultural good is traded freely, equalizing the agricultural wage in both locations, and indirectly, manufacturing wages. This equalization only occurs if agriculture is produced in both regions. We provide conditions for incomplete specialization into manufacturing in either region in Appendix 2.A. Workers consume some of the manufacturing and agricultural good, and a public good. The local government takes its budget as given, and spends it on the provision of public goods or on subsidies to firms in its region. We assume that the government subsidizes on a per-firm basis, independent of the firm's production size. We will denote the two regions as region 1 and region 2, and aggregate variables with a superscript w ("world"). To define the distribution parameters, we choose λ, v and b the denote the region 1's share in aggregate population (N), number of firms (n^w) and expenditure (E^w), respectively.

Consumers derive utility from three items: the consumption of the two types of private goods (C_a and C_m) and the consumption of government goods (G). We use a generic homogeneous government good, that can be thought of as publicly provided services, ranging from playgrounds and green areas to police services and local infrastructure. The utility function consists of two tiers. The first tier is a Cobb-Douglas function with agricultural goods, aggregate manufacturing goods and the public goods as arguments. The (lower) second tier refers to a range of goods for which the consumer has a taste of variety modeled using a Dixit-Stiglitz setup:

$$U = C_a^{1-\alpha} C_m^{\alpha} G^{\gamma};\ C_m = \left[\int_0^n c\,(i)^{(\sigma-1)/\sigma}\, di \right]^{\sigma/(\sigma-1)}. \tag{2.1}$$

In this utility function, i is an index referring to the variety of the manufacturing good, and $\sigma > 1$ measures the elasticity of substitution between varieties. The term $0 < \alpha < 1$ measures the preference for manufacturing goods in private consumption, and the term $0 < \gamma < 1$ governs preferences for government-provided goods. Potential firm profits are equally distributed over the inhabitants of the region in which they operate. The central government finances local government using a wage tax, T. The budget is hence the net wage plus the per capita firm profit, and it may be spent on the agricultural good or on the manufacturing good. The public good does not enter the private budget constraint, since it is not priced. In a later stage, we endogenize the consumption of G in the government's decisions. Using Y as the per capita disposable income, the budget constraint reads:

$$C_a + \int_0^n c\,(i)\, p\,(i)\, di = Y = (1-T)\, w + \Pi/N. \tag{2.2}$$

The demand for manufacturing goods from the first stage of optimization (i.e. maximizing the top tier of the utility function) gives that:

$$\int_0^n c(i)\,p(i)\,di = \alpha Y.$$

This constrains the budget for manufacturing goods, the utility of which is optimized by maximizing C_m. This optimization yields the demand function for manufacturing varieties that is the standard solution to the Dixit-Stiglitz setup:

$$c(i) = p(i)^{-\sigma}\,\alpha Y P^{\sigma-1},\; P = \left[\int_0^n p(i)^{1-\sigma}\,di\right]^{1/(1-\sigma)}, \tag{2.3}$$

where P denotes the (CES) harmonized price index. This term serves as an aggregate price index for manufacturing goods, taking into account the elasticity of substitution between each individual good. The manufacturing demand function intuitively states that demand for a manufacturing good decreases in its own price, but increases in the budget and the price of other manufacturing goods. The consumption of public goods, ceteris paribus, does not change demand for manufacturing goods because the utility function is unit-elastic. In the two region case, we assume iceberg transport cost, such that τ units need to be shipped for one unit to arrive in the other region. Due to the wedge of transport cost between the factory gate price and delivery price, the consumer's demand for a good is lower if the producer is more distant, given the factory gate price. The demand function for the agricultural good is derived from the maximization of the utility function (eq. 2.1) subject to the private budget constraint (eq. 2.2). It implies that share $1-\mu$ of the consumer's income is spent on agricultural products: $C_a = (1-\mu)Y$. The indirect utility function is (an affine transformation of) $V = G^{\gamma}Y/P^{\alpha}$. This shows that the consumer cares about two items: the level of public goods provided and his real income. Since the wage is a numeraire, the indirect utility function shows that real wage is higher in large markets, because less trade costs are incurred for consumption.

Firms have variable and fixed costs of production, giving rise to increasing returns to scale. The variable part of production uses labor at inverse productivity a_m. The fixed costs arise because the firm needs to buy inputs from other firms as a fixed requirement F in production. Following convention (Baldwin et al., 2003, chapter 8), the fixed factor is assembled under a Cobb Douglas technology, nesting agriculture and a constant elasticity of substitution function of manufacturing varieties with the same form and parameters as the consumer's utility function.[1] Using this technology, the firm's cost function is:

$$TC = a_m q(i)w + FP^{\alpha}, \tag{2.4}$$

where the cost of the fixed factor increases in the number of varieties that need to be imported. Aggregating the demand functions from both regions, a firm from region 1 faces the demand curve (a parallel curve exist for a firm in region 2):

$$q(i) = p\left(i\right)^{-\sigma} \alpha \left[\frac{b}{P_1^{1-\sigma}} + \phi \frac{1-b}{P_2^{1-\sigma}} \right] E^w, \tag{2.5}$$

where $\phi = \tau^{1-\sigma}$ is the freeness of trade. E^w is the aggregate expenditure on agriculture and manufacturing goods, i.e. the sum of resident's income and firms' demand for inputs from both locations. A fraction b of the expenditure stems from region 1, and the complement from region 2. Maximizing firm profits with respect to the price, using the demand curve (2.5) and cost function (2.4) under the assumption that the number of firms is large, results in markup pricing, as is standard in this model of monopolistic competition (cf. Dixit and Stiglitz, 1977):

$$p(i) = \sigma/(\sigma - 1)a_m w. \tag{2.6}$$

Following convention, we normalize the markup to the inverse productivity. Using the pricing equation and the normalization, the manufacturing price indexes for either region (eq. 2.3) raised to the power $1-\sigma$ can be written as:

$$P_1^{1-\sigma} = \left(s_n + \phi\left(1 - s_n\right)\right) n^w, P_2^{1-\sigma} = \left(\phi s_n + 1 - s_n\right) n^w, \tag{2.7}$$

where $s_n n^w = n_1$ is the share of the global number of firms that is located in region 1. With free entry and exit, the firms' profits are driven to zero. The fixed markup over marginal cost implies that a constant fraction $(1/\sigma)$ of

[1] The minimization of assembly cost is thus dual to the optimization of consumer utility: min $\int q\left(i\right) p\left(i\right) di$ s.t. $F = \left(\int q\left(i\right)^{(\sigma-1)/\sigma} di\right)^{\sigma/(\sigma-1)}$. The price of the fixed factor F is therefore equal to the aggregate consumption price index.

revenue are operating profits. In equilibrium, these cover the fixed cost minus a possible firm-level subsidy:[2]

$$\Pi = p(i)q(i)/\sigma - FP^{-\alpha} + S. \tag{2.8}$$

Finally, the local governments are the key players in our model. The local government has no discretion in the taxation of labor, which is done centrally. The central government redistributes the labor tax evenly to local governments. Therefore, the government faces a fixed budget. The policy instrument is then expenditure composition: the allocation of the government budget is a trade-off between providing public goods and subsidizing firms. The public good enters the utility function directly, while subsidies potentially attract firms, increasing the local real income. The strategic instrument of expenditure places this chapter in a tradition of industrial subsidies (e.g., Janeba, 1998) or local public inputs (Keen and Marchand, 1997; Fenge et al., 2009). However, the transfer is purely financial and not productive; therefore, increased subsidies act like a lowering of tax rates.

The local government is benevolent, and maximizes the average utility of its inhabitants. In the presence of inequalities, average utility may be a poor proxy of welfare, but since all citizens of a region face identical wages and prices, no distributional assumptions are required. The maximization problem for the local government is:

$$\max_{s} U \text{ s.t. } G + nS \leq TwN, \tag{2.9}$$

where TwN is the government budget, $S \geq 0$ is the subsidy per firm so nS reflects the total subsidy handed out. There is no possibility to set individual subsidies per firm – the policy cannot discriminate. When profits (eq. 2.8) are driven to zero, the equilibrium firm size q is smaller when the subsidy is higher. Intuitively, the operating profits, and so the firm size needed to cover the fixed cost, is lowered by the level of subsidy. The subsidy hence expands the equilibrium number of firms and decreases their

[2]As can be seen from equation 2.6, the firm-level subsidy affects only the number of firms, not their price level. A subsidy on production could be entered into the profit function as $\Pi = \big(p(i) + s\big)q(i) - a_m wq(i) - FP^{\mu}$. In that case, the labor requirement developed in the next section would become $l(i) = \frac{(\sigma-1)f}{1+s/w} + f$. A subsidy on production also reduces equilibrium labor requirements and, ceteris paribus, increases firm variety. However, the subsidy on production distorts the pricing decision and therefore eliminates closed-form solutions, which is the virtue of the "Footloose Entrepreneur with Vertical Linkages" FEVL model.

size. In a closed economy, the welfare-maximizing subsidy is generally not equal to zero, since the subsidy affects the inefficiency due to monopolistic firm behavior. The subsidy does not adress the monopolistic (markup) price distortion directly. Since the subsidy is independent of quantity, it only accrues to the fixed factor, the size of which determines the number of firms. In other words, using this subsidy, the government specifically targets growth of the number of manufacturing firms in its region. Finally, to simplify notation, we rewrite the government spending decision in a fraction of its budget. The governments spend share s of their budget (TwN) on subsidizing, so the per-firm subsidy S and the public good provision per head becomes:

$$\begin{aligned} S_1 &\equiv s_1 \frac{T\lambda}{s_n}\frac{N^w}{n^w},\ S_2 \equiv s_2 \frac{T(1-\lambda)}{1-s_n}\frac{N^w}{n^w}, \\ G &= (1-s)\,Tw/\lambda, \end{aligned} \tag{2.10}$$

where λ denotes the share of world population in region 1. We will use the budget share s rather than the absolute level of subsidy S as the policy instrument, but since they are directly related, the analysis yields the same results.

Equilibrium in the private sector

Government behavior is the key interest of this chapter. To study the strategic interaction between governments, we first define the spatial equilibrium outcomes as a function of the governments' subsidies, treating such policies as exogenous. Once we know the result of different policies, it is possible to investigate the strategies that governments use to steer outcomes. In equilibrium, the firm distribution is such that profits are non-positive in both locations. Since profits are a function of expenditure shares and of the firm distribution over regions, we proceed by expressing expenditure shares (b) in terms of firm distribution (s_n) to solve the profit equation as a function of the firm distribution s_n. This follows the same reasoning as the standard Footloose Entrepreneur Vertical Linkages model (Baldwin et al., 2003, section 8.4). By simple accounting, the expenditure originating in one region is the sum of expenditure from inhabitants and local firms buying intermediate inputs. By virtue of the zero-profit condition, a firm's expenditure on intermediates is equal to the operating profits plus the firm subsidy, so the total expenditure originating from a region is $E = (1-T)N + n\big(pq/\sigma + S\big)$, which is the sum of consumer and firm expenditure. The aggregate expenditure is obtained by adding the expenditure of the two regions:

$$E^w = (1-T)N^w + n^w p/\sigma \left(s_n q_1 + \left(1-s_n\right) q_2\right) + n^w \left(s_n S_1 + \left(1-s_n\right) S_2\right).$$

Filling out the demand function for both regions (eq. 2.5) and the expression for subsidies (eq. 2.10), the aggregate expenditure simplifies to:

$$E^w = \frac{1-T\left(1-\lambda s_1-(1-\lambda)s_2\right)}{1-\alpha/\sigma} N^w, \tag{2.11}$$

where $\lambda s_1 + (1-\lambda)s_2$ can be thought of as the size-weighted average subsidy share in the aggregate government budget. By dividing region 1's expenditure ($E_1 = (1-T)N_1 + n_1\left(pq_1/\sigma + S_1\right)$) over aggregate expenditure (eq. 2.11), the share of expenditure of region 1 can be written as:

$$\begin{aligned} b &= (1-\alpha/\sigma)\delta + \frac{\alpha}{\sigma}\left[\frac{b}{P_1^{1-\sigma}} + \phi\frac{1-b}{P_2^{1-\sigma}}\right]s_n, \qquad (2.12) \\ \delta &\equiv \frac{1-T\left(1-s_1\right)}{1-T\left(1-\lambda s_1-(1-\lambda)s_2\right)}\lambda, \end{aligned}$$

where δ reflects the policy-adjusted share of expenditure stemming from the population in region 1. The second term in the expression for b reflects that operating profits are used to buy inputs for the fixed factor. Solving for b gives the market size equilibrium as a relation between region 1's share of expenditure and region 1's share of firms 2:[3]

$$b = \delta + \frac{\phi\alpha/\sigma\left[(1-2\delta)\left(1-\phi\right)s_n\left(1-s_n\right)+s_n-\delta\right]}{\left(s_n+\phi\left(1-s_n\right)\right)\left(\phi s_n+1-s_n\right)-\alpha/\sigma\left(1-\phi^2\right)s_n\left(1-s_n\right)}. \tag{2.13}$$

In equilibrium, the share of expenditure stemming from region 1 comprises a consumer budget and firm subsidy share (the first term, δ), and firm level expenditure (the second term; the fraction in equation 2.13).

[3]Equation (2.12) differs from the standard result in the FEVL model in Ottaviano and Robert-Nicoud (2006, eq. 28). However, under symmetry in regional size and policy, the market size equilibrium condition reduces to the same equation. In that sense, the standard market size condition is a special case of this market size condition allowing for policy.

Expenditure in region 1 increases in the share of firms in region 1, because local firms prefer to buy local inputs.

Finally, the market size equilibrium condition (eq. 2.13) can be used to express the profit equation (eq. 2.8) as a function of the firms distribution s_n, eliminating expenditure b. Additionally substituting the expression of the price index (eq. 2.7) into the profit function yields the profit function as an explicit function of the distribution of firms, s_n. The complete profit function is not very insightful, so we relegate it to Appendix 2.B.1, together with its derivation. Moreover, we generally use the "core-periphery" solution, because the two stable equilibria under low transport costs involve concentration of the mobile factor. If s_n is equal to zero or 1, the profit function simplifies substantially. The three components of the profit function that remain (see eq. 2.B.1) are familiar from other NEG literature: the local presence of firms generally increases the operating profits due to a home market effect, and, second, decreases the price of the fixed factor. Third, higher per-firm subsidies increase the profits, all other things constant. Given these profit functions, the spatial equilibrium is defined as a distribution of firms (given by s_n and n^w) for which pure profits are non-negative ($\Pi_1 \leq 0$ and $\Pi_2 \leq 0$). Using these two no-positive-profit conditions, we can investigate what effect different subsidy pairs have on the spatial distribution of firms in the economy.

2.4 Strategic expenditure

By treating the level of subsidies as given in the private equilibrium, we can describe equilibria for different policy combinations. In this section, the two governments can manipulate s_1 and s_2 in pursuit of maximizing their objective function. In this model, as firms tend to cluster, multiple spatial equilibria can emerge. We shall assume that region 1 is initially the large region (hosts the majority of firms), but since the regions are symmetric in all other respects, the results may be obtained for region 2 by switching the labels.

Policy options: subsidies that relocate the agglomeration

For the equilibrium with all firms in region 1 to be stable, firm profits must be zero in region 1, and it must be unprofitable to set up a firm in region 2: $s_n = 1$, $\Pi_1 = 0$ and $\Pi_2 < 0$. Using these conditions, the number of firms (the final endogenous variable in the spatial equilibrium) under full agglomeration can be formulated. Using $\Pi_1 = 0$ and $s_n = 1$, rewriting the profit function for firms in region 1 for the number of firms gives:

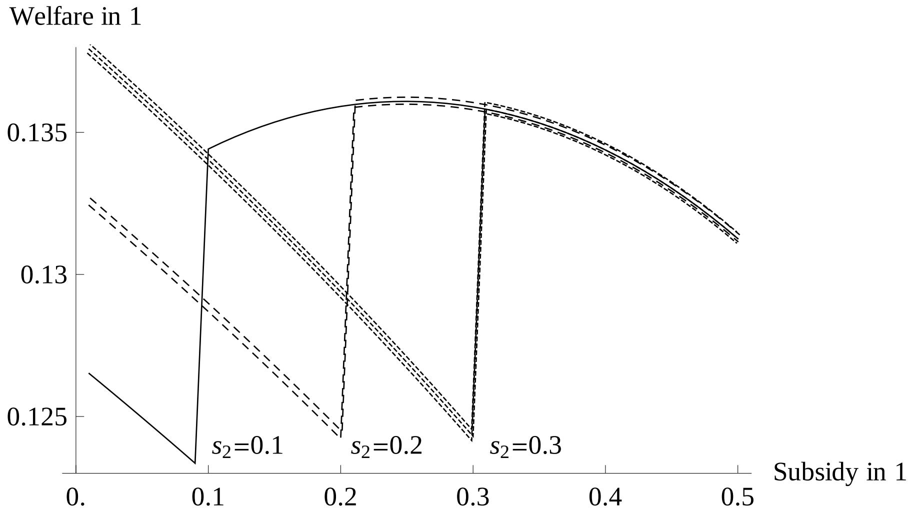
$$n^w = \left(\frac{\alpha}{\sigma - \alpha} \left(1 - T - \frac{\sigma}{\alpha} T \lambda s_1 + T (1 - \lambda) s_2 \right) \right)^{(1-\sigma)/(1-\sigma+\alpha)} . \quad (2.14)$$

This number of firms follows from the free entry of firms: with fewer firms, profits are positive, and there would be incentive to set up new firms. The expression for the number of firms permits studying the break subsidy: the minimum level of subsidy in region 2 for which a firm can profitably operate in region 2 given the initial agglomeration in region 1. Under positive scale externalities, if a first firm can operate profitably in region 2, all firms will be able to operate profitably in region 2. If a first firm moves, profits rise in region 2 and fall in region 1, so there is an incentive for all firms to move. Therefore, the subsidy budget of region 2, if successful, is not distributed over one firm, but over all firms. The subsidy per firm in region 2 (equation 2.10) then becomes $s_2 T (1 - \lambda) N^w / n^w$.

Given the number of firms and s_1, there is a subsidy s_2 in region 2 for which it becomes profitable for firms to locate in region 2: $\Pi_2 \geq 0$. There is no general closed-form solution for that break subsidy, because the expenditure share b (equation 2.13) is related non-linearly to the distribution of firms. However, since home market effects in expenditure and the vertical cost linkages cause positive externalities, the subsidy share in region 2 always needs to compensate these two effects in addition to region 1's subsidies. Hence, due to agglomeration externalities, the subsidy that region 2 needs to set to provide an incentive for the (re)location of firms is larger than the subsidy of region 1.Figure 2.1 illustrates the government considerations regarding the break subsidy. Setting a subsidy lower than the break subsidy (i.e., to the left of the discontinuities, $s_1 = 0.097$ in case $s_2 = 0.1$) will result in no firms in the region. This implies goods need to be imported. Increasing subsidies without attracting the core comes at the cost of public goods, so left of the break subsidy, welfare decreases in the subsidy. Setting a subsidy higher than the break subsidy will bring the agglomeration into the region. In that case welfare is hump-shaped: some subsidies are desirable to promote firm entry, but the opportunity costs (marginal benefit derived from public services) increase if subsidies are larger.

If the opponent subsidy rates (s_2 in this case) increase, two things happen. First, setting low subsidies and becoming a periphery becomes more

Figure 2.1: Welfare payoffs for different subsidies

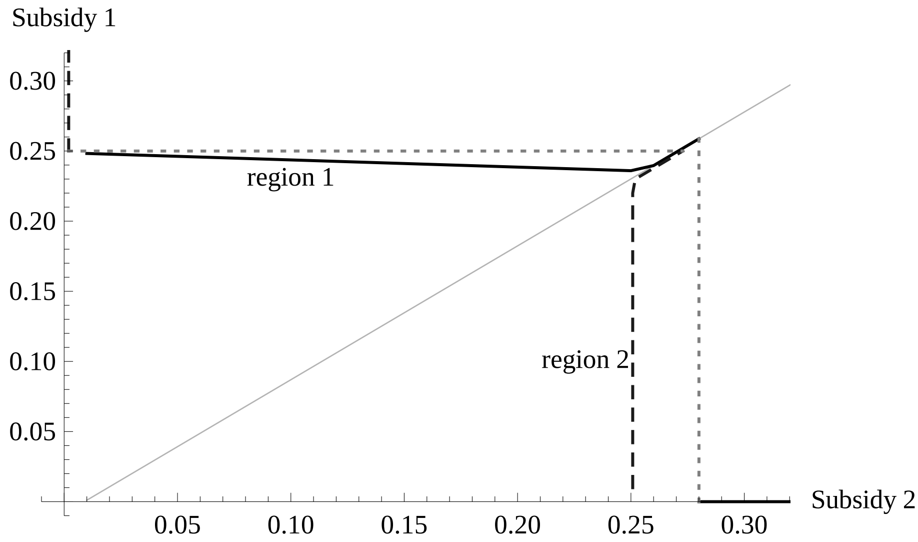


Note: Welfare payoffs in region 1 for three different subsidy shares in region 2. Single, solid: $s_2 = 0.1$; double, dashed: $s_2 = 0.2$; triple, dotted: $s_2 = 0.3$.

attractive, because the opponent subsidizes firm entry. Second, the minimum required subsidy to relocate the core – the break subsidy – rises, increasing the opportunity cost in public services, should the break subsidy be selected. In this case, if $s_2 = 0.1$, selecting a large subsidy and having the agglomeration clearly yields the highest welfare. The subsidy s_1 would be chosen to arrive at the top of the welfare curve, at about $s_1 = 0.26$. For later reference, we call this local maximum in the welfare function conditional on hosting the agglomeration s_{lo} (the "local optimal" subsidy) – it is the subsidy that the government of region 1 would set in absence of the possibility of losing the agglomeration. Clearly, if $s_2 = 0.3$, it is optimal for region 1 to set a zero subsidy: accepting to be the periphery and provide public services yields higher welfare than spending a large budget on attracting firms.

Figure 2.2 summarizes the government's strategies in best response curves. If region 2's subsidy is low (i.e., less 0.25), region 1 (solid black line) sets the local optimal subsidy, keeping the agglomeration. Conversely, if region 2's subsidy is high (i.e., exceeding 0.28), region 1 opts for a zero subsidy, providing more public goods and allowing firms to move to region 2. The same considerations shape the profile of region 2's best response set when facing very low and very high subsidies.

Figure 2.2: Reaction curves for region 1 and region 2



Note: Reaction curves for region 1 (solid) and region 2 (dashed), and the break subsidies (grey, see text). $\mu = 0.5$, $\sigma = 5$, $t = 0.3$, $s_l = 0.5$, $\phi = 0.6$.

At intermediate subsidies (between 0.25 and 0.28 for region 2), both regions play a break subsidy, marginally "out-subsidizing" the other. Even though they cannot set the local optimal subsidy (because the break subsidy is larger), it still yields higher welfare to set the break subsidy and attract the agglomeration: it avoids importing goods, even if the costs in terms of public services are higher. Note that these lines seem to coincide graphically at the break subsidy line (the grey line), but it is optimal to subsidize marginally more than the break subsidy. Compared to region 2, region 1 plays break subsidies against higher opponent subsidies. This reflects that for region 2 the opportunity cost of attracting firms is higher, since it needs to compensate the agglomeration externality in addition to the subsidy in region 1.

The game sketched above does not have a simultaneous Nash equilibrium in pure strategies. The best response to local optimal subsidies are high subsidies that take over the agglomeration. The optimal response to such high subsidies are low subsidies. However, against low subsidies, local optimal subsidies are optimal. Alternatively, it can be seen graphically that the best response functions do not coincide. Since the game has no coincidence of best responses, it needs another solution than simultaneous pure strategies. One strategy is to introduce a (Stackelberg) first-mover advantage that yields a limit-subsidy, the other is to allow for mixed responses, which we investigate in turn.

Limit subsidy

The limit-price solution of this game, following Baldwin and Krugman (2004), also holds in this model. In this version of the game, region 1 selects a subsidy first, then region 2 selects a subsidy, and finally the spatial distribution of firms materializes, yielding the payoffs. As Baldwin and Krugman show, with positive agglomeration externalities, the subgame perfect solution is that region 1 sets a subsidy such that it is optimal for region 2 to provide public goods only. Effectively, region 1 (a Stackelberg leader) acts as a limit pricer.

The indirect utility $V = G^{\gamma} Y / P^{\alpha}$, using the public good definition (2.10) and price index (2.7), is proportional to:

$$n^{w\alpha/(\sigma-1)} \left(s_n + \phi\left(1-s_n\right)\right)^{\alpha/(\sigma-1)} (1-s)^{\gamma}.$$

Under that expression, region 2 has no incentive to set competing subsidies if the welfare costs of competing are sufficiently high:

$$\left(\phi n_1^w\right)^{\alpha/(\sigma-1)} \geq n_2^{w\alpha/(\sigma-1)} \left(1-s_2\right)^{\gamma}, \tag{2.15}$$

where n_1^w is the aggregate number of firms if the agglomeration is in region 1 and n_2^w is the aggregate number of firms under a potential agglomeration in region 2 given region 2's subsidy. The left-hand side of inequality (2.15) reflects region 2's welfare after setting no subsidy, where utility from the variety of firms n_1^w is discounted by trade costs ($\tau^{1-\sigma} = \phi < 1$), because all manufacturing goods need to be imported. The right-hand side equals welfare after setting a break subsidy, so no manufacturing goods are imported, but only a share $1-s_2$ of budget can be spent on the public good. Rewriting the condition gives:

$$\left(\frac{\phi n_1^w}{n_2^w}\right)^{\alpha/(\sigma-1)} \geq \left(1-s_2\right)^{\gamma}. \tag{2.16}$$

Since the number of firms n_1^w rises faster in s_1 than in s_2 (see eq. 2.14), a higher subsidy of region 1 makes it more likely that this condition is met. If trade costs rise ($\phi = \tau^{1-\sigma}$ is an inverse measure of trade costs), so does the minimal subsidy for region 1 to deter region 2 from competing, because hosting the agglomeration becomes more attractive. Thus, the two major conclusions from Baldwin and Krugman (2004) re-emerge: first, the core (region 1) sets a positive subsidy, to which the periphery (region 2) responds with a zero subsidy. Second, as trade integration increases, the limit subsidy falls, reducing the subsidy gap.

Mixed policy responses

To understand how agglomeration affects the strategic incentives in policy competition games, ideally, simultaneous pure strategy equilibria with and without agglomeration effects should be compared. An alternative to addressing the absence of such an equilibrium is to drop the pure strategy assumption, instead of the simultaneity assumption. A mixed strategy equilibrium retains simultaneity, and thus helps to dissect the agglomeration advantages from the first-mover advantages. The main justification of the mixed strategy profile is therefore that it lists the strategic incentives if the core cannot credibly commit to its policy beforehand.

The interpretation of a mixed strategy equilibrium, however, is clearly more complicated than an equilibrium of pure strategies. The mixed strategy profile admits that more than one strategy can be an optimal response. In the sequential game, this possiblity is ruled out by construction: the limit tax ensures that no subsidies are always a best response for the smaller region. The periphery is presented with a *fait accompli*, and thus acts lethargically. The fact that this does not occur in the simultaneous case points to a strategic motive for the periphery that was assumed away in the sequential game: in mixed strategies, it is optimal for the periphery to sometimes act competitively, even if the chances of succes are low. Intuitively, a relocation of the agglomeration is not likely to occur because, as we shall see, the core is more likely to set higher subsidies. However, competitive subsidies have a strategic value for the periphery as a disciplining device toward the core. Higher core subsidies benefit peripheral consumers but this is not taken into account by the core government. If the core government faces the possibility of a competitive subsidy from the periphery, it will be forced to select higher subsidies more often, which increases firm entry and by that, welfare in the periphery. Therefore, the simultaneous game sheds light on the disciplining device as an additional element in strategic interaction, which is ignored in the sequential game.

Table 2.1 presents the payoff matrix for a 2-by-2 strategy game, where strategies from the two potential best response classes are used. The strategies are either competitive or non-competitive. The first is a strategic effort to locate the agglomeration in the government's region (s_c), while the second is an optimal response if firms could not move (the local optimal subsidy s_o in region 1 and no subsidies in region 2). Clearly, selecting two strategies limits the completeness of the strategy profile (break subsidies are optimal over an interval), but it preserves the strategic considerations. The 2-by-2 case yields analytical results; a numerical investigation into these results are presented later this section. To construct the payoff matrix, we use that the payoff (affine transformations of the indirect utility functions, discussed in section 2.4) can be written as $(1-s)^\gamma \, n^{w\alpha/(\sigma-1)}$ for the region that hosts the agglomeration, and $(1-s)^\gamma \left(\phi n^w\right)^{\alpha/(\sigma-1)}$ for the periphery, where the trade openness ϕ reflects that welfare is lower if consumption is imported against high trade costs. In the payoff matrix in Table 2.1, best responses are underlined. The Table shows no coincidence of best responses.

Table 2.1: Payoff matrix for selected strategies

		region 2	
		0	s_c
region 1	s_{lo}	$\underline{n_{lo}^{\frac{\alpha}{\sigma-1}}\left(1-s_{lo}\right)^\gamma}, \tau^{-\alpha} n_{lo}^{\frac{\alpha}{\sigma-1}}$	$\tau^{-\alpha} n_c^{\frac{\alpha}{\sigma-1}}\left(1-s_{lo}\right)^\gamma, \underline{n_c^{\frac{\alpha}{\sigma-1}}\left(1-s_c\right)^\gamma}$
	s_c	$n_c^{\frac{\alpha}{\sigma-1}}\left(1-s_c\right)^\gamma, \underline{\tau^{-\alpha} n_c^{\frac{\alpha}{\sigma-1}}}$	$\underline{n_c^{\frac{\alpha}{\sigma-1}}\left(1-s_c\right)^\gamma}, \tau^{-\alpha} n_c^{\frac{\alpha}{\sigma-1}}\left(1-s_c\right)^\gamma$

Note: Subsidy levels: c for competitive, lo for local optimal.

In an equilibrium of mixed strategies, governments choose a strategy profile such that the other government is indifferent between playing its own pure strategies. Intuitively, if one government was not indifferent, it would play its dominant strategy, leading the other government to select its best response, to which the first government's strategy is not a best response. Region 1 chooses its probability of playing a competitive strategy,

p, such that region 2 is indifferent between its pure strategies:

$$\begin{aligned} & \left(1-p\right)\tau^{-\alpha}n_{lo}^{\alpha/(\sigma-1)}+p\tau^{-\alpha}n_{c}^{\alpha/(\sigma-1)} \\ = & \left(\left(1-p\right)\tau^{-\alpha}+p\right)n_{c}^{\alpha/(\sigma-1)}\left(1-s_{c}\right)^{\gamma}, \end{aligned} \tag{2.17}$$

which are simply the weighted averages of playing 0 (left-hand side) and s_c (right-hand side). Vice versa, using q as the probability that region 2 plays a competitive strategy, indifference between pure strategies implies that:

$$\begin{aligned} & \left(1-q\right)n_{lo}^{\alpha/(\sigma-1)}\left(1-s_{lo}\right)^{\gamma}+q\tau^{-\alpha}n_{c}^{\alpha/(\sigma-1)}\left(1-s_{lo}\right)^{\gamma} \\ = & n_{c}^{\alpha/(\sigma-1)}\left(1-s_{c}\right)^{\gamma}. \end{aligned} \tag{2.18}$$

The two optimality conditions (indifference between strategies) for the mixed strategy have two unknowns, the probability that the core is competitive, p, and the probability that the periphery is competitive, q. Solving for p and q gives:

$$\begin{aligned} p &= \frac{1-\tau^{-\alpha}\left(1-s_{c}\right)^{\gamma}\left(n_{lo}/n_{c}\right)^{\alpha/(\sigma-1)}}{1-\tau^{-\alpha}\left(1+\left(1-s_{c}\right)^{\gamma}\left(\left(n_{c}/n_{lo}\right)^{\alpha/(\sigma-1)}-1\right)\right)}, \\ q &= \frac{1-\left(\frac{n_c}{n_{lo}}\right)^{\alpha/(\sigma-1)}\left(\frac{1-s_c}{1-s_{lo}}\right)^{\gamma}}{1-\tau^{-\alpha}\left(\frac{n_c}{n_{lo}}\right)^{\alpha/(\sigma-1)}\left(\frac{1-s_c}{1-s_{lo}}\right)^{\gamma}}. \end{aligned} \tag{2.19}$$

Given the best responses in Table 2.1, both governments have a positive probability of playing a competitive strategy. This is true in the explicit solution (2.19), since $\left(1-s_c\right)^{\gamma} > 1$, $n_c > n_{lo}$ and by definition of the local optimal subsidy and the competitive subsidy, $\left(1-s_c\right)^{\gamma} n_c^{\alpha/(\sigma-1)} < \left(1-s_{lo}\right)^{\gamma} n_{lo}^{\alpha/(\sigma-1)}$ (optimal subsidies that retain the agglomeration always yield higher local welfare than competitive subsidies).

Taking s_c as a limit subsidy (i.e., the highest competitive subsidy), the Stackelberg game would end up in $(s_c, 0)$: the lower-left quadrant of the Table. This is not an equilibrium in the simultaneous game, because region 1 prefers to deviate and set s_{lo}, so it has an incentive to set lower subsidies. However, if moves are simultaneous, the mixed stategy equilibrium lists an additional incentive for the periphery. If it can credibly threaten to set s_c, the core is forced to select a competitive subsidy s_c too. This implies that the number of firms is n_c instead of n_{lo}, which raises welfare in the periphery. Thus, the fact that there are goods market spillovers provides an incentive for the periphery to threaten with high subsidies, thereby internalizing some policy spillovers. Note that if the periphery was the Stackelberg follower, or subsidies had no effect on the economy, as earlier models assumed, the incentive toset a positive subsidy in the policy game would be eliminated for the periphery.

In comparison to the Stackelberg game, region 1 sets lower subsidies ($p < 1$), whereas region 2 sets higher subsidies ($q > 0$), so the policy gap is more narrow. By definition, the competitive subsidy in the 2-by-2 game is low enough to ensure that setting zero subsidies is not a dominant strategy for region 2. When the competitive subsidy s_c increases toward the limit subsidy, region 1 decreases p (see equation 2.19, where $n_{lo} \to n_c$), so it reduces the likelihood of playing a competitive subsidy. Therefore, the probably-weighted average subsidy in the mixed profile is always lower than the limit subsidy. Vice versa, with positive probability of region 1 playing a local optimal subsidy, there is incentive for region 2 to play a competitive subsidy with some probability.[4]

While region 1's subsidies are higher than region 2's on average due to agglomeration externalities, this formulation allows us to consider what the likelihood is that one of the local governments deviates from the local optimal subsidy. In particular, region 1 selects s_c more often than region 2 ($p > q$) if:

$$\phi^{\alpha/(\sigma-1)} < \frac{1+\left(\frac{n_c}{n_{lo}}\right)^{\alpha/(\sigma-1)}}{1-\left(1-s_c\right)^{\gamma}\left(\frac{n_c}{n_{lo}}\right)^{\alpha/(\sigma-1)}+\left(\frac{1-s_c}{1-s_{lo}}\right)^{\gamma}+\left(1-s_{lo}\right)^{\gamma}}. \quad (2.20)$$

Equation (2.20) shows that region 1 is more likely to play a competitive subsidy than region 2 when the trade cost are high (ϕ is low), because the value of hosting the agglomeration is larger. Likewise, the likelihood of region 1 selecting competitive strategies more often than region 2 increases in the ratio of the number of firms under a competitive subsidy compared to a local optimal subsidy (n_c/n_{lo}), see eq. (2.14). Given that $s_c < s_{lo}$, the government budget (T) unambiguously increases n_c/n_{lo}, so

[4]Region 1 playing more competitive strategies than 2 ($p > q$) can be written as $\frac{1-\phi\kappa_1}{1-\phi\kappa_2} > \frac{1-\phi\kappa_3}{1-\phi\kappa_4}$, with $\kappa_1 = \left(1-s_c\right)^{-\gamma}\left(n_l/n_c\right)^{\mu/(\sigma-1)}$, $\kappa_2 = \left(1-s_c\right)^{-\gamma}\left(n_l/n_c\right)^{\mu/(\sigma-1)} + \left(1-s_c\right)^{-\gamma}+1$, $\kappa_3 = \phi\left(\frac{1-s_c}{1-s_{lo}}\right)^{\gamma}\left(n_c/n_l\right)^{\mu/(\sigma-1)}$ and $\kappa_4 = \left(\frac{1-s_c}{1-s_{lo}}\right)^{\gamma}\left(n_c/n_l\right)^{\mu/(\sigma-1)}$. This is true if $\kappa_3 > \frac{\kappa_1-\kappa_2}{1-\kappa_2}$ and $\kappa_4 > \frac{\kappa_2\kappa_3+\kappa_2+\kappa_3-\kappa_1}{1-\kappa 1}$, which holds in eq. (2.20).

higher central taxation and budgets lead region 1 to be more competitive. The intuition is that with decreasing returns to public good consumption, a larger government budget decreases the opportunity cost of setting a competitive subsidy for both governments. Finally, region 1 is more likely to set most competitive subsidies if s_{lo} is higher, and s_c is lower, which reflects that the competitive subsidy is closer to the local optimal subsidy (since $s_{lo} < s_c$) and costs of competing are relatively low.

Finally, while the mixed game with two strategies is more insightful, assessing the game with all non-dominated strategies is more complete. Since we have no analytical expression for the payoffs following different policy-pairs, we approximate the strategy profiles numerically. We make the strategy space discrete by dividing the policy in segments. This is needed for the numerical solution, but it also guarantees the existence of a mixed strategy equilibrium (Fudenberg and Tirole, 1991, section 12.2, Dasgupta and Maskin, 1986). To approach the mixed strategy profile, we use the Lemke-Howson (see Lemke and Howson Jr., 1964, for an introduction to such algorithms) algorithm in Gambit software.

Figure 3 plots the strategy profile, i.e. the probability distribution over setting different subsidies. The density functions (panel a) are consistent with the two-strategy results: the average subsidy played in region 1 is higher than in region 2. For nearly all subsidies larger than zero, region 1 has a higher probability density. Region 1 never sets the lowest subsidy, but this is the subsidy region 2 is most likely to choose. This reflects that subsidies have a welfare effect only if there are firms in the region. Between the lowest (zero) subsidy and the highest subsidy in the support, the probability profile is generally convex. The intuition for this is that the payoff function is concave: if a government gets the agglomeration using a subsidy that balances public goods provision with subsidies on firm variety, its local welfare is highest. The mixed profile equates the opponent's expected payoff to all pure non-dominated strategies. Therefore, subsidies to which the opponent's best response yields a high payoff need to be played with relatively low probability. Changing the parameters leads to the changes predicted in the two-strategy case. Varying transport costs (panels b and c) shows that higher transport cost increase the probability weight assigned to higher subsidies. Likewise, higher subsidies are more likely to be played when changing the central tax rate from low (panel d) to high (panel e). The numerical solutions also confirm that higher transport costs between the regions and higher budgets increase the average share of the budget spent on subsidies. Finally, a higher preference for public goods increases the opportunity cost of subsidies, and so decreases the likelihood of setting high subsidies.

Harmonization and policy prescriptions

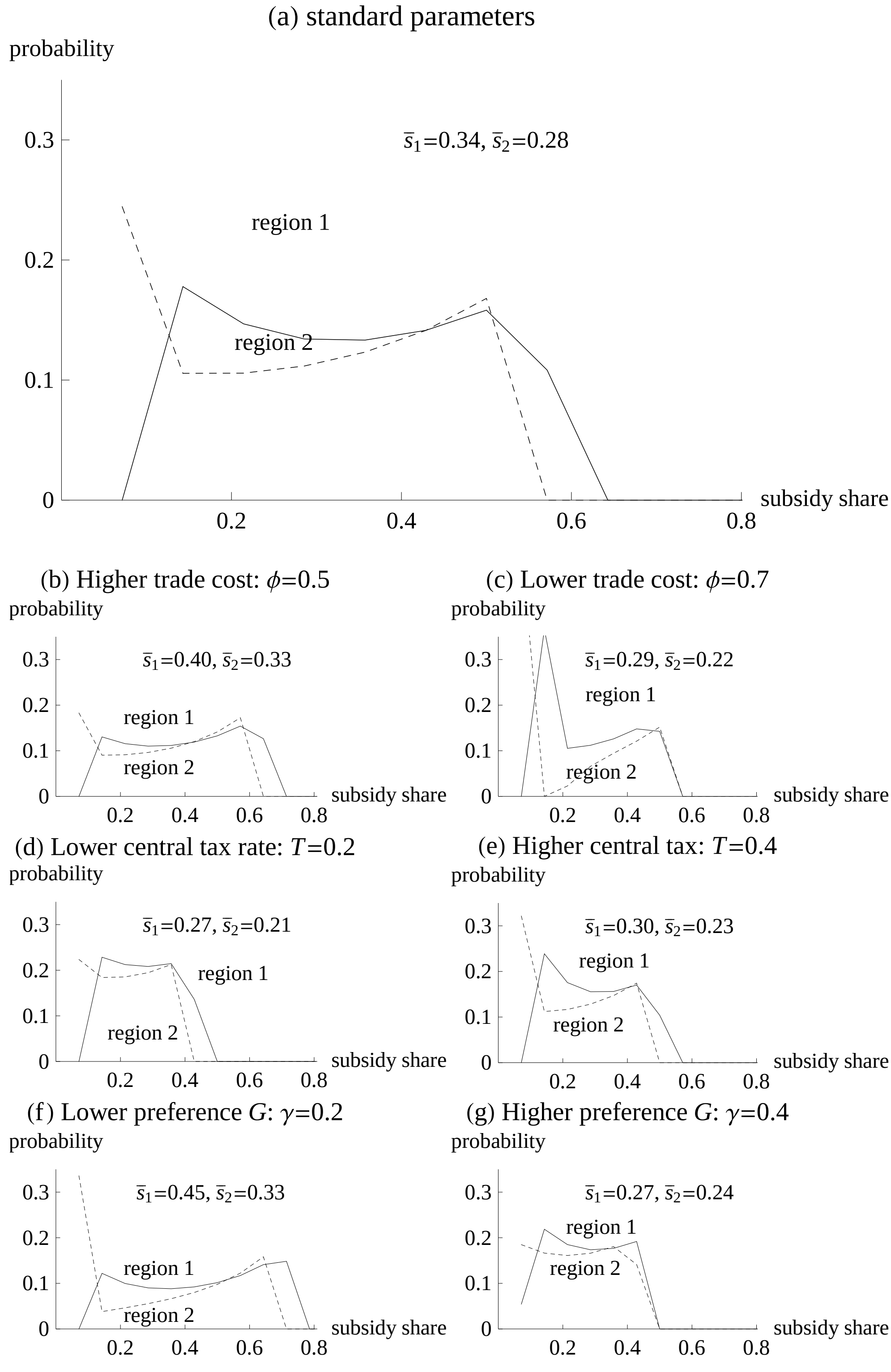
In sequential policy-setting, harmonization is never a Pareto improvement (see Baldwin and Krugman, 2004). Any subsidy between the core and periphery's Stackelberg subsidy reduces welfare in at least one: the periphery will prefer not to subsidize because it has no firms, the welfare of the core is not improved by setting no subsidies. The welfare conclusions under a Stackelberg game thus exactly mimic those of Baldwin and Krugman and related literature, which, is not surprising given the similar setup. Related literature (Baldwin and Krugman, 2004; Ludema and Wooton, 2000) concludes that instead of harmonization, a tax floor (subsidy cap in this model) is welfare improving. This does not translate to the current setup. The intuition is simple: limiting the subsidy range will induce the core to set lower subsidies and provide more public services. This translates into a smaller number of firms, which harms welfare in the periphery – the periphery is not neutral to a goods market spillover of the core's policies.[5]

In contrast to the sequential game, harmonization can be a Pareto improvement in the mixed strategy game. To develop this point, we use a harmonized subsidy $\bar{s}$. Harmonization is a Pareto improvement if both governments choose harmonization if faced with the choice between the mixed strategy game or both playing the harmonized subsidy. For region 1, it is easy to show that such a policy exists. We take advantage of the fact that in a mixed strategy equilibrium, a player's expected payoff is equal for every pure strategy. For region 1, the expected payoff of playing a competitive subsidy, and hence the expected payoff of the policy game is $n_c^{\alpha/(\sigma-1)}\left(1-s_c\right)^{\gamma}$. Since the competitive subsidy is higher than the local optimal subsidy (s_{lo}), a harmonized subsidy $\bar{s}$ between s_c and s_{lo} always improves the expected welfare of inhabitants in region 1.

For region 2, a harmonized subsidy is welfare-improving if its payoff is higher than the expected payoff of not subsidizing: $\bar{n}^{\alpha/(\sigma-1)}(1-\bar{s})^{\gamma} >$

[5] This assumes that consumers have of love of variety for manufacturing products. This occurs in the commonly used Dixit-Stiglitz model of monopolistic competition, but is not necessarily implied by that model, as Benassy (1996) shows.

Figure 2.3: Numerical mixed strategy Nash equilibria



Note: $\alpha = 0.2$, $\sigma = 5$, $N^w = 1$, $T = 0.3$, $s_l = 0.5$, $\gamma = 0.3$, $\phi = 0.6$. Lines denote the Nash equilibrium mixed strategy profile for region 1 (solid) and region 2 (dashed). Bars represent the average subsidy share.

$pn_{lo}^{\alpha/(\sigma-1)}+\left(1-p\right)n_c^{\alpha/(\sigma-1)}$ (the expected payoff of not subsidizing), where $\bar{n}$ is the number of firms resulting from policy $\bar{s}$. This requires:

$$p < \frac{1-(1-\bar{s})^{\gamma}\left(\bar{n}/n_c\right)^{\alpha/(\sigma-1)}}{1-\left(n_{lo}/n_c\right)^{\alpha/(\sigma-1)}}, \tag{2.21}$$

which shows that region 2 prefers harmonization if region 1 subsidizes little (p is the equilibrium probability of playing the higher, competitive strategy). Rewriting region 1's equilibrium probability of playing a competitive subsidy (eq. 2.19) gives:

$$p = \frac{1-\left(1-s_c\right)^{\gamma}}{1-\left(n_{lo}/n_c\right)^{\alpha/(\sigma-1)}+(\tau^{\alpha}-1)\left(1-s_c\right)^{\gamma}}. \tag{2.22}$$

Comparing p in the Nash equilibrium to the p required for region 2 to prefer harmonization (eq. 2.21) yields that region 2 prefers harmonization when:

$$\begin{aligned} &\frac{1-(1-\bar{s})^{\gamma}\left(\bar{n}/n_c\right)^{\alpha/(\sigma-1)}}{1-\left(1-s_c\right)^{\gamma}} \\ > \quad &\frac{1-\left(n_{lo}/n_c\right)^{\alpha/(\sigma-1)}}{1-\left(n_{lo}/n_c\right)^{\alpha/(\sigma-1)}+(\tau^{\alpha}-1)\left(1-s_c\right)^{\gamma}}. \end{aligned} \tag{2.23}$$

The right-hand side of this inequality is always smaller than 1 (since $\tau > 1$ and $n_{lo} < n_c$, so the denominator is larger than the numerator). From region 1's welfare problem, $(1-\bar{s})^{\gamma}\,\bar{n}^{\alpha/(\sigma-1)} > \left(1-s_c\right)^{\gamma} n_c^{\alpha/(\sigma-1)}$, so that the left hand side of the inequality is 1 if $\bar{s} = s_c$, and lower than 1 if $\bar{s} < s_c$. Hence, when $\bar{s}$ approaches s_c, the left side of the equality is 1 in the limit, while the right hand side is smaller than 1. Thus, for a $\bar{s}$ marginally lower than s_c, the inequality holds, indicating a that a harmonized subsidy can improve welfare in region 2. Therefore, compared to the mixed strategy equilibrium, a common subsidy that is marginally lower than the competitive subsidy increases expected welfare in both regions.

2.5 Conclusion

This chapter studies the welfare conclusions of policy competition under agglomeration. Agglomeration forces change the nature of the “race to the bottom”, in which governments compete to attract firms with low taxes, hence dressing down their government services. As models that introduce agglomeration show, governments of large regions can use agglomeration rents to prevent smaller regions from competing. Large regions’ governments, as a Stackelberg leader, first set a "limit" tax: a tax such that a small region, who chooses taxes second, refrains from attracting firms. The limit tax deters small regions because they need to compensate the policy of the large as well as the benefits of locating in a large region, if they want to attract firms. In that case, optimal policies differ by regional size, and harmonization is not a Pareto improvement. A policy recommendation in that case is to set tax floors or subsidy limits, as this reduces competitive pressures from smaller governments.

This chapter disentangles the strategic incentives of large regions associated with agglomeration forces and those associated with the Stackelberg (first-mover) advantage. It develops a spatial general equilibrium model of agglomeration (a footloose entrepreneur model with vertical linkages), in which governments hand out subsidies to mobile firms at the cost of public service provision. The game has no simultaneous pure strategy Nash equilibrium, so in addition to the sequential pure strategy Nash equilibrium, it develops a simultaneous mixed strategy Nash equilibrium. Throughout, the main predictions on government behavior are consistent with earlier findings: large regions set higher subsidies (the equivalent of lower taxes), and greater agglomeration advantages increase the policy gap.

The welfare conclusions, however, show that the case against harmonization rests on the assumed timing advantages in policy-setting for large regions. If policy is set at the same time instead of sequentially, there are harmonized policies that both governments would choose over the laisser-faire equilibrium. The intuition is that the credible threat of a large first-mover eliminates all incentivesfor smaller regions to compete for firms. This threat is no longer credible under simultaneous moves, and small regions get an incentive to set high subsidies. This incentive is not only to attract the agglomeration, but it also forces firm subsidies up in the core region in equilibrium. The mixed strategy thus points to competitive behavior as a disciplining device from small towards large regions, that allows them to internalize (some) policy externalities. Harmonizing policies avoids very high subsidies for the large region, but they are high enough for the small region to internalize some of the policies externalities. Therefore, if there are no first-mover advantages, harmonization can improve welfare. A subsidy floor is suggested as a welfare improvement instead of harmonization in earlier tax-competition-with-agglomeration literature. This chapter shows that if there are policy externalities (enhancing firm entry in this case), this comes at the cost of smaller regions. Therefore, a model with simultaneous policy setting and interregional policy effects produces diametrically different welfare conclusions from earlier models.

Our results imply policy harmonization may be optimal, both with and without agglomeration forces. Under agglomeration externalities, conclusions depend on whether one views policy-making as sequential or simultaneous. Additionally, the conclusions depend on whether there are interregional spillovers of policy. In the Dixit-Stiglitz model of product diversity, such effects play a natural role, although they were ignored in earlier models by assuming that the number of firms was fixed, and taxes would not affect entry.

To stress the role of policies' effects in the real economy, governments in this model have subsidies as an instrument. However, if we assumed that local governments had no central source of finance, they would optimally set negative subsidies, or effectively firm-level taxes to finance public services, which does not change our results. Arguably, other policy instruments would also fit the argument as long as they have positive external effects in other regions, such as physical intraregional infrastructure or the encouragement of R&D and productivity growth, the benefits of which do not fall within the home region exclusively.

2.A Condition for incomplete specialization

To use the wage as a numeraire, we must ensure that it is equal between the two regions. With sectoral mobility and free trade of the constant returns to scale agricultural goods, equal nominal wage occurs when the most specialized region still uses a marginal worker in agriculture. Using this observation, a condition to use the wage as numeraire is that the aggregate manufacturing labor requirement in a region where production has completely agglomerated does not exceed the labor force in that region. More formally;

$$N_{M,1} \leq \lambda N. \tag{2.A.1}$$

The manufacturing labor requirement can be found by looking at the labor input needed for equilibrium firm output, and aggregating over firms. Firm output is (following eq. 2.5)

$$q(i) = \frac{\left(\frac{\sigma}{\sigma-1} a_m w\right)}{n} \alpha \left[E_1 + E_2\right], \tag{2.A.2}$$

where we can substitute the definition for world expenditure (eq. 2.11) into the last term and aggregate over all firms, and use the normalization of the productivity:

$$N_{M,1} = nq = \frac{\alpha}{\sigma}\left[\frac{w(1-T)N + sTw\lambda N}{1-\alpha/\sigma}\right]. \tag{2.A.3}$$

Filling out this term in (2.A.1) and rewriting gives:

$$sT\lambda \leq T\lambda + \lambda(1/\alpha - 1/\sigma) - 1, \tag{2.A.4}$$

where we have assumed $w = 1$ as numeraire and the empty region has no firms to subsidize. This condition states that the maximum subsidy for which the wage can be used as a numeraire decreases in the tax rate and the preference for manufacturing goods α, and increases in the price and (i.e. decreases in market power). A typical constellation of $\lambda = 0.5$ and $\alpha \leq 0.5$ supports subsidy rates up to 100%.

2.B Operating profits as a function of the number of firms

The firm's profits for a firm in region 1 (eq. 2.8) can be written as:

$$\Pi_1 = B_1 \frac{\alpha E^w}{\sigma n^w} - FP^{-\alpha} + S, \; B_1 = \left[\frac{b}{P_1^{1-\sigma}} + \phi \frac{1-b}{P_2^{1-\sigma}} \right], \tag{2.B.1}$$

where the term B_1 describes the share of total expenditure that is spent on product produced by firms in region 1. Since the share of expenditure originating in region 1 (i.e. the workers' income and firms demand for inputs in region 1) is a function of the number of firms in region 1, we solve B_1 for s_n, inserting the expenditure share (eq. 2.13) and rewriting:

$$B_1 = \frac{\left(1-\phi^2\right)\left(1-s_n\right)s_n + \phi s_n + \phi^2\left(1-s_n\right)}{\left(1-\phi\right)^2 s_n\left(1-s_n\right)+\phi}. \tag{2.B.2}$$

Collecting terms, and using the definitions $A \equiv \left(2\delta - \phi\right)/\left(1-\phi\right) - \delta$, and $Z \equiv (1-2\delta)\left(1-\phi\right)s_n\left(1-s_n\right) - \delta$, the term B_1 can be written as:

$$B_1 = 1-(1-\nu)(\nu - A) \times \tag{2.B.3}$$

$$\left[\frac{\left(\begin{array}{c} \left(\nu(1-\nu) + \frac{\phi}{\left(1-\phi^2\right)-\left(1-\phi^2\right)\mu/\sigma}\right) \times \\ \left(\left(1-\phi^2\right)-\left(1-\phi^2\right)\mu/\sigma\right)\left(\nu(1-\nu)\left(1-\phi\right)^2+\phi\right) \end{array} \right)}{\left(\begin{array}{c} \left(\left(1-\phi^2\right)-\left(1-\phi^2\right)\mu/\sigma\right)\left(\nu(1-\nu)\left(1-\phi\right)^2+\phi\right) \\ +\left(1-\phi^2\right)\phi\left(1+\frac{\varpi+Z}{A-\nu}\right)\mu/\sigma \end{array} \right)} \right]^{-1}.$$

This term can be used to derive an analytical solution to the operating profit. To make the expression more tractable and to show the parallels to the FEVL-model without a government, we use the definition $\Psi \equiv \left(1-\phi\right)/\left(1+\phi\right)$, which gives:

$$\begin{aligned} B_1 \;=\; & 1- \\ & \frac{\left(1-s_n\right)\left(s_n - A\right)}{s_n\left(1-s_n\right)+\frac{1-\Psi^2}{4\Psi(\Psi-\alpha/\sigma)}} \times \\ & \frac{\Psi(\Psi-\alpha/\sigma)\left(s_n\left(1-s_n\right)\left(1-\phi\right)^2+\phi\right)+\Psi\phi\left(1+\frac{s_n+Z}{A-s_n}\right)\alpha/\sigma}{\Psi(\Psi-\alpha/\sigma)\left(\left(s_n\left(1-s_n\right)\left(1-\phi\right)^2+\phi\right)\right)}. \end{aligned} \tag{2.B.4}$$

Note that under symmetry of size and policy, $\left(s_n + Z\right) / \left(A - s_n\right) = -1$. The equivalent expression for B_1 in the standard model without a government of Ottaviano and Robert-Nicoud (2006) is:

$$B_1 = 1 - \frac{\left(1 - s_n\right)\left(s_n - 1/2\right)}{s_n\left(1 - s_n\right) + \frac{1-\Psi^2}{4\Psi(\Psi - \alpha/\sigma)}}. \tag{2.B.5}$$

Using symmetry in policy and share of laborers, as in Ottaviano and Robert-Nicoud, the expressions for operating profits in this chapter reduce to those of the standard model because $\delta = 1/2$ and hence $Z = -A$. If $S_1 = S_2 = 0$, i.e. if governments are absent from this model, the location choice hence reduces to that in the standard form of the FEVL model.

Equation (2.B.4) shows that in contrast to the standard NEG model, but like models of footloose entrepreneurs with vertical linkages (FEVL), the equilibrium expenditure share per region can be written as a closed-form solution using the number of firms. The same holds true for firm profits, since the price index can be expressed in terms of the firm distribution (see eq. 2.7). Finally, similar steps lead to the profit of firms in region 2, but it can easily be seen from the complementarity ($B_2 = 1 - B_1$) that expenditure and profits in region 2 are a function of s_n as well.

CHAPTER 3

AGGLOMERATION, URBAN STRUCTURE AND TAX COMPETITION: A LOCK-IN SITUATION

3.1 Introduction

Policymakers' efforts to attract firms form serious impediments to good policy. A main concern is that attracting firms with low tax rates may leave little budget to perform the government's public functions. As discussed in chapter 2, agglomeration forces may alter these concerns: government interaction changes and is potentially not harmful. The models mentioned, and the model developed in chapter 2 suggest that perfect agglomeration yields tax advantages for firms in large regions. In order to attract firms, small regions need to both undertax large regions and compensate for the agglomeration rents that they cannot offer. With sufficiently low taxes, the large region can thus eliminate small regions' incentives to attract firms altogether. However, if agglomeration is imperfect, firms spread out until profits are zero in all regions, eliminating the strategic advantage of agglomeration rents in large regions. If large regions have no such advantage, then how is policy influenced by agglomeration?

The contribution of this chapter is to shed new light on tax competition under imperfect agglomeration. In doing so, it extends insights in a recent surge in research on policy interaction under agglomeration. In this vein, specifically, Borck and Pflüger (2006) devote attention to imperfect agglomeration, concluding that results in a perfect-agglomeration setting translate into an imperfect agglomeration setting. While subscribing to their analysis, this chapter shows that subtle changes to the assumptions (especially to the original welfare function) may eliminate strategic incentives for governments that are central in the results. As a result, the tax floors prescribed in earlier literature do not generally yield welfare improvements. At the same time, the chapter advances understanding in traditional "no-agglomeration" tax competition. The policy recommendations of harmonization in that literature are substantiated with evidence of direct interaction effects in tax rates (causing, e.g., spatial correlations in policy variables). This model shows that this empirical evidence for tax competition models is also consistent with a model of agglomeration, in which harmonization will prevent optimal policies.

This chapter adds insights in situations where some cities endogenously grow large and others grow small, but not empty. Therefore, it develops an economy where New Economic Geography forces of city formation (i.e., home market effects) are countered with urban costs. These urban costs are captured by the fact that citizens' houses take up physical space, and so urban internal commutes grow larger as the city grows. This ensures that the size of small and large cities are a balance of centripetal and centrifugal forces. Apart from minor changes, the economy underlying the policy results is therefore much related to Tabuchi (1998), who pairs scale effects from the New Economic Geography in a system of cities with models of urban structure due to commuting costs.

The balance of dispersive and clustering forces that determines city size implies that the number of citizens is a result of economic trade-offs: people choose to live in small and large cities. This is unlike perfect agglomeration models (as most NEG models are), where peripheral demand is guaranteed by assumption of immobility rather than choice. Indeed, earlier results in tax competition usually require strictly dominant agglomeration forces (the NEG models) or strictly dominant spreading forces ("traditional" no agglomeration tax competition, see Baldwin and Krugman, 2004 for a distinction). This chapter, by contrast, fits into a tradition where concentrating forces between cities work against spreading forces at the city level (Henderson, 1974; Brakman et al., 1996). This is supported by the simple empirical observation that small cities exist next to big cities (not every city is New York), but data also show that income per capita and productive efficiency rise in city scale for small cities, but decrease in scale for large cities (Henderson, 1986; Au and Henderson, 2006; Mitra, 1999).

In studies of tax competition with agglomeration effects, the possibility of incomplete agglomeration rarely features. Rather, the predominant modelling strategy relies on agglomeration rents that arise with full agglomeration. This literature generally contends that the larger region is a Stackelberg leader (Baldwin and Krugman, 2004, or references in chapter 2), and chooses its tax rate such that smaller region cannot both compensate the agglomeration rents and undercut the large region's tax rate. Therefore, the large region effectively prevents smaller regions from participating in the tax game, and its behavior therefore also termed the "limit" tax. Borck et al. show that under imperfect agglomeration due to omitting income effect, the limit taxing solution remains, given a domain of trade costs.

Because Borck and Pflüger (2006) study a tax competition model under NEG with imperfect agglomeration, parallels are apparent, but two sources of diverging results can clearly be identified. Firstly, there are differences in mobility. Borck et al. use a "footloose entrepreneur" model (see Baldwin et al., 2003), in which part of the population migrates according to utility differences, while others are immobile.[1] In this chapter, by contrast, all workers are mobile. This has repercussions for the results, because migration based on welfare differences paired with benevolent government leads to a corrective force towards optimal policies. If utility equalizes via migration, then maximizing local welfare can lead to maximzal global welfare. However, the immobility of some workers, or the adoption of a government objective that is not (exactly) welfare eliminates this corrective pressure of citizen migration. In that sense, the model of Borck et al. is better interpreted as immobile workers competing to attract mobile workers ("footloose entrepreneurs"), whereas our model applies to a situation where everybody is mobile, such as intranational tax competition between jurisdictions. Therefore, both models feature the fiscal externalities of taxation, but the "voting by feet" forces are stronger in this chapter. A second difference is in the timing of government actions: Borck et al. assume that larger regions have a Stackelberg (first-mover) advantage. Borck et al. state that "a more natural way to model the tax game would be a simultaneous Cournot–Nash game" (p. 666), but choose the Stackelberg solution because a simultaneous pure strategy Nash equilibrium does not exist in their model. In this chapter, if governments explicitly pursue average welfare, a set of simultaneous Nash equilibria does exist. The welfare implications are accordingly affected: the Stackelberg solution in Borck et al. is generally not efficient, but in the current setup, simultaneous Nash equilibria can be efficient, and the Stackelberg solution is efficient.

The results of this chapter show that even if agglomeration is incomplete, it has strong impact on policy formation. The mechanism, however, is not the same as the limit tax in full agglomeration models. Rather, a preference not to relocate the partial core of both regions determines policy. Because the tax rates that will not induce major firm relocations are

[1] To avoid confusion, in the model of Borck et al., partial agglomeration is defined as partial concentration of the mobile workers.

defined relative to the other region's tax rate, a lock-in effect of policies occurs, and a wide range of simultaneous Nash equilibria is supported. This set of Nash equilibria includes the optimal taxes, but also inefficient tax rate pairs. This is surprising, since both governments optimize local welfare. Potential welfare losses are substantial: in the model, policymakers that purely maximize tax revenue may generate higher welfare.

The implications of our results are that some strategic incentives of governments engaged in tax competition may be misunderstood, and consequently, that there may be different welfare implications. The tax competition models that warn against races to the bottom produce much graver welfare conclusions than the agglomeration models that do not find races to the bottom. An empirical distinction is that the first predicts direct tax interaction effects, while the latter predicts taxation of agglomeration rents. There is both evidence of interaction effects (Bretschger and Hettich, 2002; Feld, 2000), and of agglomeration taxing (Charlot and Paty, 2007; Jofre-Monseny and Solé-Ollé, 2010; Hill, 2008; Koh et al., 2013; Brülhart et al., 2012), but the two underlying models are not theoretically consistent. The current model predicts that both effects exist, but they are not consistent with races to the bottom, nor with limit taxing of large governments. Rather, inefficiencies result through coordination issues if the optimal policies are completely defined relative to neighboring policies.

The chapter goes on to develop a model of partial agglomeration to analyze the coordination problems in tax setting. The economic equilibrium relations of that model are sketched in Section 3.3, and subsequently, policy formation is examined in Section 3.4. However, despite working with a simplest possible form to arrive at results, the model, as any NEG model with incomplete agglomeration, is analytically complex. Therefore, before presenting the more rigorous model, we first present a more intuitive discussion of the central contribution to lay out the structure of the argument, and to understand the model's motivation. This is the section we turn to next.

3.2 An intuitive exposition of the model

Assume that there are two regional economies or cities. Firms (or capital) are fully mobile, and choose to locate in either of the two regions. Governments of either region tax their local firm base with a per-firm tax t. Because there are agglomeration effects in the economy, the returns to setting up a firm in a region r depends on the share of the world number of firms s_n in that region: $r = r\left(s_n\right)$. Under positive agglomeration effects, we have that returns can increase in regional size ($dr\left(s_n\right)/ds_n > 0$ at least over some range), and we will generally have that the majority of capital ends up in one of the two regions. In an equilibrium of incomplete firm concentration, after-tax returns need to be equal across regions: $r_1\left(s_n\right) - T_1 \geq r_2\left(s_n\right) - T_2$. If we denote the excess (large over small region) agglomeration rent $\omega\left(s_n\right) = r_1\left(s_n\right) - r_2\left(s_n\right)$, then, assuming that the rent is higher in the large region, it can be seen from the equilibrium condition that the tax gap $T_1 - T_2$ cannot exceed the agglomeration rent $\omega\left(s_n\right)$; otherwise, or firms relocate and agglomerate in the other region. The "no-relocation-condition" thus holds that $T_1 - T_2 \leq \omega\left(s_n\right)$.

Now assume that governments set their tax rate to maximize local welfare, which we shall leave implicit for the moment. There is a discontinuity in its payoff function, because not satisfying the no-relocation condition will induce a strong non-continuous change in the share of capital in the region – the agglomeration shifts. This can be captured in a Kuhn-Tucker optimization problem: $\Lambda_r = V_r\left(T_1, T_2, s_n\right) + \lambda_r\left(T_1 - T_2 - \omega\left(s_n\right)\right)$, where V_r is the objective function and λ_r is the (potentially zero) multiplier on the constraint for region r not to move the agglomeration. This constraint captures the shadow cost/benefit to the welfare function of increasing the tax rate, if the agglomeration could not move. If this Kuhn-Tucker condition is not binding, the first-order conditions on V give the tax reaction functions, and governments end up in a Nash equilibrium. If governments find it optimal to shift the agglomeration (so they ignore the Kuhn-Tucker condition), they purposely violate the no-relocation condition: the large region attempts to become a small region and vice versa. This is an odd solution, as it implies unstable policies (intuitively, the agglomeration and periphery would always prefer to switch places).

The third option, which is most relevant to this chapter, is where the no-relocation constraint is binding. In that case, the government prefers not to relocate the agglomeration, even though that restricts the choice of tax rates. In particular, the large region's policymaker would prefer to set a higher tax rate, but that would make the agglomeration leave his jurisdiction. Therefore, avoiding large relocations effectively makes the large region's policymaker face a tax ceiling. The small region would set lower tax rates, if that would not make the majority of firms relocate to the small region. Effectively, preserving the agglomeration pattern poses a tax floor in the small region. If the no-relocation condition is binding ($\lambda_1 > 0$), the large region's government satisfies $T_1 = \omega\left(s_k\right) - T_2$ (which is the tax ceiling implied by the no-relocation condition). The other government faces $\lambda_2 < 0$, so $T_2 = T_1 - \omega\left(s_k\right)$. We thus get that the optimal tax in the large region 1 is the small region's tax plus an agglomeration rent, while the optimal tax in the small region 2 is the large region's tax minus the agglomeration rent.

If the no-relocation condition is binding for *both* governments, a lock-in effect occurs. If the optimal tax of one government is defined as a distance (depending on the agglomeration rent) of the other government's tax, then there are many tax pairs that classify as mutual best responses. For every t_2, an optimal response of region 1 can be identified as the ceiling of tax rates permitted by the no-relocation condition. If region 1 chooses this optimal response, the initial t_2 becomes the tax floor implied by the no-relocation condition. It is therefore also a best response, and a Nash equilibrium results. However, other pairs of Nash tax rates can be identified in the same way for different initial t_2, as long as the no-relocation conditions are binding. Essentially, the no-relocation decision becomes a first-order condition for both governments. This leaves us with one identifying equation for two endogenous tax rates. Therefore, if the partial agglomeration can occur over different tax pairs, a set of policy outcomes can occur, rather than a single tax pair.

For the situation sketched here to arise, we thus require that $\lambda_1 > 0$, $\lambda_2 < 0$. Intuitively, this implies that larger regions generally prefer higher tax rates, but regions are restricted by the fact that the majority of firms can relocate. In other words, the situation arises if the small government is restricted in taxation in the lower end: if it ran no risk of attracting the full agglomeration, it would set lower taxes (and v.v. for the large region). Also, clearly, if agglomeration is perfect, the optimization problem for an empty region becomes void, so the indeterminacy of a single optimal tax pair only occurs if agglomeration is imperfect. We put flesh on these requirements in the next section. We show that a relatively simple new economic geography model augmented with commuting costs due to urban structure replicates that i) larger regions prefer to set higher taxes on mobile firms, which is a central NEG insight, and ii) agglomeration is incomplete provided that commuting costs play large enough a role.

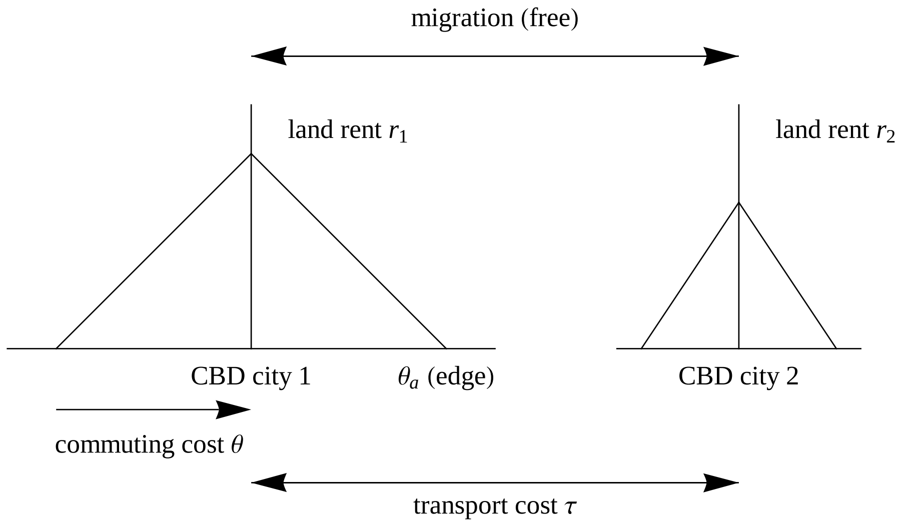
3.3 A model of two cities

This section develops the general equilibrium model used to study policy competition. The economy consists of two cities, that each take up physical space. Space is consumed by households, whereas production is carried out in the center of the city, the central business district (CBD), as in a standard urban economics model. Workers commute from their residence to the CBD of their own city. The commuting costs to other cities are assumed to be prohibitive. The products that workers produce can be shipped across cities at a transport cost, but they can be moved freely inside the city in which they are produced.

The model is closely related to Tabuchi (1998), but for analytical tractability, two simplifying assumptions are added. First, commuting comes strictly at the cost of leisure of citizens. This keeps the financial budget equal across workers, so that equal preferences for and access to government-provided goods is plausible; and it keeps the income definitions simple. Moreover, it avoids having to specify a transport sector or an additional source of demand from whoever charges the price of commuting. Secondly, there is a dispersive force emanating from commutes, so the centrifugal force of the agricultural sector, which commonly serves as a source of regional residual demand, is no longer necessary. This choice not only simplifies the model, but also simplifies the welfare- and government objective function. If immobile farmers would live next to mobile manufacturing workers, there is inherently inequality in the economy, which would require additional assumptions of political preferences about inequality. If welfare is equal across all inhabitants, such assumptions are not necessary. The nature of trade changes somewhat when excluding agriculture: instead of shipping agricultural products to the core and manufactures to the periphery, trade now balances by shipping manufactures both ways, although a wider set of manufacturing varieties are shipped to the periphery than vice versa.

The two cities are denoted with subscripts 1 and 2, and the iceberg transport cost between them as τ. We shall assume that the land around the cities is worked by farmers at an agricultural rent r_a, which is the opportunity cost of land. Furthermore, the analysis adopts a linear form of the city, and land use is symmetric around the centre. The term θ indexes the distance from the city center, so that at the CBD, $\theta = 0$. The land is owned by landowners, who consider the excess rents over the agricultural rent as income. In urban models, these proceeds usually leave the system,

Figure 3.1: The spatial organization of the economy



but not to impose an artificial urban cost, we shall assume that landowners use the income for the consumption of manufacturing goods. In general, individual firms are indexed with an i. Capital letters refer to indexes or aggregations of items. The distance from the CBD at which land use switches from residential to agricultural is termed θ_a. Furthermore, when discussing different cities, the coordinates will be subscripted, such that θ_{a2} is the coordinate of the outer edge of city 2. The total number of inhabitants is N, of which share λ live in city 1, and the complement $1-\lambda$ in city 2. A summary of the organization of the economy over the two regions is given in Figure 3.1.

We shall assume that each household consumes a house, manufacturing goods and leisure. Additionally, it derives utility from a government-provided good, G. The required house is of fixed size, so that the density is fixed. Houses located further from the CBD are associated with longer commutes to work, and are therefore less desired. The shadow costs of commutes are leisure, i.e., the worker has a labor contract of fixed length, and commuting reduces utility by diminishing his free time. Using η to denote the time cost of commuting per unit of distance, $\eta\theta$ is the total time cost of commuting to the CBD. The consumption index is C, comprising all manufacturing varieties with a constant elasticity of substitution, σ as in chapter 2. Then, the utility function is:

$$U = C^{\alpha} G^{1-\alpha} - \left(\eta\theta\right)^{\xi}; \quad C = \left[c\left(i\right)^{\frac{\sigma-1}{\sigma}} di\right]^{\frac{\sigma}{\sigma-1}}. \tag{3.1}$$

The term α captures the weight of private consumption on goods consumption, and ξ governs the elasticity of the (quasi separable) utility loss from commuting time. Assuming a unit housing density[2] and a house rent of r, the financial constraint states that the costs of housing and of consumption of manufacturing goods equals the wage: $w = r + \int p\left(i\right) c\left(i\right) di$. Dividing the first-order conditions for any two manufacturing goods i and j and simplifying gives that in the optimum of the consumer's maximization problem, it must hold that $c\left(i\right)/c\left(j\right) = \left(p\left(i\right)/p\left(j\right)\right)^{-\sigma}$. Isolating $c\left(i\right)$ in this expression and using it in the budget constraint, we can derive the demand function for a good $c\left(i\right)$. Using the definition of the CES harmonized price index $P \equiv \left(\int p\left(i\right)^{1-\sigma} di\right)^{1/(1-\sigma)}$ gives the demand function for goods produced by an individual manufacturing firm:

$$c\left(i\right) = \frac{p\left(i\right)^{-\sigma}}{P^{1-\sigma}} \left(w - r\right). \tag{3.2}$$

The financial constraint also allows reformulating the utility function in terms of prices, government services and commuting times:

$$V = \left(\frac{w-r}{P}\right)^{\alpha} G^{1-\alpha} - \left(\mu\theta\right)^{\xi}. \tag{3.3}$$

Competition on the land market implies that workers bid up prices for land that they desire more. In equilibrium, prices reflects the willingness to pay to live on a plot of land, so that no workers as an incentive to move. Since workers are identical ex ante (i.e., they have the same time and skill endowment), the land rent exactly compensates for commuting distances. Therefore, the utility of living in a location θ_i is equal to the utility of living at the border of the city, θ_a. The edge of the city is determined at the distance from the CBD where citizens bid less that the opportunity agricultural rent for land, so the boundary condition states that at the edge of the city (θ_a), the land rent is equal to r_a. Rewriting the indirect

[2] Since we have not chosen a unit of distance, only the time cost per unit of distance (θ) relative to the density matters, and so we are free to normalize either the commuting cost per unit of distance or the residential density.

representation of utility (eq. 3.3) for the bid rent and subtracting the rent at the edge of the city yields the excess urban land rent function:

$$r\left(\theta_i\right)-r_a=P\left[\left(V+\left(\mu\theta_a\right)^{\xi}\right)^{1/\alpha}-\left(V+\left(\mu\theta\right)^{\xi}\right)^{1/\alpha}\right]. \tag{3.4}$$

Consistent with urban economic theory, the land rents decrease in the distance from the CBD: citizens are willing to pay more for central plots of land. If the price index is higher, fewer manufacturing goods can be bought for a given a level of income, and so the cost of commuting relative to consumption rise, leading households to bid up land rents closer to the CBD. Equilibrium on the housing market requires that every worker is housed. With unit density, this simply implies that the size of the city is equal to the number of residents. The total differential land rent (TDR) is equal to the aggregate land rents less the agricultural opportunity cost, it is given by the integral of per unit land rents (eq. 3.4) over the city's size.

The organization of the manufacturing sector largely follows a standard Dixit and Stiglitz (1977) setup. Firms produce using labor under increasing returns to scale, due to a fixed labor requirement in production: $l = aq + f$. In addition, firms face a region-specific firm-level tax t to operate in a location. The total costs for the firms are $wl + T = wqa + wf + T$. Since the price elasticity of demand of all consumers is constant and equal to σ, and firms consider themselves too small to affect the aggregate price index, the standard markup price holds: $p = \sigma/(\sigma-1)aw$. Free entry into the industry drives pure profits to zero in the long run: $\sigma/(\sigma-1)awq - awq - wf - T = 0$. Rewriting this zero-profit condition gives the equilibrium firm size as $q = \left(f + T/w\right)(\sigma-1)/a$, and substituting this firm size in the technology function allows identification of the labor requirement: $l = f\sigma + (\sigma-1)T/w$. In absence of a tax, the labor requirement is $f\sigma$, as in the standard Dixit-Stiglitz model. The tax leads to larger firm production and firm hiring, and given a pool of workers of fixed size, to a lower number of firms that can survive in equilibrium. Taxes pose an increased fixed cost of operating in a region, and therefore act as an entry barrier for firms: they need to expand their scale to recoup taxes.

On the goods market, the aggregate demand for the good of one firm is the aggregate of individual demand curves (eq. 3.2). Since the firm needs to ship τ units of a good for one unit to arrive in the other city, a firm in city 1 faces the aggregate demand function:

$$\begin{aligned} q(i) &= p(i)^{-\sigma} MP, \\ MP_1 &= \left[\frac{Y_1}{P_1^{1-\sigma}} + \tau^{1-\sigma}\frac{Y_2}{P_2^{1-\sigma}}\right], \end{aligned} \tag{3.5}$$

where $Y_{1,2}$ is the aggregated income by city (income of landowners and workers) spent on manufacturing goods. The term MP is the market potential for the industry, or potential expenditure from both locations, which depends on the city (the term for city 1 is presented as an example). There is one wage level for which the goods market clears. The term P_1 is the standard CES-harmonized price index:

$$P_1 = \left(n_1 p_1^{1-\sigma} + n_2 \left(\tau p_2\right)^{1-\sigma}\right)^{1/(1-\sigma)}.$$

Inserting the markup price ($p = \sigma/(\sigma-1)aw$) and using the equilibrium firm size ($q = \left(f + T/w\right)(\sigma-1)/a$) in the goods market clearing condition and rewriting for the equilibrium wage rate gives:

$$w^{\sigma-1} = \frac{\left(\frac{\sigma a}{\sigma-1}\right)^{-\sigma}}{(\sigma-1)/a}\frac{MP}{\left(wf+T\right)}. \tag{3.6}$$

From this, it can be seen that keeping everything else constant, a higher tax rate has a depressing effect on wages.[3] The goods market-clearing condition shows that if taxes are higher, then firms need to sell more products to break even. This can only be achieved by asking lower prices, so if taxes rise, firms will offer lower wages. The market potential term depends on the aggregate price index and income. Using the above results, we can write the income and price definitions more explicitly. Using l_r to denote the labor requirement of firms in region r, the number of firms in region 1 is $\lambda N/l_r$: the total number of inhabitants in region 1 divided by how many workers a firm hires. Using this expression for the number of firms and the markup price, the CES-harmonized price index P_1 can be written as:

$$P_1 = N^{1/(1-\sigma)}\frac{\sigma a}{\sigma-1}\left[\frac{\lambda}{l_1}w_1^{1-\sigma} + \tau^{1-\sigma}\frac{(1-\lambda)}{l_2}w_2^{1-\sigma}\right]^{1/(1-\sigma)}. \tag{3.7}$$

[3]Implicit differentiation gives that $dw/dt = -\left(f\sigma + (\sigma-1)t/w\right)^{-1} < 0$ when keeping market potential constant.

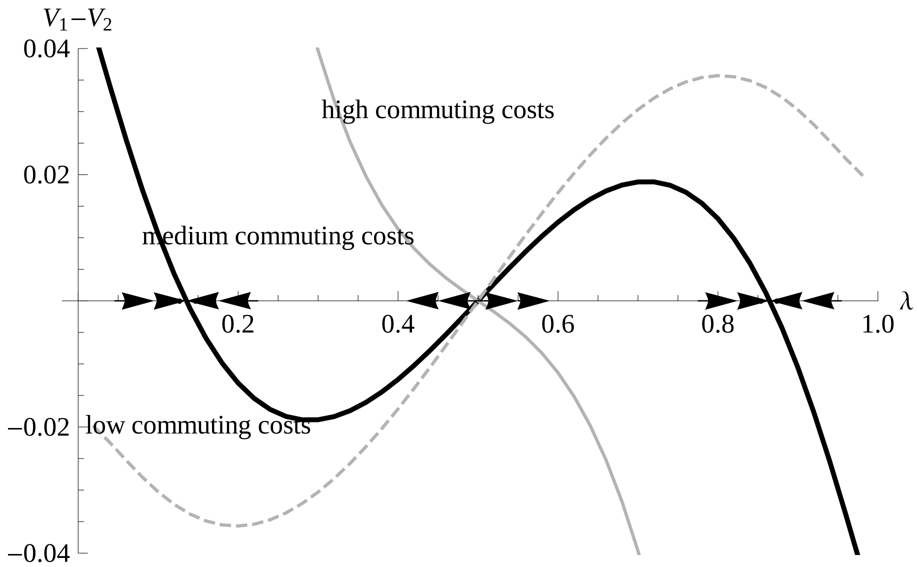
The aggregate income consists of two components: potential income out of labor andexcess land rents. The total income spent on manufacturing goods originates firstly from the consumers, who spend their income net of housing cost $(w-r)$ on manufacturing goods. Additionally, landowners spend the urban excess land rent, $r-r_a$, on manufacturing goods. Therefore, the excess land rent paid inside the city raises the demand for manufacturing goods. Since the total differential land rents plus the opportunity cost of land inside the city equal the household spending on housing, the total expenditure of one city on manufacturing goods can be written as:

$$Y_1 = \lambda N(w_1 - r_a). \tag{3.8}$$

This shows that expenditure on the manufacturing goods is the total labor income less the opportunity cost of land $(N r_a)$. The reason is that workers spend their income on land rents and manufacturing, and of that total bill of land rents, the excess land rent $(r-r_a)$ is spent on manufacturing goods by landowners.

The three-equation system for wages (eq. 3.6), price indexes (eq. 3.7) and income (eq. 3.8) is quite similar to a standard NEG model (Fujita et al., 2001, chapter 4). The first main difference is in the governments' instruments, which we take exogenous in this section and study in the next section. The second difference is in the firm labor requirements. The third difference is in the commuting costs, which add a dispersive force to the (extensively documented) three equation system. Its consequences are pictured in the medium commuting costs utility difference in Figure 3.2. The figure is included to show the theoretical possibilities of the model. Assuming that excess utility of living in one city over the other leads to immigration, it can be seen that symmetric equilibria are unstable: a small deviation from $\lambda = 0.5$ raises relative utility in the large city and starts off a migration flow into the large city. However, complete agglomeration is not an equilibrium either; if one city hosts all inhabitants, the utility of living in the empty city is higher. Instead, the curve shows that utility is equal across cities for two (asymmetric) distributions of inhabitants. The utility curve is downward sloping at those points, indicating that migration into a city reduces the relative utility, and hence that the equilibrium is stable. This is not possible in the generic model without urban structure, which is limited to equilibria of symmetry and complete agglomeration. The model yields those equilibria, however, by stressing the costs of commuting (solid grey line) so that symmetric equilibria become stable;

Figure 3.2: Spatial equilibria under different commuting costs



Note: Parameters: $N = 1$; $\sigma = 5$; $f = 1/\sigma$; $a = (\sigma - 1)/\sigma$; $\tau = 1.3$; $\rho = 5$; $\alpha = 1/2$; $\mu = 0.75$; $T_1 = T_2 = 0.01$.

or by downplaying the dispersive force of commuting costs (dashed grey line), so that full agglomeration results. Tax competition for spreading equilibria (the "classical" tax competition models) and for fully agglomerated economies are extensively discussed in the literature mentioned in the introduction. Therefore, we focus exclusively on the novel equilibria of incomplete agglomeration.

3.4 Tax setting

This section studies government behavior, given the spatial equilibrium relations derived in Section 3.3. It assumes that governments maximize the average welfare of local inhabitants. Portraying governments as benevolent follows convention in earlier tax competition literature, where the welfare function is explicitly used as an objective (see, e.g., Wilson and Wildasin, 2004, for a review). For simplicity, recent literature also uses a quadratic approximation of the welfare function. This yields a reduced-form government objective, in which tax revenue is valued, but a convex penalty function for high tax rates is included to capture the effect of tax distortions. We follow virtually all literature in this field by assuming that the government instrument is a simple lump-sum tax on firms (Borck and Pflüger, 2006; Baldwin and Krugman, 2004; Ludema and Wooton, 2000).

There are two reasons to follow the actual instead of the approximated welfare function.[4] The first reason is that the utility function identifies the exact welfare effects of taxation, apart from its impact on firms' locations decisions. In this case, primarily, taxes affect firm entry, which poses externalities in the Dixit-Stiglitz market structure. Higher taxes in one region reduce entry and thus have a negative policy externality in other regions by reducing firm variety. These can be important effects in NEG models, but are omitted in the reduced form objective, which does not incorporate love-of-variety effects. The second reason is that migration, which determines regional size, relies on the equality of utility in the two regions. Therefore, using the approximate welfare function will lead to equilibrium tax pairs that distort optimal city size if approximate and actual welfare are not equal.

The government-provided good has not been discussed extensively, because our main interest is in the effects of taxation on regional size. The government-provided good is the reason for taxation. The unit-elastic preference for the good ensures that there is an Inada-type condition on its consumption: as the consumption of the government good tends to zero, the willingness to pay for it grows infinite. Therefore, tax rates will generally be positive. We shall assume that consumption of the government-provided good is equal to the tax revenue per head – this suggests the public good is priced at 1. We shall assume that the government-provided good is rivalrous, to ensure that our result do not stem from scale advantage in government for the large region. A drawback of our model of government is that the government good is not produced and thus draws no resources, like labor, from the economy. A public production sector complicates the model without adding much insight. For robustness, we have developed the model for an alternative case: citizens have no preference for government-provided goods, and the government simply redistributes

[4]There is also a practical argument. Since partial agglomeration models do not generally have closed-form solutions, many results rely on numerics. In that case, the virtue of finding analytical solutions with the approximated welfare function disappears, while approximate and actual welfare take equal effort to calculate. One advantage that could speak for assuming taxation has no real impact is that the intuition of the no-relocation decisions could be described further, because the agglomeration rent becomes independent of taxes. However, the assumption that tax revenue has no effect on the economy also eliminates the incentives to tax.

the tax revenue lump-sum. In that case, for a welfare maximizing government, the marginal costs of taxation are equal to the marginal utility of income, instead of the marginal utility of government-provided goods. Our conclusions remain the same throughout that exercise.

Policy under benevolent local governments

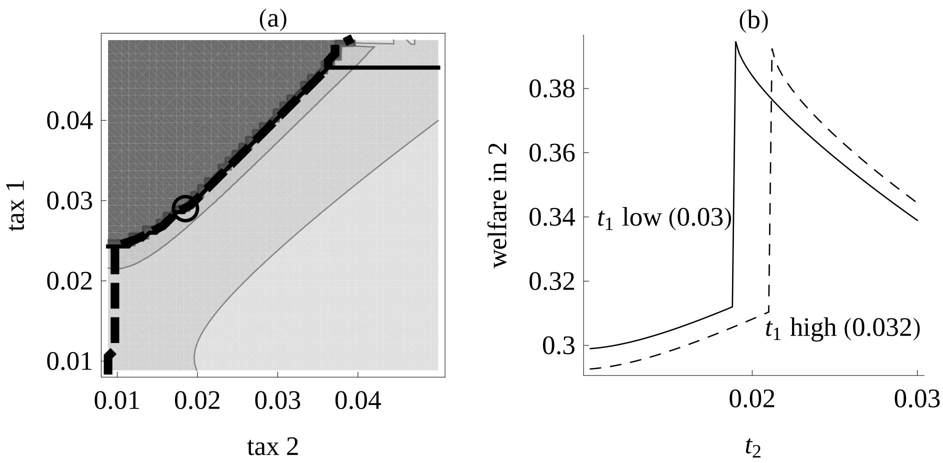
The effects of government policy on the economy can be studied given the spatial equilibrium conditions. The per-capita tax revenue is equal to the tax that each firm pays, multiplied by the local number of firms, divided by the total number of inhabitants: $T_1 n/(\lambda N)$. Inserting the expression for the number of firms, the per-head tax revenue is $G_1 = T_1/\big(f\sigma + (\sigma-1)\,T_1/w_1\big)$. The per-head revenue is independent of λ, because the number of firms and the number of inhabitants are proportional. The goal of the local benevolent policymaker is to maximize the inhabitant's average utility, given in equation (3.3). In the city's internal spatial equilibrium, utility is equal throughout the city, so we evaluate welfare (average utility) as the indirect utility at the city's edge:

$$V = \left(\frac{w-r_a}{P}\right)^{\alpha} G^{1-\alpha} - \left(\mu\theta_a\right)^{\xi}, \tag{3.9}$$

where under unit population density, θ_a is equal to half of the city's population. In the urban equilibrium, utility is equal among inhabitants in the same city. Therefore, despite not working with a representative inhabitant, a benevolent policymaker does not have to deal with issues of inequality, and if a political setup was added to the model, all inhabitants would have the same policy preferences.

In the tax-setting game, we assume that governments set taxes simultaneously. This runs counter to models based on agglomeration rents, in which the large region is given a first-mover advantage. A rationale for the first-mover advantage in the agglomeration literature is that the simultaneous pure strategy equilibrium does not exist under agglomeration rents (Borck and Pflüger, 2006, Appendix C, or chapter 2 of this thesis). The large region's limit tax evokes passivity of the small region not to attempt to take over the agglomeration. However, if the small region will not undercut the large region's tax, the optimal response for the large region is to raise taxes, and so the limit tax is not a mutual best response. Indeed as chapter 2 shows, the timing generates an advantage for the large region apart from the agglomeration effects. A second reason to choose

Figure 3.3: Best response functions and payoffs



Note: Parameters: $\sigma = 5$, $a = (\sigma - 1)/\sigma$, $f = 1/\sigma$, $N = 1$, $\rho = 5$, $r_a = 0.1$, $\mu = 0.75$. In panel (a), lighter contours indicate a higher equilibrium residential distribution λ. Best response tax rates for region 1 (solid) and region 2 (dash); the lines overlap. The welfare-maximizing tax rates are indicated by a circle.

simultaneous strategies is that it is robust to repetition and the presence of multiple players, in contrast to the Stackelberg game.

Using the welfare functions as payoffs, the tax-setting can be solved numerically. For every opponent strategy, the tax that yields the highest utility can be selected to construct a best response curve. To do this, the strategy space (the set of permissible tax rates) is divided into 50 segments, giving rise to 2500 possible tax pairs. For each of these tax pairs, the residential equilibrium is calculated. This done by varying λ (the share of residents in city 1) over 50 evenly spaced segments from 0 to 1, interpolating the function of utility difference between city 1 and 2 and calculating the long-run equilibrium. Since we assume that city 1 is initially the larger city, in case of multiple equilibria, the equilibrium with more inhabitants in city 1 is selected, which reflects the path dependency in the residential equilibrium. Unless reported otherwise, the parameter follow a fairly standard NEG parametrization: $\sigma = 5$, $a = (\sigma - 1)/\sigma$, $f = 1/\sigma$, $N = 1$, and the urban parameters are $\rho = 5$, $r_a = 0.1$, $\mu = 0.75$. These urban parameters yield imperfect agglomeration for the standard NEG parameters.

The resulting best response functions and equilibrium spatial distributions are collected in Figure 3.3 (a). In the Figure, a solid line indicates best response tax rates from the large city, t_1. Dashed lines indicate best responses for the small city 2, setting t_2. The contours denote the equilibrium residential distribution given the tax rate pair: Over most of the strategy space, there is a partial agglomeration in city 1, and the size of that agglomeration falls in tax rates in city 1 and grows in tax rates in city 2. The circle identifies the tax pair that gives the highest possible joint welfare level, which is then a first-best tax rate pair.

The most striking feature of the best response functions is that they overlap over a substantial part of the strategy space. By definition, these tax pairs are mutual best responses, and as a result, we have that there is a set of Nash equilibria, rather than a single or no Nash equilibrium. The first-best tax rate pair is one of the Nash equilibria. Also, note that the set of tax rate pairs that classifiy as an equilibrium run close to tax rate pair that would lead to strong changes in cities' residential size (as witnessed by the contour plot).

The intuition for the multitude of Nash equilibria was already provided in section 3.2, but can be reiterated here. The second panel of the Figure 3.3 provides the small city's payoff (welfare) profile for different tax rates when the taxes in the large city (1) are given. The no-relocation condition is clearly visible in this profile: for a tax around 0.02, there is a jump in the payoff profile. This reflects that if the small city sets its tax rate lower than this discontinututy, or "break tax", it will move the partial agglomeration into the small city, effectively turning it into the big city. The constraint in the government optimization problem in section 3.2 follows from this discontinuity. If larger cities desire larger tax rates, undercutting is not desirable. It will yield a large city with low taxes and a small city with high tax rates. Since migration eliminates utility differences, all inhabitants are worse off after undercutting. After the agglomeration has shifted, there is a large city with low taxes and a small city with high taxes. This yields lower welfare than before the relocation, when large cities charged high taxes, and small cities charged low taxes.

The reason that large cities set higher taxes is not related to agglomeration rents, because those rents are eliminated. As in the results of Borck and Pflüger (2006), the solution is rather found in the tax competition literature dealing with exogenous asymmetric regional size (Bucovetsky, 1991). If the city is large, raising taxes will reduce local profits, but if firms flee to the small region, they will quickly drive down returns there, too. The tax base elasticity is therefore smaller in larger regions (they influence the economy-wide prices more than the small regions does), and for that reason, the larger region will set higher taxes. A key difference with the result of Bucovetsky, however, is that the agglomeration's location is endogenous here, and the looming relocation of the core results in the multitude of policy outcomes.

Because relocating the agglomeration via low taxes is undesirable in partial agglomeration, the optimal tax rate is just at the no-relocation condition. This is the tax rate that determines the best response function. A higher tax in the large city (the dashed payoff profile) increases the minimum tax required in the small city to preserve the agglomeration: the optimal tax rate grows, because the tax floor for preserving the agglomeration changes. Thus, the discontinuity in tax payoffs is the cause of many different tax pairs that are mutual best responses. To put it slightly more formally, when both governments prefer the agglomeration not to be moved, they both satisfy the no-delocation constraint ($T_1 - T_2$ equal to the pre-tax profit premium in the large region). In that case, the constraint becomes the condition for optimal local welfare, and the equilibrium is only defined in relative terms, as long as shifts of the agglomeration are not optimal for both governments.

The interpretation of the set of Nash equilibria is that policy can be self-reinforcing. Because the efforts not to relocate the agglomeration require coordination on each other's tax rates, a lock-in effect emerges. That is, once both governments set taxes higher than first best, none will unilaterally deviate from this equilibrium. The same is true for tax rates that are lower than first best. Therefore, inefficient equilibria can sustain themselves.

The empirical patterns predicted by this model entail that large regions set larger taxes. This is shared with agglomeration models, although, as explained, the mechanics are not the same. However, any deviation in one of the city's tax rates is expected to be met by a change in the opponent's tax rate in the same direction. Interaction patterns also seem to be present, therefore. Yet, again, the mechanics are not the same as in the tax competition literature without agglomeration, where the local tax base elasticity depends on peers' taxes. Instead, the parallel movements in taxes reflect the preservation of the no-relocation condition.

Government policy inefficiency

As the first-best tax pair is a Nash equilibrium, benevolent government can jointly sustain the first-best outcome. However, there are many other tax pairs, which are clearly less efficient. Which of the tax pairs emerges, depends on the initial conditions of the game. One could assess the welfare among benevolent governments by taking the expected welfare over all possible outcomes. This requires additional assumptions on preferences, however. Rather, we discuss the range of possible welfare outcomes.

The Nash equilibrium with the highest tax pairs is numerically the least efficient of all Nash equilibria. Under our parameters, the welfare level for all inhabitants (free migration equalizes utility) is 92.4% of the first-best welfare level. In part, clearly this is due to an oversupply of the government-provided goods. However, the residential allocation is also distorted: in the first-best situation, the large city rounds up around 74% of the inhabitants, this is around 9.7 percentage points higher in the worst possible Nash equilibrium.

In contrast to most earlier literature in this tradition (such as Baldwin and Krugman, 2004; Borck and Pflüger, 2006), this study uses the actual welfare function instead of a quadratic approximation of the welfare function. In this model, it makes a fundamental difference whether one uses actual or an approximated welfare function. The reason is that migration eliminate utility differences, so utility is constant across all inhabitants. Therefore local welfare maximization is very similar to global welfare maximization (they are the same function), but each government controls only one the two instruments. Under an approximated welfare function, policymakers do not jointly optimize the welfare function. If the two policymakers do not maximize the same value, the coordination problems do not need to occur, therefore, it is interesting to compare the tax competition results under approximated and actual welfare as objectives.

The quadratic approximation to welfare assumes that governments maximize tax revenue, but penalize the tax rate with a second-order term, so the objective function is $\Gamma = G - (1/2)\, T^2$. Note that since this is an implicit welfare function, we consider tax revenue per head, not in the aggregate. This is in contrast to earlier models, which take G to be the aggregate tax revenue. In our case that would imply the government-provided good is non-rivalrous, and it would thus endow the large city with an additional advantage. Conclusions do not change between taking G to be the per-head or aggregate tax revenue, however.

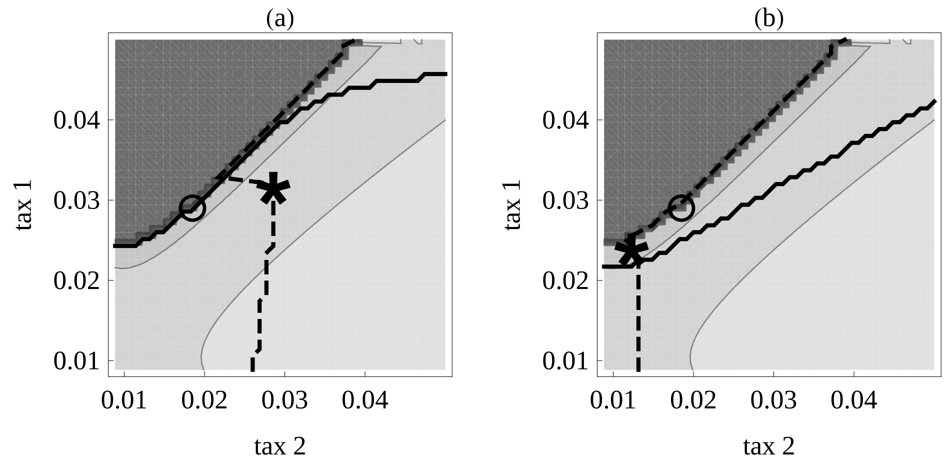
As Borck and Pflüger (2006) show, no pure strategy Nash equilibrium exists for the approximated welfare function. We follow the Stackelberg solution by allowing the large city to select its tax rate first, and then have the second city's government to select a tax rate. The game is solved backwards: city 2's payoff function is:

$$\left\{ \begin{array}{l} \Gamma_2\left(T_1, T_2|\lambda < 0.5\right) \text{ if } T_2 < \underline{T}_2 \\ \Gamma_2\left(T_1, T_2|\lambda > 0.5\right) \text{ if } T_2 > \underline{T}_2 \end{array} \right. .$$

If $T_2 < \underline{T}_2$, city 2 takes over the agglomeration, so that the majority of workers ends up there ($\lambda < 0.5$). The limit tax that the large city sets makes the smaller city indifferent between trying to take over the agglomeration with a low tax rate or remaining the (partial) periphery with a high tax rate. Therefore, the large city chooses the tax rate that solves $\max_{T_2} \Gamma_2\left(T_1, T_2 : T_2 < \underline{T}_2\right) \leq \max_{T_2} \Gamma_2\left(T_1, T_2 : T_2 > \underline{T}_2\right)$. This yields a single tax rate pair as equilibrium, because in contrast to welfare-maximizing governments, the objective function Γ does not generally equalize in the partial agglomeration. Therefore, although migration equalizes utility levels, the objectives functions diverge, and there is a single tax rate that makes the peripheral government indifferent between breaking the agglomeration pattern or not. Essentially, the first order condition of one government cannot be reformulated in the first order condition of the other: the are two separate first order conditions for local welfare with respect to the tax rate.

Based on the reduced-form welfare objective, the limit tax and the corresponding best response curves are presented in Figure 3.4 (a). At the limit tax (the star), the small region will not set a tax lower than the no-relocation condition. At 88.1% of first-best, the welfare is slightly lower than the worst-case scenario among welfare-maximizing governments. It is similarly substantially lower than first best. The residential allocation is also distorted: the large city hosts around 10.2 percentage points more inhabitants than first best. This difference can be due to the use of the approximated instead of the exact welfare function, or due to the Stackelberg equilibrium instead the simultaneous equilibrium. This underlines the difference between the approximated and actual welfare function: the simultaneous equilibrium does nto exist under the approximated welfare function. Moreover, under the exact welfare function, the Stackelberg solution leads to first best: the subgame perfect solutions allows the large city to select first-best tax rate, after which the small city also selects the

Figure 3.4: Reaction curves under approximated welfare and tax revenue as objective functions



Note: Parameters: $\sigma = 5$, $a = (\sigma - 1)/\sigma$, $f = 1/\sigma$, $N = 1$, $\rho = 5$, $r_a = 0.1$, $\mu = 0.75$. Lighter contours indicate a higher equilibrium residential distribution λ. Best response tax rates are solid for region 1 and dashed for region 2. The welfare-maximizing tax rates are indicated by a circle, limit taxes are indicated by stars.

first-best tax rate. Under the approximated welfare function, the Stackelberg solution is never first-best.

A second objective function to compare and benchmark the results is a bureaucrat's objective function. For political reasons, or out of self-interest, governments maximize the aggregate tax revenue in their city. Governments of this type should be expected to yield lower welfare than benevolent governments do, because they attempt to grow the public sector larger than is efficient. This view of government is therefore much more pessimistic about the efficiency of governments than the view of governments as maximizers of local welfare. The check on the bureaucrat's type of government is that high taxes erode the tax base because they directly drive away firms, but also reduce wage and thus repulse migrants. The objective function is close to the Leviathan government objective, in which the government consumes all tax revenue. However, by consuming all tax revenue, a Leviathan government causes a leak of funds from the general equilibrium, which is less appealing in our framework.[5] Again,

[5]Nevertheless, a Leviathan objective function yields very similar results numerically.

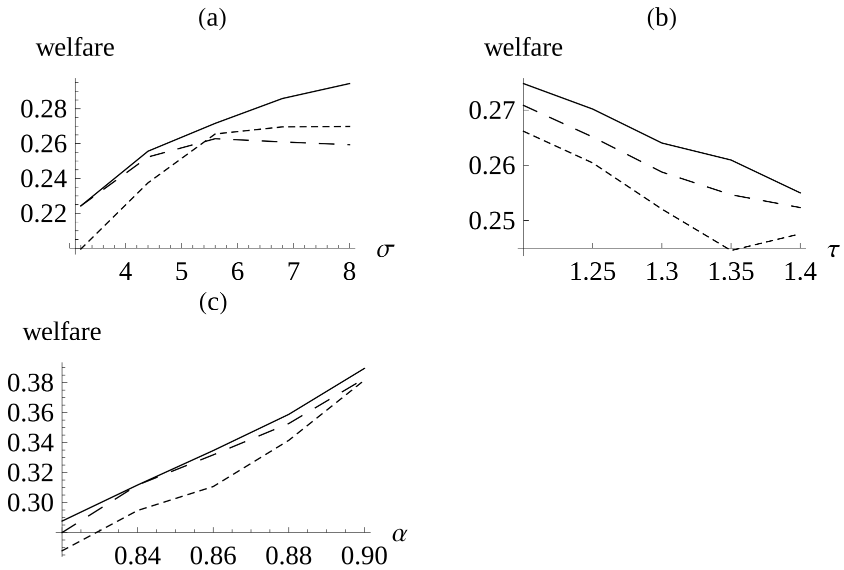
we study the tax rate that makes the large region preserve the agglomeration, although the objective function is now total revenue Tn instead of the welfare approximation.

The tax reaction functions for bureaucrat governments plotted in Fig. 3.4 (b) bear some resemblance to those of the reduced-form welfare objective in panel (a). The welfare for the equilibrium tax pair is higher, however at 98.3% of first best. From the reaction curves, under a bureaucrat's objective, the small region will switch to low tax rates even if the large region's tax rates are low. The incentive to attract a large share of firms is therefore stronger for bureaucrats than the tax distortion penalty is for governments maximizing approximated welfare (bureaucrat care about total revenue, approximated welfare takes revenue per head as an argument). The welfare among bureaucrat governments is also higher, however, than the worst outcome for benevolent governments. This suggests that the adoption of a bureaucrat's objective is potentially less harmful to welfare than the coordination issues that arise among benevolent governments. The large city is, again, larger than first best, but only by around 3 percentage points.

Sensitivity analysis

To further investigate the result that local benevolent governments' welfare may suffer from coordination problems, Figure 3.5 provides some sensitivity analyses. It varies several key parameters: the elasticity of substitution between manufacturing goods, transport costs, and the preference for government goods. In all exercises, the situation with locked-in governments (dotted line) potentially yields less than first-best outcomes. Note that the dashed-line reflects a worst-case scenario for locked-in governments, and that any outcome between the first-best and optimal welfare (solid line) is possible. The benchmark of the bureaucrat's government as an alternative source of policy inefficiency shows that over a large parameter range, the bureaucrat's equilibrium outperforms the worst-case lock-in situation. The welfare losses from coordination problems among benevolent governments are therefore possibly larger than the welfare losses of governments that maximize budget instead of welfare.

Figure 3.5: Welfare: optimal, coordination problems, and revenue maximization



Note: Optimal welfare (solid line), welfare under benevolent governments with coordination problems (dotted) and under governments that maximize tax revenue (dashed); for different parameters.

3.5 Conclusion

Government interaction in taxation has uncertain implications. Without agglomeration effects, harmful races to the bottom may emerge, while under perfect agglomeration, the picture is much less gloomy. This chapter investigates tax interactions if local governments face situations of imperfect agglomeration. This validates the critique on earlier tax competition models that agglomeration is unaccounted for. However, imperfect agglomeration also eliminates the agglomeration rents that arise with perfect agglomeration, which are responsible for many of the results in studies of tax competition with agglomeration effects.

Our results show that potential shifts of the partial agglomeration largely change strategic incentives compared to studies without scale effects. The desire to preserve the agglomeration pattern generates a no-relocation condition that both regions adhere to. This, however, defines the optimal response of both governments relative to other governments' tax rates. Therefore, many tax pairs qualify as mutual best responses. We do note that these results hold for partial agglomeration: if all mobile factors end up in one city or region, the incentive to preserve the agglomeration no longer exists (in fact, the small city no longer exists).

The multitude of equilibrium tax pairs suggests that an alternative major souce of government inefficiency may not stem from races to the bottom or government's strategic exploitation of scale effects, but from coordination failures. Because a large set of tax pairs can be sustained as best responses, a lock-in effect occurs. Once in equilibrium, there is no incentive for either government to deviate individually, whether the equilibrium is efficient or not. In our model, this leads to substantial potential welfare losses compared to optimal policies. Assuming that governments maximize budgets instead of welfare, which leads to inefficiencies due to political lack of interest in welfare (i.e., a Niskanen or bureacrat government) may yield higher welfare than the worst case lock-in scenario. Also, we show that using the approximated welfare function from earlier literature as a government objective can lead to substantially different conclusions if agglomeration is imperfect.

The policy recommendations from this chapter are rather different from other results in the tax competition literature. With the lock-in mechanism, the possibility arises that governments end up with tax rates that are undesirable, but the costs of individually setting another tax rate are high. If one of the governments commits to the first-best tax rate, the other will always select the tax that maximizes global welfare, so there may be a role for commitment devices or a central government. The role of the central government is unclear, however: if it could instate the first-best tax pair, there is no reason for the local governments to exist. Intervening in one city but not the other makes economic sense, but is legally and politically unrealistic. Traditional recommendations, like a tax floor or tax harmonization will not generally form Pareto-improvements in welfare.

The local (urban) level of analysis adopted here justifies most of the assumptions in this model, especially those on increasing commutes in jurisdictions and full mobility of all inhabitants. The commutes give rise to urban or congestion costs that present a dispersive force. It is feasible to replace centrifugal force of the commutes with housing costs (Helpman, 1998), lack of income effects (Borck and Pflüger, 2006), agricultural productivity changes (Puga, 2002) or generalized costs (Fujita et al., 2001, Chapter 14). Internal commutes form an intuitive spreading force in cities, but the results would not change much when adopting other centrifugal forces. However, full mobility of all inhabitants is a stronger assumption. Removing it, as many NEG models do by decoupling footloose entrepreneurs or capital from agricultural or low-skilled workers, may change the conclusions. When spatial allocations are still determined by utility equality of the mobile factors, then introducing the immobile factor (e.g. farmers) into the welfare equation will change the results. The reason is that governments will generally not be maximizing the same, shared welfare level because the immobile factor's welfare is not generally equal across regions. A central question for the nature of tax competition, therefore, is whether there are citizens that are absolutely immobile.

Lastly, the model complies with several empirical predictions of tax competition models. Tax competition models can be discerned by their empirical predictions: there are tax interaction effects, consistent with races to the bottom, or taxation of agglomeration, consistent with strategic use of scale. The empirical literature finds evidence of both, although the models are not generally theoretically consistent. Being a hybrid of the two, this chapter suggests that both effects can occur. However, rather than discriminating between the two models, it suggests that other mechanisms of government interaction are at play.

CHAPTER 4

URBAN POLITICS AND THE HOUSING MARKET: EXCESSIVE INDUSTRIAL POLICY

4.1 Introduction

In an ideal world, voters elect a policymaker to make optimal choices. The world is not ideal, if only because policymakers do not always select optimal tax rates or public services – as previous chapters on tax and fiscal competition showed. The mere mobility of firms may lead to overprovision of the public infrastructure that supports them (Fenge et al., 2009), and local firms, if sufficiently large, may exploit their bargaining position against local governments (King et al., 1993). If firms are aware of their impact on local public finances or employment, they could force an urban government to change policies in the firms' favor. The mere existence may thus lead policymakers with best intent to develop suboptimal policies. However, the world is not politically ideal either: the politician with the welfare-maximizing programme might never be elected.

Electoral accountability can make politicians take their voters' interest to heart. Analyses of urban politics show that citizens' influence in local policy or “voting by feet” leads local governments to make choices that maximize local well-being (e.g., Fischel, 2001). In this view, and contrary to tax competition, citizens' threats to leave or not re-elect the policymaker improves government policy instead of worsening it.

This chapter studies local politics and population mobility and reaches a conclusion opposite to "voting by feet": citizens' involvement in local politics does not lead to optimal policy. Studying policies for local business and industries, the chapter shows that elected policymakers overspend on industrial policies. It shares that conclusion with tax competition insights, but the source of inefficiency lies inside the city, not in the interaction between cities. Thus, this chapter casts doubt on the efficiency of local politics and voting-by-feet mechanisms.

In the model developed in this chapter, voters elect policymakers that are more pro-business than socially optimal. If voters' homes increase in value when employment opportunities increase, voters have a financial stake in a policy bias towards attracting industries. The homes closest to firms provide easy access to the labor market, and policies that increase local wages improve the value of such houses more than of others. If inhabitants own their home instead of renting it, mortgaging or sale allows owners to increase consumption if their home value increases. The homeowners closest to the central business district therefore prefer the largest expenditure on industrial policies. This political preference regarding changes in house prices, however, is not socially optimal.

Since homeowners consider home value in their political preferences, this chapter relates to the homevoter hypothesis (Fischel, 2001). The homevoter hypothesis holds that property gives homeowners a financial stake in their local community, and therefore motivates them to support activities that improve average well-being. The mechanisms of policy translating into houseprices are similar between the homevoter hypothesis and this chapter; and they are empirically well established. For instance, evidence shows that local public projects capitalize into house prices (Dehring et al., 2008). Local governments' choices in local policies such as education, transport and crime-fighting or the local effects of national policies are likely to impact on house values (Gibbons and Machin, 2008; Glazer and Van Dender, 2002; Hilber et al., 2011, among others). The financial interests of owning a house also lead homeowners to prefer different policies than non-homeowners do (Brueckner and Joo, 1991). Despite sharing such capitalization effects, the conclusions differ between the homevoter hypothesis and this chapter. The reason is that house price changes in the current model allow voters' political preferences to diverge from each other. The differences in political preferences render democratic voting or majority-rule voting inefficient, which is impossible with homogeneous voters.

To study the political effects of industrial policy via housing markets, the model needs (at least) two properties: i) there needs to be a government that cares about the votes of the inhabitants, and ii) inhabitants need to care about the value of their home, so they need to own it.

In the model, the government that sets industrial policies is run by an elected official, who cares about being (re-)elected. In urban economics, democratic decision-making is not a standard way to study policy formation, but it is a workhorse model in urban politics (Helsley, 2004). The relative realism of social planners and democratic government can be debated, but the preference of the median voter (who determines policies in the democratic models) is a reasonably accurate predictor of actual policy (Gramlich and Rubinfeld, 1982; Turnbull and Djoundourian, 1994;

Doi, 1999). On the more technical side, the urban model also lends itself well to the application of a median voter theorem. With homogeneous preferences and a single instrument, differences in voting behavior arise with the distance from the labor market. Therefore, political preferences have a single dimension and issues like non-single peakedness and multi-dimensional preferences are not relevant. The model of democracy builds on redistribution games among unequal voters (Romer, 1975; Meltzer and Richard, 1981). The difference between the average and median voter's preferences similarly leads to inefficient policy. In the current model, however, differences between voters are not assumed but follow from the equilibrium on the land market; moreover, the model allows free entry and exit from the jurisdiction.

To question whether democratic urban industrial policies are efficient, this chapter considers a government that sets local taxes to finance public inputs to firms. This captures the idea that urban governments often do not tax firms, but rather employ location policies like infrastructure investment, zoning and the provision of facilities. Therefore, we ignore public services like safety and education, which benefit citizens rather than firms. The focus on industrial policies draws a parallel with tax competition models (Fenge et al., 2009; Ross and Yinger, 1999), but the sources of policy bias towards firms are found inside the city, rather than in firms' mobility. Apart from industrial policies, the capitalization arguments put forward in this chapter have also been developed for zoning, property development laws and schools (Hilber and Robert-Nicoud, 2013, 2007; Hilber and Mayer, 2009), although in a different setting.

The effects of homeownership in politics only surface if voters can buy homes and sell them at a later point in time. In the model developed below, homes (or land, effectively) are durable goods. The option to retrade the durable good changes the consumer's behavior compared to ordinary (perishable) goods (Waldman, 2003). An increase in home value allows a consumer to expand consumption because it increases the lifetime wealth, or relaxes credit constraints, for instance (Campbell and Cocco, 2005). Bostic et al. (2009) and Iacoviello (2011) document such housing wealth effects, that the model below incorporates. Essentially, a house can be viewed as a financial asset, even if it is "consumed" by living in it. The homeowner's concern with the home's value can be substantial – especially since housing is the single largest item in most households' budgets – but standard urban economic models cannot explain that concern (Ortalo-Magné and Prat, 2010). In a standard renter's market (e.g., Fujita, 1989), increases in home value imply a larger rent, which is unambiguously bad news for the tenant. In a purely static framework, buying or renting a house is financially the same. In the model below, consumers possess the home they bought earlier. If housing enters the consumer's balance as wealth, house price increases benefit the owner via larger consumption options. The model below uses a reduced-form where housing wealth enters the total wealth of the consumer, so that he cares differently about house prices than a renter would. The microfoundations of this formulation follow Mankiw (1982).

The results show that in a democratic urban model with homeowners, policies are generally not efficient. The median voter tends to live closer to firms than the average voter, and therefore prefers wage-increasing policies more. Wage increases are perfectly reflected in house prices. The increased costs of the home cancel with the wealth effect of owning a home. This causes homeowners to develop different political preferences. Renters, by contrast, see any wage improvement perfectly reflected in their rents, and therefore have homogeneous preferences for the optimal policy. The model shares the conclusion of overexpenditure on industrial policies with the tax competition literature, although the underlying mechanism is entirely different. In an open city with production externalities, the results persist. The median voter has a higher than average exposure to wage improvements and prefers more public expenditure on productivity; production externalities reinforce that effect. Therefore, productive spillovers that lead to external returns to scale exacerbate the inefficieny. Increasing returns to scale in public productivity improvements reduce the inefficiency. If public expenditure has increasing returns, the median voter has an incentive to maximize the tax base. With free entry and exit of citizens in the city, the tax base is higher if the tax rate is closer to socially optimal and average utility is high.

The results also shed new light on the "renter effect" in local public finance. Oates (2005) contends that cities with a higher share of renter-occupied housing have policies that are substantially more directed toward public consumption than to industrial policies. Collecting evidence from various studies, Oates shows that higher renter shares in cities are statistically associated with larger general expenditure, general local services to citizens, police and fire-fighting services, public works, parks and recreational expenditure, and education. Oates argues that in a well-functioning housing market, the taxes that finance such public expenditure should either be paid for via taxes, or via higher housing rents, i.e., the benefits of public expenditure should be fully capitalized. Therefore, the demand for local public spending should not differ between home-owned or renter occupied housing. The observed differences may partly be explained by renters' illusion that they do not pay the land taxes, or by actual imperfect pass through of taxes into rents (Carroll and Yinger, 1994). The model in this chapter, combining urban democratic government with durable housing shows that while inefficient, it may be rational for homeowners to vote for broader support to firms than renters do.

The remainder of the chapter is organized as follows. The following section presents the setup of the monocentric city model and introduces a democratic government to the city. Section 4.3 discusses the economic and political equilibrium, and the optimality of democratic policy. Section 4.5 provides conclusions and a discussion.

4.2 Model

This section describes a monocentric city model, inhabited by N citizens. The center of the city hosts all firms in a Central Business District (CBD), while workers buy the land around the city center. The city is surrounded with agricultural land. The city is small compared to the rest of the (national) economy and we shall assume that there is a frictionless financial and housing market. The freely traded manufacturing good is the numéraire good. Consumers have a preference over housing, manufacturing goods and leisure. In what follows, the home of a voter is the plot of land that he posseses – we leave the capital ("bricks") of the house out of consideration. The distance from the CBD determines the voter's commuting costs, θ.[1] In the model, citizens vote over a productivity-enhancing policy. We shall discuss, in turn, households, firms, the market equilibrium and the political equilibrium.

Households

In this model, land is owner-occupied. Moreover, land is durable, in the sense that after living on it, it may be sold to a next owner. The model

[1] The commuting costs can be thought of as the distance from the CBD to the worker's home, multiplied with the commuting time per unit of distance. However, as commuting costs are the only metric of space inside the city relevant to the equilibrium prices, the equilibrium can be expressed in commuting costs. The specification of units of distance (which would be one of the numéraires) is hence not needed.

uses a reduced-form description of home ownership inspired from Mankiw (1982), except that consumers in this setup consume a non-durable numéraire good next to durable land.

Homeownership in this model modifies the urban economics tradition of (absentee) landlords and renters. As a result, any land value increase is appropriated by citizens, who are also voters. A mix of owner and renter-occupied housing would undoubtedly be more realistic for most cities. But the current assumption of full homeownership throughout the city is sufficient to develop the argument that voters' possesion of homes leads to different political outcomes.

Using c to define consumption of the manufacturing good, h for land consumption (the consumer's home), l_f for leisure, the utility function of a consumer is:

$$U = c^{\alpha} h^{1-\alpha} {l_f}^{\lambda}. \tag{4.1}$$

The consumer faces two constraints when optimizing the utility function. First, there is a financial constraint. The net income from working is denoted as $l_s w - T$, the labor supply multiplied with the wage rate less a lump-sum tax T. The consumer possesses a home when he makes his consumption decisions and political decisions. As the model's land market is frictionless, the land in possession can be viewed as wealth. A citizen could sell his land and use it to buy land or consumption, or in equilibrium, to buy back his own land. Therefore, the current value of the land that the citizen owns is on his balance at the moment he makes his consumption decision. The land wealth of a consumer is $h_0 r\left(\theta_0\right)$, where h_0 is the land volume acquired earlier, and $r\left(\theta_0\right)$ is the current price of that land. The citizen can relocate, so the value of the land that he owns at commuting distance θ_0, $r\left(\theta_0\right)$, does not need to equal the price of the land that he buys ($r(\theta)$) at commuting distance θ. Using the land wealth, the financial constraint is $l_s w - T + h_0 r\left(\theta_0\right) = hr(\theta) + c$. The left-hand side equals the consumer's wealth, the right-hand side equals the expenditure on land (at price $r(\theta)$) and on the numéraire consumption good. Allowing the previously bought stock of land to enter the consumer's wealth is consistent with a fully forward-looking Mankiw (1982) consumption problem, where land is a durable good. That dynamic consumption problem could be used to provide the microfoundations to our budget constraint. However, the static model makes the results easier to see. Moreover, the results from the static model hold both in and outside the steady-state in the dynamic model, so the omission of dynamics does not change the results.

The second constraint on the consumer limits the time available for work, commuting and leisure to l_t units of time. Consumers at a larger distance from the CBD face a larger commuting time, which is indexed by the term θ. The commuting time l_c is proportional to working time l_s, so that when the worker decides to work one day less, he also reduces the number of round trips by one (Verhoef, 2005). The commuting time $l_c = \theta l_s$ is therefore proportional to the time spent travelling θ and the number of trips (the labor supply). The time spent working and commuting is then $l_c + l_s = (1+\theta) l_s$. Using this, the time constraint is:

$$l_t = l_f + l_c + l_s = l_f + (1+\theta) l_s, \tag{4.2}$$

where the costs of commuting, θ, increase in the distance from the city center to the worker's home.

The choice of leisure determines the financial budget, because increased leisure reduces labor supply: $l_s w = \left(l_t - l_f\right) w/(1+\theta)$. Substituting this expression for labor income $l_s w$ into the financial constraint gives a generalized constraint that relates expenditure to income and leisure time. The generalized constraint is:

$$c + r(\theta) h = \frac{l_t - l_f}{1+\theta} w - T + r\left(\theta_0\right) h_0. \tag{4.3}$$

This generalized constraint helps to solve the consumer problem. The consumption of leisure time determines the amount of hours spent commuting and working. The first-order condition of the utility function with respect to leisure is given by:

$$\frac{dU}{dl_f} = \frac{dU}{dc} \frac{w}{1+\theta}. \tag{4.4}$$

In the optimum, the marginal utility of leisure is equal to that of consumption, multiplied by the opportunity costs of leisure, $w/(1+\theta)$. This opportunity cost reflects the rate at which marginal utility derived from leisure can be transformed into marginal utility of consumption. One unit less leisure frees up one unit of time for working and commuting. Only fraction $1/(1+\theta)$ of that time translates into labor supply (and not commuting), so financial income increases at rate $w/(1+\theta)$. The optimal consumption of leisure maximizes utility subject to the generalized budget constraint. Since the expenditure shares of the generalized budget are constant under the Cobb-Douglas formulation above, and the opportunity costs of leisure are equal to the shadow price of the time endowment, the worker consumes a fixed proportion out of his time in terms of leisure: $l_f = \lambda/(1+\lambda)\, l_t$. Consequently, the worker spends the time left apart from leisure on commuting (fraction $\theta/(1+\theta)$) and working (fraction $1/(1+\theta)$):

$$l_s = \frac{l_t - l_f}{1+\theta} = \left(1 - \frac{\lambda}{1+\lambda}\right) l_t/(1+\theta). \tag{4.5}$$

Since we have not chosen a unit of time, we are free to normalize it. Choosing the time endowment $l_t = 1+\lambda$, the effective labor supply curve can be written as:

$$l_s = (1+\theta)^{-1}. \tag{4.6}$$

The time spent working in the CBD depends inversely on the travel time from a worker's residence to the CBD. Given the labor supply, the financial wealth is $Y\left(\theta, r\left(\theta_0\right) h_0\right) \equiv w/(1+\theta) - T + r\left(\theta_0\right) h_0$. Solving for the demand functions for housing and numéraire consumption gives:

$$\begin{aligned} c &= \alpha Y\left(\theta, r\left(\theta_0\right) h_0\right), \qquad (4.7)\\ h &= (1-\alpha)\frac{Y\left(\theta, r\left(\theta_0\right) h_0\right)}{r}. \end{aligned}$$

Finally, using the demand functions in the direct utility function, the indirect utility function is:

$$\begin{aligned} V &= \zeta \frac{w/(1+\theta) - T + r h_0}{r^{1-\alpha}} = \zeta \frac{Y\left(\theta, r\left(\theta_0\right) h_0\right)}{r^{1-\alpha}}, \qquad (4.8)\\ \text{with } \zeta &= \alpha^{\alpha}(1-\alpha)^{1-\alpha}\lambda^{\lambda}. \end{aligned}$$

The term ζ is a positive constant. The indirect utility is equal to the financial wealth of a citizen, divided by the consumption price index (which includes the numéraire). The land price features in the financial wealth Y and in the price index: a higher land price increases the financial budget, but also increases the costs of living.

Production

Firms are located in the CBD, and produce using labor inputs under constant internal returns to scale. Firms perceive themselves as small and they act as price-takers on the goods market and on the labor market. Public inputs are required for production. Public input productivity is captured in the production shifter I^{ε_I}. The term I reflects the funds devoted to public inputs, transformed via a function satisfying Inada conditions. The Inada conditions imply that firms cannot operate without a minimal support from government (e.g., some roads, electricity and administration are required), but the marginal returns to using more public funds fall if more are used. This could be attributed to the selection of the most productive project first, and rules out explosive situations in which taxes fuel further productivity. The production technology is:

$$q = I^{\varepsilon_I} l, \ 0 < \varepsilon_I < 1. \tag{4.9}$$

The internal constant returns to scale yield a competitive wage on the labor market that is independent of the number of firms. The first-order condition for hiring labor, under the assumption of unit prices is:

$$w = I^{\varepsilon_I}. \tag{4.10}$$

Since public infrastructure is provided in the center of the city, the city consists of a central business district with a residential district around it.[2]

Policy

The task of the government is to raise lump sum taxes and supply public inputs for production. Lump sum taxes ensure that the policy has spatially differentiated effects: the tax is uniform but the benefits vary over space. Thus, the lump sum tax represents a redistributive policy along the distance from the CBD, or voters' centrality. The government budget is balanced. The function by which tax revenues (tax times number of tax payers) transform into public inputs, and the ensuing competitive wage rate are:

$$I = TN; \ w = (TN)^{\varepsilon_I}. \tag{4.11}$$

[2]External returns to scale can explain the formation of the Central Business District. Introducing such returns does not change the results but obscures the analytical steps towards the results. External returns play a larger role in the open-city version of the model. That version of the model (in the next section) uses a production externality.

4.3 Equilibrium and policy

The political preferences of inhabitants depend on their residential location in the city. To describe the political outcomes, we first need to establish the equilibrium on the land market, and consequently, on the labor and goods market. Given market clearing, this section studies the democratic tax rate in the city.

The model aims to compare the democratic policy to the socially optimal policy. If these involve different tax rates, they also lead to different equilibria on the labor, goods and land market. The democratic equilibrium occurs when markets are in equilibrium and there is no incentive to change policies. The land market determines the location of the median voter, and the median voter's political preferences depend on his location. A stable democratic tax rate requires that the median voter's location and his vote are consistent.

Secondly, homeowners all own a plot of land. In equilibrium, no citizen has an incentive to change his land consumption. Therefore, when evaluating policy, we shall assume that without a tax change, inhabitants will keep living on the same plots. In that case, the land in possession (that is counted as wealth), is generally consistent with the land consumption. This equilibrium assumption is arbitrary from the static perspective because we have not defined how inherited land h_0 was initially distributed. However, as long as land ownership is positive, this does not change the results.

Land, labor and goods market equilibrium and politics

The willingness to bid for land determines the price of land. Locations closer to the central business district allow for shorter commutes and are more expensive. In the land market equilibrium, nobody can improve utility by marginally moving towards or away from the CBD. The equilibrium condition balances the costs of owning land with the marginal contribution to financial wealth:

$$\begin{aligned} 0 &= \frac{dV\left(\theta, r\left(\theta_0\right)h_0\right)/d\theta}{V\left(\theta, r\left(\theta_0\right)h_0\right)} \qquad (4.12) \\ &= \frac{dY\left(\theta, r\left(\theta_0\right)h_0\right)/d\theta}{Y\left(\theta, r\left(\theta_0\right)h_0\right)} - (1-\alpha)\frac{dr(\theta)/d\theta}{r(\theta)}. \end{aligned}$$

Rearranging leads to an integrable expression:

$$\int \frac{dr(\theta)/d\theta}{r(\theta)} d\theta = \frac{1}{1-\alpha} \int \frac{dY\left(\theta, r\left(\theta_0\right)h_0\right)/d\theta}{Y\left(\theta, r\left(\theta_0\right)h_0\right)} d\theta. \tag{4.13}$$

Evaluating these integrals, along with the boundary condition that the cost of land is r_a (the returns to alternative land use) at θ_{edge} gives:

$$r(\theta) = \left(\frac{Y\left(\theta, r\left(\theta_0\right)h_0\right)}{Y\left(\theta_{edge}, r\left(\theta_{0edge}\right)h_{0edge}\right)} \right)^{1/(1-\alpha)} r_a. \tag{4.14}$$

Closer to the CBD, financial income is higher (lower commuting time leads to higher labor supply), so the willingness to pay for central lots of land is higher. The land wealth of an inhabitant at location θ is left implicit here. This does not affect the results; a discussion of the (heterogeneous) wealth effects is given in the Appendix to this chapter. The clearing condition on the housing market states that every worker is housed. The residential density is equal to the inverse of land consumption per head, $h(\theta)$. The population size is equal to the aggregate density over the city: $N = \int_0^{\theta_{edge}} 1/h(\theta)\, d\theta$. Since every head provides $(1+\theta)^{-1}$ units of labor, the corresponding aggregate labor supply equals $L_s = \int_0^{\theta_{edge}} 1/[(1+\theta)h]\, d\theta$.

Optimal and democratic tax rates

To compare the political outcomes in the model, we first benchmark the social optimal policy. If the land market is in equilibrium, citizens have no incentive to change their land consumption, and land price differentials reflect the utility differential of living in different locations. In that case, the welfare function to be optimized is the aggregate production less the taxes spent on improving productivity:

$$\int_0^{\theta_{edge}} \frac{1}{h(\theta)} \frac{w}{1+\theta} d\theta - NT, \tag{4.15}$$

or, dividing by the number of workers, the average net labor income:

$$\begin{aligned} W &= w\bar{l}_s - T, && (4.16)\\ \text{with } \bar{l}_s &= \int_0^{\theta_{edge}} \frac{1}{h(\theta)(1+\theta)} d\theta / N, \end{aligned}$$

where $\bar{l}_s$ is the average labor supply. The welfare function is optimized when the after-tax income of an average voter (i.e., one that offers the average labor supply) is maximized. Optimizing the welfare function with respect to the tax rate gives the first-order condition:[3]

$$\frac{dw}{dT}\bar{l}_s - 1 = 0. \tag{4.17}$$

This efficiency condition lists two effects of a tax increase: the resulting higher government budget increases the wage rate through productivity-enhancing inputs; and the direct reduction in net income decreases consumption. Rewriting the first-order condition for the tax rate using the government technology (eq. 4.11) gives:

$$T^* = \varepsilon \bar{l}_s w. \tag{4.18}$$

The efficient tax (denoted by an asterisk) is a fraction ε of the average labor income.

A voter that possesses land develops different political preferences. The voter's preferred policy (or bliss policy) is the tax rate that maximizes his indirect utility function. The voter does not care about average labor income, but about his personal income only. On the other hand, the voter cares about land prices, because he owns a plot of land, and intends to buy one. Optimizing the indirect utility function of the voter with respect to the tax rate gives the first-order condition:

$$\begin{aligned} \frac{dV(\theta)/dT}{V(\theta)} = \; & 0 = \frac{\frac{1}{1+\theta}\frac{dw}{dT} - 1}{w/(1+\theta) - T + rh_{-1}} \\ & + \frac{r(\theta_0)h_0}{w/(1+\theta) - T + rh_{-1}} \frac{dr(\theta_0)/dT}{r(\theta_0)} \\ & -(1-\alpha)\frac{dr(\theta)/dT}{r(\theta)}. \end{aligned} \tag{4.19}$$

The second line describes the effects of a tax change on the voter's utility via land wealth changes. If the tax increases improves the value of

[3] By the envelope theorem, the marginal change of $\bar{l}_s$ due to a change in tax rates plays no role in the welfare function. The average labor supply increases if the city is compressed through increased taxes, but that also decreases the average land consumption. Since the land rent reflects the utility differential, the consumer is neutral to moving closer to the CBD at the cost of lower land consumption. The same first-order condition is hence obtained by maximizing the sum of utility functions with the consumption of land inserted.

the land that the voter sells, the voter is better off. The third line, however, lists the opposite effect: if the tax increases the price of the land the voter buys, he is worse off. If the land market is in equilibrium, no citizen wishes to change his plot of land. In that case, it holds that the citizen's optimized expenditure on land is equal to a share $1-\alpha$ of his wealth: $hr(\theta) = h_0 r\left(\theta_0\right) = (1-\alpha)Y$. Moreover, in equilibrium, the location that a voter sells and the location that he buys virtually coincide. Therefore, the difference between $dr\left(\theta_0\right)/dT/r\left(\theta_0\right)$ and $dr(\theta)/dT/r(\theta)$ is arbitrarily small. If the shares of land wealth and land expenditure coincide and the land price changes are (virtually) equal, the second and third line from the first-order condition cancel. In equilibrium, the voter sells and buys the same plot of land, so he becomes indifferent about its price; the increased land wealth cancels with the increased land expenditure. This is the intuitive argument. We provide a more formal discussion in the Appendix of this chapter.

The voter's first-order condition (eq. 4.19) is met if $\frac{1}{1+\theta}\frac{dw}{dT} - 1 = 0$. Inserting the derivative of the wage (eq. 4.11) with respect to taxes gives:

$$T(\theta) = \frac{\varepsilon w}{1+\theta} = \varepsilon l_s(\theta) w. \tag{4.20}$$

The preferences of voters vary according to how much labor they choose to offer. Their labor supply is exclusively determined by the distance to the CBD: voters further away from the center prefer lower taxes, because they benefit less from the wage-improving effects of high taxes.

The median voter's political preferences differ from the average voter's political preference. The median voter lives closer to the CBD and supplies more labor than the average voter. Therefore, he prefers higher tax rates than is socially optimal. To see this, we compare the position of the average and median voter. Because workers further from the CBD supply less labor, the preferred tax rate strictly decreases with the distance from the CBD. Therefore, half of the citizens lives closer to the CBD than the median voter does. The sum of residential density between the CBD ($\theta = 0$) and the median voter's location θ_m must be equal to 50% of the city's population. Using inverse land consumption as the population density, the median voter's location satisfies the identity:

$$\frac{1}{2}N = \int_0^{\theta_m} \frac{1}{h(\theta)} d\theta, \tag{4.21}$$

where θ_m is the location of the median voter. The labor supply of the average worker $l_s\left(\bar{\theta}\right) = 1/\left(1+\bar{\theta}\right)$ satisfies:

$$\frac{1}{1+\bar{\theta}} = \frac{1}{N}\int_0^{\theta_{edge}} \frac{1}{h(\theta)}\frac{1}{1+\theta}d\theta. \tag{4.22}$$

Dividing equations 4.21 and 4.22 and rewriting gives that:

$$\frac{1}{2} = \frac{\int_0^{\theta_m} \frac{1}{h(\theta)}\frac{1}{1+\bar{\theta}}d\theta}{\int_0^{\theta_{edge}} \frac{1}{h(\theta)}\frac{1}{1+\theta}d\theta}. \tag{4.23}$$

The numerator is the population density weighed by average labor supply, the denominator is the aggregate labor supply. The inverse land consumption $1/h(\theta)$ is strictly downward-sloping away from the CBD: residential density decreases away from the labor market. Through the higher population density, the units of land closest to the CBD account for more labor supply. Therefore, the median voter's distance from the CBD is lower than that of the worker with the average labor supply. The position of the median voter θ_m is more central than that of the average voter, so $\theta_m < \bar{\theta}$.

Figure 4.1 shows the preference for tax rates for the median voter. The figure first assumes that the optimal tax is set, and that markets clear given this tax rate. Given this equilibrium, Figure 4.1 plots the utility function of the median voter for different tax rates. The top of the utility function is not at the optimal tax rate; but to the right of it. If the city is in an equilibrium of socially optimal policy, there is an incentive for the median voter to select a higher tax rate. For the median voter, increasing the tax rate increases his wage strongly, thus increasing his net income.

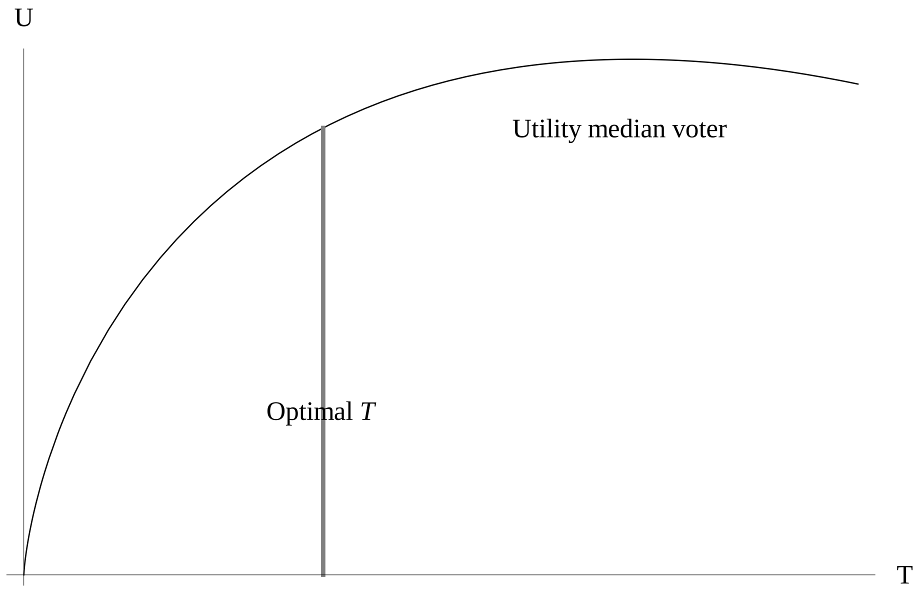
The median voter's bias from optimal tax rates depends on how far away he lives from the voter with average labor supply. The bias can be captured as $\Delta T = T\left(\theta_m\right) - T^*$. Inserting the equilibrium and optimal tax rates (eqs. 4.20 and 4.18) and using the competitive wage rate (eq. 4.11), the difference between equilibrium and optimal tax rate is:

$$\Delta T = \varepsilon_I N^{\varepsilon_I}\left[\frac{T\left(\theta_m\right)^{\varepsilon_I}}{1+\theta_m} - \frac{T^{*\varepsilon_I}}{1+\bar{\theta}\left(T^*\right)}\right], \tag{4.24}$$

where $\bar{\theta}\left(T^*\right)$ results from the optimal tax rate and θ_m from the equilibirum tax rate, $T\left(\theta_m\right)$.

Figure 4.2 provides an intuition for the political equilibrium. It provides the voter's preferred tax and the location of the median voter in a

Figure 4.1: Utility from different tax rates

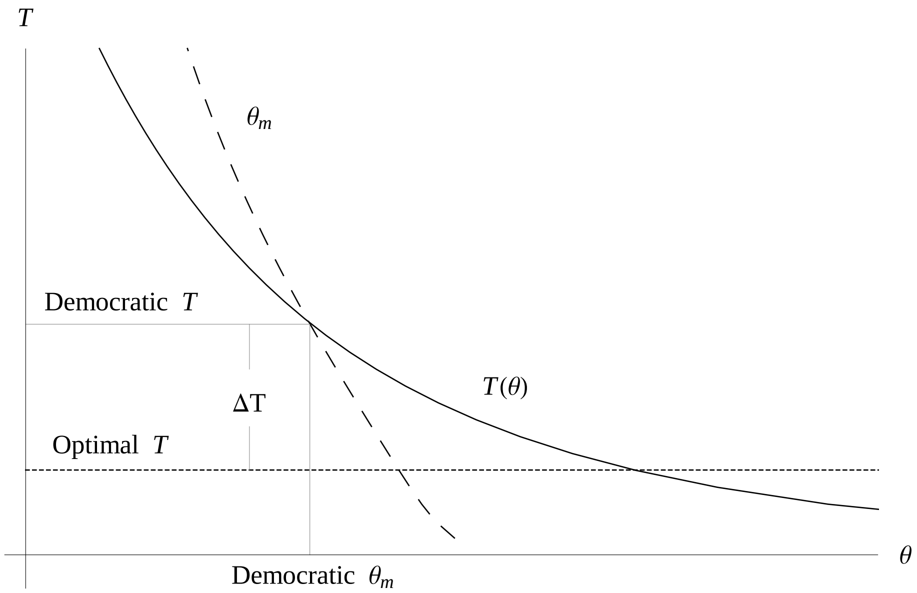


Note: Median voter's payoff to different tax rates given an equilibrium with social optimal policy.

space of the distance to CBD (θ) and the tax rate (T). For every location θ the solid line maps the tax rate that a voter from that location votes for. Only the median voter's preferred tax is instated in equilibrium, however. The dotted curve plots the correspondence between the location of the median voter and the prevalent tax rate. Given the tax rate T, markets clear and half of the population ends up more central than θ_m. As said, higher taxes reduce the distance of the median voter from the CBD because central locations give better access to higher wages and become more densely populated. In the joint market- and political equilibrium, the median voter votes for a policy that is consistent with him being the median voter. If the median voter lives closer to the CBD than the equilibrium predicts, he selects a tax rate that decentralizes the population. If the median voter lives further away, he selects a tax rate that causes residential density to grow in central areas, thus making a more central voter the median voter.

So why is homeownership essential for the political bias of voters? Suppose that voters are renters for the moment. The rent is proportional to land value, so the bid rent for land is the same as before. However,

Figure 4.2: Political and economic equilibrium



Note: $T(\theta)$ gives the tax that a voter at location θ votes for. The locus θ_m shows for every T where the median voter lives given market clearing. The political equilibrium is at the intersection of $T(\theta)$ and θ_m. The socially optimal tax rate is plotted in dots. The difference between equilibrium and optimal tax rates is ΔT.

land ownership no longer functions as wealth: the financial budget is $w/(1+\theta)-T$. For renters, the first-order condition becomes:

$$\frac{dV(\theta)/dT}{V(\theta)} = 0 = \frac{\frac{1}{1+\theta}\frac{dw}{dT}-1}{w/(1+\theta)-T+rh_{-1}} - (1-\alpha)\frac{dr/dT}{r}. \quad (4.25)$$

Differentiating the land rent (eq. 4.14) with respect to taxes gives:

$$\frac{dr/dT}{r} = \frac{1}{1-\alpha}\left(\frac{dY(\theta)/dT}{Y(\theta)} - \frac{dY\left(\theta_{edge}\right)/dT}{Y\left(\theta_{edge}\right)}\right). \quad (4.26)$$

Using this expression for $dr/dT/r$ in the first-order condition (eq. 4.25), the first-order condition becomes:

$$\frac{dV(\theta)/dT}{V(\theta)} = \frac{\frac{1}{1+\theta_{edge}}\frac{dw}{dT}-1}{w/\left(1+\theta_{edge}\right)-T} = 0, \quad (4.27)$$

which is independent of the voter's location θ. The intuition is that all voters maximize the same utility level, evaluated here at θ_{edge}. The common utility function is a result of rental market clearing: the rental market equilibrium eliminates all differences in utility, so voters maximize a common welfare function. For homeowners, the wealth effect would cancel with the rent increase. The first-order condition for homeowners and renters is the same, save the landwealth effect (the second term in eq. 4.19) The effects of a tax change in utility are therefore not uniform across space if citizens own the land.

Finally, these conclusions are derived under the assumption that citizens enter the model with a stock of land that is consistent with their demand for land. The equilibrium could be interpreted as a steady state in the sense that with no tax changes, the land market equilibrium is unchanged. The assumption of entering with an optimal amount of land is not critical to the results, however. As can be seen from the comparison between the homeowner's first-order condition (eq. 4.19) and that of the renter (eq. 4.25), the homeowner has an incentive to vote for taxes higher than the optimal tax rate as long as he owns land ($h_0 > 0$). This property of the model is important, because tax changes yield different results for voters on different distances from the CBD. For that reason, we require the results to be robust to wealth heterogeneity. If some consumers have a larger wealth than others, wealthier consumers choose to reside closer to the city's boundary (Fujita, 1989). They consume larger plots of land and poorer consumers near the center buy relatively smaller pieces of land, so wealth heterogeneity shifts the electoral mass towards the CBD. Graphically, in Figure 4.2, this heterogeneity will increase the curvature of the median voters's location (θ_m) line, shifting the intersection leftward. Differences in wealth should thus lead the location of the median voter and the average voter to diverge, and reinforce the conclusions of this chapter.

4.4 Policy in an open city

The closed city model makes an argument that the vested interests of incumbent citizens provide a source of political bias away from the optimal policies. In comparison to government interaction models, however, an important argument is missing: potential influxes of new inhabitants or production factors may change tax-setting behavior. The effects of inhabitant's mobility on the political outcomes are not clear-cut in advance. Clearly, if there are externalities to labor supply, incumbent voters may support policies that attract new residents whose presence increases the local productivity. However, if the median voter's policy is too far off the social optimum, the utility of living in the city is low and inhabitants will be chased away, rather than attracted. Thus, citizens' mobility might also discipline the median voter.

Agglomeration externalities have been completely ignored in the closed city model. In an open city, workers only enter large cities if there is a benefit to balance the large commuting costs. This section adds production externalities to provide a consistent explanation for city size.

We assume that the productive externalities take a Sheshinski (1967) form and that a scale parameter δ governs them. The total labor supply is L_s, and the externalities act as a production shifter L_s^{δ}. The production function and competitive wage rate are updated from equations 4.9-4.11 as:

$$q(i) = \left(\frac{I}{N^{\xi}}\right)^{\varepsilon_I} L_s^{\delta} l(i); \quad w = \left(\frac{I}{N^{\xi}}\right)^{\varepsilon_I} L_s^{\delta}. \tag{4.28}$$

Because the number of citizens can vary in the open city, it is now relevant to think about the scale effects in productivity. The scale externalities in public expenditure are governed by ξ. If the productivity effect is public (non-rivalrous; $\xi = 0$, public productivity equals I^{ε_I}), increases in city size imply an increase in overall productivity. In that case, there is a city-level scale effect stemming from the public provision of productive inputs. If productivity improvements are fully rivalrous ($\xi = 1$, public productivity equals $(I/N)^{\varepsilon_I}$), the productivity increases in tax per head, ruling out such scale effects. The total government budget $I = TN$ itself also depends on the number of people that pay taxes.

Firms in the competitive market are small enough to take aggregate employment L_s as given, so the productive externalities translate into higher wages. Totally differentating w reveals that the tax affects the wage rate via two channels:

$$\frac{dw/dT}{w} = \delta \frac{dL_s/dT}{L_s} + \varepsilon_I \left(\frac{1}{T} + (1-\xi) \frac{dN/dT}{N} \right). \tag{4.29}$$

First, the aggregate labor supply increases wages if the productive externality is positive ($\delta > 0$). Second, both the tax rate and the number of people that pay the tax affect the budget for public inputs, thus increasing productivity and wages. Citizens can freely enter the city, and will do so if the utility of living in the city exceeds a reservation utility $\bar{u}$ (i.e., the steady-state utility that can be obtained by living in other cities). Equally, citizens can leave if their prospective utility is higher elsewhere. Therefore, the equilibrium condition for migration is that the utility of living in the city is equal to the reservation utility (which is the maximal utility that can be obtained elsewhere). Evaluating the utility at the edge of a city, the no-migration condition reads:

$$\bar{u} = \frac{\frac{w}{1+\theta_{edge}} - T}{r_a^{1-\alpha}}. \tag{4.30}$$

The democratic tax rate is the tax rate that maximizes the median voter's utility. The median voter's optimization is the same as in the closed city; equally, he is neutral to land price changes as a land owner. However, there are two additional effects on the wage rate compared to the closed city: the externality on employment and the sensitivity of the tax base. The first-order condition of the median voter's utility function with respect to the tax rate is:

$$\delta \frac{dL_s/dT}{L_s} + \varepsilon_I \left(\frac{1}{T} + (1-\xi) \frac{dN/dT}{N} \right) - \frac{1}{wl_s\left(\theta_m\right)} = 0. \tag{4.31}$$

Rearranging for the democratic tax rate gives

$$T\left(\theta_m\right) = \frac{\varepsilon_I w l_s\left(\theta_m\right)}{1 - \varepsilon_I w l_s\left(\theta_m\right)\left[\delta \frac{dL_s/dT}{L_s} + \varepsilon_I (1-\xi) \frac{dN/dT}{N}\right]}. \tag{4.32}$$

As in the closed city, the preferred tax is a share ε_I of the median voter's labor income. The denominator of the democratic tax rate (eq. 4.32) captures the effects on labor externalities and the tax base, however. If an increase in the labor supply following tax increases is large (dL_s/dT is large), the voter benefits more from tax increases and his preferred tax rate is higher. If tax increases enlarge the tax base ($dN/dT > 0$) and the public productivity is incompletely rivalrous ($\xi < 1$), the voter has an incentive to vote for higher taxes. The opposite is true if tax increases reduce the tax base.

For a comparison with the socially optimal tax rate, like in the closed-city case, we maximize the welfare function $w\bar{l}_s - T$. In addition, if cities are perfectly governed, throughout, the term $\bar{u}$ is maximized (or city size is optimal). Therefore, under optimal policy, marginally changing the tax rate has no effect on the number of inhabitants: $dN/dT = 0$. Collecting these, the first-order condition for the welfare-maximizing tax rate is:

$$\delta\frac{dL_s/dT}{L_s} + \varepsilon_I\frac{1}{T} - \frac{1}{w\bar{l}_s} = 0, \tag{4.33}$$

or rewriting for the optimal tax rate using eq. 4.29 and $dN/dT = 0$:

$$T^* = \varepsilon w\bar{l}_s\frac{1}{1-\varepsilon w\bar{l}_s\delta\frac{dL_s/dT}{L_s}}. \tag{4.34}$$

Comparing the social optimal and the democratic tax rate (eq. 4.32) reveals two sources of inefficiency in the democratic tax rate. As before, the median voter supplies more than average labor, leading him to prefer higher public spending on productivity. A parallel sensitivity occurs in the open city (with externality) case. Productive externalities and immigration enter the social and democratic first-order condition symmetrically. However, the median voter's closer than average location to the labor market makes him more sensitive to anything that raises wages. Therefore, in addition to the direct wage-raising effect of a tax increase, possible productive externalities and tax base improvements benefit the median voter more than the average voter.

The median voter is biased towards higher taxes overall, but the scale effects from increased population lead to inefficiencies in opposite direction. Firstly, reasoning from the social optimal equilibrium, increases in the tax rate improve wages (otherwise a positive tax could not be optimal). Since the median voter benefits more than the average voter from wage increases, he has an incentive to deviate marginally from the social optimal tax rate. At the social optimal tax rate, raising taxes increases labor supply and therefore increases wage via the productive externality. The wage increases more strongly if the externality is strong (δ is large). Secondly, however, setting the tax rate higher than the social optimal tax rate also reduces equilibrium utility in the city. Reducing equilibrium utility leads to migration out of the city. Whether such migration affects the median voter (apart from the production externality) depends on whether public inputs are rivalrous. Tax-induced out-migration leads to a higher government budget per head but lower aggregate government budget. If the public inputs are not perfectly rivalrous ($\xi < 1$), out-migration following the tax increase implies a lower public productivity per head. Since the median voter's interest is maximizing productivity, lower values of ξ

provide more incentive not to set tax rates (much) higher than the social optimal tax rate. The scale effects that derive from public expenditure thus oppose the bias that the scale effect on aggregate employment generates. The democratic bias is therefore larger with scale effects via productive externalities, but lower with scale effects via public expenditure.

4.5 Conclusion

This chapter studies the degree of pro-business policies in cities. It shows that democratic voting in a city leads to policies that are inefficiently geared towards business. The median voter lives closer to the labor market (the CBD) than the average citizen and therefore overly supports large expenditure on firm productivity. If the median voter were a renter, such excessive expendiure would raise his rent. As a homeowner, however, the capitalization is also a wealth improvement, which cancels the increased costs of homes. Too large public expenditure occurs for cities with a fixed size of population, so that internal re-allocations on the housing market are responsible for suboptimal policies.

In open cities, where production externalities determine city size, the median voter benefits more than average from productivity effects of migration. Migration combined with productive benefits of scale thus exacerbate the median voter's political distortion. Scale effects in policymaking work in opposite direction. If the government's productivity improvements are non-rival, the median voter has incentive to maximize aggregate tax revenue. Optimal tax rates attract most inhabitants and therefore maximize the tax base, encouraging the median voter to vote for taxes closer to the optimal rate.

Compared to a standard urban setup, the modifications towards durable housing and democratic government are essential in the argument. They are not common in the literature, but we do not view them as unrealistic in light of the respective rivalling traditions of absentee landlords and benevolent policymakers in urban models. Our results suggest that the role of housing and government require significant consideration when studying local public finance. The results also comply with public finance literature that studies voting among heterogeneous voters. However, here, the differences between voters are not assumed but follow endogenously; and the differences persist if voters are allowed to leave jursidictions with poor policy.

Predicting policies that are too pro-business, this model shares results with the tax competition literature. However, it shows that the mobility of firms of such models is not required to explain inefficiently expensive industrial policies. The chapter also shares the foundations of voting-by-feet and homevoter arguments, but demonstrates that their welfare message of efficient policies can be reversed. The policy recommendations of the current model are less clear-cut: policies that avoid capitalization run against both beneficial effects of homeownership behavior as discussed in the homevoter hypothesis, and they run against efficient public services known from Tiebout or voting-by-feet approaches. The lump sum tax studied in the model represent a redistributive policy between centrally and less centrally located inhabitants. A tax on land value could reduce the degree of redistribution, and therefore the median voter's incentive to deviate from the optimal tax rate.

The result that homeownership leads to more firm-oriented policies also ties in with the "renter effect" in local public finance; the empirical observation that cities with lower shares of renters generally set more pro-business policies. This chapter possibly explains why cities of homeowners support firms, but residents are homeowners by assumption. In reality, cities clearly exhibit a mix of owner- and renter-occupied housing. To ignore the underlying economic behavior that determines the mix of homeowners and renters makes it likely to miss a deeper explanation of the renter effect.

Finally, the analytical advantages of the monocentric city come at the cost of realism. In actual cities, firms can locate in areas other than the physical center, or form multiple centers. However, a key insight is that unequal access to the labor market combined with higher population density around employment centers leads to policy inefficiencies. Since both high land prices near economic centers and commuting costs are pervasive phenomena, it is quite possible that similar policy inefficiencies occur in models of polycentric cities or of multiple regions.

4.A Wealth heterogeneity in land prices and political preferences

The model in the main text leaves the effect of (land) wealth differences implicit. We further examine the role of wealth heterogeneity in the results here. We first show that marginal tax changes preserve the voters' ordering in the city, and then that under wealth heterogeneity, the land price effects still cancel from the voter's first-order condition for taxes.

Wealth heterogeneity makes it hard to solve for land prices explicitly, although the equilibrium generally does exist (Fujita and Smith, 1987). Households are ordered by the steepness of their bid rent; the steepness is affected by their wealth position. An individual's bid rent is:

$$r^b(\theta) = \max \frac{w/(1+\theta) - T + r\left(\theta_0\right)h_0 - c^*}{h^*}, \tag{4.A.1}$$

where stars denote optimized demand. To study the effect of tax changes on the slope of the bid rent at any point in the city, we first write the change in bid rent over space as:

$$\frac{dr}{d\theta} = \frac{-w}{(1+\theta)^2}\frac{1}{h^*}. \tag{4.A.2}$$

Differentiating with respect to taxes and dividing by the slope to get relative changes in the slope gives:

$$\frac{d\left[dr^b/d\theta/r^b\right]/dT}{\left[dr^b/d\theta/r^b\right]} = \frac{dw/dT}{w} - \frac{dY\left(\theta, h_o r\left(\theta_0\right)\right)/dT}{Y\left(\theta, h_o r\left(\theta_0\right)\right)}. \tag{4.A.3}$$

The term $dw/dT/w$ is constant over space. The relative change in net labor income is:

$$\frac{dY\left(\theta, h_o r\left(\theta_0\right)\right)/dT}{Y\left(\theta, h_o r\left(\theta_0\right)\right)} = \alpha \frac{\frac{dw}{dT}/(1+\theta) - 1}{w/(1+\theta) - T}, \tag{4.A.4}$$

which is positive in the center of the city and negative at the edge of the city (assuming that the policy is redistributive). As a result, a marginal tax increase makes the bid rent steeper in central locations and flatter in decentral locations. Since inhabitants are ordered by steepness of the bid-rent, a tax increase maintains the order of the citizens along the distance to the CBD and the commuting costs θ.

Second, the voter's first-order condition requires an expression of how land prices change with tax changes. Differentiating the expression for land prices (eq. 4.14) and dividing bij the land price gives the relative change in land prices:

$$(1-\alpha)\frac{dr(\theta)/dT}{r(\theta)} = \frac{\frac{dw/dT}{1+\theta}-1+\frac{dY(\theta)}{dr(\theta_0)h_0}\frac{dr(\theta_0)h_0}{dT}}{Y(\theta)} - \frac{\frac{dw/dT}{1+\theta_{edge}}-1+\frac{dY(\theta_{edge})}{dr(\theta_{0edge})h_{0edge}}\frac{dr(\theta_{0edge})h_{0edge}}{dT}}{Y(\theta_{edge})}. \tag{4.A.5}$$

The land price at θ changes in $r(\theta_0)h_0$, the land wealth of the buyer. All else constant, if a citizen has higher wealth, he bids up the land price of his location. The land price at θ is what a voter pays, but the land price at θ_0 is what he receives. By the same logic:

$$(1-\alpha)\frac{dr(\theta_0)/dT}{r(\theta_0)} = \frac{\frac{dw/dT}{1+\theta_0}-1+\frac{dY(\theta_0)}{dr(\theta_{00})h_{00}}\frac{dr(\theta_{00})h_{00}}{dT}}{Y(\theta_0)} - \frac{\frac{dw/dT}{1+\theta_{edge}}-1+\frac{dY(\theta_{edge})}{dr(\theta_{0edge})h_{0edge}}\frac{dr(\theta_{0edge})h_{0edge}}{dT}}{Y(\theta_{edge})}, \tag{4.A.6}$$

where θ_{00} denotes the land plot owned by the buyer of θ_0.

The two expressions for land price changes solve the voter's first-order condition (eq. 4.19) for the tax that maximizes his welfare. However, in an equilibrium where land possession and land consumption are equal (or arbitrarily close), the first-order condition simplifies considerably. First, the term $r(\theta_0)h_0/(w/(1+\theta)-T+r(\theta_0)h_0)$ is equal to $1-\alpha$; land wealth represents a fixed share of total wealth. Second, the tax preserves the order of atomistic citizens, whose land consumption is atomistically small. Marginal relocations are achieved by trading land with the direct neighbors, so θ is arbitrarily close to θ_0. Using these in the first-order condition, the first-order condition simplifies to:

$$
\begin{aligned}
\frac{dV\left(\theta_m\right)/dT}{V\left(\theta_m\right)} &= 0 && \text{(4.A.7)}\\
&= \frac{\frac{dw/dT}{1+\theta}-1}{w/(1+\theta)-T+r\left(\theta_0\right)h_0} \\
&\quad +(1-\alpha)\left[\frac{\frac{dY\left(\theta_0\right)}{dr\left(\theta_{00}\right)h_{00}}\frac{dr\left(\theta_{00}\right)h_{00}}{dT}-\frac{dY(\theta)}{dr\left(\theta_0\right)h_0}\frac{dr\left(\theta_0\right)h_0}{dT}}{Y\left(\theta_0\right)}\right] \\
&= \frac{\frac{dw/dT}{1+\theta}-1}{w/(1+\theta)-T+r\left(\theta_0\right)h_0} \\
&\quad +(1-\alpha)\left[\frac{dr\left(\theta_{00}\right)/dT}{r\left(\theta_{00}\right)}-\frac{dr\left(\theta_0\right)/dT}{r\left(\theta_0\right)}\right]. && \text{(4.A.8)}
\end{aligned}
$$

The land price and the change in land prices are continuous over space, θ. Because θ_{00} and θ_0 are arbitrarily close (with fixed ordering of citizens according to wealth, they are atomistic neighbors), the term $\frac{dr\left(\theta_{00}\right)/dT}{r\left(\theta_{00}\right)}$ is arbitrarily close to $\frac{dr\left(\theta_0\right)/dT}{r\left(\theta_0\right)}$, so $(1-\alpha)\left[\frac{dr\left(\theta_{00}\right)/dT}{r\left(\theta_{00}\right)}-\frac{dr\left(\theta_0\right)/dT}{r\left(\theta_0\right)}\right]$ tends to zero. As a result, the first-order condition is satisfied if $dw/dT/(1+\theta)-1=0$, which is equal to the first-order condition presented in the text.

CHAPTER 5

IMPROVING INTERCITY INFRASTRUCTURE: WHAT HAPPENS TO POPULATION AND JOBS?

5.1 Introduction

For most people and firms, connectedness makes cities attractive places to live, operate or work. Connectedness allows for easier access to other markets, both for selling and for buying. Similarly, it allows people to supply labor elsewhere, or to live elsewhere and commute back. Commutes are on the rise, both in number and in length. In Europe, it is no exception to cross a regional NUTS-2[1] region border to work (OECD, 2005); in the US, over 8% of commuters traveled out of their own metropolitan area, and intercity commuting flows grew nearly three times as fast as internal flows the last decades (Pisarski, 2006). Such commuting flows have substantial effects on the urban landscape, for instance leading to "jobs-housing imbalances" (Levine, 1998). As argued and documented by Glaeser and Kohlhase (2003) and Anas (2004), the costs that most shape the urban structure may not (only) be the falling freight transport cost, but also the cost of moving people. Changes in infrastructure affect not only travel times and transport costs, but also change commuting flows, residential patterns and the location of economic activity. It is therefore no surprise that governments actively guide commuting behavior with financial policies and infrastructure.

Improved infrastructure is often considered to benefit cities, but are improved connections an unmixed blessing? There are clear arguments to assume that they are. In their public role in production systems, they may improve local productivity (i.e., an Aschauer-type of argument, see Fernald, 1999). Similarly, lowering distance impediments is likely to increase options to buy goods and to sell goods, thus increasing market access and local incomes (Fujita et al., 2001, and references therein). Over the past decades, larger road networks have increased city growth in population as well as employment (Duranton and Turner, 2012).

[1]NUTS-2 regions typically contain 0.8–3 million inhabitants, and should cover a regional labor market (for the Netherlands, the regions are smaller for most other countries in Europe).

However, a series of recent empirical studies suggests that the regional effects of building new roads depend on what those roads connect. Connecting to highways or railways may lead to population losses, as population decentralization (i.e., away from large city centers) is likely to occur (Baum-Snow, 2010, 2007). Counties in the US that gained access to main roads have seen significant increases in economic activity (Michaels, 2008), but the growth of counties situated directly at new roads comes at the cost of adjacent counties, that lose economic activity (Chandra and Thompson, 2000). Indeed, employment can respond differently than population size does: in contrast to the population declines found by Baum-Snow, Duranton and Turner (2012) document increased employment resulting from more highways in a city. This is consistent with causal positive effects of infrastructure investments on travel distances (Duranton and Turner, 2011). Similarly, Baum-Snow et al. (2012) document that for Chinese cities, radial and ringroad highways decentralize population; and radial and ring railways, as well as ring highways, decentralize production.

This empirical evidence suggests that the urban models that are used to evaluate costs and benefits of infrastructural projects may be inaccurate. The differential effects of infrastructure on industries and population point to a decoupling of labor supply and residence. As many people do not commute to the employment center nearest to their house (Pisarski, 2006; Aguilera, 2005), the employment and residential size of regions can diverge. Therefore, apart from reducing commuting time and lowering trade costs, infrastructure investments can have stong relocation effects. Firms and people alike will rethink their preferred places of operating, working and living. By assuming industrial locations fixed, urban models can only explain changes in residential patterns, but cannot provide much insight into why employment changes location.

This chapter aims to provide a better understanding of the effects of infrastructure investments by studying firm mobility in addition to residential mobility. It proposes a model of endogenous agglomeration in a two-region setting, where firms as well as citizens are footloose. Therefore, it explains the population size of cities as well as the distribution of employment by examining the location choices of firms and the residential and commuting choices of workers. It can thus provide a theoretical framework to study the empirical differences in regional population and employment effects of interregional infrastructure.

In the model, the location choices of firms and residents and the commuting flow surface as the result of several intuitive tradeoffs. Agglomeration effects cause workers to commute into larger regions until the wage differences no longer compensate commuting costs. Due to a lower access to the labor market, the smaller region is less desirable to live in. This is reinforced by the costs of importing goods from larger regions. Both costs of goods shipment and of commuting reduce land prices in peripheral regions because it increases prices and reduces labor market access there.

The results show that lowering commuting costs will generally increase the size of the commuting flow, and make peripheral regions more attractive to live in, and the central region more attractive to work in. Improved commuting infrastructure therefore raises employment density in large regions but decentralizes population. Whether it centralizes or decentralizes jobs (in absolute numbers, instead of per head) depends on, among others, the quality of the existent infrastructure. Lowering the transport costs predominantly decreases consumer prices in peripheral regions, decentralizing both jobs and people. Therefore, either type of infrastructure investment is likely to increase the number of jobs per head in large regions, and shrink it in the periphery.

Not many other models incorporate both the mobility of workers and firms at the same time. An exception are New Economic Geography (NEG) models, in which Borck et al. (2010) introduce the possibility to commute. In their model, home market effects drive agglomeration and commuting, while the absence of income effects forms the crucial spreading force to prevent peripheries from emptying. In our model, there is a generalized agglomeration force, and the presence of land and non-tradable goods assures that people will choose to live in the periphery. For commuting costs between zero and infinity, the model of Borck et al. has no closed-form solutions, so the effects of infrastructure need to be inferred from numerical simulations. In addition to commuting from small to large regions, and in contrast to our chapter, the model of Borck et al. predicts that the net commuting flow can also run from large to small regions.

The foundations of agglomeration effects in the present chapter are based on an input-sharing model (cf. Ethier, 1982 and Abdel-Rahman and Fujita, 1990), that yields closed-form solutions. By varying the geographical range of externalities, the agglomeration model is sufficiently general to encompass both a Dixit-Stiglitz monopolistic market structure (which yields the home-market effects of the NEG) and pure Marshallian externalities (i.e., localized production externalities). Only for the particular Dixit-Stiglitz version of our model do we reach equilibria that are not unique and not closed-form; other formulations of agglomeration effects lead to a single, solvable residential and industrial distribution over regions. In other words, the agglomeration effects from NEG are very specific in that they cannot be achieved if any other type of externality operates (or more strictly, if the Dixit-Stiglitz love-of-variety effects are not equal to the elasticity of substitution between firms).

The following section develops a short-run model, where households are immobile (section 5.2) to study firms' location decisions and worker's labor supply decisions. It also discusses the parallels to the Dixit-Stiglitz setup and the long run, in which residential choice is free. Section 5.3 discusses the implications of reducing transport and commuting costs. The last section concludes.

5.2 Model

This section lays out the structure of the economy, followed by a more detailed discussion of households' and firms' behavior. Consumers and producers locate in one of the two regions in the economy. Households consume land for housing, a tradable good, and a non-tradable or local good. Both regions are endowed with a stock of land. Landowners are tied to their land and have the same preferences as workers. Workers live in one region but may choose to supply their labor in the other region, incurring a loss of utility by commuting. Within each region, commuting costs and transport costs are zero, so there is a single land price in the region. In the long run, workers can change residence. Producers of the tradable good acquire their inputs from local input-producing firms and assemble their good using a technology that has a constant elasticity of substitution between the inputs. Intermediate firms produce the inputs using labor under increasing returns to scale. The inputs and the local good cannot be traded across regions, but the tradable good can be transported at iceberg transport cost, requiring τ units to ship for one unit to arrive. When discussing the consumer side of the model, we refer to delivered prices.

Households

Consumers derive utility function from three items: consumption of the traded and local good C_t and c_{nt}, housing h. They suffer a potential utility loss caused by commuting, captured by the term θ which reflects the share of leisure that is lost travelling. Consumers have Cobb-Douglas utility over their consumption of goods and housing:

$$U = C_t^{\alpha} c_{nt}^{\mu} h^{1-\alpha-\mu} (1-\theta); \quad C_t = \left(C_{t,1}^{\varepsilon} + C_{t,2}^{\varepsilon}\right)^{1/\varepsilon}. \tag{5.1}$$

The parameters α and μ measure preferences (effectively budget shares) for the traded and non-traded good. The consumption of tradables is a composite of the consumption of tradable goods produced in region 1 and in region 2. The parameter ε governs the substitutability between goods from different origins, so it could be interpreted as the strength of the Armington assumption implicitly made – if ε is very high, consumers are indifferent about the origin of the good, if ε is low, they prefer variation in the consumption of tradables. As the inputs from which the tradable is constructed are not generally the same across regions, the term ε can capture to what extent this regional difference in upstream firms translates into different final products. Nevertheless, assuming that $\varepsilon = 1$ makes the products homogeneous, which reflects that the variety of inputs strictly lowers the producers' costs.

The utility function yields a unit-elastic housing demand, like in Helpman (1998). The commuting costs are captured by θ, which is zero if a worker supplies labor in his residential region, but positive if the worker commutes to another region. The term $1-\theta$ is interpreted as the share of leisure left after commuting, and the choice to commute is part of the maximization problem. This utility maximization problem does not lead the worker to change his supply of hours worked if his commuting time changes (because leisure enters unit-elastically in the utility function). Although the independence of labor supply and commuting time is simplifying, it is not too far besides reality (Gutiérrez-i-Puigarnau and Van Ommeren, 2010) and could be explained, for instance, by only allowing full-time jobs. The functional form also allows the expenditure shares on housing and consumption goods to be unaffected by the commuting decision. There is no direct financial cost to commuting, so the entire wage is spent on housing and consumption goods:

$$\begin{aligned} w &\geq P_t C_t + p_{nt} c_{nt} + p_h h, \\ P_t &= \left(P_{t,1}^{\varepsilon/(\varepsilon-1)} + P_{t,2}^{*\varepsilon/(\varepsilon-1)}\right)^{(\varepsilon-1)/\varepsilon}, \end{aligned} \tag{5.2}$$

where P_t, p_{nt} and p_h are the (delivered) prices of the traded and local good and housing rental rate. The star denotes that the consumer deals with delivered price for goods from the other region; in equilibrium, trade costs will drive a wedge between factory door price and delivery price. P_t is the harmonized price index of tradable goods C_t, which follows from the optimization of C_t with respect to the two goods, with the budget restriction that $P_t C_t = P_{t,1} C_{t,1} + P^*_{t,2} C_{t,2}$. The tradable good is consumed at home, even if the consumer works elsewhere. The income of landowners per unit of land is p_h. Landowners' demand functions are equal to the workers' demand functions, albeit that their income is generated by land rent rather than wage. The utility function and budget constraint give rise to the demand functions:

$$\begin{aligned} h &= \frac{\left(1-\alpha-\mu\right)w}{p_h}, \quad c_{nt} = \frac{\mu w}{p_{nt}}, \quad c_t = \frac{\alpha w}{P_t} b, \\ b_1 &\equiv \frac{P_{t1}^{\varepsilon/(\varepsilon-1)}}{P_{t1}^{\varepsilon/(\varepsilon-1)} + P_{t2}^{*\varepsilon/(\varepsilon-1)}}. \end{aligned} \tag{5.3}$$

Given the homotheticity of the demand function and the absence of commuting costs in the budget constraint, the expenditure shares are independent of the commuting decision; commuters, non-commuters and landowners allocate the same share of their budget to each good, and their consumption decision is unaffected by commuting.[2] For later analysis, it is useful to describe the indirect utility function by reintroducing the demand functions into the direct utility function. The indirect representation of the utility function can be written as (an affine transformation of):

$$V = \frac{w\,(1-\theta)}{p_h^{1-\alpha-\mu} p_{nt}^{\mu} P_t^{\alpha}}. \tag{5.4}$$

Firms

There are three types of firms: the intermediate firms that produce inputs; producers of the tradable good (assemblers of inputs); and producers of a local (non-transportable) good. Input producers and local goods producer use labor exclusively, the tradable goods producers use inputs exclusively.

[2]This assumes that commuters only consume the non-tradable at home. Allowing them to consume some of the local good at their workplace does not affect the property of the wage ratio sloping from infinity at no commuting to zero at full commuting, which turns out to define the commuting equilibrium. The equilibrium outcomes therefore do not change qualitatively due to this assumption.

This setup closely follows the differentiated inputs models for single regions of Ethier (1982) and Abdel-Rahman and Fujita (1990). However, the current model is a two-region model, where we assume that the shipping costs of inputs between regions are prohibitive, to capture that benefits of input variety are localized. Because the input set can differ per region, we allow goods to be imperfect substitutes, as detailed in Section 5.2.

One might argue that assembly requires labor, which this model does not assume. However, not much stands in the way of interpreting some of the inputs as an assembly workforce. Moreover, solving the model using both inputs and labor as arguments in a Cobb-Douglas production function for assemblers proves to make no difference for the results. Likewise, we assume that the non-traded sector faces constant returns to scale, which makes them different from manufacturing. Solving the model assuming that the non-traded sector also uses inputs does not change the results qualitatively either.

The tradable good producers operate under perfect competion. They assemble input quantities $y\left(i\right)$, the individual variety of which is indexed by i, to assemble a final good Y (the consumption of which is given by C_t), using a CES technology:

$$Y_r = \left(\int_0^{n_r} y\left(i\right)^{\frac{\sigma-1}{\sigma}} di\right)^{\frac{\sigma}{\sigma-1}}, \tag{5.5}$$

where n_r is the number of intermediate firms in region r. Note that while there are increasing returns to the variety of inputs, the production function of the tradables producers exhibits constant returns to scale if the set of inputs is given. The price of intermediates is $p(i)$. The profit function of a representative tradable goods producer is the aggregate revenue minus the expenditure on inputs: $\Pi_r = P_{t,r} Y_{t,r} - \int p\left(i\right) y\left(i\right) di$. The first-order condition with respect to $y(i)$ yields the demand for intermediate good i:

$$y\left(i\right) = \frac{p\left(i\right)^{-\sigma}}{\int_0^{n_r} p\left(i\right)^{1-\sigma} di} P Y_r, \tag{5.6}$$

which shows that demand for inputs is a fraction of the aggregate production and the fraction depends negatively on the price of the own variety, but positively on a harmonized average price of inputs. Filling out the demand function in the final goods producers' zero profit condition gives an expression for the aggregate input price index, or the price of the final good:

$$P_{t,r} = \left(\int_0^{n_r} p\,(i)^{1-\sigma}\, di \right)^{\frac{1}{1-\sigma}}. \tag{5.7}$$

This is a standard CES harmonized price index. It shows that when there are more intermediate firms in the region, the assembler can acquire a higher variety of inputs, and the consumer price of the final good is lower. In particular, when the input prices are symmetric ($p\,(i) = p \forall i$), the price index is $P = n^{1/(1-\sigma)} p$. The positive exponent on the number of input producers implies a positive scale externality.

The intermediate firms only employ labor, but require a fixed amount of labor in the fixed stage of production. This upfront requirement is the cause of increasing returns to scale in intermediate production. The total labor requirement for production of inputs is $l\,(i) = a_m y\,(i) + f$, where f is the fixed labor requirement and a_m is the inverse marginal labor productivity. The total costs amount to $wl(i)$. Intermediate good suppliers set prices to maximize profits. Facing the final goods producers' demand curve, they set a markup price, which is familiar from other CES demand models. Taking the first-order condition of the profit function $p\,(i)\,y\,(i) - wl\,(i)$ with respect to prices, and rearranging gives: $p\,(i) = \sigma/(\sigma - 1)\, a_m w$. The markup price over marginal cost combined with a zero profit condition on intermediate producers implies a constant firm size, which is a standard result in monopolistic competition models à la Dixit-Stiglitz. The intermediate producers' firm size is $y\,(i) = y = f\,(\sigma - 1)/a_m$. The corresponding labor requirement for individual intermediate firms is σf.

The third and last set of firms, the producers of the local good, operate under perfect competition, employing workers to produce at constant returns to scale: $y_n = a_n l_n$. Their profit-maximizing price equates to marginal cost: $p_n = w/a_n$. The producers of local goods employ the same workers as the intermediate producer, and with intersectoral labor mobility, the wage will equalize between the two sectors if they coexist in the same region.

Equilibrium in the short run

In the short-run equilibrium, the goods, labor and housing market clear, given a distribution of households over the regions. Workers select the region in which they supply their labor, so the commuting decision is part of the short run equilibrium. The two regions of the economy are indexed by 1 and 2. The share of total population L^w living in region 1 is λ, and the share working in region 1 is κ. Consequently, there is a commuting flow from region 2 to region 1 of size $\kappa-\lambda$ (all employed in region 1 less those that live region 1) if $\kappa > \lambda$. From the equilibrium conditions, commuting occurs only one way, so in the following section, we label regions such that the commuting flows runs from region 2 to 1 ($\kappa \geq \lambda$, or labor supply in region 1 is no lower than the number of residents in region 1).

In region 1, the clearing condition for the housing market implies that all available land is inhabited: $H_1 = \left(1-\mu-\alpha\right)\left(\lambda L^w w_1 + p_{h,1}H_1\right)/p_{h,1}$ (the right-hand side is the sum of demand of workers and landowners). Rewriting the clearing condition gives the equilibrium land price:

$$p_{h,1} = \frac{1-\mu-\alpha}{\mu+\alpha}\frac{\lambda L^w w_1}{H_1}. \tag{5.8}$$

The equilibrium land price is determined by the preferences for housing, by the aggregate labor income $\lambda L^w w_1$, and negatively by the supply of housing. Multiplying both sides of equation (5.8) also shows that the total houseowners' income is proportional to the aggregate labor income in the region. The clearing condition and equilibrium land price for region 2 are:

$$\begin{aligned} H_2 &= \left(1-\mu-\alpha\right)\frac{L^w\left((\kappa-\lambda)w_1 + (1-\kappa)w_2\right) + p_{h,2}H_2}{p_{h,2}}, \qquad (5.9) \\ p_{h,2} &= \frac{1-\mu-\alpha}{\mu+\alpha}\frac{L^w\left((\kappa-\lambda)w_1 + (1-\kappa)w_2\right)}{H_2}. \end{aligned}$$

The clearing condition in region 2 is slightly different from that in region 1, because the potential commuters $(\kappa-\lambda)$ earn the wage of region 1, whereas the non-commuters $(1-\kappa)$ earn the wage of region 2; therefore, the expression doe aggregate labor income is slightly longer.

In equilbirum on the local goods market, the supply of goods is equal to the quantity demanded. Using that $p_n q_n = w l_n$, this gives the pair of equilibrium conditions:

$$l_{n,1} = \frac{\mu}{\mu+\alpha}\lambda L^w;\ l_{n,2} = \frac{\mu}{\mu+\alpha}L^w\left((\kappa-\lambda)\,w_1/w_2 + (1-\kappa)\right). \quad (5.10)$$

In region 1, a fixed share of the inhabitants works in the non-traded sector. This occurs because the demand for non-tradables goods is unit-elastic, and their price is proportional to the wage in region 1. In region 2, citizens' average wages are higher than the local wage (because commuters earn a higher wage in region 1), so the non-tradable is relatively cheap. If we define $\nu_1 = l_{n1}/(\kappa L^w)$, and $\nu_2 = l_{n2}/((1-\kappa)\,L^w)$ as the regional share of employment in the non-tradable sector, the clearing conditions can be rewritten as:

$$\nu_1 = \frac{\mu}{\mu+\alpha}\frac{\lambda}{\kappa};\quad \nu_2 = \frac{\mu}{\mu+\alpha}\left(1 + \frac{\kappa-\lambda}{1-\kappa}\frac{w_1}{w_2}\right), \quad (5.11)$$

which shows again that commuting into region 1 decreases non-tradables employment in region 1 but increases it in region 1. The share of workers employed in the tradables' input sector is the complement of the share of workers employed in the local goods sector. Therefore, the pool of workers producing inputs for tradables is $\left(1-\nu_1\right)\kappa L^w$ in region 1 and $\left(1-\nu_2\right)(1-\kappa)\,L^w$ in region 2. Moreover, with a worker requirement of $f\sigma$, the number of intermediates firms is $\left(1-\nu_1\right)\kappa L^w/\left(f\sigma\right)$ in region 1 and $\left(1-\nu_2\right)(1-\kappa)\,L^w/\left(f\sigma\right)$ in region 2. Inserting the equilibrium worker allocations $(\nu_1,\ \nu_2)$ gives:

$$\begin{aligned} n_1 &= \left(\kappa - \frac{\mu}{\mu+\alpha}\lambda\right)\frac{L^w}{f\sigma}, \quad (5.12)\\ n_2 &= \left((1-\kappa)\frac{\alpha}{\mu+\alpha} - \frac{\mu}{\mu+\alpha}(\kappa-\lambda)\,w_1/w_2\right)\frac{L^w}{f\sigma}. \end{aligned}$$

The asymmetry in the expressions for the number of firms stems from the fact that we only consider commuting into region 1, but if there is no commuting ($\kappa = \lambda$), both regions devote share $\alpha/\left(\mu+\alpha\right)$ of the labor force to the production of tradables.

The demand for individual inputs is determined indirectly by the demand for the final tradable good. In equilibrium, the aggregate demand for the tradable good C_t from each region is equal to the region's supply Y. Since assemblers of inputs are assumed to have no market power, they charge the marginal cost of delivering a good, so the effective price in another region is the home market price, multiplied by the transport cost. Given the income of landowners, the aggregated demand functions for the tradable good become:

$$\begin{aligned} C_{t1} &= \frac{\alpha}{\mu+\alpha}\frac{L^w}{P_{t1}}\left(\lambda w_1 b_{11} + \left((\kappa-\lambda)w_1 + (1-\kappa)w_2\right)b_{21}\right), \quad (5.13) \\ C_{t2} &= \frac{\alpha}{\mu+\alpha}\frac{L^w}{P_{t2}}\left(\lambda w_1 b_{12} + \left((\kappa-\lambda)w_1 + (1-\kappa)w_2\right)b_{22}\right). \end{aligned}$$

The term $\alpha/\left(\mu+\alpha\right)$ capture the overall expenditure share of consumers on tradables, P_{t1} is the local price of tradables, and the terms in parentheses capture the income, corrected for the relative prices in the regional expenditure share of tradables, b (derived in in equation 5.2). Hence, b_{12} is effectively the share of tradable goods expenditure that consumers in region 1 spend on goods from region 2. Inserting the CES price indexes, the explicit expressions for those expenditure shares are:

$$\begin{aligned} b_{11} &= \frac{P_{t1}^{\varepsilon/(\varepsilon-1)}}{P_{t1}^{\varepsilon/(\varepsilon-1)}+\left(\tau P_{t2}\right)^{\varepsilon/(\varepsilon-1)}}; \quad b_{21} = \frac{\left(\tau P_{t1}\right)^{\varepsilon/(\varepsilon-1)}}{\left(\tau P_{t1}\right)^{\varepsilon/(\varepsilon-1)}+P_{t2}^{\varepsilon/(\varepsilon-1)}}; \quad (5.14) \\ b_{12} &= \frac{\left(\tau P_{t2}\right)^{\varepsilon/(\varepsilon-1)}}{P_{t1}^{\varepsilon/(\varepsilon-1)}+\left(\tau P_{t2}\right)^{\varepsilon/(\varepsilon-1)}}; \quad b_{22} = \frac{P_{t2}^{\varepsilon/(\varepsilon-1)}}{\left(\tau P_{t1}\right)^{\varepsilon/(\varepsilon-1)}+P_{t2}^{\varepsilon/(\varepsilon-1)}}, \end{aligned}$$

and it is readily verified that $b_{11} = 1 - b_{12}$ and $b_{21} = 1 - b_{22}$.

The wage that clears the market for tradables can be obtained by inserting the demand for the tradable good (eq. 5.13) in the demand equation for intermediates (eq. 5.6) and equating it to the supply ($y = f\,(\sigma-1)/a_m$). Rewriting that expression (in implicit form) for the wage rate gives:

$$\begin{aligned} w_1 &= c\left(\frac{\lambda w_1 b_{11} + \left((\kappa-\lambda)w_1 + (1-\kappa)w_2\right)b_{21}}{P_{t1}^{1-\varepsilon}}\right)^{1/\sigma}, \quad (5.15) \\ c &\equiv \left(\frac{L^w}{y}\right)^{1/\varepsilon}\frac{\alpha-1}{\alpha a_m}. \end{aligned}$$

where the term $\lambda w_1 b_{11} + \left((\kappa-\lambda)w_1 + (1-\kappa)w_2\right)b_{21}$ reflects expenditure on the tradable good by local producers. In this expression, c is a parametric constant. The parallel equation for region 2 is:

$$w_2 = c\left(\frac{\lambda w_1 b_{12} + \big((\kappa-\lambda)\,w_1 + (1-\kappa)\,w_2\big)\,b_{22}}{P_{t2}^{1-\sigma}}\right)^{1/\sigma}.$$

These two equations show that the local wage rate in inputs for tradables goods increase in the expenditure on the locally produced tradabe good.

The above presumes that producers of intermediates and tradables are active in both regions. If this is not the case, the producer in one region will capture the full expenditure on tradables. Given the scale externality, this occurs in the large region. Concentration of the tradables and intermediates sector in one region occurs if n_2 tends to zero (eq. 5.12). Rewriting $n_2 \leq 0$ gives the following condition for concentration of the tradables sector in region 1:

$$\frac{(1-\kappa)\,w_2}{\mu} < \frac{(\kappa-\lambda)\,w_1}{\alpha}. \tag{5.16}$$

The intuition for this inequality is that in equilibrium, the commuters' earnings (relative to total expenditure) must exceed the local earnings in non-tradables (relative to total expenditure): in order for the tradables production to concentrate, working in the tradables sector in another region yields more revenue per head than working in the local non-tradables sector (or setting up a local tradables firm).

The labor market clears if all workers are employed, and no worker has an incentive to change jobs. Since local goods, transportable goods and housing are all consumed in the place of residence, the commuting decision is not affected by their prices. From the indirect utility function, it can be seen that commuting yields higher utility than not commuting if $w_1(1-\theta) > w_2$, that is, if the nominal wage in region 1 compensated for commuting cost is higher than the nominal wage in region 2. Vice versa, commuting is undesirable if $w_1(1-\theta) < w_2$, and workers are indifferent about commuting if this condition holds with equality.

In case producers of tradables operate in both regions, the wage ratio is obtained by dividing the wage in the tradables sector of region 1 by that of region 2. Inserting the definition for the price index (eq. 5.7) and the number of firms (eq. 5.12) in that ratio and rewriting, the wage ratio is:

$$\frac{w_1}{w_2}|_{n_2>0} = \frac{\lambda w_1 b_{11} + \big((\kappa-\lambda) w_1 + (1-\kappa) w_2\big) b_{21}}{\lambda w_1 b_{12} + \big((\kappa-\lambda) w_1 + (1-\kappa) w_2\big) b_{22}} \times \frac{\alpha(1-\kappa) + \mu(\lambda-\kappa) w_1/w_2}{\alpha\kappa + \mu(\kappa-\lambda)}. \tag{5.17}$$

The first fraction in this wage ratio reflects the ratio of expenditure faced by producers in region 1 over region 2. The second fraction is the ratio of the number of intermediates firms operating in region 1 and 2, respectively. Therefore, the wage ratio is simply the ratio of expenditure in the regions per head employed.

The other possibility is that the production of tradables/intermediates concentrates in the large region. In that case, region 2 has no intermediate production, and the relative wage rate is determined by the full concentration of thetradable goods sector in region 1 (region 2 specializes in non-tradables). Inserting the price index $P_{t,1}^{1-\sigma} = n_{t1} p_{t1}^{1-\sigma}$ in the market clearing condition for tradables (eq. 5.15) under concentration ($b_{11} = b_{21} = 1$) and isolating the wage ratio yields:

$$\frac{w_1}{w_2}|_{n_2=0} = \frac{(1-\kappa)\alpha}{(\kappa-\lambda)\mu}. \tag{5.18}$$

This term is decreasing in κ, i.e., a higher labor supply in the region that hosts the tradable production reduces the wage rate in that region. It can be seen from the employment share in local goods in region 2 (ν_2, eq. 5.11) that as more people commute to region 1 (κ grows), the share of employment in local goods rises in region 2, until it fully specializes in non-tradables ($\nu_2 = 1$) and tradables production concentrates in region 1, in which case equation 5.18 holds. As a result, when tradables are not concentrated, more commuting into the large region may drive up wages in the large region (eq. 5.17), but also concentrates the production of tradables further into the larger region. Once tradables production is fully concentrated in the large region, more commuting always drives down the wages in the large region (eq. 5.18).

The interregional elasticity of substitution

The interregional elasticity of substitution is paramount in the spatial organization of the economy, because it determines the expenditure shares on tradables, b, in the process of concentration. When the input prices are symmetric in the tradables goods' final prices (eq. 5.7), the expenditure shares b_{11} can be written as:

$$b_{11} = \frac{n_1^{\varepsilon/(\varepsilon-1)/(1-\sigma)} p_1^{\varepsilon/(\varepsilon-1)}}{n_1^{\varepsilon/(\varepsilon-1)/(1-\sigma)} p_1^{\varepsilon/(\varepsilon-1)} + n_2^{\varepsilon/(\varepsilon-1)/(1-\sigma)} \left(\tau p_2\right)^{\varepsilon/(\varepsilon-1)}}. \tag{5.19}$$

The central role of this expenditure share is best explained by examining under what circumstances it is profitable to set up a firm. Suppose that there are initially no firms in region 1, and the first infinitesimally small firm is set up. Being the only firm, it can round up the entire expenditure share b on its region, while hiring labor at the regional wage rate. Given that the expenditure share for goods from region 2 is close to unity, we can study the potential earnings of a first firm by taking the derivative of the price index $n_1^{\varepsilon/(\varepsilon-1)/(1-\sigma)} p_1^{\varepsilon/(\varepsilon-1)}$ in the expenditure share (eq. 5.19) with respect to the number of firms in region 1: $\varepsilon/(\varepsilon-1)/(1-\sigma) \times n_1^{\varepsilon/(\varepsilon-1)/(1-\sigma)-1}$. Evaluated at the right limit of $n_1 = 0$,

$$\lim_{n \to 0^+} n_1^{\varepsilon/(\varepsilon-1)/(1-\sigma)-1} \begin{cases} 0 \text{ if } \varepsilon > 1 - 1/\sigma \\ \infty \text{ if } \varepsilon < 1 - 1/\sigma, \end{cases}$$

the expenditure on a first firm tends to zero if $\varepsilon > 1 - 1/\sigma$. This inequality states that the elasticity of substitution (ε) between different regions' final products is larger than the elasticity of substitution between inputs $(1 - 1/\sigma)$. In that case, the scale externalities in the region with many firms dominate the love-of-variety for goods from different regions, and entry in the small region is not profitable. Vice versa, if $\varepsilon < 1 - 1/\sigma$, the earnings for the entrant tends to infinity. In this case, willingness to pay for a good from that region is so high, that scale externalities can be foregone in the large region, to supply this good from the small region. A concentration of tradables producers can only take place if $\varepsilon \geq 1 - 1/\sigma$, otherwise, it is always profitable to set up a peripheral firm. If this condition holds with equality, the numerator and denominator in the expenditure share b become linear in the number of firms, and it is uncertain whether firms can profitably set up (marginal earnings could be anywhere between zero and infinity). We return to this special case below. It should be noted, finally, that the limit of the expenditure shares is determined by the interplay between consumer preferences and agglomeration externalities on the number of firms. Prices are treated as a positive constant (wages in the non-traded sector restrict their range). By the small firm assumption usually made in Dixit-Stiglitz models, firms do not respond to hte individual behavior of other firms and so the markup is constant, but this is inconsistent with letting n run to 0 asymptotically. However, this does not reverse the entry decision: the potential entrant can act as a pure monopolist, as in the above reasoning, or as a markup-pricer, in which case the entrant can (easily) outbid other firms on the labor market.

Commuting flows in the short-run equilibrium

If there is potential for concentration and there is commuting, tradables production must be concentrated in equilibrium. Commuting in cannot occur in a stable spreading equilibrium, because it implies that a marginal increase in the commuting flow increases the returns to commuting, thus reinforcing the commuting flow and the concentration of tradables. This leads a full concentration of tradables production. Only in that case will an increase in the number of commuters reduce the incentive to commute. This equilibrium can be seen graphically in Figure 5.1 where as the commuting flow grows, the relative wage of commuters first rises and then falls, and a stable equilibrium occurs only when an increase in the commuting flow decreases incentives to commute. More formally, if spreading allows commuting ($\left(w_1/w_2|n_2 > 0\right) \geq 1/(1-\theta)$), then the wage difference encourages more commuting ($d\left(w_1/w_2\right)/d\kappa > 0$, in eq. 5.17), until the tradables concentrate. Once tradables firms are concentrated, commuters' wages reduces in the number of commuters ($d\left(w_1/w_2\right)/d\kappa < 0$, in eq. 5.18), until the returns from commuting (higher wages) equal the leisure cost of commuting: $\left(w_1/w_2|n_2 = 0\right) = 1/(1-\theta)$ in equation 5.18. The labor supply distribution in the commuting equilibrium is therefore determined by $w_1/w_2 = 1/(1-\theta)$ under concentration of tradables (eq. 5.18). Rewriting this condition for the labor supply concentration implies:

$$\kappa_c = \frac{\lambda + (1-\theta)\,\alpha/\mu}{1 + (1-\theta)\,\alpha/\mu}. \tag{5.20}$$

The term κ_c is the equilibrium labor supply in region 1 if there is commuting. It decreases in the commuting costs θ, and increases in the residential concentration. Also, the large region's labor supply intuitively increases in the share of budget spent on tradables, but decreases in the share spent on non-tradables.

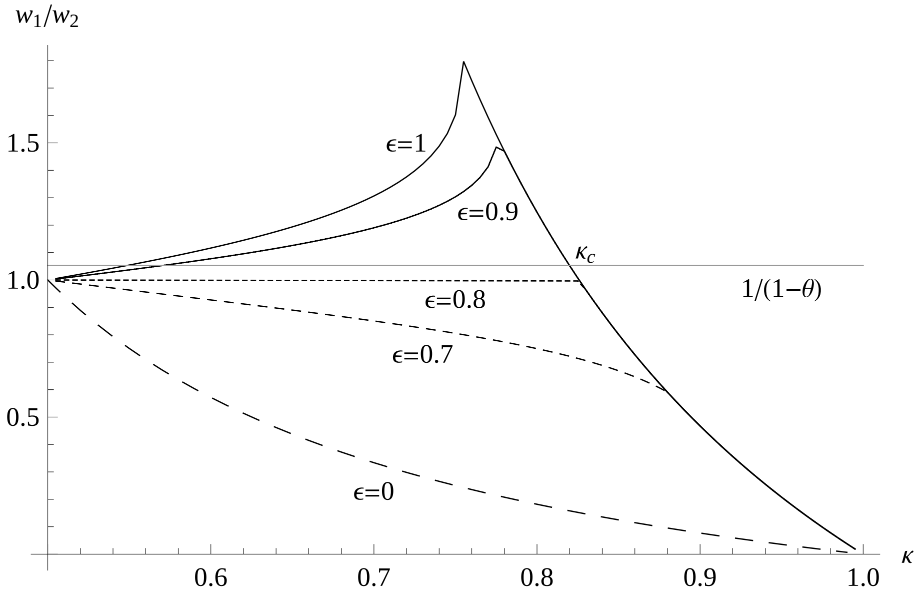
Another implication of eq. (5.20) is that if there is potential for tradables concentration ($\varepsilon > 1 - 1/\sigma$), the labor distribution κ_c is the only possible commuting equilibrium. Moreover, the equilibrium is unaffected by ε, so the equilibrium commuting flow (given that it exists) is insensitive to the the substitutability of goods from different regions. The intuition is that under perfect concentration of tradables production, producers of tradables do not compete with producers from other regions, so ε is irrelevant. This is captured in the wage patterns plotted in Figure 5.1, where under concentration, all values of ε imply the same wage ratio, and by extension, the same equilibrium labor supply κ_c.

The equilibrium size of the commuting flow may be the same over the parameter range that permits tradables concentration ($1-1/\sigma < \varepsilon < 1$), but the existence criteria need not be the same. For commuting to be feasible, the maximum wage ratio (which reflects returns to commuting) needs to exceed commuting cost. For potential commuting equilibria, the wage ratio is upward-sloping for smaller κ, when tradables producers are dispersed, and downward-sloping for larger κ, when tradables are concentrated. Therefore, if the commuting flow for which wages compensate commuting cost under dispersion is smaller than the commuting flow for which wages compensate commuting costs under tradables concentration, commuting if feasible. More formally, a commuting equilibrium is possible when:

$$\left(\kappa : \frac{w_1}{w_2}|_{n_{2>0}} = 1/(1-\theta)\right) < \left(\kappa : \frac{w_1}{w_2}|_{n_{2=0}} = 1/(1-\theta)\right) = \kappa_c, \quad (5.21)$$

which intuitively states that the labor supply concentration for which commuting becomes feasible must be smaller than the maximum labor supply concentration that permits a commuting equilibrium (κ_c). A higher θ signifies a higher cost of commuting, so that renders the commuting equilibrium less likely (by reducing κ_c and increasing the required labor concentration for commuting). A higher Armington elasticity (higher ε) will affect the left-hand side of this condition: the relative wage increases in κ if tradables are not perfectly concentrated, and a higher ε increases expenditure on the concentrating tradables sector, thus making it more likely that commuting costs are recovered (see the presence of b in eq. 5.17). This is also seen in Figure 5.1: exploiting the scale externalities is easiest if commuters do not shift away consumption quickly from a good that is supplied more (if ε is high). The right-hand side of inequality (5.21), which deals with a concentrated tradables sector, by contrast, does not depend on ε (see eq. 5.18). Therefore, in the example of Figure 5.1, both $\varepsilon = 1$ and $\varepsilon = 0.9$ lead to the same commuting equilibrium at

Figure 5.1: Relative wage as a function of labor distribution



Note: Parameters: $\varepsilon = 5$, $\sigma = 0.3$, $\mu = 0.3$, $\tau = 1.05$, $\theta = 0.05$.

κ_c, where the wage differential exactly compensates the commuting costs (the horizontal grey line at $1/(1-\theta)$). For $\varepsilon = 0.8$, however, a commuting equilibrium is not feasible, because the relative wage does not exceed $1/(1-\theta)$ at any point.

The above discussion suggests that any value for ε between $1 - 1/\sigma$ and 1 will result in the same equilibrium, so it is worthwhile to study the case $\varepsilon = 1$, for which the model yields closed-form results. The reason the commuting equilibrium does not depend on ε is that there is no interregional competition in tradables in the commuting equilibrium. Under a competitive tradables market, producers from the large region 1 charge the iceberg transport cost over their home price in region 2: $P_{t2} = \tau P_{t1}$. By logic of the trade balance, the reverse (goods shipping in the same direction as the commuting flow) cannot occur: since local goods and housing cannot be used to pay for tradables, commuting into region 1 implies shipping of goods into region 2. Rewriting the pricing condition $P_{t2} = \tau P_{t1}$ using the definition of the price for the tradable good (eq. 5.7), the implied wage rate (i.e., expression for eq. 5.17 in the commuting equilibrium) must be:

$$\frac{w_1}{w_2}|_{n_2>0} = \tau^{-1}\left(\frac{\left(\alpha+\mu\right)\kappa-\mu\lambda}{\alpha\left(1-\kappa\right)-\mu\left(\kappa-\lambda\right)w_1/w_2}\right)^{1/(\sigma-1)}. \tag{5.22}$$

This shows that the relative wage increases in the local employment κ. A lower σ reflects stronger agglomeration externalities, so it magnifies the returns to concentration of employment. In this case, setting the above wage equation equal to $1/(1-\theta)$, and inserting the solution in the condition for existence (eq. 5.21) gives an explicit expression for the commuting equilibrium to exist:

$$\left(\frac{\tau}{1-\theta}\right)^{\sigma-1}\left(\alpha+\mu\lambda/(1-\theta)-\kappa_c\right) < \left(\mu+\alpha\right)\kappa_c-\mu\lambda. \tag{5.23}$$

This states that a lower spatial frictions (trade and commuting costs) make a commuting equilibrium more likely, as do larger scale externalities (lower σ).

Special case: Dixit-Stiglitz

The analysis so far shows that closed-form commuting equilibria can exist when $\varepsilon > 1-1/\sigma$, and commuting never occurs when $\varepsilon < 1-1/\sigma$. Results are uncertain when the $\varepsilon = 1-1/\sigma$. Under that assumption, the elasticity of substitution between goods from different regions (the "Armington" elasticity) is equal to the elasticity of substitution between intermediates in the production function of the producer of tradables. In that case, the expenditure share b (eq. 5.19) becomes:

$$\begin{aligned} b_{11} &= \frac{P_{t1}^{1-\sigma}}{P_{t1}^{1-\sigma}+\left(\tau P_{t2}\right)^{1-\sigma}}, \\ P_{t1}^{1-\sigma}+\left(\tau P_{t2}\right)^{1-\sigma} &= n_1 p_{t1}^{1-\sigma}+\tau^{1-\sigma}n_2 p_{t2}^{1-\sigma}, \end{aligned} \tag{5.24}$$

where the second line shows the denominator explicitly. Using this in the market-clearing condition for inputs (eq. 5.6), the demand for individual varieties of the intermediates producers is:

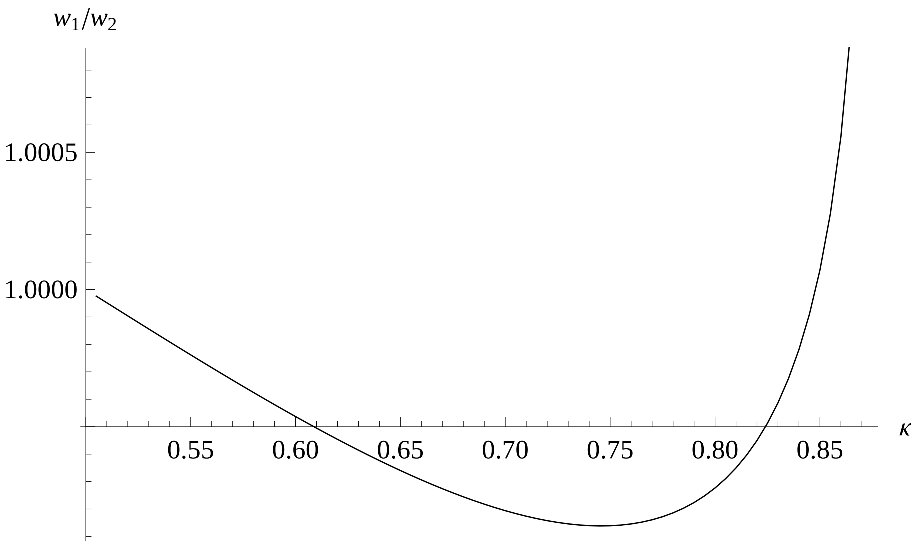
$$\begin{aligned} y &= p(i)^{-\varepsilon}\frac{\alpha}{\alpha+\mu}L^w \\ &\times\left(\frac{\lambda w_1}{\left(n_1 p_{t1}^{1-\sigma}+\tau^{1-\sigma}n_2 p_{t2}^{1-\sigma}\right)}+\tau^{1-\varepsilon}\frac{\left((\kappa-\lambda)w_1+(1-\kappa)w_2\right)}{\left(\tau^{1-\sigma}n_1 p_{t1}^{1-\sigma}+n_2 p_{t2}^{1-\sigma}\right)}\right). \end{aligned} \tag{5.25}$$

This equation is mathematically equivalent to the goods market clearing condition of a standard two-region Dixit-Stiglitz model of monopolistic competition. It features the constant elasticity of demand (via the term $p(i)^{-\sigma}$). The other terms constitute the market potential (income divided by $P^{1-\sigma}$, the CES price index to the power $1-\sigma$). The terms $n_1 p_{t1}^{1-\sigma} + \tau^{1-\sigma} n_2 p_{t2}^{1-\sigma}$ and $\tau^{1-\sigma} n_1 p_{t1}^{1-\sigma} + n_2 p_{t2}^{1-\sigma}$ in equation (5.25) are equal to the harmonized price index in constant-elasticity-of-substitution models over two regions raised to the power $1-\sigma$. Using this, equation 5.25 is the same as used in a standard New Economic Geography model (compare Fujita et al., 2001). Clearly, the assumptions used in New Economic Geography are rather different (the intermediate products presented in this chapter represent final products in the NEG model), but mathematically, the Dixit-Stiglitz results are a limiting case of this formulation if $\varepsilon = 1 - 1/\sigma$. Therefore, the present model behaves the same as if we assumed labor was employed in the standard manufacturing setup in NEG (compare Fujita et al., 2001), and a non-tradable good was added. Also note that all feasible NEG parameters ($1 < \sigma < \infty$) can be translated to an equivalent $0 < \varepsilon < 1$.

The parametrization $\varepsilon = 1 - 1/\sigma$ allows the wage pattern to become non-monotonic. It renders the numerator and denominator of the expenditure share b linear in the number of firms n. Therefore, it is uncertain whether the wage in a region where the tradables sector nearly vanishes ($n \to 0$) goes to infinity or zero; it could be either. Since the parametrization yields a system of non-linear equations, it is complex to establish the closed-form expression for the wage ratio, and it is not generally possible to say what determines the commuting patterns (as long as transport costs are present). The results derived under the assumption $\varepsilon = 1$ are part of the possible outcomes, but we cannot exclude other solutions. Figure 5.2 provides a numerical example for which $\varepsilon = 1 - 1/\sigma$. The Figure confirms that this parametrization gives rise to patterns of relative wage that are non-monotonic when tradables are not concentrated. Also, marginally changing the value of ε (adding or subtracting 0.00001, a relative change of 0.00125% in either way) removes the non-monotonicity in the wage pattern.

Theory does not provide much guidance for the choice of elasticity of substitution of goods between different regions. If ε is very high, any price increases in the final good lead to strong losses of sales in the region. Therefore, the returns to increased input variety are very localized. If ε is lower, scale increases and price reductions of tradables in one region

Figure 5.2: Relative wage with equal inter-firm and inter-regional substitution elasticity



Note: parameters for the numerical solutions are $\varepsilon = 5$, $\gamma = (\varepsilon - 1)/\varepsilon$, $\sigma = 0.6$, $\mu = 0.2$, $\tau = 1.03$.

only partially translate into lower sales elsewhere: ε lower than 1 implies an imperfect change in consumer's budget following interregional price difference changes. In that sense, scale effects are less localized when ε is lower. The lowest interregional elasticity of substitution in tradables that still allows agglomeration, $\varepsilon = 1 - 1/\sigma$, is consistent with a Dixit-Stiglitz formulation.

In the context of our model, the Dixit-Stiglitz (D-S) case, that yields non-closed forms, requires a very specific parametrization. Even if assuming the D-S case, the model converges to the general solution under $1 - 1/\sigma < \varepsilon < 1 - 1/\sigma$ if other sources of centripetal force are present. Suppose that individual firms' price depend on the aggregate employment or firm population, say $p(i, n)$. A fairly common log-linear formulation to allow employment scale effects (governed by δ) on the final price effect of employment would be $P = n^{1/(1-\sigma)} p(n) = n^{1/(1-\sigma)+\delta} p$. A δ other than zero could occur through a variety of mechanisms of learning, sharing or matching (Duranton and Puga, 2004). Following the above model, $\delta \neq 0$ (the equivalent to $\varepsilon \neq 1 - 1/\sigma$ in the previous section) would eliminate the non-closed form solutions from the model in the D-S case. The scale effects responsible for this can also stem from the consumer side. The

Dixit and Stiglitz (1975) working paper version, and later Benassy (1996), argued for a utility function of the form $n^{\delta}\left(\int \left(c\left(i\right)/n\right)^{(\sigma-1)/\sigma} di\right)^{\sigma/(\sigma-1)}$. The motivation is that this disentangles the love-of-variety effects (based on the number of firms) from the elasticity of substitution between goods captured in σ. The short (popular) form assumes that $\delta = 0$, which implies a very specific relationship between the elasticity of substitution and the love of variety effects. Allowing for a more general form (setting δ not zero), therefore, will also eliminate the specific CP equilibria. We do note that Borck et al. (2010), and other NEG-based models do not display this sensitivity love-of-variety effects, because they assume the number of firms to be exogenous, thus eliminating love-of-variety effects.

Residential allocation and the long run equilibrium

The short-run version of this model already has mobility in production factors. Therefore, part of the location decision, especially those of firms and on labor markets, are already present in the short-run equilibrium. However, additional long-term dynamics occur in this model: workers can choose to change residence. We shall assume that the migration dynamics are driven by the relative utility of living in either location - agents relocate to another region if that yields them a higher utility. Stable equilibria therefore involve utility equality ($V_1 = V_2$) and local stability ($d\left(V_1/V_2\right)/d\lambda$ < 0), consistent with the standard myopic migration dynamics in NEG models: $d\lambda = \left(V_1 - V_2\right)\lambda\left(1-\lambda\right)$. In the analysis, we distinguish between the complementary cases in which the commuting equilibrium is relevant (κ_c) and those in which there is no commuting: $\kappa = \lambda$.

The relative utility of living in region 1 is the ratio of indirect utility functions (eq. 5.4):

$$\frac{V_1}{V_2} = \frac{w_1}{w_2}\left(\frac{p_{h,1}}{p_{h,2}}\right)^{\mu+\alpha-1}\left(\frac{P_{t,1}}{P_{t,2}}\right)^{-\alpha}\left(\frac{p_{nt,1}}{p_{nt,2}}\right)^{-\mu}, \tag{5.26}$$

and there is relocation of households into region 1 if $V_1 > V_2$, and vice versa.

Furthermore, we can rule out two constellations because they are inconsistent with the short-run equilibrium. As discussed in the short-run equilibrium,[3] first, if there is no commuting, then the production of tradables cannot be concentrated in one region. Second, if there is a commuting flow, then the production of tradables must be concentrated.

In an equilibrium that involves commuting, we know from the short run equilibrium equations that $w_1(1-\theta) = w_2$ (commuting equilibrium) and $\tau P_{t1} = P_{t2}$ (competitive pricing). Inserting the price of houses (eq. 5.8) and the prices of local goods ($p_{nt1}/p_{nt2} = w_1/w_2$) into the utility ratio function gives:

$$\frac{V_1}{V_2} = \left(\frac{H_1}{H_2} \frac{\kappa_c - \lambda + \left(1-\kappa_c\right)(1-\theta)}{\lambda} \right)^{1-\mu-\alpha} (1-\theta)^{\mu-1} \tau^{\alpha}. \qquad (5.27)$$

Like in the short-run worker allocation, the long-run residential distribution is unaffected by the substitutability of tradable goods from different regions, ε; the tradables sector in region 1 rounds up all expenditure in that sector, and the tradables sector in region 2 does not affect the labor market. Accordingly, like in the short run, the particular assumptions made about the Armington elasticity do not lead the equilibrium to change, only its stability conditions are affected, because ε governs the degree to which scale increases lead to substitution into the other region's good.

Clearly, the equilibrium that involves commutes is most interesting for the purpose of this analysis. Nevertheless, there is another equilibrium that does not involve commuting, so the distributions of employment and residents must be equal: $\kappa = \lambda$. Inserting the prices of local goods and houses in the relative utility function, the residential equilibrium without commuting occurs when:

$$\frac{V_1}{V_2} = \left(\frac{w_1/P_{t1}}{w_2/P_{t2}} \right)^{\alpha} \left(\frac{H_1}{H_2} \frac{1-\lambda}{\lambda} \right)^{1-\mu-\alpha} = 1 \text{ and } \frac{d\left(V_1/V_2\right)}{d\lambda} < 0 \qquad (5.28)$$

is equal to 1, and its derivative with respect to λ is negative. The relative wage is determined on the market for tradables (eq. 5.17). Under

[3] To quickly reiterate: without commuting, tradables production cannot be concentrated. This can be seen from eq. (5.18) (which shows the concentrated wage rate), in which the relative wage in region 1 tends to infinity if there is no commuting ($\kappa \to \lambda$). Second, under commuting, the production of tradables is concentrated. If it were not concentrated, the returns to commuting would increase in the scale of the commuting flow (given that some workers already commute), and the equilibrium is unstable. In other words, as long as $w_1(1-\theta) \geq w_2$ and the tradables sector is not concentrated, more workers have an incentive to commute.

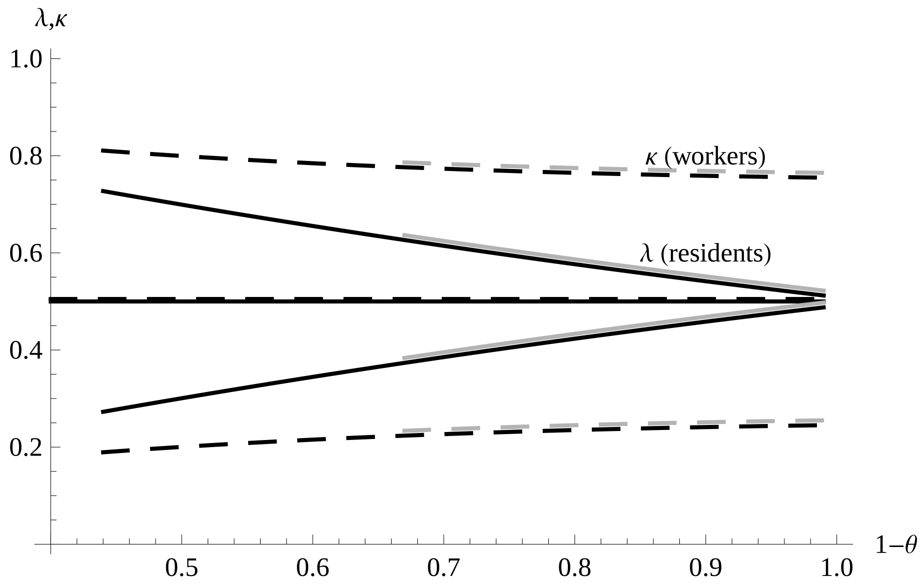
the assumption that final goods are homogenous ($\varepsilon = 1$), inserting the equilibrium wage and price indexes and rewriting for $V_1 = V_2$ (no migration because utility is equal across regions), the equilibrium residential distribution is

$$\lambda = \frac{\left(H_1/H_2\right)^{\frac{1-\mu-\alpha}{1-\mu-\alpha\sigma/(\sigma-1)}}}{1+\left(H_1/H_2\right)^{\frac{1-\mu-\alpha}{1-\mu-\alpha\sigma/(\sigma-1)}}}. \tag{5.29}$$

This suggests that if there are no commutes, relative housing supply exclusively determines regional size. Size is insensitive to spatial frictions: commuting costs do not matter because nobody commutes, trade costs do not matter because trade is symmetric. For the utility ratio to be downward sloping in λ, we require that $1-\mu < \alpha\sigma/(\sigma-1)$, which is akin to the no-black-hole condition found in NEG models. This condition states that expenditure on tradables and increasing returns in tradables must be low relatively the preferences for housing, otherwise the scale effects make the economy collapse into one region.

Figure 5.3 provides a graphical representation of the long-run equilibrium. It shows a bifurcation diagram, plotting the long-run equilibrium distributions of labor supply (dashes) and residents (solid) as a function of the commuting costs. If commuting costs are high ($1-\theta$ is low, left in the graph), commuting is not feasible and only the symmetric equilibrium is stable. When commuting costs become sufficiently low, a residential concentration can be sustained (the solid line deviates from 0.5), accompanied by an even stronger concentration of jobs (dashed): people commute into the large region. This is reflected in the distance between the residential distribution and in the more outward labor supply distribution (the amplitude of the dashed line is larger than of the solid line). Once this asymmetric equilibrium is established, further decreases in commuting costs make the residential equilibrium more symmetric, reducing residential concentration. Even if commuting costs become very low, the symmetric equilibrium remains locally stable, because under equal wages, there is no way to recoup the positive costs of commuting. Finally, while the black lines refer to the equilibrium where final goods are homogenous ($\varepsilon = 1$), the grey lines permit imperfect substitutability between goods from different regions ($\varepsilon < 1$). As shown before, this does not affect the equilibrium distributions, because the tradables sector fully concentrates

Figure 5.3: Bifurcation patterns



Note: Bifurcation patterns for residential (solid) and labor supply (dash) distribution, for $\gamma = 1$ (black) and $\gamma = 0.85$ (grey). Parameters: $\varepsilon = 5$, $\sigma = 0.3$, $\mu = 0.3$, $\tau = 1.05$.

in the asymmetric equilibrium.[4] However, with imperfect substitution, the returns to concentration are lower because consumers value the loss of consumption from the smaller region more heavily. Therefore, the stability conditions are affected: if the tradables from different regions are less close substitutes (ε is lower), then the equilibrium is less likely to exist. This can be seen from Figure 5.3, where the asymmetric equilibrium for lower substitutability ($\varepsilon = 0.85$, grey lines) requires lower commuting costs to exist than under perfect homogeneity of tradables.

5.3 How do infrastructure investments change the economy?

To understand how infrastructure developments change regions, this section collects and summarizes the model's predictions regarding changes in spatial frictions. The interregional interactions are impeded by two types

[4] In the Figure, a value of 0.01 was added to the equilibrium values of κ and λ in case $\gamma = 0.85$, to avoid losing the lines visually due to overlap with the lines in case $\gamma = 1$. They are, in fact, on the same coordinates.

of barriers: commuting costs and the costs of transporting goods. In order to avoid assumptions on the relation between the two, we shall study them in turn.

In the short run, where the residential distribution is fixed, the employment distribution was given by $\kappa_c = \left(\lambda + (1-\theta)\,\alpha/\mu\right) / \left(1 + (1-\theta)\,\alpha/\mu\right)$ (equation 5.20). This term decreases in θ, so reducing the commuting costs leads to larger labor supply in the large region in equilibrium. When commuting costs fall, the higher wage in the large region makes it feasible for more people in the smaller region to cover the commuting costs. Moreover, writing the jobs per head by dividing κ_c by λ as:

$$\left(1 + (1-\theta)\,\alpha/\mu/\lambda\right) / \left(1 + (1-\theta)\,\alpha/\mu\right),$$

shows that employment density rises in the large region, and falls in the small region if θ reduces. The transport costs do not affect the labor supply distribution. The cause of this is that lowering transport costs proportionally decreases the tradables' price asked in the other region, and therefore does not affect relative wages.

In the long run, the residential distribution was implicitly given by utility equality (equation 5.27). Implicitly differentiating the utility ratio with respect to the distance frictions (around the equilibrium where $V_1/V_2 = 1$) yields:

$$\frac{d\lambda}{d\,(1-\theta)} = \frac{-(1-\lambda)\,\lambda\left(1-\theta\left(1-\mu\right)\right)\alpha}{(1-\theta)\left(1-\mu-\alpha\right)\left(\mu+\alpha\,(1-\theta)\right)} < 0, \tag{5.30a}$$

$$\frac{d\lambda}{d\tau} = \frac{(1-\lambda)\,\lambda\alpha}{\left(1-\mu-\alpha\right)\tau} > 0, \tag{5.30b}$$

which states that the population size in the large region, λ, decreases in the commuting costs and trade costs ($1-\theta$ is a negative measure of commuting costs), so reductions in commuting and transport costs decrease the residential concentration in the large region 1.

When spatial barriers in commuting and transport are reduced, the equilibrium number of inhabitants in the large city reduces. Two price effects are the reason for this. First, the price of tradables is lower in the smaller region, making it a more attractive place to reside. Secondly, if commuting costs fall, it is easier to take advantage of the lower prices of houses and non-traded goods in the smaller region while still working in the high-wage region – the labor market access of smaller regions increases. Therefore, in the commuting equilibrium, a reduction of spatial frictions disperses residences rather than concentrate them.

Because reducing commuting costs unambiguously raises the incentive to commute, and pushes population outwards, infrastructure investment will centralize employment density in the long run: jobs per head go up in the large region, and down in the small region. However, the long-run effect on *absolute* employment in both regions is unclear: with fewer residents but more commuters, does the large region effectively get more or fewer jobs? The long-run effects are given by differentiating the equilibrium labor supply in the large region (eq. 5.20) with respect to commuting costs, while taking the effects on the long-run residential equilibrium into account:

$$\frac{d\kappa_c/d\,(1-\theta)}{\kappa_c} = \frac{d\lambda/d\,(1-\theta)+\alpha/\mu}{\lambda+(1-\theta)\,\alpha/\mu} - \frac{\alpha/\mu}{1+(1-\theta)\,\alpha/\mu}. \tag{5.31}$$

Using the expression for $d\lambda/d\,(1-\theta)$, (5.30a), and rewriting gives that:

$$\begin{aligned}\frac{d\kappa_c/d\,(1-\theta)}{\kappa_c} &= \frac{\alpha}{\mu}\frac{1-\lambda}{\big(\lambda+(1-\theta)\,\alpha/\mu\big)\big(1+(1-\theta)\,\alpha/\mu\big)} \\ &\quad\times\left(1-\lambda\frac{1-\theta\big(1-\mu\big)}{(1-\theta)\big(1-\mu-\alpha\big)}\right).\end{aligned} \tag{5.32}$$

While this is a long result, the sign of the effect of trade costs on employment takes the sign of the third term in the product; the first two are positive. Rewriting the third term, the effect of commuting cost reductions on employment in the large region is positive if:

$$\theta < \frac{1-\lambda-\mu-\alpha}{1-\lambda\big(1-\mu\big)-\mu-\alpha}. \tag{5.33}$$

This shows that the effect of improving commuting connections on absolute centralization in the long run can go either way. The effect depends on the current commuting costs: lowering commuting costs decentralizes jobs if commuting costs are sufficiently high, but centralizes jobs if commuting costs are already low. For reductions of transport costs for goods, the results are unambiguous: since they only decentralize population, by equation (5.20), employment inherently also decentralizes.

5.4 Conclusion

This chapter presents an analysis of commuting flows in the presence of agglomeration externalities. The general equilibrium model with markets for land, labor and goods predicts that firms may colocate to reap the benefits of agglomeration, while attracting commuter flows. Workers trade off higher wages with commuting costs when choosing where to work. In the long run, when choosing locations, the benefits of living close to high-wage jobs are balanced with the benefits of lower prices in the periphery. Some workers will choose to live and work in the periphery due to cheaper land and the presence of non-tradable goods. Therefore, the commuting equilibrium involves a larger city with a high employment density, while the smaller city has a relatively large share of residents and lower jobs per head.

In a commuting equilibrium, improving infrastructure generally disperses residents at a faster pace than jobs. Lower commuting costs allow workers to further exploit scale externalities in the large region, increasing labor supply there, while in the long run, it also permits workers to flee the high prices of houses and non-traded goods in the large region. Relative to population, improving commuting connections always centralizes employment, but in absolute terms, the number of jobs in large cities may grow or decline. Integration of the two labor markets leads to job centralization in the absolute sense only if commuting costs are already very low. Goods' transport costs reductions do not affect commuting incentives directly, but in the long run, they facilitate living in smaller regions. Therefore, goods market integration unambiguously decentralizes population and jobs. These results hold for an economy that has already endogenously developed size differences. If not, both reductions in transport costs and in commuting costs increase the likelihood of the economy moving symmetry between regions to agglomeration, which unambiguously centralize jobs and people.

The model's predictions match with residential decentralization following infrastructure investment documented in the empirical literature. It is also consistent with the ambiguous empirical results on job decentralization (e.g., Duranton and Turner, 2011, 2012; Baum-Snow et al., 2012), although it is impossible to say whether that is for the right reasons. These insights could help explain why infrastructure investment has mixed impacts on regions. The literature on European Structural Funds has already made the argument that increased infrastructure investments may reduce incentives for firms to settle in a region, because improved connectivity makes it easier to supply from outside the region (see, e.g., Puga, 2002). That insight on firm relocation is formalized in the present chapter. In addition, this chapter argues that there is a workers' perspective, which suggests that peripheral regions will see reductions in the jobs per head.

Firms in this chapter are not modelled as users of land, suggesting that they value access to markets and inputs over land. If firms were to use land, then given the scale externalities, they would bid up land prices in the large region. This increases workers' incentives to relocate in search of lower land prices, thus increasing the size of commuting flows. Therefore, from that perspective, the present chapter puts a lower bound on the asymmetry in regional size and in the residential or industrial specialization fo regions.

The microfoundations of agglomeration effects in this chapter are quite different from the Dixit-Stiglitz assumptions that underlie New Economic Geography models. Yet, they can be made equivalent with a knife-edge assumption that the elasticity of substitution between goods from different regions is equal to the elasticity of substitution between its inputs. Under this assumption, non-closed form solutions arise that may yield other results than those described so far. However, these solutions are no longer relevant if any other agglomeration forces play a role. They also disappear if the specific assumption on the relation between elasticity of substitution and love of variety is dropped from the model (see Dixit and Stiglitz, 1975, and Benassy, 1996, for a discussion on the limited generality of that assumption). Therefore, within the context of the model, the uncertainty over infrastructure effects that arise with the non-closed form equilibria in NEG do not appear most germane.

CHAPTER 6

TAX COMPETITION WITH COMMUTERS AND AGGLOMERATION

6.1 Introduction

Factor mobility hinders good policymaking. As we saw in chapters 2 and 3, financial capital and firms may flow out of countries to avoid high taxes, causing governments to cut tax rates and undersupply public services. At the level of cities or regions, capital and firm mobility may be relevant, but another form of factor mobility is at least equally plausible: workers can commute, as we studied in chapter 5. Travel-to-work times of over an hour are no exception in many countries, so workers also have a substantial geographical range to consider job opportunities (OECD, 2005). As a result, it is possible for cities and regions to design policies that attract workers but not residents, or vice versa. Not much is known, however, about how desirable the policies are that governments set if workers' labor supply is mobile.

The aim of this chapter is to investigate the policy obstacles that arise if workers can commute. To that extent, this chapter sets up a tax competition model in which workers can commute and migrate, and in which agglomeration forces shape cities, as they do in chapter 5. Agglomeration effects play a central role for two reasons. First, they cause wage differences across cities, and therefore lead to commuting between cities.[1] Therefore, agglomeration forces allow differentiation between policy instruments targeted at workers and instruments targeted at inhabitants, while describing commuting choices endogenously. Secondly, as a natural consequence of inserting agglomeration as a commuting motive, the strategic incentives for governments change. As the tax competition literature points out (Baldwin and Krugman, 2004), agglomeration has marked effects on government behavior: it introduces city-scale considerations in tax-setting and renders tax harmonization undesirable. This is in sharp

[1] This means we ignore commuting motives based on search/matching or worker heterogeneity motivations, in which case our model is less relevant. Since these typically do not lead to the asymmetric commuting flows in the current model, their effects on optimal policy are likely to differ substantially.

contrast to a large literature that shows that without agglomeration effects, harmonization is welfare-enhancing. To the best of our knowledge, this chapter is the first to examine tax competition within a framework where agglomeration effects drive both commuting and residential mobility.

The possibility to commute decouples labor supply and residence, and so provides governments with a strategic choice of tax instruments. Effectively, commutes are based on nominal wage differences, while residential choices are based on differences in the utility of living in different locationd. Taxes that affect nominal wages and utility levels differently will have disparate impacts on the number of inhabitants and the number of workers. Therefore, depending on whether they impact commuters or residents, some forms of taxation may be preferred over others. Empirically, local governments rely much more heavily on land and property taxation than central governments do. According to Guo (2009), a plausible reason is that commuting makes labor supply more elastic, so local governments prefer to tax land and property, which are less elastically supplied.[2] In contrast to earlier literature, the setup presented below allows for migration next to the commuting decision. Joint migration and commuting decisions have different implications for welfare. On the one hand, it is plausible that governments gear their policies toward local business in an attempt to attract commuters. On the other hand, voting with their feet, workers put pressure on local governments to provide balanced (welfare-maximizing) policies. Lee (2002) shows that if both capital and labor are allowed to relocate, the effects of increased mobility on welfare are ambiguous; capital mobility distorts taxes, higher labor mobility leads to tax rates closer to the social optimum.

A related literature considers the effects of commuting on tax competition explicitly (Braid, 2000, 1996; Guo, 2009). In these models, workers choose to live in a jurisdiction in a metropolitan area, and subsequently, can travel to work in other jurisdictions inside the same metropolitan area costlessly. Government policy can involve a mix of source- and residence-based taxes,[3] but when the number of local jurisdictions grows large, governments avoid residence-based (wage) taxes altogether. Guo shows that the welfare conclusions depend on how the benefits of tax revenue are distributed: governments generally underprovide public goods, but if public

[2]We note that in many countries, local governments cannot levy labor taxes by institutional design.

[3]Depending on the separability of labor and capital in the production function.

inputs improve production, which also benefit commuters, governments can achieve a Pareto-efficient outcome. Additionally, Guo studies the effects of asymmetry among jurisdictions, which fosters commuting flows. In contrast to this chapter, however, the size difference are assumed ex ante, while in our model, they are endogenous, and form a main incentive to commute. In addition, the present chapter allows for residential mobility across jurisdictions in the long run, which changes the policy insights.

The agglomeration effects in our model also relate the chapter to tax competition models using increasing returns, that have been developed in a New Economic Geography setting. These models show that the resulting scale effects change the strategic incentives for governments. The intuition is that agglomeration externalities tie firms to larger regions. Smaller regions competing to attract firms not only need to offer lower taxes, they also need to compensate firms for the foregone benefits of locating in the larger market. Exploiting this insight, governments of large regions can set higher taxes without having firms leave their region, undermining "races to the bottom". The current chapter extends this discussion by studying the threat of a race-to-the-bottom under scale effects. Moreover, it expands the residential mobility of that literature with the possibility to commute. The ensuing options of instrument choice are studied in tax competition models, but not in the context of scale effects.

The motive to commute in this chapter derives from an agglomeration externality based on an Ethier-type (1982) variety of inputs, that can be interpreted as a model of sectoral specialization, or of returns to a diverse set of inputs (Abdel-Rahman and Fujita, 1990). Nevertheless, as shown in chapter 5, the model can easily be extended into a fairly general representation, of which the New Economic Geography setup (based on Dixit-Stiglitz) is a limiting case. The current form allows analytical results, but retains a large number of properties of the New Economic Geography setup. Therefore, we expect that the results obtained in this model will be similar to those in a model that uses NEG forces of agglomeration. Borck et al. (2010) study a comparable NEG model. The present chapter's model allows for a wage premium as a commuting motive, where commuters travel against commuting costs from small regions to large regions. However, as increased demand on land markets drives up rents, the residential choice involves living in the larger region, versus migrating outward to face higher commuting costs but lower house prices. Since some goods in the economy cannot be traded, the smaller region always employs workers in equilibrium, and the local wage (relative to the higher wage in the large region) determines the equilibrium size of the commuting flow.

Lastly, the governments' instruments require discussion. We assume that the government attempts to maximize the welfare of its inhabitants, excluding the potential commuters that travel into their region. One of the government's tasks is to provide services directly to its inhabitants. Additionally, the government contributes to the level of public inputs, which improves local productivity. These expenditures must be financed from a tax on land. We follow Guo (2009) in letting the government affect local productivity, instead of using labor taxes, as much of the earlier literature does. The reason is that most local governments have no discretion regarding labor taxation. In Europe, hardly any region sets wage taxes, and in the U.S., even among the few local governments that are allowed to tax labor, labor taxes are uncommon, and account for a small part of revenue in those places where they are applied (Braid, 1996). However, the intuition behind the results should not be very different. In equilibrium, the government could raise the level of local services to citizens by reducing the expenditure on public inputs and thus lowering the local wage. Under a labor tax, the government service level would be raised by raising taxes, which reduces the wage, so similar opportunity costs would be in place. Lastly, since this chapter extends earlier literature on agglomeration-driven commutes with residential mobility, it permits studying whether land market instruments, such as mortgage interest deductions, can improve welfare. In particular, towards the end of the chapter, we insert a central government that provides land market subsidies.

Our results show that government support to local firms encourages commutes that increase welfare; races to the bottom do not occur in this setting. However, because commuters do not take into account their external productivity effects, the labor concentration is too low from a global perspective, and local governments have no incentive to correct this inefficiency. In the long run, when residents can choose the region that they prefer to live in, migration forces governments to take commuting inefficiencies into account. Government policies are second best, however, because the local instruments cannot correct for the inefficiently small commuting flow. Depending on whether the size distortion in commuting leads to residential overconcentration or employment underconcentration, a central government could improve welfare by a housing subsidy scheme that is regressive respectively progressive in the house price.

The chapter is organized as follows. The next section sets up a spatial equilibrium model to describe the residential choices, commutes and regional sizes. Section 6.3 studies the behavior of welfare-maximizing governments given the equilibrium equations outlined in section 6.2. Section 6.4 then introduces a central government to the model and discusses its policy. Section 6.5 concludes.

6.2 Model

The model presented here largely follows the model presented in chapter 5, in which most assumptions are discussed more extensively. The main differences with chapter 5's model are that there is a government service, G, a public input, A, and an ad valorem tax on housing t, while the model is simplified by setting $\varepsilon = 1$ (final goods from different regions are perfectly substitutable). As shown in chapter 5, assuming that final tradable goods are homogeneous ($\varepsilon = 1$) will yield closed-form solutions while not affecting the agglomerated spatial equilibria, or their intuition.

The economy consists of two regions. Households can choose to reside and work, potentially commuting to another region. The main agents are governments, workers and firms involved in the production of tradable goods or non-tradable goods. This section solves for the spatial equilibrium for given government policies. The next section examines the role of governments that maximize local average utility.

Households are endowed with one unit of labor, which they use to obtain an income, and have unit-elastic preferences over consumption goods, housing and time losses due to commuting. Their utility function is:

$$U = CG(1-\theta), \text{ with } C = C_t^{\alpha} C_{nt}^{\mu} h^{1-\mu-\alpha}. \tag{6.1}$$

Since the government service and the commuting costs are not priced, all income is spent on consumption goods: $w = P_t C_t + P_{nt} C_{nt} + rh + trh$, where t is an ad valorem tax on housing. Commuting takes the form of a leisure loss (assuming labor contracts cannot be adjusted at the intensive margin), so that θ is the fraction of time lost in commuting when travelling to another region. The corresponding demand curves are:

$$C_t = \frac{\alpha w}{P_t},\ C_{nt} = \frac{\mu w}{P_{nt}},\ h = \frac{\left(1-\mu-\alpha\right) w}{r}. \tag{6.2}$$

As a result, the indirect utility function is (an affine transformation of) $V = wG(1-\theta) / \left(P_t^{\alpha} P_{nt}^{\mu} r^{1-\mu-\alpha}\right)$.

The production of the tradable good is carried out by firms that assemble a set of inputs according to a CES production function:

$$Y_r = A\left[\int_0^{n_r} y(i)^{\frac{\sigma-1}{\sigma}}\, di\right]^{\frac{\sigma}{\sigma-1}}, \tag{6.3}$$

in which subscript r denotes the region, n_r is the number of intermediates producers and $y(i)$ denotes the individual input of variety i. The term A is affected by expenditure on public inputs, in an Aschauer (1989) formulation of government effects on productivity. They can be viewed as the the soft and physical infrastructure that allows firms to operate, such as road-, communications- and energy infrastructure, but also efficiency of the local administration and assistance in acquiring locations, issues which are not explicitly modelled here. Also, given that the model is static, it is not helpful to talk of public capital, unless it would depreciate in one period.

Given the constant returns to scale for a fixed set of inputs, we treat the assemblers as a representative firm under perfect competition. Optimizing the profit function $\Pi = PY - \int_0^n p(i)\, y(i)\, di$ with respect to individual varieties $y(i)$, the demand for individual varieties is:

$$y(i) = \frac{p(i)^{-\sigma}}{\int_0^{n_r} p(i)^{1-\sigma}\, di} PY, \tag{6.4}$$

The assemblers' revenues PY are based on the final goods' price. If total expenditure on tradables is equal to E, introducing the demand function for intermediates back into the production function yields that in a zero-profit equilibrium,

$$E = YA^{-1}\left(\int_0^{n_r} p(i)^{1-\sigma}\, di\right)^{1/(1-\sigma)}.$$

In equilibrium, as expenditure equals the product of consumption and its price, the price of the tradable good is

$$P_t = A^{-1}\left(\int_0^{n_r} p(i)^{1-\sigma}\, di\right)^{1/(1-\sigma)}.$$

Improvements in the public input hence decrease the price of the tradable by making the assembler more productive.

Producers of inputs face a constant marginal productivity, but increasing returns to scale due a fixed labor requirement f:

$$l(i) = a_m y(i) + f, \tag{6.5}$$

where a_m is the unit labor requirement. We assume that the input producer is too small to consider his impact on the price index, and given the constant elasticity in the demand for the good, the input producer charges a markup over marginal costs $p(i) = \sigma / (\sigma - 1) a_m w$. Under this markup, and under zero profits in equilibrium, the firm's use is σf. The public infrastructure does not affect the input producers directly, but increased expenditure on public inputs increases demand for inputs, allowing them to bid up wages. If the labor supply is endogenous, this will lead to an increase in the number of firms (the equilibrium employment size of each firm is unaffected).

The producers of the local good operate under perfect competition under a linear technology $y_n = a_n l_n$, employing only labor. The corresponding first-order condition implies that $p_n = w / a_n$.

To describe the allocation of jobs and inhabitants, we define λ as the share of the population L that lives in region 1, while share κ works in region 1. The commuting flow into region 1 is therefore described by $\kappa - \lambda$. We assume that each region is endowed with a stock of land, H_r. Land is taxed at an ad valorem tax rate T. The after-tax proceedings of housing rent after taxes are spent on local goods, thus giving the house an interpretation of the land value and local services required for living. The government revenue equals TrH. From the demand functions and the fixed supply of land, we have the following market-clearing land rents:

$$\begin{aligned} r_1 &= \frac{\left(1-\mu-\alpha\right)\lambda L w_1}{H_1}\left(1+T_1\right), \\ r_2 &= \frac{\left(1-\mu-\alpha\right)L\left((\kappa-\lambda)w_1+(1-\kappa)w_2\right)}{H_2}\left(1+T_2\right). \end{aligned} \tag{6.6}$$

The equilibrium land rent increases in the preference for housing $(1-\mu-\alpha)$, and in the local income, while it decreases in supply. The tax that a supplier of houses needs to pay drives up house rents.

Spatial equilibrium allocations

Given the above preferences and production structure, different spatial equilibria can occur, as explained in chapter 5. We shall focus on the equilibrium most relevant to our case, which is characterized by an agglomeration of tradables producers in one region, with higher residential density in that region and commutes into that region. One of the regions will endogenously be the larger one, which we shall label region 1. The details of this equilibrium and its stability condition are elaborated on in Chapter 5.

Free mobility between sectors implies that the tradables sector pays the same wages as non-tradables do. In equilibrium, a share ν of all workers in region 1 works in non-tradables. This share is determined by the market-clearing condition that expenditure of local landlords and workers equals the pay to workers in the non-tradable sector: $\nu\kappa L w_1 = P_{nt1}Q_{nt1} = \mu\lambda L w_1 + r_1 H_1$. Using the expression for clearing on the housing market (eq. 6.6) and isolating the share of workers in the non-traded sector gives:

$$\nu = \left(\mu + \left(1-\mu-\alpha\right)/\left(1+t_1\right)\right)\frac{\lambda}{\kappa} = \varphi\left(t_1\right)\frac{\lambda}{\kappa}. \tag{6.7}$$

where $\varphi\left(t_1\right) \equiv \mu + \left(1-\mu-\alpha\right)/\left(1+t_1\right)$. This shows that the share of employment in non-tradables depends on the size of its consumers (local inhabitants) relative to the workers, the preference for non-tradables, and local taxes that reduce the demand from home-owners for local services. Subsequently, the wage in the large region relative to that of the small region is determined by the clearing condition of the tradable goods markets. Given that all input suppliers face the same technology and demand, they act symmetrically. The demand for an individual input supplier is $y = p^{-\sigma}/\left(A^{-1}np^{1-\sigma}\right)PY$. This clearing condition for intermediates can be rewritten as:

$$npy = A\alpha L\left(\kappa w_1 + (1-\kappa)w_2\right),$$

by realizing that the expenditure $PY = \alpha L\left(\kappa w_1 + (1-\kappa)w_2\right)$. Using the fact that the number of firms under a constant labor requirement equals $n = (1-\nu)\kappa L/\left(f\sigma\right)$, the markup price $p = \sigma/(\sigma-1)\,a_m w$, and the fixed size of firms $f\,(\sigma-1)/a_m$, the term npy is equal to $(1-\nu)\kappa w_1$. Using this in the clearing condition for intermediates, along with the equilibrium share of workers in non-tradables production (ν) and re-organizing for the relative wage yields:

$$\frac{w_1}{w_2} = \frac{A\alpha(1-\kappa)}{\kappa(1-\alpha A) - \varphi\left(t_1\right)\lambda}. \tag{6.8}$$

This shows that, keeping other things equal, a larger supply of labor in the large region (i.e., a larger commuting flow) drives down wages relative to the small region, while improvements in public inputs improves the local relative wage.

In the short run, we assume that while residences are fixed, workers can choose where to supply their labor, potentially by commuting. We assume that workers only consume in their home region. The indirect utility function then shows that for workers in the small region (region 2) to be indifferent between working in the home region and travelling to the large region, the wage premium must satisfy that $w_1(1-\theta) = w_2$. If wages are larger than permitted by this commuting condition, more workers would commute, driving down the wage premium from commuting, and vice versa. Using this equilibrium condition for commutes and the relative wage expression (eq. 6.8), the large region's employment in equilibrium is given by:

$$\kappa = \frac{(1-\theta)\alpha A + \varphi\left(t_1\right)\lambda}{1-\theta\alpha A}. \tag{6.9}$$

The equilibrium employment share in the large region is positively affected by the local residents (who increase demand for non-tradables) and the productivity in the tradables sector: raised productivity increases wages, which attracts workers.

In equilibrium, the size of the commuting flow is lower than socially optimal. We formally show this in Appendix 6.A, but the intuition is simple. The clustering of workers stems from the external productivity improvements that workers exert on each other. A worker choosing to commute has private costs in terms of time, and a private benefit derived from the higher wage earned in the larger region. His choice does not take into account, however, that his commute has a productivity benefit to other workers employed in the tradables sector, so fewer workers commute than is optimal.

In the long run, additionally, workers may choose to change residence. The stable long-run equilibrium occurs if the utility of living in either region is equalized, and small deviations do not lead to further reallocation.

More formally, using migration dynamics $d\lambda = \left(V_1 - V_2\right)$, the equilibrium requires that $V_1 = V_2$, and $d\left(V_1/V_2\right)/d\lambda < 0$. This requires setting up the ratio of indirect utility functions. The utility ratio, however, is considerably simplified by two equilibrium conditions: one is the commuting equilibrium condition, which states that $w_1/w_2 = 1/(1-\theta)$, and the second holds that competitive assemblers of tradables charge competitive prices throughout the regions, which implies that $\tau P_{t1} = P_{t2}$. Finally, using the expression for land rents and inserting the equilibrium commuting flow yields $r_1/r_2 = \lambda/\left(\frac{(1-\theta)}{1-\theta\alpha A} - \lambda\left(1 - \frac{\varphi\theta}{1-\theta\alpha A}\right)\right)H_2/H_1$. Using these, the short-run equilibrium utility ratio can be written as:

$$\frac{V_1}{V_2} = \frac{G_1}{G_2}\frac{w_1}{w_2}\left(\frac{P_{t1}}{P_{t2}}\right)^{-\alpha}\left(\frac{P_{nt1}}{P_{nt2}}\right)^{-\mu}\left(\frac{r_1}{r_2}\right)^{-(1-\mu-\alpha)} \tag{6.10}$$

$$= \frac{G_1}{G_2}(1-\theta)^{\mu-1}\tau^{\alpha} \tag{6.11}$$

$$\left(\times\frac{H_2\left(1+T_1\right)}{H_1\left(1+T_2\right)}\frac{\lambda}{(1-\kappa)(1-\theta)+\kappa-\lambda}\right)^{-(1-\mu-\alpha)}.$$

Isolating the share of inhabitants of large region on the left-hand side, the residential equilibrium is characterized by:

$$\frac{\lambda}{1-\lambda-(1-\kappa)\theta} = \left(\frac{G_1}{G_2}\frac{\tau^{\alpha}}{(1-\theta)^{1-\mu}}\right)^{1/(1-\mu-\alpha)}\frac{H_1\left(1+T_2\right)}{H_2\left(1+T_1\right)}. \tag{6.12}$$

Since the left-hand side of this equation is upward-sloping in the equilibrium share of residents in region 1, the population of region 1 is larger, keeping other things constant, if it provides relatively more public goods, the land supply is larger, and the taxes on land are lower. Both higher transport costs and higher commuting costs lead to a larger population living in region 1. Given that we are looking at an agglomerated equilibrium, the reasoning is as follows: higher commuting costs discourage commuters, so given a large labor demand in the large region, fewer people will locate elsewhere to commute back into the large region. For the transport costs, a consumer price effect drives population centralization: given that tradables are produced in the large region, living in the smaller becomes less attractive if those goods need to be imported at higher costs.

The spatial equilibrium sketched so far is conditional on the governments' policies (taxes, government services and public inputs). The equilibrium conditions show the effect of government choices on the economy, and therefore allow discussion of government behavior, which we turn to next.

6.3 Government policy

The governments in this model are faced with the problem of selecting tax rates (T), and consequently, to allocate the tax revenue to productive public inputs (A) or local public services to the citizens (G). Hence, the constraint that the government faces is that spending on public services and public inputs cannot exceed the tax revenues. The objective of the government is to maximize average local welfare of its inhabitants, and by this assumption, the government is not interested in the commuters travelling into their region, except for the productive effects that they might have on local inhabitants.[4]

We compare the situation in the short run, where the governments do not take into account residential mobility (because residents are fixed, or governments are not sufficiently forward-looking to anticipate residential mobility) with a situation in which governments take into account all forms of mobility. The two situations differ in that utility need not be equal across locations in the short run. This potentially yields very different results, because policy may affect local welfare through the reduction of prices and in the short run, utility differences can arise. In the long run, by contrast, inhabitants are expected to eliminate any utility differences, which implies an optimization of the joint utility function with respect to the two tax rates (given that the spatial concentration of workers is imperfect).

The analysis is simplified by the fact that one of the two wage rates acts as a numeraire, and there is a fixed relation with the other region's wage rate, irrespective of which wage is chosen as a numeraire. Finally, to ease notation, we shall define the relative change in a variable due to tax rate change with a hat, such that $dA/dt/A = \hat{A}$.

[4]The local government cares about current inhabitants, not about potential entrants. A justification is that only current inhabitants vote over the local policies. Assuming potential or actual inhabitants in the government objective function does not affect the equilibria, however, and therefore the results are the same.

Policy with immobile residents but commuting options

In the short run, workers can commute but not relocate, so local governments are concerned with the local organization of public services and inputs, and potentially, how their choices affect the commuting equilibrium. Average welfare is given by a linear transformation of the indirect utility function:

$$V = \frac{Gw}{P_t^{\alpha} p_{nt}^{\mu} r^{1-\alpha-\mu}}, \tag{6.13}$$

which equals the real wage multiplied with government services.[5] For the small region's government (region 2), the function to be maximized is:

$$\frac{G_2 w_2}{\left(\tau A^{-1} n^{1/(1-\sigma)} w_1\right)^{\alpha} \left(a_n w_2\right)^{\mu} r_2^{1-\mu-\alpha}}. \tag{6.14}$$

Realizing that in equilibrium, $w_1 = w_2/(1-\theta)$, the wage rate cancels altogether from this fraction, which reflects that it acts as a numeraire. Taking out the wage rate and constant terms independent of policy that are proportional transformations of the welfare function, and inserting the expression for the number of intermediates producers, the maximand is:

$$\frac{G_2}{\left(A^{-1} \frac{\left(\kappa-\varphi\lambda\right)}{1-\theta}^{1/(1-\sigma)}\right)^{\alpha} \left(\frac{(\kappa-\lambda)/(1-\theta)+(1-\kappa)}{H_2}\left(1+T_2\right)\right)^{(1-\mu-\sigma)}}. \tag{6.15}$$

ch6eq:ckearinghousingmarket), and therefore reduces welfare, keeping other things constant. The government of region 2 cannot directly invest in public inputs in production in the large region. Therefore, it is also unable to affect the size of the commuting flow (i.e. it cannot manipulate eq. 6.9). Optimizing local welfare (eq. 6.15) with respect to the local tax t_2 therefore implies that government 2's optimal tax policy adheres to:

$$\left(1+t_2\right)\left(1-\mu-\alpha\right) = \hat{G}_{2,2}. \tag{6.16}$$

The left-hand side reflects the relative decrease in the utility level due to tax increases solely, because land prices increase. The relative marginal benefits in terms of utility due to increases of public goods services are

[5] If the worker commutes, the function needs to be multiplied with $(1-\theta)$, while taking the other region's wage as the wage earned. In equilibrium, however, utility is constant across commuters and non-commuters living in the same region.

captured on the right-hand side. Therefore, the first-order condition balances the relative marginal utility costs of public funds with the marginal benefits of public services. The most important conclusion is that in the short run, the small region's government is unaffected by mobility issues; as the government does not manipulate the size of the commuting flow $(\kappa - \lambda)$. The intuition is that the commuting decisions in the short run are unaffected by local land taxes, and therefore the government has no policy instrument to manipulate them.

Using the same expressions (cancelling the wage rate and terms independent of policy), the welfare function that region 1 maximizes is:

$$\frac{G_1}{\left(A^{-1}\left(\kappa - \varphi\lambda\right)^{1/(1-\sigma)}\right)^{\alpha}\left(\lambda\left(1+T_1\right)\right)^{(1-\mu-\alpha)}}. \tag{6.17}$$

The term $\kappa - \varphi\lambda$ can be written as $\kappa\left(1-\varphi\lambda/\kappa\right) = \kappa\left(1-\nu\right)$, and is therefore proportional to the number of intermediate firms in tradables production, n (see the discussion of eq. 6.8). The first-order condition of average welfare in region 1 with respect to the tax rate, T_1, implies:

$$\left(1+T_1\right)\left(1-\mu-\alpha\right) = \hat{G}_1 + \alpha\hat{A} + \frac{\alpha}{\sigma-1}\hat{n}. \tag{6.18}$$

Compared to region 2's first-order condition, the large region takes two additional effects into account. Firstly, they are able to directly affect productivity, possibly improving local welfare if the return to local public inputs is sufficiently steep. Secondly, the government is able to affect the size of the commuting flow, which leads to extra employment in the increasing-returns-to-scale sector. (Note that only the fraction $\left(\kappa - \mu\lambda\right)/\kappa$ ends up in the tradables sector, as the non-tradables sector will also raise labor demand if local income increases.) The expression for the equilibrium employment share in the large region shows that this effect runs primarily through the local inputs that shift productivity.[6]

Since the equilibrium labor supply in the large region is too low relative to the social optimum (see Appendix 6.A), the large region's government will provide a positive amount of local inputs, as that increases local employment and brings the number of input producers closer to the social optimum. Assuming that the returns to scale in public input and service provision are constant, we will generally have that the large region provides public inputs, which need to be financed from (typically a combination of) higher land taxes and lower services to citizens.

While the large region's government corrects part of the market failure that causes insufficient entry into the tradables sector, it is questionable whether this leads to an efficient allocation over the sectors. To investigate this, we set up a social welfare function of all inhabitants as the average of all individuals' utility levels: $\bar{V} = \lambda V_1 + (1-\lambda) V_2$. We do not distinguish between commuters and workers in region 2, because in the short-run (and the long-run) equilibrium, the commuters into region 1 have the same utility as workers in region 2. The social welfare can be optimized with respect to the land taxes in region 1 and 2. Because the tax of region 2 only affects region 2's government services, region 2's tax rate that maximizes local welfare also maximizes global welfare. For region 1, however, this exercise gives the efficiency condition that the overall relative change in welfare due to tax changes ($d\bar{V}/dt_1/\bar{V}$) is to be equal to zero:

---

[6]Differentiating the employment share with respect to the local inputs gives that $\frac{d\kappa/dt_1}{\kappa} = \frac{dA/dt_1}{A} \frac{\sigma A\left(1-\theta\left(1-\varphi\lambda\right)\right)}{\left((1-\theta)\sigma A+\varphi\lambda\right)(1-A\sigma\theta)}$, which gives a positive association between the local public inputs and local employment, provided that $A < (\sigma\theta)^{-1}$. For completeness, we note that:

$$
\begin{aligned}
\hat{n} &= \frac{d\kappa/dt - \lambda d\varphi\left(t_1\right)/dt}{\kappa - \varphi\lambda} = (1-\nu)\left(\hat{\kappa} - \nu\hat{\varphi}\right) \\
&= (1-\nu)\left(\left(\frac{(1-\theta)\sigma A}{(1-\theta)\sigma A + \varphi\lambda} + \frac{\theta\sigma A}{1-\theta\sigma A}\right)\hat{A} + \left(\frac{\varphi\lambda}{(1-\theta)\sigma A + \varphi\lambda} - \nu\right)\hat{\varphi}\right), \\
\hat{\varphi} &= \frac{-1/(1+t)}{1+(1+t)\mu/\left(1-\mu-\sigma\right)} < 0,\ \hat{A} > 0,
\end{aligned}
$$

so that an equilibrium increase in the tax rate (given the government's first-order conditions) increase the number of firms in tradables production.

$$0 = \lambda V_1 \left(\hat{G}_1 + \alpha \hat{A} + \frac{\alpha}{\sigma - 1} \hat{n} - \left(1 + T_1\right)\left(1 - \mu - \alpha\right) \right) \qquad (6.19)$$
$$+ (1 - \lambda) V_2 \left(\alpha \hat{A} + \frac{\alpha}{\sigma - 1} \hat{n} \right).$$

Rearranging gives the global socially optimal tax rate of region 1 (given that no tax revenue can be transferred across locations)

$$\left(1 + T_1\right)\left(1 - \mu - \alpha\right) = \hat{G}_1 + \frac{1}{s_w} \left(\alpha \hat{A} + \frac{\alpha}{\sigma - 1} \hat{n} \right), \qquad (6.20)$$

where $s_w = \lambda V_1 / \left(\lambda V_1 + (1 - \lambda) V_2 \right)$ region 1's share of population weighted by utility (if utility levels were equal, it was simply the population share). The tax of region 1 in equilibrium (eq. 6.18) is therefore generally smaller than the tax set to maximize global welfare. Namely, if the share of utility-weighted population of region 1 is lower than 1 ($s_w < 1$), the right-hand side of equation (6.20) is larger than the right hand side region 1's first-order condition in equation (6.18). As a result, the equilibrium marginal cost of public funds is higher in the global welfare-maximizing case, and the tax rate is higher. The government of region 1 hence sets a lower tax rate than optimal because he does not take into account two external effects of his policy: an increase in the public inputs A directly reduces the price of final goods in region 2, and indirectly increases productivity through migration (n changes), which also benefits region 2.

Thus, as a result in the situation where residence is fixed, we have that the mobility of labor does not lead the government of the small region to adapt its tax rates. Financed from a land tax and resulting in services consumed by the local inhabitants, the small region's government's choices have no effect on the commuting decision, and so the government does not take into account labor supply mobility. The large region, by contrast, attempts to attract workers by facilitating them through public inputs, making them more productive. The reason is not that the government has direct financial benefits from a larger labor force, but rather, it adresses a market failure that arises due to agglomeration externalities. Thus, while the large region's policymaker spends resources in attracting workers, the behavior resembles policy competition, but there is no true interaction. The support of firms is generally too low, rather than too high (as the sub-optimally equilibrium taxes in tax competition models suggest). Because the large region's government fails to take into account its policy effect on the small region, the provision of public inputs is welfare-improving, but will generally be lower than globally socially optimal.

Policy with fully mobile residents

In the long run, in addition to deciding where to work, residents can also decide to choose where to live. As a consequence, local governments need to incorporate the effects of policy changes on the residential distribution of people. A key change with respect to the situation of residential immobility is that the small government affects the residential distribution using land taxes, and is therefore able to change the commuting flows.

To consider local welfare-maximizing taxes, we first need to derive how the residential distribution responds to tax setting. Differentiating the residential allocation (eq. 6.12) in relative terms with respect to taxes gives that:

$$\hat{\lambda} - \frac{\theta\hat{\kappa} - \lambda\hat{\lambda}}{1-\lambda-\theta\,(1-\kappa)} = \hat{G}_1/\left(1-\mu-\alpha\right) + 1/\left(1+T_1\right). \tag{6.21}$$

The left-hand side of this equation increases in λ. This, this intuitively shows that in the long run, ceteris paribus, regions host more inhabitants if they provide more public services, but fewer if they set higher land taxes. In addition to its effects on the land market, relocation has an effect on the employment distribution: if the large region grows larger residentially, more non-traded goods are demanded, and hence, more labor is required. Whether the commuting flow grows, depends on whether the labor demand grows faster than the number of residents.

Using the residential response to tax changes (eq. 6.21), it is possible to determine both governments' first-order conditions. The larger region's government, respecting the commuting equilibrium ($w_1/w_2 = 1/(1-\theta)$) and the residential equilibrium, sets a relative first-order condition with respect to taxes:

$$\left(1+t_1\right)\left(1-\mu-\alpha\right) = \hat{G}_1 + \alpha\hat{A} + \frac{\alpha}{\sigma-1}\hat{n} - \left(1-\mu-\alpha\right)\hat{\lambda}. \tag{6.22}$$

Compared to residential immobility, mobility generates two extra effects that the government takes into account: relocation may change the local employment with consequences for local productivity (third term on the right-hand side), and immigration increases demand on the housing market, which increases local rents (last term). Using the changes in the equilibrium commuting flows (eq. 6.9), and rearranging gives the following first-order condition:

$$\left(1-\mu-\alpha\right)\left(\left(1+T_1\right)-\hat{\lambda}\right)-\hat{G}_1=\frac{\alpha}{\sigma-1}\hat{n}+\alpha\hat{A}. \tag{6.23}$$

The right-hand side of this equation represents the pure productivity effects of the policy, due to increased public inputs and and increased labor supply. An optimizing government balances these benefits with the left-hand side: the cost of taxation (increased land prices due to taxes and immigration), net of public services provided (relative changes in G).

Whether region 1's private policies also coincide with its tax rate in a social optimal situation (respecting the commuting equilibrium) depends on the extent to which it takes into account its effects on the other region. Social welfare is measured as average utility $\bar{V}=\lambda V_1+(1-\lambda)V_2$. Optimizing this with respect to region 1's tax rate gives that:

$$\begin{aligned}\frac{d\bar{V}/dT_1}{\bar{V}} &= d\lambda/dT_1\left(V_1-V_2\right)+\lambda dV_1/dT_1 \\ &\quad +(1-\lambda)\,dV_2/dT_1=0, \\ \frac{dV_1/dT_1}{V_1} &= \frac{1-\lambda}{\lambda}\frac{dV_2/dT_1}{V_2}.\end{aligned} \tag{6.24}$$

The last step uses that utility levels are equal across regions in equilibrium. This implies that the relative effect of region 1's tax rate in region 1 is equal to that in region 2, weighted by the inverse relative population shares. The intuition is that region 1's local welfare maximization is socially optimal if region 1 (acts as if it) takes into account policy spillovers in region 2. Thus, to check if region 1's policies optimize global welfare, we first need to evaluate the external policy effects as the relative change in welfare in region 2 following a tax change in region 1:

$$\frac{dV_2/dT_1}{V_2}=\frac{\alpha}{\sigma-1}\hat{n}+\alpha\hat{A}+\left(1-\mu-\alpha\right)\frac{-\lambda\hat{\lambda}+\theta\hat{\kappa}}{1-\lambda-\theta\,(1-\kappa)}. \tag{6.25}$$

The welfare change in region 2 consists of the same productivity change as in region 1, plus a term that increases in the share of population in region 1. That (last) term reflects that with fewer people living in region 2, their average land consumption is higher. Inserting the mobility condition (eq. 6.21) into region 1's first-order condition shows that indeed, region 1's equilibrium tax rate has equal effects in either region (it satisfies eq. 6.25), and therefore maximizes welfare.

The result that two benevolent governments produce a maximum in global welfare is not surprising. The mobility condition states that every worker migrates to wherever utility is highest, so given the local-welfare maximizing government objectives, workers will allocate themselves according to a distribution that yields the highest utility, conditional on the fact that the common utility level $\bar{V}$ is smooth and locally concave. The equilibrium is still second-best, however, because the inefficiency associated with the smaller-than-optimal equilibrium commuting flow is not adressed.

6.4 Central government intervention on the land market

The previous section concluded that local government do not have the instruments to correct the size of the commuting flow. Therefore, subsidies on commuting may be welfare-improving. That is also extensively documented in related literature (see Borck and Wrede, 2009).

An unstudied aspect in related literature, however, is that workers not only commute, but can also migrate. The model therefore permits analyzing whether the land market policies of a central government can improve welfare. An intuitive argument for such benefits is that subsidizing housing leads to further concentration into the large city and thus internalizes the agglomeration externality. Whether or not with the aim of affecting agglomeration patterns, many countries already apply some form of ownership subsidies or mortgage rate tax deductions, and their use is debated (e.g., Glaeser and Shapiro, 2002), but not much is known about their effects in the current context.

To investigate the effect of land/housing policies in the model, a number of modifications is required. First, while assuming that housing demand is unit-elastic is not too far off empirical evidence, the fixed supply of housing in our model looks less realistic. We do not view this as a problem because for mobility considerations, the situation with fixed housing supply explains the consumer's behavior rather well (i.e., all additional demand ends up in higher house prices, and density increases due to smaller lot size). However, if we want to examine the role of housing market instruments, ignoring the supply side could omit important effects from the model. Therefore, we introduce a housing supply function, where increasing the city's physical size becomes more expensive, the larger the city is. Second, associated with non-constant marginal costs of housing, it no longer looks realistic to let excess rent flow into the non-tradables sector: the increasing marginal costs of housing supply can hardly be attributed to a constant-returns to scale sector. Therefore, we shall assume that the excess rents are redistributed to all inhabitants via the central government. This redistribution is not realistic in the model described so far (in the first place because there was no central government), but developing the model of last section under this alternative assumption does not change the conclusions.

Assume that the government can increase the physical size of the city, where the total costs of making housing available have a fixed elasticity with respect to the total size: $R = H^{\eta}/\eta$. This way, the marginal cost of supplying land are $r = H^{\eta-1}$. We shall assume that the government develops this land, and finances the development by taxing inhabitants. To avoid that the financing of the housing subsidy drives the welfare results, we shall assume that the commuting decisions are neutral to the financing; i.e. that central government transfers are proportional to the wage.[7] To leave the optimal subsidizing scheme implicit, we assume that the ratio of rents net of subsidies to the actual rent is S, so that in the optimum, S captures the ratio at which housing prices are subsidized. The local government takes the centrally set subsidy policy as given. The land market equilibrium then states that $rHS = \left(1-\alpha-\mu\right) w \left(1-T_c\right)$, where T_c is the central tax. Equating the marginal price to the costs of developing land suggests that $H^{\eta} = \left(1-\alpha-\mu\right) w \left(1-T_c\right)/S$. Then, the equilibrium land rent is:

$$r = \left(\left(1-\alpha-\mu\right) w \left(1-T_c\right)/S\right)^{1-1/\eta}.$$

If η is higher, the marginal costs of supplying land increase quickly. If η tends to infinity, the model converges to the result under fixed land supply; the cost of supplying new land become infinite and the supply is completely inelastic. The respective utility functions, cleared of wages are (see the derivation in section 6.3, updated with the land rent from this section):

[7] The transfer could also be viewed as a reduction in the wage tax that a central government sets.

$$
\begin{aligned}
V_1 &= \frac{G_1\left(1-T_c\right)}{\left(A^{-1}\left(\kappa-\mu\lambda\right)^{1/(1-\sigma)}\right)^{\alpha}\left(\lambda/\left(1+T_1\right)/S_1\right)^{(1-\mu-\alpha)}}; \qquad (6.26)\\
V_2 &= \frac{G_2\left(1-T_c\right)(1-\theta)}{\left(A^{-1}\left(\kappa-\mu\lambda\right)^{1/(1-\sigma)}\right)^{\alpha}\left(\frac{(1-\kappa)(1-\theta)+\kappa-\lambda}{\left(1+T_2\right)S_2}\right)^{(1-\mu-\alpha)}}.
\end{aligned}
$$

Realizing that n is now proportional to $\kappa-\mu\lambda$ instead of $\kappa-\varphi\lambda$ (because excess land rent no longer flows back into the non-tradables), and T and S (the central government's tax and subsidy rate) are taken as given by the local governments, their first-order conditions are still:

$$
\begin{aligned}
\left(1+T_1\right)\left(1-\mu-\alpha\right) &= \hat{G}_1+\alpha\hat{A}+\frac{\alpha}{\sigma-1}\hat{n}-\left(1-\mu-\alpha\right)\hat{\lambda}, \qquad (6.27)\\
\left(1+T_2\right)\left(1-\mu-\alpha\right) &= \hat{G}_2+\left(1-\mu-\alpha\right)\frac{-\lambda\hat{\lambda}+\theta\hat{\kappa}}{1-\lambda-\theta\left(1-\kappa\right)},
\end{aligned}
$$

so, in relative changes, the first-order conditions for local governments do not change. Moreover, we know that these first-order conditions lead the first region's government to internalize the effects of its policy on the second region because any utility differences are eliminated through migration. Therefore, instead of optimizing global welfare, the central government can optimize the welfare function in region 1, with the constraint that welfare is equal between the two regions. The second constraint, the balanced budget constraint that the central government faces is that the central wage tax needs to finance the subsidies on housing in both regions. Optimizing the welfare in region 1 with respect to both subsidy rates, while taking into account the effect on the residential pattern gives the two first-order conditions on use of land subsidies in regions 1 and 2 for the central government:

$$
\begin{aligned}
0 &= \frac{-dT_c/dS_1}{1-T} - \frac{\alpha}{1-\sigma}\frac{dn}{d\lambda}\frac{\lambda}{n}\frac{d\lambda/dS_1}{\lambda} \\
&\quad + \left(1-\mu-\alpha\right)\left(1-1/\eta\right)\left(1-\frac{d\lambda/dS_1}{\lambda}\right), \\
0 &= \frac{-dT_c/dS_2}{1-T} - \frac{\alpha}{1-\sigma}\frac{dn}{d\lambda}\frac{\lambda}{n}\frac{d\lambda/dS_2}{\lambda} \\
&\quad + \left(1-\mu-\alpha\right)\left(1-1/\eta\right)\left(-\frac{d\lambda/dS_2}{\lambda}\right).
\end{aligned} \tag{6.28}
$$

The terms dT_c/dS_1 and dT_c/dS_2 are identified from the balanced budget constraint of the central government. They capture that increasing the subsidy decreases welfare, because the taxes that finance them go up. In a closed economy, housing subsidies are neutral or distortive to welfare, such that $-dT_c/dS_1/(1-T) + \left(1-\mu-\alpha\right)\left(1-1/\eta\right) < 0$ (i.e. in terms of utility, the returns to lowering the land price are lower than the cost of raising the required funds). The other terms capture the effect of the subsidy on the local number of firms (productivity), and on the land rents; directly via the price of land, and indirectly via the migration that they cause. Adding the two first-order conditions, optimal central government policy satisfies:

$$
\begin{aligned}
0 &= \frac{-\left(dT_c/dS_1 + dT_c/dS_2\right)}{1-T} \\
&\quad - \left(\frac{\alpha}{1-\sigma}\frac{dn}{d\lambda}\frac{\lambda}{n} + \left(1-\mu-\alpha\right)\left(1-\frac{1}{\eta}\right)\right)\left(\frac{d\lambda/dS_1}{\lambda} + \frac{d\lambda/dS_2}{\lambda}\right) \\
&\quad + \left(1-\mu-\alpha\right)\left(1-\frac{1}{\eta}\right).
\end{aligned} \tag{6.29}
$$

By assumption, $-dT_c/dS_1/\left(1-T_c\right) + \left(1-\mu-\alpha\right)\left(1-1/\eta\right) \leq 0$, and since a subsidy in region 2 invariably requires central taxation, the central tax rises in subsidies and $-dT_c/dS_2/\left(1-T_c\right) < 0$. Thus, we have that optimal policy requires that:

$$
\left(\left(1-\mu-\alpha\right)\left(1-1/\eta\right) - \frac{\alpha}{\sigma-1}\frac{dn}{d\lambda}\frac{\lambda}{n}\right)\left(\frac{d\lambda/dS_1}{\lambda} + \frac{d\lambda/dS_2}{\lambda}\right) < 0
$$

The first term of this product captures the effects of residential migration towards region 1. The two channels are the effect on house prices ($\big(1-\mu-\alpha\big)\big(1-1/\eta\big)$) and the effects on the variety of firms in tradables production. The second term of the product is the net effect of the subsidies on (re)location behavior. This migration response can be inferred from totally differentiating the spatial equilibrium constraint, which holds that $V_1 = V_2$. We detail this in Appendix 6.C. Taking local policies as constant, this yields that $d\lambda/dS_1/\lambda = 1/\big(S_1(1-\zeta)\big)$ and $d\lambda/dS_2/\lambda = -1/\big(S_2(1-\zeta)\big)$, where $0<\zeta<1$ is a term that depends on κ and λ. The elasticity of the firm base with respect to the number of residents is positive, which we also show in Appendix 6.C. Given this positive elasticity, the optimal design of the policy is ambiguous.

The structure of the optimal policy depends on the sign of $\frac{\alpha}{1-\sigma}\frac{dn}{d\lambda}\frac{\lambda}{n} + \big(1-\mu-\alpha\big)\big(1-1/\eta\big)$, which, in turn, depends on the elasticity of housing supply. If η is very high (housing supply is inelastic), it can be positive, in which case we require that $\frac{d\lambda/dS_1}{\lambda} + \frac{d\lambda/dS_2}{\lambda} = \big(1/S_1 - 1/S_2\big)/(1-\zeta)$ be negative. Thus, under inelastic housing supply, an optimal policy needs to increase the percentage of subsidy in the value of housing, such that inhabitants of the large region benefit more from the policy not only in absolute terms (the subsidy is higher in absolute terms), but also in relative terms (the percentage decrease in housing prices needs to be larger in the large region). Vice versa, if housing supply is elastic and η is small, such a regressive scheme (land with higher value gets a higher percentage subsidy) is welfare-decreasing. The critical η for which the optimal subsidy scheme turns from progressive to regressive is given by $\eta^* = \big(1-\mu-\alpha\big)/\Big(1-\mu-\alpha\Big(1+\frac{dn}{d\lambda}\frac{\lambda}{n}/(\sigma-1)\Big)\Big) > 1$. At that point, increasing residential size in one region balances the (social) returns in terms of firm variety with the price increases. Because $\eta^* > 1$, even if the costs of supplying housing are convex ($\eta > 1$), a regressive subsidy scheme might be optimal: the difference between private and social benefits of labor concentration are sufficiently large. Should η be equal to this critical value, then the optimal scheme is to provide subsidies that are proportionally neutral, so they do not affect the residential distribution. However, if the housing subsidies do not improve welfare in the closed economy, this suggests the optimal scheme is to provide no subsidies.

Because land supply is no longer inelastic in this setup, subsidies by a central government can be welfare improving. Clearly, an optimal subsidy that increases the size of the large region corrects for too few inhabitants in the large region in the market equilibrium. The intuition is that house prices rise quickly in the size of population if η is large (inelastic housing supply), so agglomeration externalities cannot be internalized. Since agglomeration externalities are unambiguously positive in the model, that does not explain why subsidies that spread inhabitants can also be optimal. The key to this insight is that the suboptimal size of the commuting flow ($\kappa - \lambda$) can have two sources: a too small number of workers in the large region (κ) or a too large number of inhabitants in the small region ($1-\lambda$). Therefore, a too small commuting flow implies that if the worker allocation is optimal, then there is insufficient residential concentration. Or, vice versa, if the number of inhabitants in either region is optimal, the labor supply in the large region is too low. Whether the commuting inefficiency mostly translates to insufficient worker concentration or over-dispersed residents depends on the housing supply elasticity. Because η governs the rate at which house prices rise in concentrated areas, high η hampers agglomeration gains from concentration via strong house price increases, while low η fosters concentration, by contrast. The land supply elasticity for which no policy is desired (η^*) thus balances the effect effect of too small commutes between residential overconcentration and employment underconcentration. For elasticities lower than η^*, the dominant inefficiency is residential overconcentration; for elasticities higher than η^* (including previous section where η was infinite), the dominant inefficiency is a lack of employment concentration.

6.5 Conclusion

This chapter studies whether commuting options lead governments to set suboptimal policies. If workers are residentially immobile but have the possibility to commute, policy is generally not efficient. Agglomeration externalities in the large region lead to expenditure on public inputs by the large region's government. However, the policies that attract workers to the agglomeration are financed by the large region, but benefit inhabitants of both the large region and the small region. Therefore, the large region will spend less than is socially optimal on public inputs to attract commuters. This is in contrast to the standard tax competition literature, which concludes that the support for mobile factors is too large, rather than to small.

If workers are residentially mobile, they migrate to where utility is highest. In that case, the large region fully internalizes the spillovers of its policy, because workers will migrate into the large region until the spillover is internalized. Still, this equilibrium is second best. This can be traced back to two inefficiencies. First, none of the policies adress the issue that commuters ignore the agglomeration externalities that they confer to peers. The second is that the residential distribution for which the policy spillovers are internalized need not be the optimal residential allocation. In particular, it is possible that welfare is improved if fewer people lived in the smaller region, and contribute to the productive public inputs in the large region. The inefficiency thus indirectly stems from the requirement that local government budgets need to be balanced: ideally, small region governments would finance part of the public inputs in the large region.

While the model is closely related to traditional tax competition models, the results diverge substantially. In particular, one could follow the tax competition literature in the presence of two instruments (e.g., Razin and Sadka, 1991), and assume that the supply of capital in other regions is similar to mobile labor supply through commutes. In that case, one would expect governments to provide more than optimal public inputs (the equivalent of undertaxation of capital) financed from high taxes on land (Gordon and Hines Jr., 2002). In the current model, the provision of public inputs is inefficiently low, rather than inefficiently high. This difference can be traced back to two externalities that are absent in models of capital taxation. First, there are good market spillovers: the small region benefits from increased productivity elsewhere. Second, in contrast to a situation with fully mobile capital, the commuting costs lead to a suboptimal allocation of labor, which is not accounted for by the large region's policymaker. A last difference is that as the regions are endogenously asymmetric, harmonization of public inputs or tax rates is generally not desirable (which is a recommendation of many tax competition models).

Given that our model is one of the first to study residential mobility jointly with commuting, the chapter investigates whether a central government intervention in the housing market can improve welfare. The optimal design depends on specific parameters, of which the housing supply elasticity is an essential one. If the housing supply is inelastic, it is optimal to let the subsidy as a percentage of housing costs increase in the housing price. The reverse holds if housing can be elastically supplied. The reason is that if the costs of housing do not rise quickly in the quantity, the market equilibrium compensates part of the smaller-than-optimal commuting flow by increasing the residential size of the large city. If the housing supply is inelastic, however, residents will spread and the main inefficiency due to suboptimal commuting is insufficient employment concentration.

Finally, the particular set of assumptions that lead to these conclusions is generalizable in a number of dimensions. First, the microfoundations of the agglomeration externality are a localized diversity effect in input production. It would be feasible, however, to consider the resulting reduced form as a Sheshinski formulation of Marshallian externalities or general scale effects. Moreover, chapter 5 shows that the spatial equilibria are consistent with a large set of equilibria that a new economic geography model would yield. Second, the way in which local governments attract commuters is to build infrastructure that improves their productivity. This is arguably a more realistic setup for local governments that can generally not tax labor. However, since the decision to supply local public inputs has an opportunity cost in terms of higher taxes on land (the immobile factor) or supplying fewer public services to citizens, the trade-off is similar to providing financial incentives (subsidies or lower taxes) to stimulate local employment. Last, while the policy conclusions critically depend on the agglomeration externality, we note that another important externality is missing in the chapter: transport congests with increasing number of commuters. Congestion externalities have no immediate impact on local policy, except that they offset the desirability to agglomerate via commuting. Therefore, the model seems readily generalizable to a case where the congestion effects work against the agglomeration benefits. In that case, our results seem to hold if the agglomeration externalities are larger than congestion effects, but might be affected otherwise.

6.A The optimal size of the commuting flow

To see that the market equilibrium commuting flow is smaller than socially optimal, we first use that the utility functions of individuals that work and live in region 1 are related to those that work in region 1 but live in region 2, and to those that live and work in region 2. Define the consumer price index as $\bar{P}_r = P_{tr}^{\alpha} P_{tr}^{\mu} r_r^{1-\mu-\alpha}$. Then, from the indirect utility function, we have that the utility of a commuter into region 1 is equal to the utility of a worker/resident of region 1, multiplied by a factor $(1-\theta)\bar{P}_1/\bar{P}_2$. Similarly, the utility level of a worker/resident in region 2 is equal to the utility level of a worker/residents in 1, multipled with $w_2/w_1\left(\bar{P}_1/\bar{P}_2\right)$. Inserting the production structure of traded and non-traded goods gives that the social welfare is optimized when the following function is optimized:

$$S = \left(\kappa - \varphi\lambda\right)^{\alpha\sigma/(\sigma-1)} \left(\varphi\lambda\right)^{\mu} \left(\begin{array}{c} \lambda + (\kappa-\lambda)(1-\theta)\bar{P}_1/\bar{P}_2 \\ +(1-\kappa)w_2/w_1\left(\bar{P}_1/\bar{P}_2\right) \end{array} \right). \quad (6.A.1)$$

The first argument, $\left(\kappa - \varphi\lambda\right)^{\alpha\sigma/(\sigma-1)} \left(\varphi\lambda\right)^{\mu}$, weighs the production in traded and non-traded goods in the utility function. The second argument weighs the different groups in the economy: a share λ of inhabitants live and work in region 1 (and the utility is proportional to the first argument of the product), share $\kappa-\lambda$ commutes, and the residual lives and works in region 2.

Optimizing this function in relative terms with respect to the employment location κ (given that λ is constant in the short run, this also optimizes the commuting flow) yields that $dS/d\kappa/S = 0$ when:

$$\frac{\alpha\sigma}{\sigma-1}/\left(\kappa-\varphi\lambda\right) + \frac{\bar{P}_1}{\bar{P}_2} \frac{(1-\theta)-w_2/w_1}{\left(\begin{array}{c} \lambda + (\kappa-\lambda)(1-\theta)\bar{P}_1/\bar{P}_2 \\ +(1-\kappa)w_2/w_1\left(\bar{P}_1/\bar{P}_2\right) \end{array} \right)} = 0 \quad (6.A.2)$$

The private equilibrium entails that $w_1/w_2 = 1/(1-\theta)$. Setting this cancels the second term in the efficiency condition. Since the first term is strictly positive (given that $\sigma > 1$), the market commuting equilibrium is not efficient. To satisfy efficiency, we need that $(1-\theta) - w_2/w_1 < 0$, or rewriting, that $w_1/w_2 < 1/(1-\theta)$. Because the large region's wage schedule is downward-sloping in labor supply, the social optimal size of the commuting flow must be larger than the flow in the market equilibrium.

6.B The optimal residential allocation

Taking the residential distribution (λ) as the second argument to optimize the social welfare function gives a second efficiency condition for the long run. Again, for commuters, the welfare function is equal to V_1 times $\tau^{-\alpha}(1-\theta)\left(G_2/G_1\right)\left(P_1/P_2\right)$, and for workers in region 2, this is term is equal to $\tau^{-\alpha}\left(w_2/w_1\right)\left(G_2/G_1\right)\left(P_1/P_2\right)$. The welfare function and corresponding efficiency condition with respect to λ is:

$$S = \left(\kappa-\mu\lambda\right)^{\alpha\sigma/(\sigma-1)}\left(\mu\lambda\right)^{\mu}\lambda^{-\left(1-\mu-\alpha\right)} \times\begin{pmatrix} \lambda+(\kappa-\lambda)\tau^{-\alpha}(1-\theta)G_2/G_1\left(\bar{P}_1/\bar{P}_2\right) \\ +(1-\kappa)\tau^{-\alpha}w_2/w_1G_2/G_1\left(\bar{P}_1/\bar{P}_2\right)\end{pmatrix}. \tag{6.B.1}$$

$$0 = -\mu\frac{\alpha\sigma}{\sigma-1}/\left(\kappa-\mu\lambda\right)-\left(1-\alpha-2\mu\right)/\lambda + \frac{1-\tau^{-\alpha}(1-\theta)G_2/G_1\left(\bar{P}_1/\bar{P}_2\right)}{\begin{pmatrix} \lambda+(\kappa-\lambda)\tau^{-\alpha}(1-\theta)G_2/G_1\left(\bar{P}_1/\bar{P}_2\right) \\ +(1-\kappa)\tau^{-\alpha}w_2/w_1G_2/G_1\left(\bar{P}_1/\bar{P}_2\right)\end{pmatrix}}. \tag{6.B.2}$$

In the market equilibrium, equal utility across regions implies that the term $\tau^{-\alpha}(1-\theta)G_2/G_1\left(\bar{P}_1/\bar{P}_2\right)$ is equal to unity, and so the second term in the efficiency condition is zero. Therefore, the residential allocation in the market equilibrium is equal to the social optimum in the specific case that $-\mu\frac{\alpha\sigma}{\sigma-1}/\left(\kappa-\mu\lambda\right)-\left(1-\alpha-2\mu\right)/\lambda=0$, which boils down to a particular parameter constellation. The first part of this term is negative, because $\kappa-\mu\lambda>0$. The second term $\left(1-\alpha-2\mu\right)/\lambda$ reflects the consumption of land per head and the production of the local good. Thus, under the most plausible parameter sets, where μ is not prohibitively large, the first term is negative, and the second term is positive in the social optimum. The latter requires that the utility level would be higher in the larger region. Noting that this would lead to migration in the market equilibrium, this implies that as long as $-\mu\frac{\alpha\sigma}{\sigma-1}/\left(\kappa-\mu\lambda\right)-\left(1-\alpha-2\mu\right)/\lambda<0$, the large city is residentially too large in the market equilibrium, compared to the social optimum.

6.C Subsidy effects on the number of firms and residents

In Section 6.4, we deferred the definition of the term ζ because its derivation distracts from the central argument. The term ζ captures a partial derivative of the spatial equilibrium condition with respect to the residential distribution. The ratio of utility functions in the presence of the central government is given by:

$$V_1 = \frac{G_1(1-T)}{G_2(1-\theta)\tau^{-\sigma}} \left(\frac{\lambda/\left(1+t_1\right)/S_1}{\frac{(1-\kappa)(1-\theta)+\kappa-\lambda}{\left(1+t_2\right)S_2}} \right)^{\left(1-1/\mu\right)\left(1-\mu-\alpha\right)}, \tag{6.C.1}$$

where our interest is in the implicit differentiation to obtain $d\lambda/dS_1$ ($d\lambda/dS_2$ is obtained in a similar way). The derivative with respect to λ, including the effect via the employment distribution κ, is equal to:

$$\begin{aligned} &\frac{d\lambda}{\lambda} - \frac{d\left((1-\kappa)(1-\theta)+\kappa-\lambda\right)}{(1-\kappa)(1-\theta)+\kappa-\lambda} \\ &= \frac{d\lambda}{\lambda} - \frac{d\kappa\theta/(1-\theta)-d\lambda}{(1-\kappa)(1-\theta)+\kappa-\lambda} \end{aligned} \tag{6.C.2}$$

From the equilibrium expression for κ, we have that:

$$\frac{d\kappa}{\kappa} = \frac{\frac{\mu}{1-\theta\alpha A}d\lambda}{\frac{(1-\theta)\alpha A+\mu\lambda}{1-\theta\alpha A}} = \frac{\mu\lambda}{(1-\theta)\alpha A+\mu\lambda}\frac{d\lambda}{\lambda}. \tag{6.C.3}$$

Collecting terms, the net relative change in the residential distribution, including the employment distribution changes are:

$$\begin{aligned} \frac{d\lambda}{\lambda} - \frac{d\left((1-\kappa)(1-\theta)+\kappa-\lambda\right)}{(1-\kappa)(1-\theta)+\kappa-\lambda} &= \\ \frac{d\lambda}{\lambda}\left(1-\frac{\kappa\theta/(1-\theta)\frac{\mu\lambda}{(1-\theta)\alpha A+\mu\lambda}-\lambda}{(1-\kappa)(1-\theta)+\kappa-\lambda}\right) &= (1-\zeta)\frac{d\lambda}{\lambda}. \\ \text{with } \zeta \equiv \frac{\kappa\theta/(1-\theta)\frac{\mu\lambda}{(1-\theta)\alpha A+\mu\lambda}-\lambda}{(1-\kappa)(1-\theta)+\kappa-\lambda} & \end{aligned} \tag{6.C.4}$$

As a result, keeping other policies constant, the spatial equilibrium condition implies that:

$$\begin{aligned} 0 &= \left(1-1/\mu\right)\left(1-\mu-\alpha\right)\left((1-\zeta)\frac{d\lambda}{\lambda}-\frac{dS_1}{S_1}\right), \\ \text{so } \frac{d\lambda/dS_1}{\lambda} &= \frac{1}{(1-\zeta)S_1}. \end{aligned} \tag{6.C.5}$$

Additionally, we used that the elasticity of the number of firms with respect to local population was positive. The number of firms is proportional to $\kappa - \mu\lambda$. Inserting the expression for the equilibrium κ yields:

$$\kappa - \mu\lambda = \frac{\alpha A\left(1-\theta+\theta\mu\lambda\right)}{1-\theta\alpha A}. \tag{6.C.6}$$

Totally differentiating and writing in relative terms results in:

$$\frac{d\left(\kappa-\mu\lambda\right)}{\kappa-\mu\lambda} = \frac{\lambda\theta\mu}{\left(1-\theta+\theta\mu\lambda\right)}\frac{d\lambda}{\lambda}, \tag{6.C.7}$$

where $1 > \lambda\theta\mu/\left(1-\theta+\theta\mu\lambda\right) > 0$, and therefore the net effect of an increased number of residents on the firm base is positive (though less than proportional).

CHAPTER 7

POLICY COMPETITION FOR HIGH-SKILLED WORKERS

7.1 Introduction

Over the last decades, cities with more high-skilled workers have developed faster than their less skilled counterparts (Glaeser and Saiz, 2004; Glaeser and Shapiro, 2003). At the same time, urban amenities like culture, natural surroundings, and nightlife have played a major role in attracting high-skilled workers (Adamson et al., 2004; Glaeser et al., 2001). The benefits of hosting high-skilled workers might trickle down by creating additional jobs, improving overall productivity, or increasing the range of goods available in a city. If supplying amenities can promote the commendable effects of high-skilled workers on a city, policmakers' expenditure might be directed toward such aims. For instance, the success of revitalization projects in the Spanish city of Bilbao, with the Guggenheim museum as a cornerstone,[1] led up to a series of significant European Structural Funds investments in culture-based urban development plans. If urban amenities help the economic development of towns and cities, then city governments can justify subsidizing of museums, clean air, opera houses and green areas. But are such policies, then, desirable?

The purpose of this chapter is to discern whether urban policies are optimal when governments interact and workers of different skills can migrate between cities. While there is evidence that urban amenities can play an important, if not crucial role in urban development plans, it is restrictive to look at a city in isolation. City interactions in policy, trade, and migration of different types of workers suggest that studying the advantages of amenity-based policies requires a broader picture, including peer cities. In doing so, we broaden the scope of the debate on government-provided cultural and other amenities by introducing the spatial implications of such decisions. Furthermore, we extend insights in the literature

[1] The programme was perceived so succesful that the museum was namesake for the "Guggenheim effect": a European commissioner remarked that it was "a good example of investment in bricks accompanied by investment in people rejuvenating a city and a region... an example of a complex process for the benefit of knowledge creation, knowledge exploitation and vision" (European Union Press release IP/08/875).

of tax competition for financial capital towards (local) governments that are likely to care about the types of inhabitants and firms that they attract. The new insights compared to the tax competition literature stem from the fact that workers are heterogenous and incorporate policy preferences in their migration choices.

A central insight in the chapter is that a city's more diverse labor skill mix nurses production or consumption, but complicates policymaking: policy preferences are also diverse. If inhabitants are very homogeneous, the policy can be tailored well to citizens' specific preferences. This comes at the cost of productive specialization, or having to import goods. If the effects of policy efficiency dominate, the model allows for endogenous sorting of workers into high and low-skilled cities. By contrast, symmetric cities emerge if the decreasing returns to productive specialization or the transport costs are high. The skill differences introduced in this chapter and the sorting patterns and endogenous specialization that arise as a result turn out to be crucial to the type of distortions that enter policy decisions.

The results show that under symmetry, even with general welfare in mind, cities bias their policies towards high-skilled workers, thus overproviding the public goods that high-skilled workers prefer. high-skilled workers weigh more heavily in the policymakers' decisions, through their higher incomes, and productivity effects conferred to others. Therefore, the bias occurs despite the fact that low-skilled workers are perfectly mobile and can migrate away from the policies that they dislike. If cities specialize in low or high skills, they become more efficient in policy-making, because policy preferences are relatively uniform across inhabitants. The cost of policies that increase specialization (lower productivity, higher transport costs) are taken into account for the own city, but not for partner cities. Therefore, if cities specialize, they specialize too much. Trade in differentiated goods lowers the costs of city specialization, thus allowing for efficient policy. Nevertheless, the patterns of specialization are not efficient.

The next section briefly sketches the context of this chapter. Section 7.3 outlines the assumptions and equilibrium conditions of a two-city, two-skill model. It studies the implications of strategic tax setting among city governments, and considers the effect of different production externalities on the policy outcomes. Section 7.4 includes multiple tradable goods into the model, so the effects of intercity trade can be discussed. Additionally, it extends the earlier results to many cities and skill groups. Section 7.5 concludes.

7.2 Amenities in the competition for high-skilled workers

High-skilled workers are a sought-after population group among policy-makers. Different researchers identify productivity effects of high-skilled workers (Rauch, 1993; Glaeser and Saiz, 2004), leading to growth in the city's employment and its size. A related literature finds that there are external returns to education: working with better-educated peers increases one's own productivity (see Moretti, 2004 for an overview). Similarly, high-skilled workers carry out most of the productivity-enhancing R&D and have the highest absorptive capacity to benefit from technological advances (Furman et al., 2002). Additionally, the effect of high-skilled workers on population growth may also work through other channels than increasing productivity. Shapiro (2006) finds that roughly a third of the population growth effects is due to changes in the quality of life. Such externalities might include differences in the type of firms attracted to the location, provision of public goods and lower crimes rates. Given these effects, it is not surprising that attracting high-skilled workers, and the industries that employ them, is a traditional policy objective (Malecki, 1981).

High-skilled workers can be attracted by different strategies, however. If strategies are targeted well, they primarily attract high-skilled workers. Less specifically targeted policies, such as subsidies on housing or lower taxes might benefit low- and high skilled workers alike, such that the policy will attract both, leaving the worker skill composition unchanged. high-skilled workers appear to be particularly sensitive to several amenities, and providing those amenities plausibly attracts high-skilled workers. Dalmazzo and de Blasio (2011) document that Italian high-skilled workers benefit disproportionally from local public goods like transport, health and schooling and from cultural amenities like museums. Similarly, Florida (2002b) reports that U.S. workers with a bachelor's degree or higher are especially attracted to areas with specific merits like cultural and nightlife amenities (Florida's "coolness" of a city). In fact, even if the creative class or *bohemiens* do not directly affect productivity, they can attract people who do improve productivity via cultural channels. Thus, the argument can be extended that the cultural policies eventually improve economic performance (Florida, 2002a). Falck et al. (2011) show that in Germany, the presence of Baroque-era opera houses attracted workers with high levels of human capital, and subsequently, that these workers have increased local growth rates. These insights are, presumably, not new to most policymakers: surveys of local US governments show that many have devised plans to develop their cities using cultural amenities (Grodach and Loukaitou-Sideris, 2007).

Since this chapter examines cities' competition to attract a mobile production factor, there are clear links with the tax competition literature for (financial) capital. In that literature, governments bias policy towards capital (i.e., tax them too little) to avoid tax base erosion. However, in this chapter's context, the production factor to be attracted are workers (not capital), and the instrument is the provision of amenities, rather than (low) tax rates. The model that Buettner and Janeba (2009) use to structure their empirical study produces similar results to the tax competition literature, because governments balance the positive externalities of high-skilled mobile workers with the cost of attracting them via cultural policies. Maximizing the welfare of low-skilled ("non-creative") workers, governments in this model treat high-skilled workers like financial capital, although their appeal is in productive externalities, not tax revenue. As low-skilled workers are assumed to be immobile, cities provide too many public goods, to the liking of mobile highly educated workers, but to the disfavor of lesser educated, immobile workers, harming overall welfare.[2]

Although related, a number of features sets this study apart from the tax competition literature. These follow inevitably from shifting the focus from capital to different groups of people. As inhabitants care about real wages and local policy instead of financial returns, policymakers need to select different instruments. As a result, the central problem in policy formation changes from attaining sufficient tax revenue into efficient expenditure decisions. The government's main concern is then with the composition and size of the local economy, rather than with its stock of capital. We argue that especially for local (urban) governments, a view based on expenditure decisions and policy devised to target specific groups may be more accurate than a view based on financial capital mobility. Another intuitive consequence of workers caring about their "quality of life" is that a Tiebout-motive enters: workers, irrespective of their type, can vote with their feet. Such mobility adds an extra tension compared to tax competi-

[2]We do not mean to criticize the model of Buettner and Janeba, as for its purpose of motivating empirical analyses, it is rather effective.

tion: attracting specific workers through specialized policies may induce others to leave. Indeed, the possibility for skill differences to arise between cities endogenously is not found in tax competition models, but matters greatly for the type of inefficiencies that arises from policy competition.

7.3 Competition for high-skilled workers in a two city setting

We study a two-by-two economy, where workers of two skill types inhabit one of two cities. Workers consume local housing, a numeraire consumption good, and a local amenity provided by the government. Workers migrate freely to where the utility of living is highest. Overall, high-skilled workers, K, are more productive than low-skilled workers L, and (consistent with that) high-skilled workers earn a higher wage. The uniform preferences of both types of workers are given by a Cobb-Douglas function over the consumption of housing h, consumption of a numeraire good c, and government-provided amenities m:

$$U = h^{1-\alpha} c^{\alpha} m^{\gamma}. \tag{7.1}$$

Workers maximize utility subject to their budget constraint. They face a housing rent, r; the consumption good is a numeraire; and the amenity provided by the local government is financed from taxes. The budget constraint reads that income after taxes equals expenditure on housing and consumption: $w - T \geq rh + c$. Maximizing utility subject to the private budget constraint yields the standard Cobb-Douglas demand functions:

$$h = (1-\alpha)(w-T)/r; \; c = \alpha(w-T). \tag{7.2}$$

Using these demand curves generates an indirect utility function $V = \zeta(w-T) r^{-\alpha} m^{\gamma}$, where ζ is a parametric constant. Assuming that there are K high-skilled workers and L low-skilled workers present, the equilibrium land rent can be written as $r = (1-\alpha)\left(Kw_k + Lw_l\right)/H$, where H is the aggregate supply of land in the city. We shall assume that the land market clears at the equilibrium price for which all land is inhabited; and land owners use the rent to consume the numéraire good.

We assume that the government finances the amenities with lump sum head taxes. Paired with Cobb-Douglas preferences for government-provided amenities, workers with higher incomes will prefer cities where the lump-sum taxes are higher. This aligns with empirical findings that high-skilled workers have political preferences for amenities and are attracted to regions where the government supplies relatively many amenities.[3]

Using head-tax financing of publicly provided goods, a worker's preferred tax is the tax that maximizes his utility. To avoid results driven by scale effects, we shall assume that the level of government services is proportional to the tax levied per head: $(K+L)m = (K+L)p_g T$, where the implicit price of government services, p_g, is constant. From the unit-elastic preferences, workers ideally want a policy that transforms a fixed share of their income into the amenity. However, since the tax is raised lump sum, the preferred tax (that amounts to constant fraction of income) is higher for workers with a higher wage. Optimizing the indirect utility function with respect to T and rewriting gives the worker's preferred tax rate:

$$T = w - 1/\left(\gamma \hat{m} - \alpha \hat{r}\right),$$

where a hat denotes the relative change in a variable due to changes in the tax rate ($\hat{x} = dx/dT/x$). The second term, which lists changes in the amenity and land rent, is equal for all inhabitants (taxes are increased until the marginal utility of amenities become small or land rent increases grow large). As other terms are constant across skill individuals, the heterogeneity in preferred tax rates is due to variations in individuals' wages. This ensures that as long as high-skilled workers earn more, they prefer a higher tax rate to finance amenities, which is effectively a consequence of the lump-sum way of financing them.

Assume that the government acts as a welfare-maximizer in a closed city, and optimizes welfare $KV_k + LV_l$ by choosing the tax rate T. The government's first-order condition for welfare with respect to taxes satisfies $(1-\alpha)(K+L)/(Q-T(K+L)) = \gamma/T$. The left-hand side is the cost of taxation, of which fraction α tranlates into lower property prices;

[3] This is not to say that lump-sum taxation most accurately describes local government financing. To capture the luxury nature of amenities, a Stone-Geary could be used, but it yields similar results under slightly more involved analytical steps. An alternative is to assume outright that highly-skilled workers have stronger preferences for amenities (i.e., heterogeneity in γ), for which Buettner and Janeba (2009) provide an empirical justification. An advantage of the lump sum model over the heterogeneous parameter is that the preference for amenities becomes endogenous: as the spatial organization of the economy changes, so do workers' incomes and therefore their preferences for amenities. This effect is eliminated in the heterogeneous-preference model.

the right-hand side reflects the benefits in terms of amenities ($\hat{m} = 1/T$). Rewriting the first-order condition for the optimal tax rate gives:

$$T = \frac{L}{K+L} w_l + \frac{K}{K+L} w_k - 1/\big(\gamma/T - \alpha\hat{r}\big). \tag{7.3}$$

This simply states that the optimal tax is the city's average wage less a term that captures the effects on utility of taxes via amenities and land rents. The latter is equal for both skills. As a result, in a city of mixed skills that sets taxes to maximize average welfare, high-skilled workers face a tax that is lower than they ideally prefer, and consequently, a lower level of amenities than they prefer is provided. The reverse holds for low-skilled workers: they prefer a tax rate lower than the city tax rate.

To discuss how skills mix in the city, it is crucial to understand how workers with different skills complement or substitute for each other. A production function with constant elasticity between the two skills allows for different degrees of substitutability that workers might have in production:

$$Q = A(K^\rho + L^\rho)^{1/\rho}. \tag{7.4}$$

If the parameter ρ is high (i.e., the close to one), low and high-skilled worker are relatively easy to substitute in production. The producers that employ the worker produce the numeraire good, and they take the aggregate productivity term A as given. Consequently, the first order conditions for hiring either type of worker reveal the wage (i.e., the marginal productivity):

$$\begin{aligned} w_l &= A(L^\rho + K^\rho)^{1/\rho - 1} L^{\rho-1} = Aa_L; \\ w_k &= A(L^\rho + K^\rho)^{1/\rho - 1} K^{\rho-1} = Aa_K. \end{aligned} \tag{7.5}$$

We will set up a two-city system, to study which workers end up where, and how policymakers respond to worker mobility. Different spatial equilibria can emerge in the two-city model, but some situations can be excluded beforehand. First, if one city is completely empty, its land rents are zero. As can be seen from the indirect utility function, the potential utility of living in an empty city tends to infinity, in that case. Therefore, regions are never empty in equilibrium. A similar Inada-type condition holds for the mixing of skills inside a city. If low and high-skilled workers are imperfect substitutes ($\rho < 1$), then if a city only houses one specific type of worker, the potential wage for the other type will tend to infinity (as can be seen from the wage equations, letting the employment of the worker group tend to zero). Therefore, the scarcity of skills will always lead workers to migrate to a city that has none of their specific skill. Therefore, we can exclude cities that are completely empty or completely specialized in one skill from the possible outcomes.

While perfectly specialized constellations are not possible, imperfect specialization may well be possible. The worker's tradeoff is between settling in a city that has high wages for his skill and settling in a city that has good policies. As the government sets a tax that is an average of what the different skill groups in the city prefer, the tax rate could be perfectly tailored if workers sort into cities according to skill. However, that would lead to inefficiency of the input mix of firms, and hence to lower wages. This balance determines whether cities specialize in a skill type, or host the average skill distribution.

We assume that the government maximizes the average welfare of its inhabitants. Whereas in the closed-city case, the tax effects on production and the tax base are neutral due to the immobility of workers, this is no longer true in the open city. When workers are allowed to migrate, increases in the tax rate lead high-skilled workers to migrate into a city, and low-skilled workers to migrate away. As a consequence, given changing local inputs, the local production might change, and the tax base might change, both due to worker mobility. Maximizing the welfare function $KV_k + LV_l$ with respect to the tax rate therefore gives a first-order condition as a composite of the closed-city efficiency condition and the sum of mobility effects:

$$0 = \underbrace{\left(1-\beta\right)\left(\frac{-(K+L)}{Q-t\,(K+L)}\right)+\gamma/t}_{\text{closed economy F.O.C.}} \tag{7.6}$$
$$\underbrace{+\left(1-\beta\right)\frac{dQ/dt-t\,(dK/dt+dL/dt)}{Q-t\,(K+L)}}_{\text{mobility effects}}.$$

The mobility effects collect the terms that arise compared to the closed economy due to the mobility of workers, i.e., $dK/dT \neq 0$, $dL/dT \neq 0$. These effects capture that the mobility response of workers has an effect on production (the skill mix changes) and on the government budget (the tax base changes). Realizing that in equilibrium under zero profits, $Q = w_k K + w_l L$, the mobility effects can be rewritten into:

$$(1-\alpha)\left(dK/dT\left(w_k - T\right) + dL/dT\left(w_l - T\right)\right)/\left(Q - T\left(K + L\right)\right).$$

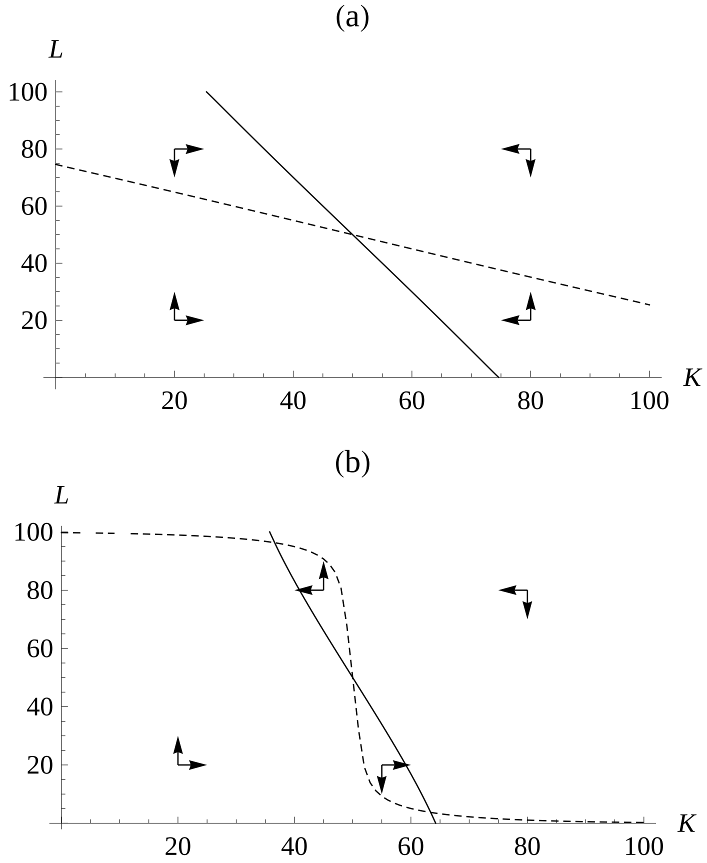
This shows that compared to the closed economy first-order condition, the government calculates the marginal effect on the average after-tax income, where high-skilled labour is favoured because their higher income increases the average income in the economy.

In a symmetric economy, low- and high-skilled workers are evenly distributed over the two cities. As a result, the globally first-best policy, which ignores the mobility effects, is equal to the closed-economy tax rate: it is as if solving a large single region problem (where the externalities associated with migration following tax changes are not relevant). Therefore, if the mobility effect in the first-order condition is positive, competing governments set a tax rate that is higher than globally socially optimal.

Examining the welfare effects of worker migration following tax rate increases reveals that in a symmetric economy, where skill shares and size of the cities are equal, the mobility effects are positive. Keeping all else constant, in equilibrium, the migration that a local tax increase causes, improves local welfare. Therefore, incorporating the consequences of worker migration, governments set higher taxes than they would in a closed economy. The result can be obtained by taking the derivative of the production function (eq. 7.4) and totally differentiating spatial equilibrium conditions to obtain dQ/dT, dK/dT and dL/dT. Together, these can be used to sign the mobility effects. The derivations are delegated to Appendix 7.B, but the mobility response to tax rates is quite intuitive. Due to the policies that they prefer, high-skilled workers are attracted by higher taxes, but low-skilled workers are driven away. As high-skilled workers are scarce, they are more productive than low-skilled workers (and low-skilled workers' wages improve from attracting high-skilled workers), so attracting them provides a positive net effect on aggregate production.

The higher than optimal equilibrium tax rates suggest that the mobility of workers of different skills leads to policy competition in the sense of Wilson (1986): a joint common reduction in the provision of amenities would make cities as a whole better off. Such a policy intervention would leave the distribution of low- and high-skilled workers unchanged, and policy would reflect average preferences in the city. This result does not rest on the mobility of high-skilled workers only, as low-skilled workers are equally mobile. However, the impact that high-skilled workers have

Figure 7.1: Skill sorting across cities



Note: K and L are the share of a mass of 100 high and low-skilled workers in city 1, respectively (the complement lives in city 2). Lines indicate utility equality for high-skilled (solid) and low-skilled (dash) workers.

on the average welfare is larger (not only are they more productive, they also make the low-skilled incumbent more productive), and therefore their mobility is weighed more heavily by the policymaker.

There are, however, distributional consequences from a common lowering of the taxes and amenity provision. As the laissez-faire solution and the first-best solution have the same distribution of workers, the production structure and wages are the same in either situation. However, since high-skilled workers prefer tax rates strictly higher than first-best, reducing the equilibrium rates to welfare-maximizing rates reduces welfare of the high-skilled, but improves it for the low-skilled.

The model, however, also allows for skill-sorting across cities. The conditions for this to occur are discussed in Appendix 7.A. Intuitively, a (partial) sorting equilibrium requires that individuals have strong preferences about the level of amenity-provision; in that case, living in a more like-minded city yields high payoffs in terms of policy-efficiency. Sorting implies that one city emerges with a (relative) majority of high-skilled workers, and one city with a majority of low-skilled workers. Figure 7.1 show the symmetric and sorting equilibria.[4] It plots the spatial equilibrium conditions for low (dashed) and high skilled workers (solid) in a space of distributions of a mass of a hundred low and high skilled workers over the two regions. North of the spatial equilibrium condition for low skilled workers (dash), the utility of living in region 2 is higher than that of living in region 1 for low skilled workers; thus inducing a force to migrate to city 2 (i.e., a downward movement in the graph). For high skilled workers, distributions East of the solid line indicate that the utility of living in region 2 is higher. The upper panel of Figure 7.1 shows a situation where the symmetric situation is stable. The lower panel shows the case where mixing of skills is unstable: small deviations in north-western or south-eastern direction lead into the basin of attraction of one of the asymmetric equilibria.

The intuition for this sorting pattern is different from the symmetric situation. Suppose that the elasticity of substitution between low- and high-skilled workers is nearly perfect. This allows one region to specialize in high-skilled production, requiring only little low-skilled labor. The policies of that high-skilled region can then be tailored to the high-skilled worker. This does not attract low-skilled workers in general, except that low-skilled wages grow very large if very few of them are employed in the high-skilled region. Therefore, some low-skilled workers always forego their preferred policies in the low-skilled region to earn a larger wage in the high-skilled region.

Whether the taxes in specialized cities are optimal depends, as before, on the net externalities that each cities conveys onto the other city. Since cities are now asymmetric, the comparison to a closed economy no longer yields the welfare-maximizing policies. Instead, we check the sign of the policy externalities that a central planner would take into account, but the local policymaker does not. This exercise is carried out in Appendix 7.B,

[4]In this case, the difference was obtained by varying the preference for amenities, γ. The exact conditions for which the two equilibria occur are discussed in Appendix 7.A.

and shows that the policy externality of raising taxes in the high-skilled city in the low-skilled city is negative. Vice versa, the the policy externality in the high-skilled of the less skilled city raising its taxes is positive. The intuition is as follows: a city balances the benefits of specialization in terms of better tailoring of its policies with the costs of specialization in terms of productive inefficiency. However, if the high-skilled city raises its taxes, it draws high-skilled workers from the low-skilled city, forcing the low-skilled city into further specialization. Keeping the other city's tax rates constant, the benefits of specialization (preferred policy homogeneity) are therefore fully internalized, while part of the costs of specialization (the productive specialization of the other city) are not internalized. As a result, compared to a planner, the high-skilled city sets its tax rates too high, while the low-skilled city sets its tax rates too low: both cities favor their local abundant factor more than is efficient, and cities become too specialized.

Amenity provision under production externalities

So far, we have studied the strategic behavior of governments assuming that there are no productive externalities. Possibly, welfare conclusions change if there are production externalities. If there are productive effects that neutralize the mobility effects, then the equilibrium tax rates might be efficient. Therefore, this subsection questions whether there are externalities for which harmonization does not improve welfare in both cities, i.e., whether there could be externalities that justify the equilibrium tax rates.

To deal with a fairly general variety of production externalities, we assume that the aggregate productivity takes the form:

$$A = \left(K^{\psi} + L^{\psi}\right)^{\delta/\psi}. \tag{7.7}$$

This firstly allows us to eliminate all externalities by setting $\delta = 0$. Secondly, it comprises a variety of production externalities found in the literature. If ψ is relatively low, there are returns to diversity; cities with perfectly mixed skill groups will be more productive. The higher ψ is, the higher the returns to specialization into one of the two skills. Moreover, when setting ψ equal to ρ, the elasticity of substitution in the production function, the productivity is simply an aggregate scale effect in production, so that aggregate production is raised to the power $1 + \delta$ (but since

A is taken as given by producers, they act as if they are facing constant returns to scale).

The first-order condition of welfare $KV_k + LV_l$ with respect to the local tax rate for a government that takes into account the effects of tax on production externalities is modified to:

$$\underbrace{(1-\alpha)\left(\frac{-(K+L)}{Q-T(K+L)}\right)+\gamma/T}_{\text{closed economy F.O.C.}}+\underbrace{(1-\alpha)\frac{\frac{dQ}{dT}-T\left(\frac{dK}{dT}+\frac{dL}{dT}\right)}{Q-T(K+L)}}_{\text{mobility effects}} \tag{7.8}$$
$$+\underbrace{(1-\alpha)\frac{Q/A\left(\frac{dA}{dK}\frac{dK}{dT}+\frac{dA}{dL}\frac{dL}{dT}\right)}{Q-T(K+L)}}_{\text{production externality}}=0.$$

This simply extends the earlier first-order condition (eq. 7.6) with the production externality. Since the positive mobility effects lead to overtaxation, a negative effect of the tax rate on welfare via the production externality could neutralize the inefficiency. Note that the production externality is effectively also a mobility effect, because in a closed economy, it would not be part of the government's consideration: production externalities require migration to play any role in the government's consideration. If the production externality can justify the equilibrium tax rates, it needs to neutralize the positive mobility effects, and so it must be negative. To study whether this can occur, we write the marginal effects of high- and low-skilled labor on aggregate productivity as:

$$\frac{dA}{dK}=\delta\frac{A}{K^{\psi}+L^{\psi}}K^{\psi-1};\ \frac{dA}{dL}=\delta\frac{A}{K^{\psi}+L^{\psi}}L^{\psi-1}. \tag{7.9}$$

If $\psi=\rho$, in which case the externality simply operates on aggregate production, the contribution of each factor to aggregate productivity is equal to fraction δ times the wage, which simply reinforces the mobility effect (i.e., the mobility effect is multiplied by $1+\delta$). Such an externality on the scale of production will therefore not be consistent with cities setting efficient tax rates in equilibrium. A more plausible candidate is then that there are returns to diversity. The most extreme case can be modelled by having ψ approach 0. Taking the marginal contributions to productivity of both factors, and inserting them in the production externality gives that the sign of the production externality takes the sign of:

$$\frac{dA}{dK}\frac{dK}{dT}+\frac{dA}{dL}\frac{dL}{dT}=\delta Q\left(\frac{dK}{dT}/K+\frac{dL}{dT}/L\right). \tag{7.10}$$

Among symmetric cities, this factor is unlikely to be negative: high-skilled workers respond positively to tax rate changes ($(dK/dT)/K > 0$), while low-skilled workers respond by migrating away. As discussed in Appendix 7.B, the scarcer factor is more responsive to tax changes. For L to be scarce does not appear realistic, especially if high-skilled workers could do low-skilled tasks. This conforms to intuition: if diversity encourages productivity, then having fewer of the (abundant) low-skilled workers improves productive efficiency. For asymmetric cities, however, the results depend on the relative scarcity of high-skilled workers. If the concentration of high-skilled workers in the the high-skilled city is larger than a perfect mix (50%), lowering the taxes increases diversity through inward migration of low-skilled and outward migration of high-skilled. In this case, the production externalities of raising the tax are negative, and it is possible that equilibrium taxes are efficient or even too low. Clearly, in the low-skilled city, with an absolute abundance of low-skilled workers, a diversity externality increases tax rates toward the optimum. If, however, high-skilled workers are relatively scarce and do not form a majority in the most highly skilled city, the externality increases taxes both in the low and high-skilled city, so there is no justification for the equilibrium tax rates.

The last extreme is the situation where production externalities depend on the specialization of the workforce, i.e., the case where ψ is very high. Assuming that low-skilled workers are abundant, symmetric cities would lower taxes to stimulate low-skilled immigration and propagate high-skilled emigration. A growth strategy based on attracting low-skilled workers runs counter to all empirical evidence, and would no longer hold already in case high-skilled workers were able to perform low-skilled tasks. Among specialized cities, again, the effect of the specialization externality depends on the relative abundance of high-skilled workers. If high-skilled workers form a majority in the high-skilled city, this encourages higher taxes in the high-skilled city and increases city specialization, rather than decrease it, which cannot lead equilibrium taxes to be optimal.

Table 7.1 summarizes the results regarding tax-setting for different types of externalities and different urban constellations. Symmetric cities set too high tax rates unless there are strong returns to specialization, which was argued to be unrealistic in this setting. In a specialized equilibrium, low-skilled cities will generally set too low taxes unless there are returns to diversity, in which case the outcome is uncertain. Returns to diversity thus potentially provides a force that corrects the policy bias toward high-skilled workers. The high-skilled city generally sets its taxes higher than optimal. However, if externalities are not neutral with respect to the ratio of wages (i.e., $\psi \neq \rho$, total returns favor diversity differently than private returns do), there is room for doubt. If high-skilled workers form a majority in the high-skilled city, then an externality based on diversity favors attracting low-skilled workers, so taxes are reduced, possibly justifying the equilibrium tax rates. However, if high-skilled workers do not form a majority, diversity-based externalities still promote higher taxes to attract high-skilled workers.

Table 7.1: Optimality of equilibrium taxes under production externalities

externality	symmetry	specialization	
		low-skilled	high-skilled
none	$T > T^*$	$T_l^n < T_l^*$	$T_h^n > T_h^*$
diversity	$T > T^*$	$> T_l^n, \lesseqgtr T^*$	$> T_k^n, > T^*$ $< T_k^n, \lesseqgtr T^*$ if K majority
scale	$T > T^*$	$T_l^n < T_l^*$	$T_h^n > T_h^*$
specialization	$\lesseqgtr T^*$	$< T_l^n, < T^*$	$< T_k^n, \lesseqgtr T^*$ $> T_k^n, > T^*$ if K majority

Note: T^*: optimal tax, T^n: equilibrium tax under no externalities
K is assumed not to be the majority, unless stated otherwise

Summing up, production externalities do not generally justify the equilibrium tax rates that cities set. An exception could be strong returns to the diversity of skills, but only under the condition that high-skilled workers have an absolute majority in the high-skilled city.

7.4 Trade and policy-driven specialisation

The previous section argued that skill heterogeneity can be a source of policy inefficiencies, using the bare minimum of requirements for skill difference to play a spatial role: two regions and two production factors. However, contemplating a more reasonable setting with more cities and skill types, it seems less realistic that every city hosts every skill. A sensible, if not central issue in specialization is the possibility to trade goods. In particular, if there were two sectors of production, we would see sectoral specialization per city, and given the possibility of factor price equalization, it is no longer clear-cut which of the results so far survives. If final goods can be traded, it is possible for workers of a single skill to cluster in a city to enjoy optimal policies, while trading their specialized output against a more general consumption basket from other cities.

An extension into many skills and cities hence suggests a role for trade in the specialization of cities, in contrast to the two-city-two-skill case. Additionally, it shows that (part of) our results translate into a more realistic world where skill and productivitity can be viewed as continuous. Apart from the generalization, this addition also generates conceptual differences. In a many-region world, policy-makers are small compared to the rest of the world, and they see the externalities of their policies wash out in the global economy. In particular, in the two-region world studied above, the policymaker can anticipate the effects of his policy (via migration) on the productivity and wages in other regions; but in a many-region world, the policymaker can be assumed to take prices in other regions as given. Furthermore, inserting the trade model from this section in a two-city/two-skill economy, trade would yield specialization as a discrete process: the cities are either both specialized or exactly symmetric. The model in this section provides additional understanding as to which cities specialize and in which skill they specialize; or, alternatively, which groups of workers prefer to specialize and which prefer to mix. A corollary is that the preference to co-locate with particular skill groups can now also be attributed to the demand side: vicinity to high-skilled workers makes for better consumption options.

The extension of the model builds on a static version of the Acemoglu and Ventura (2002) model of trade in intermediates in an Armington world (where every country produces a single differentiated good), but allows for worker mobility and endogenous specialization.

Utility is the same Cobb-Douglas function over housing and consumption as in section 7.3 (eq. 7.1), subject to the same budget constraint. Instead of a numeraire good, however, consumers now have a preference over a range of different goods. These goods are produced by workers with different skills, and they are imperfect subsitutes, captured by an elasticity of substitution σ in the following consumption index:

$$C = \left[\int c(z)^{(\sigma-1)/\sigma} \, dz \right]^{\sigma/(\sigma-1)}, \tag{7.11}$$

where the term z is an index for the different skills that workers may possess, associated with different productivity levels, $a(z)$. If the elasticity of substitution is high (σ is high), goods produced by workers of different skills are similar, and a more productive worker easily outcompetes workers of lesser skills. If the elasticity of substitution is low (σ is low), consumers do not easily substitute between goods produced by workers of different skill, and so low-skilled workers face less fierce competition from high-skilled workers (compared to when the elasticity of substitution is high).

We shall assume that workers are spread over a large number of cities, k. Access to products inside the city is free, but it is costly to import goods from outside the city. We shall make the assumption that the costs of obtaining goods from other cities take the form of an iceberg transport cost, τ, that is constant across all city pairs. This ignores the exact geography (it is already hard to think of city constellations of four equidistant cities that have this property) but incorporates the idea of distance frictions well enough to see the argument. Using the transport costs in the perfect price index gives $P = \left[\int \tau(z) w(z)^{1-\sigma} / a(z) \, dz \right]^{1/(1-\sigma)}$, where $\tau(z)$ is 1 if the products of skill group z are locally produced, and τ if imported. We assume that there is a fixed number of cities of equal size in terms of land supply, and that the land market clears in the same manner as in the last section. The only difference is that there is now a large distribution of worker skills, and each skill type potentially earns a different wage. Taking this into account, land market clearing satisfies: $r = (1-\alpha) \int l(z)(w(z) - T) \, dz / H$, where $l(z)$ is the local employment of workers of type z.

Workers supply their labour in a competitive market, and are therefore paid their (real) marginal product. The productivity varies, however, per skill: the aggregate production per skill type depends on z as $q(z) = l(z) a(z)$. In a competitive market, the price is $p(z) = w(z) / a(z)$.

To study the effects of specialization on policy, we shall examine two polar cases: one where all cities are symmetric, and one where all cities are specialized. In the original Acemoglu and Ventura (2002) model, the

Armington assumption (goods are differentiated by location of production) ensures a specialized world. In this modification, however, goods are differentiated by skill, and workers are mobile, so specialization is endogenous.

In the symmetric case of unspecialized cities, solving the demand function, and equating city-level demand to supply gives the goods market clearing condition:

$$l(z)\,a(z) = \frac{(w(z)/a(z))^{-\sigma}}{P^{1-\sigma}}\alpha Y_k, \tag{7.12}$$

where Y_k is the city's aggregate after-tax income. Rewriting the market clearing condition for $w(z)/a(z)$ and inserting that into the price index gives the aggregate resource constaint that $\left[\int (l(z)\,a(z))^{(\sigma-1)/\sigma}\,dz\right]^{\sigma/(\sigma-1)} = Y_c/P$. Since output and consumption are related by an undetermined price level, we can set the harmonized price level as the numeraire, as Acemoglu and Ventura do.

The market-clearing wages and the price level define the welfare in the system of cities when inserted into the aggregated indirect utility function $W = \int l(z)(w(z) - T)\,T^{\gamma} r^{-\alpha} dz$. The land rent is proportional to income, and from the clearing condition (eq. 7.12), the equilibrium wage is $w(z) = l(z)^{-1/\sigma}\,a(z)^{(\sigma-1)/\sigma}\left(\alpha Y_k\right)^{1/\sigma}$. Using these in the welfare function gives that:

$$W = \chi\left(\left[\int (l(z)\,a(z))^{(\sigma-1)/\sigma}\,dz\right]^{\sigma/(\sigma-1)} - \int l(z)\,T\,dz\right)^{\alpha} T^{\gamma}, \tag{7.13}$$

where χ is a positive constant. In words, this term simply states that the city's welfare is equal to aggregate production less taxes, multiplied by a term that captures returns from public goods, T^{γ}. The land market ensures that cities are always inhabited; if not, land would be free and attract citizens ($0 < \alpha < 1$, so utility tends to infinity if the city becomes deserted).

If migration is allowed, $l(z)$ is endogenous, and we need to specify the individual utility levels to determine where workers want to live. Rewriting the market clearing condition under a numeraire perfect price index gives that

$$w(z) = a(z)^{(\sigma-1)/\sigma}\,l(z)^{-1/\sigma}\,Y_k^{1/\sigma}. \tag{7.14}$$

This shows that the wage depends positively on the own productivity and the city's income, but the supply of labor of a particular skill bids down the wage for worker of that skill. The elasticity of wages with respect to labor supply is $-1/\sigma$. Note that if skills are easily substitutable ($\sigma \to \infty$), the limit of the worker's wage is $a(z)$, so only the productivity matters to the worker, as all workers produce the same good. Conversely, low substitutability ($\sigma \to 1$, the exponent on $a(z)$ tends to zero, the exponent on $l(z)$ to -1) stresses that skill scarcity, rather than productivity leads to high wages. Using the individual expression for wages, the utility level of a worker of skill type z is given by:

$$V(z) = \chi \left(a(z)^{(\sigma-1)/\sigma} \, l(z)^{-1/\sigma} \, Y_c^{1/\sigma} - T \right)^{\alpha} T^{\gamma}. \tag{7.15}$$

If we assume both that the marginal worker is small compared to the city population, and the city is small relative to all other cities, it can be seen that abundance of the worker's own skill reduces his utility of living in the city: a larger presence of his own skill group $l(z)$ reduces utility. This is akin to the terms of trade effect in Acemoglu and Ventura, which ensures a stable world distribution of income. In the Acemoglu and Ventura model, imperfect substitutability with other countries' products ensures that there is demand for a country's good; in this model, it is the scarcity of skills that ensures a stable skill distribution within cities. Therefore, it exerts a "mixing" force: it spreads workers of the same skill over different cities. We assume that workers can migrate to obtain a reservation utility $\bar{V}(z)$ elsewhere in the group of cities (i.e., in equilibrium, through migration, a common utility applies for workers of a particular skill type). This leads to a free-migration condition as workers of type z will enter any city that yields the highest utility: $\bar{V}(z) = V(z) \, \forall z$. Totally differentiating the free-migration conditions (so that in eq. 7.15, V must be constant over cities) gives a static migration response to tax policy changes:

$$\frac{dl(z)}{dT} / l(z) = \sigma \left(\frac{\gamma}{T} \left(1 - \frac{T}{w(z)} \right) - 1 \right). \tag{7.16}$$

The left hand side of this term is a quasi-elasticity of labor supply (through migration) with respect to tax rates per skill level z. A lower wage reduces the elasticity of z-type population with respect to taxes, so low-wage worker grow less fast in number when taxes increase than high-skilled workers do, or might even reduce in number. This matches the results from Section 7.3: high levels of government services attract high income workers and chase away low-wage workers. This effect is magnified by the elasticity of substitution between skills σ: if it is high, high-skilled workers easily take over jobs from low-skilled workers, so low-skilled workers cannot exploit the uniqueness of their skills to retain higher wages. Note that the wage can be high because of scarcity or productivity; so relatively unproductive workers might be attracted to high tax rates, as long as their skill is not abundantly supplied.

This free-migration condition, the competitive equilibrium, and the welfare function allow studying policy among symmetric cities. In an open city, the first order condition for a policymaker maximizing welfare (eq. 7.13) by varying tax rates is given by:

$$\frac{dW/dT}{W} = \gamma/T + \alpha\frac{Q\Phi - \int l(z)\,T\left(\frac{dl(z)/dT}{l(z)} + 1/T\right)dz}{Q - \int l(z)\,T\,dz}, \tag{7.17}$$

$$\text{with } \Phi \equiv \frac{\int (l(z)\,a(z))^{(\sigma-1)/\sigma}\frac{dl(z)/dT}{l(z)}dz}{\int (l(z)\,a(z))^{(\sigma-1)/\sigma}\,dz}.$$

Like in the two-city case, the direct utility returns to providing amenities (γ/T) are balanced with the cost of providing them, and there are two mobility effects: changes in city productivity (captured by $Q\Phi$) and budget effects (captured in $\int l(z)\,T\frac{dl(z)/dT}{l(z)}dz$). The term Φ is the average relative change in labor supply through migration across skill groups, weighted by the productivity of skill groups. Using the wage rate (e. 7.14), weighing the elasticity of supply with $(l(z)\,a(z))^{(\sigma-1)/\sigma}$ in Φ is the same as weighing it with $l(z)\,w(z)$, which simply reflects the skill group's income (which is equal to their marginal product).The policymaker, like in the two-region case, assigns heavier weights to more productive skill groups. In a symmetric equilibrium, the optimal policy can be found by considering the closed-economy policy. For the closed-economy policymakers, the optimal policy is the tax that maximizes the welfare function (eq. 7.13) under the restriction that there is no migration ($dl(z)/dT = 0$). The first-order condition is:

$$\frac{dW/dT}{W} = \gamma/T - \alpha\frac{\int l(z)\,T\,(1/T)\,dz}{Q - \int l(z)\,T\,dz}.$$

Comparing this to the first-order condition of a policymaker in equilibrium (eq. 7.17), the open city policymakers sets higher tax rates than optimal if $Q\Phi - \int l(z)\,T\frac{dl(z)/dT}{l(z)}dz > 0$. The term $Q - \int l(z)\,T\,dz$ is simply the aggregate after-tax income, which is positive. The term $dl(z)/dT/l(z)$, which appears both in $Q\Phi$ (Φ is weighted average of it) and in $\int l(z)\, T \frac{dl(z)/dT}{l(z)} dz$, is higher for z associated with higher wages. The term Φ weighs the mobility response of high-wage skill groups more heavily, so since $Q > \int l(z)\, T dz$, it also holds that $Q\Phi > \int l(z)\, T \frac{dl(z)/dT}{l(z)} dz$: open city policymakers set higher tax rates than the optimal tax rate. In an equilibrium of symmetric cities, the prospect of attracting high-skilled workers with higher taxes raises equilibrium taxes above the optimal taxes. Since the elasticity of substitution between skills magnifies the migration responses, high substitution exacerbates the inefficiency. This is intuitive, because it implies that the policymaker's bias is purely directed toward productivity; diversity is not worth pursuing if low-skilled workers are easily substituted for. Moreover, the migration effect relies on the workers' wages, not (only) on his skill type. In fact, as can be seen from the expression for wages, $w(z) = l(z)^{-1/\sigma} a(z)^{(\sigma-1)/\sigma} \left(\alpha Y_k\right)^{1/\sigma}$, even if they are less productive, workers with scarce skills may choose to live in high-income cities (and are attracted to high-skilled taxes), as long as their relative scarcity compensates their low productivity. Therefore, the policymaker does not care about high-skilled workers per se, but about skills compensated for the supply that skill.

A second spatial configuration is the outcome in which cities specialize in one factor. This is effectively the case studied in the international trade literature by assuming that each country produces its own vintage good or input. In the present case, there is good reason to specialize: clustering workers of the same type allows for more efficient policies. There is, however, a cost due to the trade model incorporated in this section: the goods produced by workers of different skills need to be imported. In particular, following common assumptions (Acemoglu and Ventura, 2002), we shall assume that there are many skill levels and even more cities, so that effectively, cities are allowed to fully specialize and export their product while importing all other products.

If the economy is specialized, clearing on the goods market implies that the quantity of goods sold to other regions, $l(z)\, a(z)/\tau$, is equal to the demand from other regions:$(\tau w(z)/a(z))^{-\sigma}/\tau^{1-\sigma}\alpha \int Y_k dk$ (where $\alpha \int Y_k dk$ is the aggregate expenditure on consumption goods over all cities k). If this clearing condition holds for all cities, the value of imports equals that of exports, and so trade is balanced. The unique wage rate that satisfies goods market clearing is:

$$w(z) = l(z)^{-1/\sigma} a(z)^{(\sigma-1)/\sigma} \alpha \left[\int Y(k)\, dk \right]^{1/\sigma}.$$

Using this, the utility for a representative inhabitant, and hence a city's welfare under specialization is given by:

$$V(z) = V(c) = \left(l(z) \frac{w(z) - T(c)}{\tau} \right)^{\alpha} T(c)^{\gamma}. \tag{7.18}$$

Since the wage rate is constant across inhabitants, setting an optimal tax rate is easier for the policymaker; the first-order condition of the above utility function with respect to the tax rate satisfies:

$$\alpha \left(\frac{dl(z)/dT}{l(z)} + \frac{dw(z)/dT - 1}{w(z) - T} \right) + \gamma/T = 0.$$

The concave welfare function in regional size ensures that if cities are of the same skill, they are also of the same size. Additionally, if all cities set locally optimal policies, then the willingness to pay of a worker of type z for land in a city specialized in his skill is higher than for any other city. Thus, in equilibrium, the marginal effect of tax changes on wages and city size is zero: $dl(z)/dT = 0$, $dw(z)/T = 0$. Using these observations, the equilibrium tax rate in a city k specialized in z simplifies to:

$$T(k(z)) = \frac{\gamma}{\alpha + \gamma} w(z),$$

which is intuitive: under unit elastic preferences for the government-provided amenity, ideally, workers prefer to spend a fixed share of income on public goods. Thus, we get that under a fully specialized system of cities, policymakers behave optimally; in equilibrium, marginal tax rate changes do not lead to migration that distorts tax-setting.

It is left to compare welfare under specialized versus diversified cities. Since the number of workers of each skill is equal, whether they allocate in diversified or specialized cities, we can compare aggregate welfare. Define:

$$\Upsilon(z) \equiv \frac{w(z) - T}{w(z)} T^{\gamma/\alpha} \tag{7.19}$$

as the effective ratio between a worker's wage and utility. The more the tax rate faced in the home city diverges from the preferred tax rate, both upward and downward, the lower Υ. Therefore, $\Upsilon(z)$ is a measure of how efficient the policy in the city of residence is for a specific worker relative to his preferred policy. Using this definition, we can compare welfare levels across the symmetric and specialized cities.[5] Subtracting welfare in specialized cities (eq. 7.18 aggregated over cities) from welfare in diversified cities (eq. 7.13) gives:

$$\int \left(\bar{\Upsilon}(z)\, a(z)^{(\sigma-1)/\sigma}\, l(z)^{-1/\sigma}\, Y_k^{1/\sigma}\right) dz \qquad (7.20)$$
$$-\int (1/\tau)\left(\Upsilon^*(z)\, a(z)^{(\sigma-1)/\sigma}\, l(z)^{-1/\sigma}\left[\int Y_k dk\right]^{1/\sigma}\right) dz,$$

in which $\bar{\Upsilon}(z)$ is consistent with the tax burden for workers in the diversified equilibrium, and $\Upsilon^*(z)$ is consistent with the worker's preferred tax. Any tax inconsistent with $\Upsilon^*(z)$ reduces welfare from the worker's perspective. Using this definition for the efficiency of taxes, rewriting equation 7.20 (where Y_k is constant across symmetric cities) shows that welfare among diversified cities is higher if:

$$\tau > \frac{\int \Upsilon^*(z)\, a(z)^{(\sigma-1)/\sigma}\, l(z)^{-1/\sigma}\, dz}{\bar{\Upsilon}(z)\int a(z)^{(\sigma-1)/\sigma}\, l(z)^{-1/\sigma}\, dz}. \qquad (7.21)$$

The left-hand side of this inequality form the iceberg transport costs, the right-hand side is a measure of policy inefficiency due to population heterogeneity. There may be one type of worker in the symmetric cities who prefers the equilibrium tax rate ($\bar{\Upsilon}(z) = \Upsilon^*(z)$), but for all others, equilibrium welfare must be lower than if their preferred policy was chosen. Therefore, the right-hand term is larger than 1 and increasing in worker heterogeneity. As a result, in equilibrium, specialized cities are preferred over diversified cities if there is large heterogeneity in productivity levels or in the scarcity of skills, or if trade costs are sufficiently low. In fact, if there were no trade costs, cities would always specialize. Secondly, the results in this section show that the equilibrium tax rate is higher than socially optimal if cities are diversified but not if cities are specialized. Therefore, there must exist a range of transport costs for which the aggregate welfare of a system of fully specialized cities is between that of

[5]Rewriting using this definition represents a homothetic transformation that preserves citizens' preferred tax rates. However, the transformation is not linear, so this only allows examining the direction of the effects on welfare, not the magnitude.

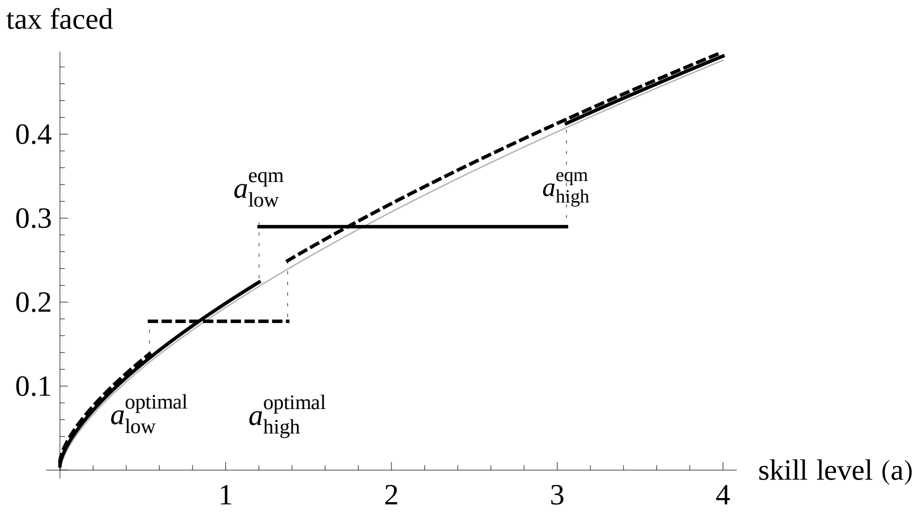
a system of diversified cities with competitive taxes, and that of diversified cities with optimal (harmonized) taxes. If mixed-skill cities would set optimal taxes, its citizens would balance the costs of policy heterogeneity (but less costly imports) in the mixed-skill city with the costs of shipping (but perfect policy) in the specialized city. However, because prospective utility is lower in the mixed-skill city in equilibium due to distorted taxes, citizens will choose to locate in specialized cities sooner than is optimal.

Without specifying further distributional properties of skills, systems of cities that are partially specialized are difficult to study, and the non-linearity of optimal policy in mixed-skill cities complicates closed-form solutions. Figure 7.2 provides a numerical example, however, based on unit density over the skill group ($l(z) = 1 \forall z$), and assuming that productivity is uniformly distributed. Partial specialization occurs if some workers prefer to live in cities of mixed skills, while others choose to live in specialized cities. Since mixed skill cities' tax rates are rarely extreme, it is likely that workers in the middle of the skill distribution find the tax rate of mixed-skill cities acceptable, while very low- or high-skilled workers prefer specialized cities: to them, a tailored tax rate of a specialized city is more likely to weigh against incurring transport costs.[6] Given this reasoning, the figure that emerges from Figure 7.2 is intuitive; only average-skilled workers between a_{low}^{eqm} and a_{high}^{eqm} remain in mixed-skill cities in equilibrium. Therefore, they are in symmetric cities, and their policymakers set the same tax rates. Workers with either very high or very low skills choose to live in cities that specialize in their skill, where everybody has the same preference for tax rates: policy aligns perfectly with their preferred policy (the thin gray line in the Figure).

A comparison between the optimal and equilibrium outcomes confirms the analytical results. The first-best tax pattern is drawn as the dashed profile in Figure 7.2. Unlike in the equilibrium case, optimal policy in mixed-skill cities does not take into account the mobility (migration) effects, and therefore optimal taxes are lower than equilibrium taxes. Since the policy bias is towards higher skilled labor, mixed-skill cities are more skilled than is optimal: in equilibrium, high-skilled workers that would optimally specialize, live in a mixed-skill city; low-skilled workers that would optimally live in a mixed-skill city choose to live in specialized cities. This can be

[6] Indeed, we get a strictly concave parabola of preference to live in mixed cities along the skill distribution. However, this is not to say that this equilibrium is necessarily the only one.

Figure 7.2: City specialization and equilibrium tax rates



Note: Taxes faced by workers of different skill types. Workers with skill levels between a_{low} and a_{high} live in mixed-skill cities. Lines indicate the worker's preferred tax (gray), the equilibrium tax faced (solid), and the optimal tax in the urban system (dash). Parameters: $\varepsilon = 3$; $\gamma = 0.5$; $\beta = 0.5$; $\tau = 1.1$.

seen graphically, as the lowest skill type that prefers to live in a mixed-skill city is lower under optimal policy ($a_{low}^{optimal}$) than in the equilibrium (a_{low}^{eqm}). The same is true for the highest skill type ($a_{high}^{optimal} < a_{high}^{equilibrium}$). Also, checking for the optimal policy in equilibrium mixed skill cities (i.e., cities with skills ranging from a_{low}^{eqm} to a_{high}^{eqm}) shows that welfare in those cities could be improved by commonly lowering the tax rate when keeping the population fixed.

Up to this point, the number of cities was assumed fixed. An interesting extension to this model could be to endogenize the number of cities. This would require additional assumptions, however. In particular, due to the unit-elastic preference for houses, newly constructed (empty) cities have zero land rents, and therefore yield infinite utility of living (in the limit). Without "lumpiness" in city development, one would therefore expect a large number of one-man cities. This could be remedied by having land developers found new cities with a minimum size threshold. That, however, would need modifications beyond the setup of this chapter.

A final important point worth noting with respect to the two-city case is that tradability of goods has eliminated an Inada-type condition on diversity: the wage of workers of a scarce skill type does not tend to infinity, because their product can be imported. Therefore, the policy externalities in terms of specialization in other regions have vanished, which changes the conclusions about the efficiency of specialization. A second point is that positive trade costs can provide the policymaker with another plausible objective. If goods are freely traded, as in section 7.3, the average productivity of a worker is a policy objective. The inclusion of trade costs provides a natural motivation: the presence of producers of the most desired products avoid transport costs to get them, thus benefitting low- and high-skilled workers.

7.5 Conclusion

Urban amenities, among which culture, museums, green areas, schooling, and transport infrastructure can be centerpieces of cities' growth strategies, and urban policymakers are aware of those possibilities. In particular, if urban facilities attract high-skilled workers, the benefits might "trickle down" to lesser skilled workers through increased productivity and local externalities. Even if evidence suggests that amenities are effective instruments for urban revitalization and growth, assessing their desirability requires a broader geographical scope: a plausible source of growth in one city is the departure of firms and productive workers from other cities. This chapter studies urban strategies based on amenities that attract high-skilled workers, given the interactions with other cities.

In general, the results imply that pursuing amenity-based strategies is rational, but not advisable from a broader perspective. The desirability of hosting high-skilled workers causes a policy bias towards them, even if workers of different skill have equal opportunities to relocate to their preferred locations. This always occurs in cities of mixed skills. If workers sort according to skill, cities will bias their policies toward the skill type that is locally relatively abundant, because the benefits (efficient policies) are fully internalized, but the cost of specialization are not: specialization in one city induces specialization elsewhere. As a result, if cities specialize, they become overly specialized. These conclusions are generally robust to different types of productive externalities that workers might confer upon eachother. Trade of consumption goods allows cities to better specialize, eliminating some inefficiencies, but mixed skill cities remain inefficient; consequently, the patterns of specialization are not optimal.

The results cast doubt on the desirability of using cultural public goods to attract human capital and create smart cities. The argument that cultural spending trickles down throughout the city is paired with welfare improvements for the highly skilled, but the argument leads to larger amenity provisions than the lesser skilled prefer, thus leading to lower welfare to the lesser skilled and lower welfare in the aggregate. Pareto improvements seem hard to achieve; government budgets on culture could rather be spent on allowing cities to specialize, which requires integrating them. This, however, is paired with stronger segregation, not all effects of which are taken on board in this chapter. Also, given that many (especially cultural) amenities are significantly subsidized, we need to stress that this model only focuses on the allocation of skills across cities, and ignores the many other external effects of such goods (Throsby, 1994).

An unstudied, but potentially important issue that the chapter ignores, is that workers are born with a certain skill type. Possibly, skill is endogenous, in the sense that low-skilled workers might develop skills if that pays off sufficiently. Taking these into account is beyond the scope of this chapter, but could provide richer insight into the efficiency of education as a growth strategy.

7.A Stability of diversified cities

The wages of low and high skilled workers are determined by their respective marginal productivity:

$$\begin{aligned}\frac{dQ}{dK} &= A(K^{\rho}+L^{\rho})^{1/\rho-1}K^{\rho-1} \\ &= Q\frac{K^{\rho-1}}{(K^{\rho}+L^{\rho})}, \\ \frac{dQ}{dL} &= Q\frac{L^{\rho-1}}{(K^{\rho}+L^{\rho})}.\end{aligned} \tag{7.A.1}$$

The average wage adds up to $\bar{w}=Q/(K+L)$. If the government sets a tax rate equal to $T=\xi\bar{w}$, where $\xi=\gamma/\big(\gamma+1-\alpha\big)$, then the indirect utility functions can be simplified to:

$$\begin{aligned}V_k &= \frac{Q^{1-\alpha+\gamma}}{(K+L)^{\gamma}}\left(\frac{K^{\rho-1}}{(K^{\rho}+L^{\rho})}-\frac{\xi}{K+L}\right), \\ V_l &= \frac{Q^{1-\alpha+\gamma}}{(K+L)^{\gamma}}\left(\frac{L^{\rho-1}}{(K^{\rho}+L^{\rho})}-\frac{\xi}{K+L}\right).\end{aligned} \tag{7.A.2}$$

If we assume that utility difference lead to migration flows towards the city with higher utility we can specificy the migration dynamics as: $\dot{K}=V_k^1-V_k^2=\omega_k$, $\dot{L}=V_l^1-V_l^2=\omega_l$, where city is denoted in the superscript and ω is the utility difference. Linearizing this system gives:

$$\begin{bmatrix}\dot{K}\\ \dot{L}\end{bmatrix}=\begin{bmatrix}d\omega_k/dK & d\omega_k/dL\\ d\omega_l/dK & d\omega_l/dL\end{bmatrix}\begin{bmatrix}K\\ L\end{bmatrix}. \tag{7.A.3}$$

This system is Lyapunov stable if the eigenvalues of the Jacobian are negative. The characteristic equation is given by:

$$\begin{aligned}\lambda^2-\big(d\omega_k/dK+d\omega_l/dL\big)\lambda+d\omega_k/dK*d\omega_l/dL & \\ -d\omega_k/dL*d\omega_l/dK &= 0.\end{aligned} \tag{7.A.4}$$

Taking a Taylor expansion around symmetry ($L_1=L_2=1; K_1=K_2=1$) and solving for the roots gives:

$$\lambda_1 = \frac{2^{1/\rho-1-\gamma}}{\left(1-\alpha+\gamma\right)^{1+\alpha-\gamma}}\left(2^{1-1/\rho}\left(1-\alpha\right)\right)^{\gamma-\alpha} \tag{7.A.5}$$
$$\times\left(\left(1-\alpha\right)\left(\alpha-4\gamma\right)+\gamma\left(1-\gamma\right)\right)\lesseqgtr 0,$$
$$\lambda_2 = \frac{2^{1/\rho-\gamma}}{\left(1-\alpha+\gamma\right)^{1-\alpha+\gamma}}\left(2^{1-1/\rho}\left(1-\alpha\right)\right)^{-\gamma+\alpha}\left(\rho-1\right)\leq 0$$

The second root is always negative, because $\rho < 1$ (that is, low- and high-skilled workers are imperfect substitutes). The first root takes the sign of the term $(1-\alpha)\left(\alpha-4\gamma\right)+\gamma\left(1-\gamma\right)$. Solving this leads to a quadratic equation that shows the term is positive if:

$$\alpha < (1/2)\left(1+4\gamma-\sqrt{1-4\gamma+12\gamma^2}\right).$$

The second root ($(1/2)\left(1+4\gamma+\sqrt{1-4\gamma+12\gamma^2}\right)$) is larger than one for $\gamma > 0$. Intuitively, it can be well explained that α must be large relative to γ for the symmetric equilibrium to be stable. If the preference for public goods is high, workers prefer to sort in those areas where similar workers locate, because the policies are close to their preferred policies. By contrast, starting from symmetry, if the preference for land is high, the increasing land prices following migration will offset any returns from better policies. The second root shows that the symmetric equilibrium becomes unstable if $\rho > 0$. The interpretation behind that is that this implies increasing returns to specialization in one factor, in which case wages rise if factors specialize.

7.B The net mobility effect

As can be seen from eq. (7.6), the mobility effects on the tax rate are positive if $dQ/dT - T\,(dK/dT + dL/dT)$ is positive. Using that $dQ/dK = w_k$, this can be rewritten into $dK/dT\left(w_k - T\right) + dL/dT\left(w_l - T\right)$. To obtain the effect of tax rates changes on the size of the worker pool, we use that in equilibrium, $V_1 - V_2 = \omega = 0$. Totally differentiating gives that $dK/dT = -(\partial\omega/\partial T)/(\partial\omega/\partial K)$.

Taking these implicit derivatives, multiplying by the after-tax wage $(w-T)$ and simplifying yields:

$$\begin{aligned}
\frac{dK}{dT}\left(w_k-T\right) &= \frac{1-\alpha\frac{w_k-T}{\bar{w}-T}+\gamma\frac{w_k-T}{T}}{2\left(\frac{dw_k/dK}{w_k}\frac{w_k}{w_k-T}-\frac{\alpha}{L+k}\frac{w_k-T}{\bar{w}-T}\right)}; \qquad (7.B.1)\\
\frac{dL}{dT}\left(w_l-T\right) &= \frac{1-\alpha\frac{w_l-T}{\bar{w}-T}+\gamma\frac{w_l-T}{T}}{2\left(\frac{dw_l/dK}{w_l}\frac{w_l}{w_l-T}-\frac{\alpha}{L+k}\frac{w_l-T}{\bar{w}-T}\right)}.
\end{aligned}$$

Moreover, we use that:

$$\begin{aligned}
\frac{dw_k/dK}{w_k} &= \left(1-\rho\right)\left(Q^{-\rho}-1/K^{2-\rho}\right); \qquad (7.B.2)\\
\frac{dw_l/dL}{w_l} &= \left(1-\rho\right)\left(Q^{-\rho}-1/L^{2-\rho}\right).
\end{aligned}$$

If high-skilled labor is more scarce than low-skilled labor, the percentage change in the wage rate due tax changes is smaller for high-skilled than for low-skilled wages. As a result, the numerator of $dK/dT\left(w_k-T\right)$ is larger than that of $dL/dT\left(w_l-T\right)$, and its denominator is smaller. The intuition for the numerator is that K is scarce, so tax policy puts more weight on low-skilled workers. The utility function is strictly concave in the tax rate, as can be seen from the expression for preferred tax rates, and as the policy is further from the high-skilled preferred tax rate than from the low skilled preferred tax rate, tax rate changes have a greater effect on high-skilled utility. The effect of the denominator consists of a wealth effect (wage changes), and derived from that, a land rent effect. Land rents are proportional to average after-tax income, and therefore changes in the income of one factor are diluted in the average income when the other factor's wage is unchanged. Jointly, these imply that $dK/dT\left(w_k-T\right)+dL/dT\left(w_l-T\right)$ is larger than zero.

If cities are specialized, the above algebra does not hold. A central policymaker would maximize $W=KV_k+LV_l+K^*V_k^*+L^*V_l^*$. The local policymaker's first order condition states that $d\left(KV_k+LV_l\right)/dT=0$, so that the externality is captured by $d\left(K^*V_k^*+L^*V_l^*\right)/dT$. Inserting the equilibrium land rent, this other region's welfare effects are summarized by $T^{*\gamma}\alpha^{-\alpha}\left(L^*\left(w_l^*-T^*\right)+K^*\left(w_k^*-T^*\right)\right)^{\alpha}$. Since the other region's tax T^* is held constant from the perspective of the local policymaker, there are only effects of production via worker relocation: policy is unchanged. Therefore, aggregate externality must take the sign of:

$$dQ^*/dT - T^*\left(\frac{dK^*}{dT} + \frac{dL^*}{dT}\right). \tag{7.B.3}$$

Rearranging the first order condition of the high-skilled city's policymaker, we have that:

$$\begin{aligned} & (Q - t\,(K + L))\,\gamma/t - \beta\,(K + L) \\ = \;& \beta\,(dQ/dt - t\,(dK/dt + dL/dt)). \end{aligned} \tag{7.B.4}$$

The left-hand side of this expression represents the utility gains of supplying the government less the costs, and it is positive in equilibrium. Therefore, the term $dQ/dT - T\,(dK/dT + dL/dT)$ needs to be positive. This is intuitive: since higher taxes attract more high-skilled workers, policy becomes more efficient due to increased homogeneity. Since taxes are below the level preferred by the high-skilled worker, there must be a cost to balance the benefits of full specialization, which are the costs of overspecialization in production. Since the mobility conditions are based on the equality of:

$$\frac{(w - T)\,T^{\gamma}}{r^{\alpha}} = \frac{(w^* - T^*)\,T^{*\gamma}}{r^{*\alpha}}, \tag{7.B.5}$$

tax increases applied in one region must have opposite effects on migration if applied in the other region. Therefore, the term $dQ^*/dT - T^* \times (dK^*/dT + dL^*/dT)$ needs to be negative. The intuition is that the benefits of homogeneity in terms of policy are fully appropriated by the home city, but the cost in terms of productive overspecialization are shared by the two cities: increased specialization in one city (the city-level cost of homogenous policy) implies that the other cities is becomes more specialized in the other factor, which the first city does not take into account. Therefore, taxes are too high in the high-skilled city. To specialize in the low skilled city, however, the policymaker sets lower instead of higher taxes, and therefore his equilibrium taxes are too low.

CHAPTER 8

CONCLUSION

What can be concluded about the relationship between space and policy competition? The six core chapters that constitute this thesis have documented that various spatial and geographical relations between nations and cities affect policy formation: space matters significantly for policy competition, and as a consequence, for economic welfare. By highlighting relations through trade, labor, housing, migration and policy effects, this study moves beyond the focus on factor mobility as a source of problems faced by peer governments that need to develop policies.

The results largely suggest that the mobility of people and capital do not need to stand in the way of good policy, especially if agglomeration forces shape the economy. However, that does not mean that there is no role for central governments (or other checks on policy interactions). The possibilities of commuting, migration and trade still complicate policy making: often, we should expect that rational governments, even with the interest of their citizens in mind, arrive at suboptimal policies.

Departing from the tax competition literature (e.g., Wilson, 1991), this thesis suggests races to the bottom may not be common, or at least less so than many earlier studies suggest. Under agglomeration, policymakers may not even attempt to tie capital and firms to their region by offering lower taxes than their colleagues. However, while the absence of races to the bottom features in the most recent studies on tax competition, this thesis shows that agglomeration forces may bring about yet other policy problems, as not only fiscal externalities may occur, but other spillovers via goods markets, commuting and migration are relevant too. Instead of the laissez-faire or tax floors that are advocated in that literature (see, e.g., Baldwin et al., 2003, for an overview), it is argued here that government performance, at times, can be improved by intervention from central governments or institutions. However, the scope for improvement depends on the geography, and is therefore subtle. This role of geography is central in several parts of the thesis. Chapter 2, for instance, suggests that the need to harmonize policies, which is a key policy conclusion in earlier literature, depends on whether that policy has economic effects outside the jurisdiction. Chapter 3 shows that if cities tend to agglomerate imperfectly, the main worry may not be tax competition, but rather an inefficient policy "lock-in" that stems from the desire not to attract or drive away firms. Similarly, chapters 5 and 6 show that the mobility of labor (commuting) does not bias policies in the sense that governments compete to attract economic activity. Instead, the commuting costs between regions result in governments exerting too little effort to attract workers, because they mat not be able correct the suboptimal size of the commuting flow.

8.1 A summary of the results

Several chapters provide answers to the central question of the thesis – whether spatially connected governments choose optimal policies. However, the different aspects stressed throughout the text give rise to several, rather than a single answer to this question. This subsection provides a brief summary of the main results.

Chapter 2 takes up the question whether government policies require harmonization if economies tend to agglomerate. Harmonization was argued against in earlier agglomeration-based literature, because it runs against the difference in policy preferences that arise with size differences when regions agglomerate. The chapter shows, however, that when policies have externalities (i.e., a policy has favorable effects in nearby regions but those benefits are not considered by the policy-setter), harmonization can be an improvement. Given significant goods market integration, policies (subsidies) that promote the start-up of new firms benefit surrounding areas: neighboring consumers have access to a larger market. If there are no advantages in the timing of policy-formation, a harmonization is shown to exist that benefits all citizens.

Chapter 3 suggests that if economies tend to agglomerate incompletely, the main source of inefficiency is a policy coordination problem caused by the geography of cities. Incomplete agglomeration occurs because if the city grows, more houses are required and internal commutes increase. The desire to preserve the agglomeration leads the government's optimal tax to be the tax of the other region, with a compensation for agglomeration rents added or subtracted. The two tax rates are consequently perfectly dependent. Therefore, a lock-in effect occurs: many policies, including inefficient ones can be sustained as a best response (Nash) equilibrium.

Chapter 4 further examines the role of housing in policies. It shows that home-ownership biases political preferences away from the social optimum, if future policies are anticipated in the house price (i.e., capitalization of industrial policies). As a result, if policy is set democratically, voters will weigh a one-shot increase in their houseprice in the vote, and therefore excessively support pro-business policies. Expected immigration to fill future jobs puts a similar pressure on citizen's house prices, and therefore also leads to too extensive industrial policies. The model not only helps to understand the stylized fact that cities with more homeowners gear their policies more towards firms, it also suggests that these homeowners' cities' policies may not be efficient.

Chapter 5 examines the effects of intercity infrastructure investments. Adequate connectedness is often viewed as an asset for a city, but recent empirical evidence suggests that such investments often cause inhabitants to leave large cities, while firms may be either drawn towards large cities or leave them. The chapter shows that augmenting infrastructure leads workers to exploit lower houseprices and lower goods prices in peripheries while still travelling to the core region to work. In the model, population decentralization (and a decentralization of jobs per head) therefore follows inevitably from improving connectedness. In absolute terms, firms will also decentralize (move from larger to smaller regions), unless the commuting costs are already low.

Chapter 6 analyzes the formation of government policy if workers commute as they do in chapter 5. Larger regions' governments attract commuters by supporting their local firms (which indirectly improves local wages). Yet, in contrast to tax competition models, their efforts to attract activity are too small, rather than too large. Because the social benefits of commuting (agglomeration externalities) are higher than their private costs (commuting costs), the commuting flow is too small. Local governments only partly correct for the suboptimal commutign size, because they only attempt to adress the adverse effects to their own region. Worker migration forces policies to a second-best situation: policies are optimal given the inefficiency of commutes.

Chapter 7, finally, examines whether governments' efforts to attract highly skilled workers are efficient. Governments spend substantial public funds to supply goods that highly skilled workers like, such as cultural amenities (e.g., museums and concert halls). Because highly skilled workers contribute disproportionately to the local economy and improve the fate of the local lowly skilled workers, local government go at (too) great length to attract them. If some cities specialize in low-skilled production and others in high-skilled production, however, cities will overspecialize.

The reason is that cities fully enjoy the benefits of specialization (policy is tailored to local preferences) but bear the costs partly (productive specialization in the lowly skilled cities inevitably leads to productive specialization in the highly skilled city). Even if final goods are easily traded, which reduces the costs of city specialization, rational local government policies will lead to inefficient specialization patterns.

8.2 Scope for higher-level government intervention

On the regional and city level, the most feasible policy recommendations imply interventions from higher governments. Whether any form of centralization is desirable, depends much on whether people can migrate: in that case they may eliminate inefficiencies through voting by feet. Under immobility, even under agglomeration, policy harmonizations can improve welfare (chapter 3). If people relocate to their preferred cities and regions, harmonization is less desirable, but restricting at least one government could make all citizens better off (chapter 4). However, while there is a clear economic foundation, it is questionable whether politically and legally, it is feasible to intervene in one local government but not the other. A clear case for central government action is when cities compete to attract high-skilled workers: the economic contributions of the highly skilled can lead cities to bias their policies towards them. In that case, harmonization will improve welfare by forcing policies to be better tailored to the average person. Such a policy change is inherently political, too, as it redistributes between low- and high-skilled workers. If citizens sort into the (skill-specialized) cities whose policies they enjoy, cities typically become too lowly or highly skilled. If this occurs, harmonization defies the logic of policy tailoring: it would imply imposing high-skilled policies in low skilled cities and vice versa. However, as cities overspecialize, the model of chapter 7 suggests that some degree of policy convergence is desirable.

To counter problems stemming from commutes, a central government might improve the fate of all of its inhabitants by interfering in the housing market, for instance with a mortgage interest rate dedcution. Commuters ignore the productivity effects they confer upon peers, so fewer than the socially optimal number of workers commutes. If housing supply is inelastic (prices rise fast when the agglomeration grows), the suboptimally small commuting flow translates mainly to employment underconcentration (because workers prefer to live in the cheaper periphery). If housing supply is elastic, workers can afford living near their job, so the small commuting flow results in a residential overconcentration. A progressive housing subsidy (a higher percentage for higher house value) is justified for an inelastic housing market, as that concentrates residences, while an inelastic housing market calls for regressive housing subsidies. The housing subsidy thus needs to vary over regions. Moreover, the analysis does not take into account worker inequality other than interregional; so even if significant simplifications are made, the optimal design of housing subsidies is not easy.

Housing markets themselves are a source of policy bias in chapter 7. It argues that if different houses have different access to the labor market and if they are owned instead of rented, democratically elected urban governments are unable to select welfare-maximizing policies. Even without interaction with other cities, cities may spend too much on industrial policies. In this case, however, it is uncertain whether centralization will improve policy: if voters live in symmetric cities, they might elect equally biased national governments.

8.3 Methodological contributions

Looking at several expressions of geography in policy competition not only yields policy insights, but also requires contemplating methodological issues.

A recurring theme in the thesis is the ambiguous (chastening or spoiling) effect of people and firm mobility on policymakers. This ambiguity comes about because many models are a hybrid of two workhorse public finance models. The first is the classical tax competition model where policymakers end up with suboptimally low tax rates in attempts to retain or attract firms. The second is a strand of literature that focuses on people's mobility and political power, unequivocally leading to a corrective force on the local policymaker. It comprises voting by feet, Tiebout (preference) sorting, urban politics and the homevoter hypothesis. While the abstractions' appeal is to make the mechanisms understandable, one is not likely to encounter one of these models in pure form in reality. Throughout the thesis, and especially in chapters 3, 6 and 7, a mix of these models appears. On the one hand, this is due to a spatial equilibrium condition: people migrate to wherever utility is highest, therefore validating and reinforcing good policies by residential choice. However, if firms can be attracted with lower taxes, there is still a strategic consideration for tax-setters:

lower taxes lead to a larger group of firms that can be taxed. So, even if people leave places with bad policies, if all governments engage in tax competition, utility is simply lower everywhere. What this means for people's location choices probably depends on where the effects of tax competition work out particularly negatively: citizens will avoid living in such locations. How harmful the effects of tax competition are depends, again, on geography. In chapter 3, the tax base elasticities (how easily firms leave high-tax areas) determine the relocation due to tax competion; in chapter 6, commuting costs cause distortions in policy and location choice.

The empirical literature studying how mobility of firms and people affects policy formation cannot usually disentangle whether policy enhances or deteriorates welfare under mobility. Since the tax competition literature and the voting-by-feet/local politics literature evolved rather separately, it is hard to identify discriminating hypotheses: both can be consistent with the observations that policy variables correlate with neighboring policy variables. Moreover, within the tax competition literature, the dust has not settled on whether the classical or the agglomeration-based models are empirically more relevant. Explanatory power of policy variables of nearby jurisdictions for local policy is consistent with the classical policy competition literature, but potentially also with agglomeration-based models or voting by feet (see , e.g., chapter 3), even if confounding spatial phenomena, like market potential, can be accounted for. In short, thinking of the applicability of different types of policy competition of the thesis reveals that the literature is in need of empirical scrutiny, especially identifying discriminating hypotheses to distinguish models that predict the emergence of socially optimal policy from "races to the bottom".

The thesis also contributes to advances in the literature that recognize that agents are not representative, but inherently different from one another. Differences among locations, and the way policy affects them run centrally throughout the thesis (particularly chapters 2, 5 and section 6.4). Chapter 7 explores the ramifications of a skill difference that is non-trivial, because it modifies workers' behavior. These differences turn out to shape cities and cause policy inefficiencies, and are therefore a significant departure from the homogenous economic agent that is typically assumed. While such insights are present in the labor market literature, they have a more recent tradition in spatial economics (Zenou, 2009). Although insights in the effects of firm heterogeneity in spatial (trade) models are quickly developing (Melitz, 2003, is a seminal article), the thesis has paid less attention to firm heterogeneity. A number of studies have already examined this issue, however (e.g., Davies and Eckel, 2010).

The portrayal of government in this thesis largely follows a tradition in public finance, where government are benevolent entities. Chapter 4, by contrast, opens up the black box of policy formation and studies a democracy. It shows that in a spatial setting, a democratic government can fail to produce welfare-maximizing policies, because citizens (voters) have different political preferences as they live in different locations. The rational behavior of voters and politicians is consistent with rational behavior of all other parties in the model. Indeed, this ties in with a broader development in public economics, or rather political economics (Acemoglu and Robinson, 2006; Drazen, 2000; Persson and Tabellini, 2000) that emphasize the positive aspects of policy formation.

8.4 Topics for future research

One caveat that applies to the two-region models in this thesis is that it is uncertain how their results translate to a situation of many regions. It shares this concern with virtually all NEG models: even a three-region setup becomes difficult to solve with pencil and paper (see extensive discussion in Brakman et al., 2001, chapter 4, Fujita et al., 2001, chapter 6 and Combes et al., 2011, chapter 4). The reason for setting up a two-region model is clear: it captures the local economy (intraregional), its interaction with a neigboring economy (interregional), and the economy-wide effects of policy changes and behavior. Translating the results to a setting with many regions is sometimes possible if governments and citizens think of all other regions as a single, global economy (or a "representative" neighbor economy). In that case, relations with all other regions can be summarized easily: interregional effects are the same for all regions. This was the assumption of choice before the advent of NEG models, and the results for multiple regions in chapter 7 rely on it (where all cities have the same transport costs between them). However, in NEG models, where trade costs form a hallmark ingredient, assuming equal interactions between all regions is not feasible: not all regions are connected equally, nor are there "representative neighbors", because regions diverge in size. If transport costs matter realistically, then occurences nearby should matter more than those far away. This re-introduces asymmetric interregional interactions and eliminates closed form solutions. Thus, many-region results that accurately capture the effects of transport (and other geographical) costs need to rely on numerical simulations, which may impede the development of these theories.

Another, and possibly more serious obstacle to extending the results in this thesis to a multi-region setup stems from the strategic interaction between governments. If the optimal response of one government depends on the choices of other governments, a large multilateral system emerges. Therefore, the computational issues from multi-region location models also arise for governments. For governments, an additional complication arises when studying interaction compared to firms or residents. Firms and inhabitants are typically assumed to be small when they select location or trade, so they have no strategic incentives. Governments, on the other hand, need to understand how other governments respond to their choices, to formulate their optimal policies. From the classical work of Nash (1950), we know that at least one equilibrium exists if the number of governments and their possible set of policies are finite. However, the strategic interaction leads to additional computational issues, and even computationally, results are hard to establish (see Stengel, 2010 for issues in the computation of n-player strategic equilibria). Nevertheless, a better insight into both the spatial and strategic interregional relations among more than two regions is an interesting theme for follow-up research. Not only could it show the extent to which the two-region results generalize, it also helps identifying real-life predictions of the models.

A third issue that calls for reservation in applying the models' recommendations in actual policy is that the transport costs in the models are still stylized. In nearly all chapters, a single parameter (τ or θ) has captured the notion that it somehow takes effort to transport goods or to commute. In that sense, this thesis is clearly grounded in economics: other disciplines, like geography, have an arguably more sophisticated treatment of space and distance. The formulations used in the thesis (and more widely in the field of economics) could ignore important transport costs configurations (McCann, 2005): for instance, if setting up the trade channel costs much more than shipment does, firms' behavior might change. More importantly, there are probably many spatial barriers that are ignored here: information sets, perceptions and even culture may change when moving along a few hundred kilometers. These may be lost when reducing transport costs to a proportional loss of goods or time.

A next matter worth consideration is the role in the welfare of the Dixit and Stiglitz (1977) model of monopolistic competition that is applied extensively throughout the thesis. As noted at several points (especially chapters 2 and 5), this setup helps to depart from perfect competition, but also has welfare implications due to love-of-variety effects, which are not necessarily related to the elasticity of substitution (Benassy, 1996). A less appealing feature of the Dixit-Stiglitz model is that all consumers always consume all goods in the market, no matter how distant or expensive those goods are to the consumer. As a result, local policies that increase the local number of producers have external effects in other regions. The presence of such external effects may be more realistic than their complete absence, which follows if the number of firms is fixed (which is assumed in the footloose entrepreneur and footloose capital models commonly used as an analytically tractable form of NEG). Linear models of imperfect competition (see, e.g., Melitz and Ottaviano, 2008) feature choke-off prices: product varieties are not consumed if they are too expensive (or their trade costs are too high). This limits the scope of external effects, because the goods of new entrants are not necessarily consumed elsewhere: they do not contribute to utility if they are very expensive to consume. Additionally, those models can explain pro-competitive effects: one benefit of firm entry may be that the average markup of firms falls, which does not occur in the workhorse Dixit-Stiglitz model. Therefore, although the Constant Elasticity of Substition-setup is the predominant choice in the literature related to this thesis, more realistic welfare results might be obtained by other models of imperfect competition.

Lastly, although the political perspective was listed as an advance in the insights of government behavior, the models of politics employed are still rather simple. To be able to understand what governments do, it could prove over-simplifying to reduce their role to suppliers of generic government services. In part, chapters that deal with productive inputs (chapter 4, among others) allude to this simplistic portrayal, but governments also serve as an insurance to shocks (Epifani and Gancia, 2009), and perform various other tasks, such as solving information asymmetries, coordination and redistribution (Myles, 1995), that are ignored here. Similarly, the median voter theorems that are employed work fine for one-dimensional political preferences, but in reality, politics deals with preferences in many dimensions, discourse, coalitions, strategy and probably marketing. An appealing step forward could be a tie to the literature on lobbying for protection into the tax competition literature.

Altogether, this thesis has attempted to explain how space affects policy competition. It provides some guidance in the design of policies and institutions that could help remedy the problems that arise when different economies jointly need to decide on what their government does. Yet, when taking stock, one might also conclude that a complete portrayal of how governments interact is far out of grasp. If anything, the large differences in results from adding different building blocks to tax competition models in this thesis demonstrates how stylized the workhorse models are. One way forward is to start pairing theoretical predictions with empirical work with good identification strategies. The contribution of such work to our understanding of policy competition could be high in a literature that is mostly abstract and theoretical to date.

NEDERLANDSE SAMENVATTING

Uit angst voor buitenlandse legers begon het Nederlandse Departement voor Oorlog in 1880 met de bouw van de Stelling van Amsterdam. Met een radius van vijftien kilometer rondom het centrum van de stad werd een droge gracht gegraven. Met de stelling van Amsterdam kon de stad Amsterdam met één bevel tot eiland worden gemaakt. Zover is het nooit gekomen: Amsterdam, noch andere grote Nederlandse steden zijn een eiland geworden. Het beeld van zo'n mogelijke isolatie helpt wel om voor te stellen hoe groot het economisch belang van de buitenwereld is voor de stad: Amsterdam zou er beduidend anders uitzien als het een eiland was, volkomen afgesloten van de buitenwereld. Amsterdamse fabrieken en winkels konden dan hun waar niet kwijt aan de rest van de wereld, en Amsterdammers zouden geen inkopen doen buiten de stad, laat staan bij Amazon. Multinationals zouden zich niet in Amsterdam vestigen. Forensen zouden niet dagelijks uit Utrecht en Almere stromen, en Amsterdammers zouden niet meer naar de Flevopolder verhuizen.

Bestuurders van Amsterdam moeten er rekening mee houden dat Amsterdam géén geïsoleerd eiland is. Ze moeten bij het uitdenken van hun beleid bedenken wat de interactie is met andere steden en regio's. Een groei in bedrijvigheid trekt banen aan, dus er zou een grotere forensenstroom ontstaan, en immigratie. Maar huizenprijzen zouden ook stijgen, en misschien andere inwoners de stad doen verlaten. Het bouwen van nieuwe musea in Amsterdam trekt misschien hoogopgeleiden aan, maar verlaagt het opleidingsniveau van de bevolking in Utrecht, als de nieuwe hoogopgeleide Amsterdammers daar vandaan komen.

De verbindingen tussen steden en regio's bemoeilijken het voeren van optimaal beleid. De mobiliteit van bedrijven tussen landen, bijvoorbeeld, leidt tot de discussie of Nederland een belastingparadijs is. Een lage Nederlandse belasting trekt veel bedrijven aan, waardoor de belastingopbrengsten stijgen. Goed nieuws voor Nederlandse burgers, maar minder voor de VS, waar de belastinginkomsten dalen. Nederlandse politici, die zich minder zorgen maken over de Amerikaanse staatsfinanciën, houden dus geen rekening met de buitenlandse effecten van hun beleid. Daardoor is de Nederlandse belasting gunstig vanuit Nederlands perspectief, maar te laag vanuit een mondiaal perspectief. De VS zou haar belastingvoet ook kunnen verlagen in een poging de uitstroom van bedrijven tegen te gaan. Dit fenomeen heet belastingconcurrentie, en de klassieke remedie is de harmonisatie van belastingvoeten: dan blijven belastingvoeten hoog en zijn beide landen beter af. Ook ander beleid wordt verstoord door interacties met andere steden of regio's. Als Amsterdam denkt de lokale welvaart te kunnen verhogen door kantoorruimte te subsidiëren, dan hebben inwoners van Utrecht baat bij de extra bedrijvigheid en toegenomen werkgelegenheid. Maar Amsterdam zal kantoren subsidiëren tot het punt waar het haar eigen inwoners niets meer oplevert, en dus voorbijgaan aan mogelijke gunstige of ongunstige effecten die inwoners van Utrecht per saldo ondervinden. Daarmee zou de subsidie uit maatschappelijk oogpunt te laag of te hoog kunnen zijn.

Ruimtelijke interacties als handel, woon-werkverkeer, verhuizingen en bedrijfsvestiging beïnvloeden dus beleid. Dit proefschrift gaat in op zulke relaties tussen steden, en bestudeert of en hoe ze leiden tot verstoringen van optimaal beleid vanuit maatschappelijk perspectief. Een centrale vraag is of overheden uit zichzelf sociaal optimaal beleid kiezen, gegeven hun interacties met andere overheden. Om die vraag te beantwoorden is een duidelijk beeld van ruimtelijke relaties nodig. De ruimtelijke economie speelt daarom een grote rol in het proefschrift.

Agglomeratiekrachten in het bijzonder blijken een sterk effect te hebben op de manier waarop overheden met elkaar omgaan. Als bedrijven sterk de neiging hebben zich dicht bij elkaar te vestigen, leiden kleine belastingverschillen niet tot verplaatsingen van bedrijven. In een belangrijk artikel laten Paul Krugman en Richard Baldwin zien dat bedrijven in een dergelijke situatie de voordelen van een grote markt niet snel opgeven voor een iets gunstigere belasting. Belastingconcurrentie treedt daarom niet op en harmonisatie is daarom niet wenselijk. Sterker nog, als grote steden het liefst andere belastingen kiezen dan kleine steden, levert harmonisatie welvaartsverlies op. Het model van een ruimtelijke economie dat dit argument ondersteunt, de "New Economic Geography", is het theoretische bouwblok van een groot deel van dit proefschrift.

Het tweede hoofdstuk van dit proefschrift laat zien dat agglomeratiekrachten de wenselijkheid van beleidsharmonisatie niet wegnemen. Als er er handel in gedifferentieerde goederen is, en overheden kiezen tegelijkertijd hun beleid, dan verhoogt harmonisatie de welvaart. De gelijktijdige keuze van beleid dwingt de overheid van de grote regio om er rekening mee te houden dat de de overheid van de kleinere regio bedrijvigheid probeert aan te trekken. De grotere overheid subsidieert daarom bedrijvigheid, wat gunstig is voor inwoners in de kleinere regio. Als het beleid in de grote regio tot stand komt voorafgaand aan de beleidsvorming in de kleine regio, dan verstrekt de beleidsmaker in de grote regio subsidies om te voorkomen dat er bedrijvigheid verdwijnt uit zijn regio. Harmonisatie van beleid is dan ofwel slecht voor de kleine regio (omdat de uit de grote regio te importeren productie minder wordt gesubsidïeerd), ofwel slecht voor de grotere regio (omdat de subsidie hoger is dan nodig om de lokale bedrijvigheid te behouden).

Het derde hoofdstuk beschouwt het effect van de interne structuur van steden. In tegenstelling tot het originele model nemen steden in dit hoofdstuk fysieke ruimte in. Daardoor leidt stedelijke groei tot toegenomen werkgelegenheid en bedrijvigheid, maar ook tot hogere interne kosten van woon-werkverkeer. Als economische activiteit zich concentreert, proberen groot en klein gegroeide steden de *status quo* te behouden. Om de concentratie van bedrijven te behouden kiest de kleine stad niet voor te lage belastingvoeten en de grote stad niet voor te hoge belastingvoeten: er is een minimaal verschil tussen de twee belastingvoeten. Echter, er is een hele set aan belastingvoeten die voldoet aan dat verschil. Het is dus mogelijk dat de steden samen een optimaal beleid voeren. Maar het is ook mogelijk dat er een "lock-in" tot stand komt op een verre van optimaal punt: beide overheden voeren suboptimaal beleid, maar geen van de steden kan daar van afwijken zonder een verplaatsing van bedrijvigheid te veroorzaken.

De ruimtelijke structuur van de stad speelt een centrale rol in hoofdstuk 4, net als in hoofdstuk 3. Inwoners bezitten een huis, en reizen vanaf hun huis naar de centrale werkplek. Huizen met eenvoudige toegang tot de arbeidsmarkt (centraal gelegen huizen) zijn gewilder, en daarom duurder. Het hoofdstuk beschouwt vervolgens democratische beleidsvorming binnen de stad rondom de steun aan bedrijven. Als bedrijven met publiek geld worden gesteund, worden ze productiever, en neemt het lokale loon toe. Het hogere loon doet huizenprijzen stijgen, met name op centrale locaties van waaruit het eenvoudiger is naar de werkplek te reizen. Het model laat zien dat als kiezers de huiswaardestijging meenemen in hun stemgedrag, democratisch beleid niet tot optimale uitkomsten leidt. Het huiswaarde-effect geeft kiezers een kans om eenmalig de waarde van hun bezit te verhogen. Een zelfde verstoring van het democratische beleid treedt op als nieuwe inwoners zich in de stad kunnen vestigen. Een verhoging van de steun aan bedrijven verhoogt het loon en trekt inwoners aan, waardoor de huizenprijzen stijgen ten gunste van de huidige huizenbezitters.

Hoofdstuk 5 ontwikkelt een ruimtelijk model waarbij een andere ruimtelijke interactie optreedt: woon-werkverkeer tussen steden. Het laat zien dat een betrekkelijk algemeen model (waar de eerder genoemde "New Economic Geography" ook onder valt) kan verklaren waarom kleine en grote steden naast elkaar bestaan, en waarom er een forensenstroom richting grote steden optreedt. Het model biedt ook een verklaring voor de wisselende empirische resultaten over de effecten van infrastructuurinvesteringen op bevolking en werkgelegenheid. Infrastructuur die reiskosten verlaagt, maakt het mogelijk om in de goedkopere, kleine stad te wonen en in een grotere stad te werken tegen een hoger loon. Omdat bevolkingsgroei ook de vraag naar lokale diensten en producten vergroot, kan de werkgelegenheid zich zowel concentreren in de grotere stad als verspreiden over de steden. Reiskostenverlagingen voor forensen verschuiven bedrijvigheid vooral naar grote steden als er al een goede infrastructuur aanwezig is. Een vermindering van de transportkosten voor goederen spreidt bedrijvigheid juist richting kleinere steden, van waaruit ze tegen lagere kosten hun product eenvoudig naar een grotere markt verschepen.

De effecten van woon-werkverkeer op beleidsconcurrentie staan centraal in het zesde hoofdstuk. Het model bouwt voort op het forensenmodel van hoofdstuk 5, maar voegt overheden toe die bedrijven aan kunnen trekken met lokale investeringen in productiviteit. Omdat inwoners vrij kunnen verhuizen naar de stad waar ze het beste af zijn, worden overheden geprikkeld om op sociaal optimaal niveau te investeren in lokale productiviteit. Ze worden daar in gehinderd doordat de woon-werkstroom niet optimaal is. Een forens reist naar grotere steden omdat zijn productiviteit en daarom zijn loon daar hoger is. Hij neemt niet in beschouwing dat zijn aanwezigheid in de grote stad ook andere werknemers productiever maakt, en daarom nemen minder werknemers de beslissing om te reizen dan maatschappelijk optimaal zou zijn. Vervolgens stelt het hoofdstuk de vraag of een subsidie op huizen, zoals de hypotheekrenteaftrek, werknemers zo kan prikkelen dat hun verhuis- en reisgedrag richting het sociale optimum verandert. Afhankelijk van de aanbodelasticiteit van huizen kan het optimaal zijn huizen op duurdere of juist op goedkopere locaties te subsidiëren. Het is dus niet eenvoudig het optimale beleid te formuleren.

Het zevende en laatste inhoudelijke hoofdstuk gaat er van uit dat overheden proberen een specifieke groep mensen en bedrijven aan te trekken. Een voorbeeld daarvan is het bouwen van een museum als het Guggenheim in Bilbao (Spanje) om hoogopgeleide werknemers aan te trekken. In het model hebben overheden de beschikking over een beleid dat hoogopgeleide werknemers in sterkere mate aantrekt dan laagopgeleide werknemers. Als hoogopgeleiden zich gelijk over steden verspreiden, leidt het gedrag van de stadsbestuurders er toe dat er meer dan optimaal wordt uitgegeven om hoogopgeleide werknemers aan te trekken. Als alle steden gemeenschappelijk de uitgaven aan "hoogopgeleiden-beleid" verlagen, dan verhoogt dat de algehele velvaart. Maar dat kan niet zonder herverdeling: het is goed voor laagopgeleiden, maar slecht voor hoogopgeleiden. Als hoogopgeleide werknemers hun beleid erg belangrijk vinden, kan het ook voorkomen dat ze zich concentreren in de steden die beleidsmatig sterk op hoogopgeleiden zijn gericht. Dan specialiseren steden zich in hoog- of laagopgeleide productie. Als hoogopgeleiden uit een laagopgeleide stad worden aangetrokken naar de hoogopgeleide stad, raakt de laagopgeleide stad nog lager opgeleid. Als beleidsmakers zich op de lokale welvaart richten, nemen ze zulke specialisatie-effecten buiten hun grenzen niet in beschouwing. Steden specialiseren zich dan verder dan optimaal is: de stad van hoogopgeleiden is té hoogopgeleid, en de stad van laagopgeleiden is té laagopgeleid. Als het mogelijk is producten tussen steden te verhandelen, kunnen mensen van verschillend opleidingsniveau en inkomen zich concentreren in de steden die voor hun optimaal beleid voeren. De keerzijde daarvan is dat veel goederen getransporteerd moeten worden. Een verlaging van de transportkosten is daarom welvaartsverhogend: het stelt steden in staat te specialiseren en daarmee beter beleid te voeren.

De centrale vraag van het proefschrift - hoe de ruimtelijke organisatie van de economie beleidsconcurrentie beïnvloedt - heeft geen eenduidig antwoord. De interacties die maken dat steden en regio's geen eilanden zijn, zijn ook een bron van beleidsproblemen. Migratie leidt tot prikkels voor lokale overheden om het beleid te formuleren dat de lokale welvaart maximaliseert, omdat inwoners anders wegtrekken. Dat betekent echter niet dat de combinatie van lokale beleidskeuzes maatschappelijk optimaal is. Als mensen niet hetzelfde opleidingsniveau hebben, of als mensen forensen, is te verwachten dat beleidsmakers die de welvaart van hun eigen inwoners in gedachten hebben, niet de beste keuzes maken uit maatschappelijk perspectief. Ook het aanbod van bedrijven in verschillende regio's is van belang in de welvaartsconclusies: het bepaalt of beleidsharmonisatie gewenst is en of steden zich gunstige sorteerprocessen kunnen permitteren.

Omdat stedelijke bestuurders en regionale overheden onder een nationale overheid vallen, kan een overkoepelende autoriteit ingrijpen in het gedrag van overheden. De gewenste ingrepen van die autoriteit hangen in de verschillende modellen van dit proefschrift af van de ruimtelijke interactie tussen lokale overheden. Harmonisatie van belasting en uitgaven (gelijktrekken tussen steden) leidt onder sommige omstandigheden tot welvaartswinst: als er voldoende gehandeld wordt en overheden gelijktijdig hun beleid bepalen (hoofdstuk 2), of als overheden van soortgelijke steden proberen hoogopgeleide werknemers aan zich te binden (hoofdstuk 7). Als steden en regio's niet gelijk zijn, bijvoorbeeld door agglomeratie of door specialisatie, dan kan een geharmoniseerd beleid juist verstorend werken. Hoofdstuk 3 laat zien dat grote steden liever een hogere belastingvoet kiezen dan kleine steden. Als steden gespecialiseerd zijn in één type arbeid, zoals in hoofdstuk 7, is harmonisatie ook niet optimaal. In hoofdstuk 6 blijkt dat het verschil in belastingvoet tussen kleine en grote steden te klein is, in plaats van te groot. Samenvattend, er kan reden zijn voor het ingrijpen van een centrale overheid in het beleid van haar eigen lagere overheden, maar de instrumenten en de mate waarin ze gebruikt moeten worden hangen sterk af van de ruimtelijke context. In die zin breidt dit proefschrift de inzichten van de literatuur over belastingconcurrentie uit voorbij de aanbeveling om te harmoniseren. Agglomeratieeffecten, migratie, handel en woonwerktverkeer hebben elk hun eigen effect op beleidsinteracties.

Ruimtelijke relaties via handel, forensenstromen en verhuizingen zijn gezond: ze stellen mensen in staat op hun favoriete plek te wonen, de beste werkgever te zoeken, en uit een breed aanbod goederen en diensten te kiezen. De problemen die ruimtelijke relaties opleveren voor beleidsmakers zijn een belangrijk bijproduct. Soms is een goed ontsloten en verbonden stad moeilijker te besturen. Op verschillende plekken laat dit proefschrift zien dat de groeiende integratie van steden en regio's ander beleid vergt, of zelfs andere bestuursstructuren. Een verdedigingsgracht zou Amsterdam tot een onaangenaam eiland maken – goed bestuur stelt steden juist in staat isolatie te verminderen om vruchtbare verbindingen aan te gaan.

BIBLIOGRAPHY

Abdel-Rahman, H. M. and M. Fujita (1990). Product variety, Marshallian externalities, and city sizes. *Journal of Regional Science 30*, 165–183.

Acemoglu, D. and J. Robinson (2006). *Economic Origins of Dictatorship and Democracy*. Cambridge University Press.

Acemoglu, D. and J. Ventura (2002). The world income distribution. *Quarterly Journal of Economics 117*(2), 659–694.

Adamson, D. W., D. E. Clark, and M. D. Partridge (2004). Do urban agglomeration effects and household amenities have a skill bias? *Journal of Regional Science 44*(2), 201–224.

Aguilera, A. (2005). Growth in commuting distances in French polycentric metropolitan areas: Paris, Lyon and Marseille. *Urban Studies 42*(9), 1537–1547.

Anas, A. (2004). Vanishing cities: what does the new economic geography imply about the efficiency of urbanization? *Journal of Economic Geography 4*(2), 181–199.

Aoyagi, M. (1996). Reputation and dynamic Stackelberg leadership in infinitely repeated games. *Journal of Economic Theory 71*(2), 378–393.

Aschauer, D. A. (1989). Is public expenditure productive? *Journal of Monetary Economics 23*(2), 177–200.

Au, C.-C. and V. J. Henderson (2006). Are Chinese cities too small? *Review of Economic Studies 73*(3), 549–576.

Baldwin, R., R. Forslid, P. Martin, G. Ottaviano, and F. Robert-Nicoud (2003). *Economic Geography and Public Policy*. Princeton University Press.

Baldwin, R. and P. Krugman (2004). Agglomeration, integration and tax harmonisation. *European Economic Review 48*(1), 1–23.

Baum-Snow, N. (2007). Did highways cause suburbanization? *Quarterly Journal of Economics 122*(2), 775–805.

Baum-Snow, N. (2010). Changes in transportation infrastructure and commuting patterns in US metropolitan areas, 1960–2000. *American Economic Review 100*(2), 378–382.

Baum-Snow, N., L. Brandt, J. V. Henderson, M. Turner, and Q. Zhang (2012). Roads, railroads and decentralization of Chinese cities. *Mimeo, University of Toronto*.

Behrens, K. and F. Robert-Nicoud (2009). Krugman's papers in regional science: The 100 dollar bill on the sidewalk is gone and the 2008 Nobel prize well-deserved. *Papers in Regional Science 88*(2), 467–489.

Benassy, J.-P. (1996). Taste for variety and optimum production patterns in monopolistic competition. *Economics Letters 52*(1), 41–47.

Borck, R. and M. Pflüger (2006). Agglomeration and tax competition. *European Economic Review 50*(3), 647–668.

Borck, R., M. Pflüger, and M. Wrede (2010). A simple theory of industry location and residence choice. *Journal of Economic Geography 10*, 913–940.

Borck, R. and M. Wrede (2009). Subsidies for intracity and intercity commuting. *Journal of Urban Economics 66*(1), 25–32.

Bostic, R., S. Gabriel, and G. Painter (2009). Housing wealth, financial wealth, and consumption: New evidence from micro data. *Regional Science and Urban Economics 39*(1), 79–89.

Braid, R. M. (1996). Symmetric tax competition with multiple jurisdictions in each metropolitan area. *American Economic Review 86*(5), 1279–1290.

Braid, R. M. (2000). A spatial model of tax competition with multiple tax instruments. *Journal of Urban Economics 47*(1), 88–114.

Brakman, S., H. Garretsen, R. Gigengack, C. van Marrewijk, and R. Wagenvoort (1996). Negative feedbacks in the economy and industrial location. *Journal of Regional Science 36*(4), 631–651.

Brakman, S., H. Garretsen, and C. Van Marrewijk (2001). *An Introduction to Geographical Economics: Trade, Location and Growth*. Cambridge University Press.

Brakman, S., H. Garretsen, and C. van Marrewijk (2002). Locational competition and agglomeration: The role of government spending. *CESifo working paper no. 775*.

Bretschger, L. and F. Hettich (2002). Globalisation, capital mobility and tax competition: theory and evidence for OECD countries. *European Journal of Political Economy 18*(4), 695–716.

Brülhart, M., M. Jametti, and K. Schmidheiny (2012). Do agglomeration economies reduce the sensitivity of firm location to tax differentials? *Economic Journal 122*(563), 1069–1093.

Brueckner, J. K. and M.-S. Joo (1991). Voting with capitalization. *Regional Science and Urban Economics 21*(3), 453–467.

Bucovetsky, S. (1991). Asymmetric tax competition. *Journal of Urban Economics 30*(2), 167–181.

Buettner, T. and E. Janeba (2009). City competition for the creative class. *Mimeo, UC Louvain*.

Campbell, J. and J. Cocco (2005). How do house prices affect consumption? Evidence from microdata. *NBER working paper no. 11534*.

Carroll, R. J. and J. Yinger (1994). Is the property tax a benefit tax? The case of rental housing. *National Tax Journal 47*(2), 295–316.

Chandra, A. and E. Thompson (2000). Does public infrastructure affect economic activity? Evidence from the rural interstate highway system. *Regional Science and Urban Economics 30*(4), 457–490.

Charlot, S. and S. Paty (2007). Market access effect and local tax setting: Evidence from French panel data. *Journal of Economic Geography 7*(3), 247–263.

Combes, P.-P., T. Mayer, and J.-F. Thisse (2011). *Economic Geography: The Integration of Regions and Nations*. Princeton University Press.

Commendatore, P., I. Kubin, and C. Petraglia (2008). Productive public expenditure in a new economic geography model. *Economie Internationale 114*, 133–160.

Cruz, J. J. (1975). Survey of Nash and Stackelberg equilibrium strategies in dynamic games. In *Annals of Economic and Social Measurement*, Volume 4, pp. 339–344. National Bureau of Economic Research.

Dalmazzo, A. and G. de Blasio (2011). Amenities and skill-biased agglomeration effects: Some results on Italian cities. *Papers in Regional Science 90*(3), 503–527.

Dasgupta, P. and E. Maskin (1986). The existence of equilibrium in discontinuous economic games, i: Theory. *Review of Economic Studies 53*(1), 1–26.

Davies, R. and C. Eckel (2010). Tax competition for heterogeneous firms with endogenous entry. *American Economic Journal: Economic Policy 2*(1), 77–102.

Dehring, C. A., C. A. Depken II, and M. R. Ward (2008). A direct test of the homevoter hypothesis. *Journal of Urban Economics 64*(1), 155–170.

Dixit, A. K. and J. E. Stiglitz (1975). Monopolistic competition and optimum product diversity. *University of Warwick working paper no. 64*.

Dixit, A. K. and J. E. Stiglitz (1977). Monopolistic competition and optimum product diversity. *American Economic Review 67*(3), 297–308.

Doi, T. (1999). Empirics of the median voter hypothesis in Japan. *Empirical Economics 24*(4), 667–691.

Drazen, A. (2000). *Political Economy in Macroeconomics*. Princeton University Press.

Duranton, G. and D. Puga (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson and J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics: Cities and Geography*, Volume 4, Chapter 48, pp. 2063–2117. Elsevier.

Duranton, G. and A. Rodríguez-Pose (2005). When economists and geographers collide, or the tale of the lions and the butterflies. *Environment and Planning A 37*(10), 1695–1705.

Duranton, G. and M. A. Turner (2011). The fundamental law of road congestion: Evidence from US cities. *American Economic Review 101*(6), 2616–2652.

Duranton, G. and M. A. Turner (2012). Urban growth and transportation. *Review of Economic Studies 79*(4), 1407–1440.

Epifani, P. and G. Gancia (2009). Openness, government size and the terms of trade. *Review of Economic Studies 76*(2), 629–668.

Ethier, W. J. (1982). National and international returns to scale in the modern theory of international trade. *American Economic Review 72*(3), 389–405.

Falck, O., M. Fritsch, and S. Heblich (2011). The phantom of the opera: Cultural amenities, human capital, and regional economic growth. *Labour Economics 18*(6), 755–766.

Feld, L. P. (2000). Tax competition and income redistribution: An empirical analysis for Switzerland. *Public Choice 105*(1–2), 125–164.

Fenge, R., M. von Ehrlich, and M. Wrede (2009). Public input competition and agglomeration. *Regional Science and Urban Economics 39*(5), 621–631.

Fernald, J. G. (1999). Roads to prosperity? Assessing the link between public capital and productivity. *American Economic Review 89*(3), 619–638.

Fischel, W. A. (2001). *The Homevoter Hypothesis. How Home Values Influence Local Government Taxation, School Finance, and Land-Use Policies.* Harvard University Press.

Florida, R. (2002a). Bohemia and economic geography. *Journal of Economic Geography 2*(1), 55–71.

Florida, R. (2002b). The economic geography of talent. *Annals of the Association of American Geographers 92*(4), 743–755.

Forslid, R. and G. I. Ottaviano (2003). An analytically solvable core-periphery model. *Journal of Economic Geography 3*(3), 229–240.

Fudenberg, D. and J. Tirole (1991). *Game Theory*. MIT Press.

Fujita, M. (1989). *Urban Economic Theory: Land Use and City Size*. Cambridge University Press.

Fujita, M., P. Krugman, and A. J. Venables (2001). *The Spatial Economy: Cities, Regions, and International Trade*. MIT Press.

Fujita, M. and T. E. Smith (1987). Existence of continuous residential land-use equilibria. *Regional Science and Urban Economics 17*(4), 549–594.

Fujita, M. and J.-F. Thisse (2002). *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth*. Cambridge University Press.

Fujita, M. and J.-F. Thisse (2009). New economic geography: An appraisal on the occasion of Paul Krugman's 2008 Nobel prize in economic sciences. *Regional Science and Urban Economics 39*(2), 109–119.

Furman, J. L., M. E. Porter, and S. Stern (2002). The determinants of national innovative capacity. *Research Policy 31*(6), 899–933.

Garretsen, H. and R. Martin (2010). Rethinking (new) economic geography models: Taking geography and history more seriously. *Spatial Economic Analysis 5*(2), 127–160.

Gibbons, S. and S. Machin (2008). Valuing school quality, better transport, and lower crime: evidence from house prices. *Oxford Review of Economic Policy 24*(1), 99–119.

Glaeser, E. and J. Kohlhase (2003). Cities, regions and the decline of transport costs. *Papers in Regional Science 83*(1), 197–228.

Glaeser, E. L., J. Kolko, and A. Saiz (2001). Consumer city. *Journal of Economic Geography 1*(1), 27–50.

Glaeser, E. L. and A. Saiz (2004). The rise of the skilled city. *Brookings-Wharton Papers on Urban Affairs 1*, 47–105.

Glaeser, E. L. and J. M. Shapiro (2002). The benefits of the home mortgage interest deduction. *NBER Working Paper no. 9284*.

Glaeser, E. L. and J. M. Shapiro (2003). Urban growth in the 1990s: Is city living back? *Journal of Regional Science 43*(1), 139–165.

Glazer, A. and K. Van Dender (2002). How congestion pricing reduces property values. *Institute of Transportation Studies University of California working paper no. 02-1*.

Gordon, R. and J. Hines Jr. (2002). International taxation. In A. J. Auerbach and M. Feldstein (Eds.), *Handbook of Public Economics*, Volume 4, Chapter 28, pp. 1935–1995. Elsevier.

Gordon, R. H. (1992). Can capital income taxes survive in open economies? *Journal of Finance 47*(3), 1159–1180.

Gramlich, E. M. and D. L. Rubinfeld (1982). Micro estimates of public spending demand functions and tests of the Tiebout and median-voter hypotheses. *Journal of Political Economy 90*(3), 536–560.

Grodach, C. and A. Loukaitou-Sideris (2007). Cultural development strategies and urban revitalization. *International Journal of Cultural Policy 13*(4), 349–370.

Guo, J. (2009). Tax competition for commuters. *Regional Science and Urban Economics 39*(2), 148–154.

Gutiérrez-i-Puigarnau, E. and J. N. Van Ommeren (2010). Labour supply and commuting. *Journal of Urban Economics 68*(1), 82–89.

Helpman, E. (1998). The size of regions. In D. Pines, E. Sadka, and Y. Zilcha (Eds.), *Topics in Public Economics*, pp. 33–54. Cambridge University Press.

Helsley, R. W. (2004). Urban political economics. In J. V. Henderson and J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Volume 4 of *Handbook of Regional and Urban Economics*, Chapter 54, pp. 2381–2421. Elsevier.

Henderson, J. V. (1974). Optimum city size: The external diseconomy question. *Journal of Political Economy 82*(2), 373–388.

Henderson, J. V. (1986). Efficiency of resource usage and city size. *Journal of Urban Economics 19*(1), 47–70.

Hilber, C. A., T. Lyytikäinen, and W. Vermeulen (2011). Capitalization of central government grants into local house prices: Panel data evidence from England. *Regional Science and Urban Economics 41*(4), 394–406.

Hilber, C. A. and C. Mayer (2009). Why do households without children support local public schools? Linking house price capitalization to school spending. *Journal of Urban Economics 65*(1), 74–90.

Hilber, C. A. and F. Robert-Nicoud (2013). On the origins of land use regulations: Theory and evidence from US metro areas. *Journal of Urban Economics 75*, 29–43.

Hilber, C. A. L. and F. Robert-Nicoud (2007). Homeownership and land use controls: a dynamic model with voting and lobbying. *London School of Economics and Political Science, Geography and Environment Department working paper no. 119*.

Hill, B. C. (2008). Agglomerations and strategic tax competition. *Public Finance Review 36*(6), 651–677.

Iacoviello, M. (2011). Housing wealth and consumption. *Federal Reserve Board Washington discussion paper no. 1027*.

Janeba, E. (1998). Tax competition in imperfectly competitive markets. *Journal of International Economics 44*(1), 135–153.

Jofre-Monseny, J. (2013). Is agglomeration taxable? *Journal of Economic Geography 13*(1), 177–201.

Jofre-Monseny, J. and A. Solé-Ollé (2010). Tax differentials in intraregional firm location: Evidence from new manufacturing establishments in Spanish municipalities. *Regional Studies 44*(6), 663–677.

Keen, M. and M. Marchand (1997). Fiscal competition and the pattern of public spending. *Journal of Public Economics 66*(1), 33–53.

Kempf, H. and G. Rota-Graziosi (2010). Endogenizing leadership in tax competition. *Journal of Public Economics 94*(9–10), 768–776.

Kind, H. J., K. H. M. Knarvik, and G. Schjelderup (2000). Competing for capital in a 'lumpy' world. *Journal of Public Economics 78*(3), 253–274.

King, I., R. P. McAfee, and L. Welling (1993). Industrial blackmail: Dynamic tax competition and public investment. *Canadian Journal of Economics 26*(3), 590–608.

Klemm, A. (2010). Causes, benefits, and risks of business tax incentives. *International Tax and Public Finance 17*, 315–336.

Koh, H.-J., N. Riedel, and T. Böhm (2013). Do governments tax agglomeration rents? *Journal of Urban Economics 75*, 92 – 106.

Krogstrup, S. (2008). Standard tax competition and increasing returns. *Journal of Public Economic Theory 10*(4), 547–561.

Krugman, P. R. (1991). Increasing returns and economic geography. *Journal of Political Economy 99*(3), 482–499.

Lee, K. (2002). Factor mobility and income redistribution in a federation. *Journal of Urban Economics 51*(1), 77–100.

Lemke, C. E. and J. Howson Jr. (1964). Equilibrium points of bimatrix games. *Journal of the Society for Industrial and Applied Mathematics 12*, 413–423.

Levine, J. (1998). Rethinking accessibility and jobs-housing balance. *Journal of the American Planning Association 64*(2), 133–149.

Ludema, R. D. and I. Wooton (2000). Economic geography and the fiscal effects of regional integration. *Journal of International Economics 52*(2), 331–357.

Malecki, E. J. (1981). Science, technology, and regional economic development: Review and prospects. *Research Policy 10*(4), 312–334.

Mankiw, N. (1982). Hall's consumption hypothesis and durable goods. *Journal of Monetary Economics 10*(3), 417–425.

Marshall, A. (1890). *Principles of Economics*. Macmillan.

McCann, P. (2005). Transport costs and new economic geography. *Journal of Economic Geography 5*(3), 305–318.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71*(6), 1695–1725.

Melitz, M. J. and G. I. P. Ottaviano (2008). Market size, trade, and productivity. *Review of Economic Studies 75*(1), 295–316.

Meltzer, A. H. and S. F. Richard (1981). A rational theory of the size of government. *Journal of Political Economy 89*(5), 914–927.

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the interstate highway system. *Review of Economics and Statistics 90*(4), 683–701.

Mitra, A. (1999). Agglomeration economies as manifested in technical efficiency at the firm level. *Journal of Urban Economics 45*(3), 490–500.

Moretti, E. (2004). Human capital externalities in cities. In J. V. Henderson and J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics: Cities and Geography*, Volume 4, pp. 2243–2291. Elsevier.

Myles, G. (1995). *Public Economics*. Cambridge University Press.

Nash, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences of the United States of America 36*(1), 48–49.

Oates, W. E. (2005). Property taxation and local public spending: The renter effect. *Journal of Urban Economics 57*(3), 419–431.

Oates, W. E. and R. M. Schwab (1988). Economic competition among jurisdictions: Efficiency enhancing or distortion inducing? *Journal of Public Economics 35*(3), 333–354.

OECD (2005). *Employment Outlook*. OECD.

Ortalo-Magné, F. and A. Prat (2010). Spatial asset pricing: A first step. *CEPR discussion paper no. 7842*.

Ottaviano, G. I. and F. Robert-Nicoud (2006). The “genome” of NEG models with vertical linkages: A positive and normative synthesis. *Journal of Economic Geography 6*(2), 113–139.

Persson, T. and G. Tabellini (2000). *Political Economics: Explaining Economic Policy*. MIT Press.

Pisarski, A. (2006). *Commuting in America III: The Third National Report on Commuting Patterns and Trends*. Transportation Research Board.

Puga, D. (2002). European regional policies in light of recent location theories. *Journal of Economic Geography 2*(4), 373–406.

Rauch, J. E. (1993). Productivity gains from geographic concentration of human capital: Evidence from the cities. *Journal of Urban Economics 34*(3), 380–400.

Razin, A. and E. Sadka (1991). International tax competition and gains from tax harmonization. *Economics Letters 37*(1), 69–76.

Romer, T. (1975). Individual welfare, majority voting, and the properties of a linear income tax. *Journal of Public Economics 4*(2), 163–185.

Ross, S. and J. Yinger (1999). Sorting and voting: A review of the literature on urban public finance. In P. Cheshire and E. S. Mills (Eds.), *Handbook of Regional and Urban Economics: Applied Urban Economics*, Volume 3, Chapter 47, pp. 2001–2060. Elsevier.

Shapiro, J. M. (2006). Smart cities: Quality of life, productivity, and the growth effects of human capital. *Review of Economics and Statistics 88*(2), 324–335.

Sheshinski, E. (1967). Tests of the "learning by doing" hypothesis. *Review of Economics and Statistics 49*(4), 568–578.

Stengel, B. (2010). Computation of Nash equilibria in finite games: Introduction to the symposium. *Economic Theory 42*(1), 1–7.

Tabuchi, T. (1998). Urban agglomeration and dispersion: A synthesis of Alonso and Krugman. *Journal of Urban Economics 44*(3), 333–351.

Throsby, D. (1994). The production and consumption of the arts: A view of cultural economics. *Journal of Economic Literature 32*(1), 1–29.

Turnbull, G. K. and S. S. Djoundourian (1994). The median voter hypothesis: Evidence from general purpose local governments. *Public Choice 81*(3–4), 223–240.

Verhoef, E. T. (2005). Second-best congestion pricing schemes in the monocentric city. *Journal of Urban Economics 58*(3), 367–388.

Waldman, M. (2003). Durable goods theory for real world markets. *Journal of Economic Perspectives 17*(1), 131–154.

Wang, Y.-Q. (1999). Commodity taxes under fiscal competition: Stackelberg equilibrium and optimality. *American Economic Review 89*(4), 974–981.

Wilson, J. D. (1986). A theory of interregional tax competition. *Journal of Urban Economics 19*(3), 296–315.

Wilson, J. D. (1991). Tax competition with interregional differences in factor endowments. *Regional Science and Urban Economics 21*(3), 423–451.

Wilson, J. D. and D. E. Wildasin (2004). Capital tax competition: bane or boon. *Journal of Public Economics 88*(6), 1065–1091.

World Bank (2011). *World Development Indicators (Edition: 2011)*. World Bank.

Zenou, Y. (2009). *Urban Labor Economics*. Cambridge University Press.

Zodrow, G. R. and P. Mieszkowski (1986). Pigou, Tiebout, property taxation, and the underprovision of local public goods. *Journal of Urban Economics 19*(3), 356–370.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

522 B.KAYNAR, Rare Event Simulation Techniques for Stochastic Design Problems in Markovian Setting
523 P. JANUS, Developments in Measuring and Modeling Financial Volatility
524 F.P.W. SCHILDER, Essays on the Economics of Housing Subsidies
525 S. M MOGHAYER, Bifurcations of Indifference Points in Discrete Time Optimal Control Problems
526 C. ÇAKMAKLI, Exploiting Common Features in Macroeconomic and Financial Data
527 J. LINDE, Experimenting with new combinations of old ideas
528 D. MASSARO, Bounded rationality and heterogeneous expectations in macroeconomics
529 J. GILLET, Groups in Economics
530 R. LEGERSTEE, Evaluating Econometric Models and Expert Intuition
531 M.R.C. BERSEM, Essays on the Political Economy of Finance
532 T. WILLEMS, Essays on Optimal Experimentation
533 Z. GAO, Essays on Empirical Likelihood in Economics
534 J. SWART, Natural Resources and the Environment: Implications for Economic Development and International Relations
535 A. KOTHIYAL, Subjective Probability and Ambiguity
536 B. VOOGT, Essays on Consumer Search and Dynamic Committees
537 T. DE HAAN, Strategic Communication: Theory and Experiment
538 T. BUSER, Essays in Behavioural Economics
539 J.A. ROSERO MONCAYO, On the importance of families and public policies for child development outcomes
540 E. ERDOGAN CIFTCI, Health Perceptions and Labor Force Participation of Older Workers

541 T.WANG, Essays on Empirical Market Microstructure
542 T. BAO, Experiments on Heterogeneous Expectations and Switching Behavior
543 S.D. LANSDORP, On Risks and Opportunities in Financial Markets
544 N. MOES, Cooperative decision making in river water allocation problems
545 P. STAKENAS, Fractional integration and cointegration in financial time series
546 M. SCHARTH, Essays on Monte Carlo Methods for State Space Models
547 J. ZENHORST, Macroeconomic Perspectives on the Equity Premium Puzzle
548 B. PELLOUX, the Role of Emotions and Social Ties in Public On Good Games: Behavioral and Neuroeconomic Studies
549 N. YANG, Markov-Perfect Industry Dynamics: Theory, Computation, and Applications
550 R.R. VAN VELDHUIZEN, Essays in Experimental Economics
551 X. ZHANG, Modeling Time Variation in Systemic Risk
552 H.R.A. KOSTER, The internal structure of cities: the economics of agglomeration, amenities and accessibility.
553 S.P.T. GROOT, Agglomeration, globalization and regional labor markets: micro evidence for the Netherlands.
554 J.L. MÖHLMANN, Globalization and Productivity Micro-Evidence on Heterogeneous Firms, Workers and Products
555 S.M. HOOGENDOORN, Diversity and Team Performance: A Series of Field Experiments
556 C.L. BEHRENS, Product differentiation in aviation passenger markets: The impact of demand heterogeneity on competition
557 G. SMRKOLJ, Dynamic Models of Research and Development
558 S. PEER, The economics of trip scheduling, travel time variability and traffic information
559 V. SPINU, Nonadditive Beliefs: From Measurement to Extensions
560 S.P. KASTORYANO, Essays in Applied Dynamic Microeconometrics
561 M. VAN DUIJN, Location, choice, cultural heritage and house prices

562 T. SALIMANS, Essays in Likelihood-Based Computational Econometrics
563 P. SUN, Tail Risk of Equidity Returns
564 C.G.J. KARSTEN, The Law and Finance of M&A Contracts
565 C. OZGEN, Impacts of Immigration and Cultural Diversity on Innovation and Economic Growth
566 R.S. SCHOLTE, The interplay between early-life conditions, major events and health later in life
567 B.N.KRAMER, Why don't they take a card? Essays on the demand for micro health insurance
568 M. KILIÇ, Fundamental Insights in Power Futures Prices
569 A.G.B. DE VRIES, Venture Capital: Relations with the Economy and Intellectual Property
570 E.M.F. VAN DEN BROEK, Keeping up Appearances
571 K.T. MOORE, A Tale of Risk: Essays on Financial Extremes
572 F.T. ZOUTMAN, A Symphony of Redistributive Instruments