City Growth: The Role of Knowledge Spillovers and Entrepreneurship

Dissertation zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft

eingereicht an der Fakultät für Wirtschaftswissenschaften der Universität Regensburg

vorgelegt von:

Dirk Assmann

Berichterstatter:

Prof. Gabriel S. Lee, Ph.D., Universität Regensburg

Prof. Dr. Kristof Dascher, Universität Regensburg

Tag der Disputation:

25. Februar 2016

Abstract

City Growth: The Role of Knowledge Spillovers and Entrepreneurship

by Dirk Assmann

My thesis highlights potential drivers of the observed extent and heterogeneity of urban growth and further investigates the role of knowledge spillovers and entrepreneurship. My research contributions are as follows: First, I set up a simplified version of the Monocentric City Model to discuss how potential key drivers affect subsequent urban growth. I summarize previous empirical findings to examine the model's theoretical predictions. Second, I present a search-theoretic model of urban face-to-face interactions in which I further investigate the role of local knowledge spillovers for urban growth. It is the first theoretical urban model that explicitly distinguishes between two types of knowledge spillovers, the transmission and creation of knowledge. The model analysis suggests that an inefficient knowledge exchange in urban face-to-face interactions causes inefficient city sizes as agglomeration forces do not reach their optimal extent. Third, I investigate the causal relation between local entrepreneurship and subsequent growth. I argue that the perceived positive connection is partially due to reverse causality as entrepreneurs self-select into fast growing cities. I propose the 19th century patent activity across the US as an instrument for the distribution of entrepreneurial activity observed in the US today. It turns out that the entrepreneurship's influence on subsequent growth diminishes substantially once the issue of endogeneity is taken into account.

Acknowledgements

First of all I thank my beloved wife Franziska for her personal support and her extremely helpful comments on my thesis. I thank my long-standing friend Johannes Stiller for his contributions to chapter 3. I am very grateful for the encouragement and advices of my first supervisor Gabriel Lee during the last seven years. I also thank Kristof Dascher for the spontaneous take-over of the job as second supervisor for my thesis and Lutz Arnold for his helpful comments on chapter 3. Finally, I want to thank my parents for supporting me throughout my entire education.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis examines the key drivers of urban growth. I summarize previous theoretical and empirical findings on the importance of a set of potential city growth determinants. I extend these findings by investigating agglomeration economies due to knowledge spillovers and the impact of the city's entrepreneurial activity regarding subsequent city growth. The two main findings of my investigation are as follows: First, my work on urban knowledge spillovers in a search-theoretic environment suggests that agglomeration forces do not reach their optimal extent as city residents perform face-to-face interactions with an inefficient range of people. Second, my empirical analysis on the connection between entrepreneurship and subsequent urban growth indicates that the entrepreneurship's influence diminishes substantially once reverse causality is taken into account. My empirical strategy is to instrument modern rates of local entrepreneurship across US cities by the 19th century US patent activity.

The development of city population over the last century highlights why the topic of city growth is worth to be further investigated. While the share of US population living in metropolitan areas was only 28 percent in 1910, the share increased to 56 percent in 1950 and reached 81 percent in 2014. In 1950 only 2 US metropolitan areas had a population of 5 million residents and more (New York City, NY and Chicago, IL). Their combined population of 18.4 million accounted for 12.2 percent of the entire US population. In 2000 the US had 9 metropolitan areas with more than 5 million residents accommodating almost 30 percent of the US population. The increase in urban population was observable across the majority of US states. While in 1950 only 15 US states exhibited a population that was predominantly metropolitan, the number of states increased to 37 in 2000. US Census Data also indicate that population patterns have changed within cities. The share of the metropolitan population living within the central city was 75 percent in 1910. In 1950 the share decreased to 59 percent and further decreased to 38 percent in 2000. Thus beside substantial city growth one can observe a flight of metropolitan population to the suburbs.[1] The European Union exhibits city growth rates comparable to the US. While the urbanization rate within the EU was roughly 61 percent in 1960, it increased to 75 percent in 2014.[2] Also the share of population living in urban agglomerations of more than 1 million residents increased across the EU from 16 percent in 1960 to 18 percent in 2014. China had an urbanization rate of 16 percent in 1960, which increased to 54 percent in 2014. The percentage of China's population in urban agglomerations of more than 1 million residents rose from 8 percent in 1960 to 23 percent in 2014.[3] While the growing importance of cities is apparent from the data, it is unclear which cities benefit most and which cities are not able to take advantage of rising urbanization rates. Since 1950 the cities of St. Louis, MO and Detroit, MI lost over 60 percent of their population, while during the same time Houston, TX and San Jose, CA exhibited growth rates of almost 300 and 900 percent, respectively.

My thesis explains part of the observed extent and heterogeneity of city growth. Chapter 2 summarizes potential city growth determinants as presented in the literature. I first introduce the Monocentric City Model to discuss theoretical implications in a unified framework. The drivers I focus on are the city's provision of transportation infrastructure, the city's endowment with amenities, the city's extent of agglomeration economies and the city's supply of entrepreneurial activity. Chapter 3 presents theoretical work, which examines the role of local knowledge spillovers for the extent of agglomeration economies. The transmission and creation of knowledge during urban face-to-face interactions crucially depends on the meeting partners' diversity of knowledge types. The choice of an inefficient range of people to interact with results in inefficient city sizes as agglomeration forces do not reach their optimal extent. Chapter 4 contains empirical work on the relation between local entrepreneurship and subsequent urban growth. Contrary to previous results, I find that the impact of local entrepreneurship becomes insignificant and substantially diminished as soon as reverse causality is taken into account by Instrumental Variable regressions. I use the distribution of 19th century patents across the US as exogenous source of variation for modern rates of local entrepreneurship in US cities. Chapter 5 summarizes the main findings of the thesis and concludes.

[1] All information on the US are derived from the Census Special Report "Demographic Trends in the 20th Century" from 2002. Information on the US urbanization rate in 2014 are taken from the World Development Indicators on databank.worldbank.org.

[2] E.g. Germany: Urbanization rate of 71 percent (in 1960) increased to 75 percent (in 2014), France: 62 to 79, Italy: 59 to 69, Spain: 57 to 79, United Kingdom: 78 to 82.

[3] All information on the European Union and China are drawn from the World Development Indicators on databank.worldbank.org.

Chapter 2

Determinants of City Growth

2.1 Introduction

In this chapter, I discuss a set of city growth determinants which are able to explain the observed extent and heterogeneity of growth rates across cities.[1] The presented drivers are defined by Duranton and Puga (2014) and further examined in this chapter. In section 2.2, I introduce a simplified version of the Monocentric City Model. I apply the model to derive theoretical implications for each of the presented growth determinants. Those determinants encompass the city's quality of transportation infrastructure (section 2.3), the city's endowment with amenities (section 2.4), the city's extent of agglomeration economies (section 2.5) and the city's supply of entrepreneurial activity (section 2.6). After establishing the model's theoretical implications concerning each of the discussed determinants, I present empirical evidence on the theoretical implications. In section 2.7, I summarize further potential explanations for city growth prevalent in the literature. Section 2.8 outlines the main findings of this chapter.

2.2 The Monocentric City Model

The origin of the Monocentric City Model goes back to the work of Alonso (1964), Mills (1967) and Muth (1969). I present a simplified version which is closely related to the work of Brueckner (1987) and Duranton and Puga (2015). The model is capable of making theoretical predictions about different drivers' impact on urban growth. These predictions can then be tested empirically.

[1]See chapter 1 for illustrative data on city growth in the US, Europe and China.

2.2.1 The Household Sector

Consider a positive real line. A segment $[0, \bar{u}]$ of that real line is covered by a monocentric city, with the city boundary $\bar{u}$ being endogenously determined by the model. All jobs are exogenously located at the Central Business District (CBD) at point $u = 0$. The city is populated by N identical and freely mobile individuals earning a wage $w > 0$ for producing a composite good x. Preferences of an individual living at distance u from the CBD are represented by the strictly quasi-concave utility function

$$V(u) = V[x(u), H(u)], \tag{2.1}$$

with $\frac{\partial V}{\partial x} > 0$ and $\frac{\partial V}{\partial H} > 0$. $x(u)$ denotes consumption of the composite good, available at price $p = 1$ everywhere in the city, and $H(u)$ denotes the lot size of housing with $P(u)$ being the price per housing unit at distance u from the CBD. Individuals at distance $u \in [0, \bar{u}]$ incur transportation costs of tu (with $t > 0$) to go to work. An individual at location u faces the following maximization problem:

$$\max_{x(u), H(u) \geq 0} \; V[x(u), H(u)] \quad \text{s.t.} \qquad x(u) + P(u)H(u) \leq w - tu, \tag{2.2}$$

with expenditures on the composite good $x(u)$ and on housing $P(u)H(u)$ on the left-hand side of the budget constraint and income w net of transportation costs tu on the right-hand side. Let $v[P(u), w - tu]$ be the indirect utility function for an individual at location u with $\frac{\partial v}{\partial P} < 0$ and $\frac{\partial v}{\partial (w-tu)} > 0$. $v(\cdot)$ is the maximum-value function to maximization problem (2.2), i.e. $v[P(u), w - tu] = V[x^*(u), H^*(u)]$ with $x^*(u)$ and $H^*(u)$ being the utility maximizing quantities of consumption and housing at u.
A residential equilibrium within the city requires that all city residents achieve the same utility level $\bar{V}$ in order to eliminate any relocation incentives, which can be stated as

$$v[P(u), w - tu] = \bar{V} \qquad \forall \; u \in [0, \bar{u}]. \tag{2.3}$$

Total differentiation of identity (2.3) with respect to u and applying Roy's Identity[2] yields the condition for a residential equilibrium within the city[3]:

$$P'(u) = -\frac{t}{H^*(u)} < 0. \tag{2.4}$$

Gradient (2.4) implies that the price of housing $P(u)$ decreases in distance u to the CBD. This is due to the fact that utility equalization within the city requires individuals to be compensated for higher transportation costs at larger distances to the CBD. As housing

[2] Roy's Identity says that $-\frac{\partial v[P(u), w-tu]/\partial P(u)}{\partial v[P(u), w-tu]/\partial (w-tu)} = H^*(u)$.
[3] Duranton and Puga (2015) denote (2.4) as Alonso-Muth condition.

becomes cheaper, individuals can substitute into more housing consumption in order to keep utility constant at $\bar{V}$.[4] Thus housing consumption increases with distance u to the CBD:

$$H'(u) > 0. \tag{2.5}$$

Gradient (2.4) together with gradient (2.5) imply that housing prices $P(u)$ decrease convexly in distance u to the CBD.

2.2.2 The Construction Sector

The city hosts a perfectly competitive construction industry with firms using a constant returns to scale technology to produce $F(u)$ units of housing floorspace at distance u from the CBD. The inputs to production are land $L(u)$ and capital $K(u)$. The production process at location u is represented by

$$F(u) = F[K(u), L(u)], \tag{2.6}$$

with $F(\cdot)$ being strictly quasi-concave and $\frac{\partial F}{\partial L} > 0$ and $\frac{\partial F}{\partial K} > 0$. The rental price of capital r is constant and exogenous, the rental price of land $R(u)$ is endogenous and depends on location u within the city. As the rental price of capital r is exogenously given, it can be omitted as an argument in the floorspace's unit cost function. Constructing one unit of floorspace at u bears costs of $c = c[R(u)]$ with $\frac{\partial c}{\partial R} > 0$. As the construction industry is perfectly competitive, each firm makes zero profits independent of location u, implying that

$$P(u) = c[R(u)] \quad \forall\ u \in [0, \bar{u}], \tag{2.7}$$

with $P(u)$ being the revenue and $c[R(u)]$ being the cost per unit of housing. Total differentiation of identity (2.7) with respect to u and applying Shephard's Lemma[5] yields the locational equilibrium condition for firms in the construction industry:

$$R'(u) = P'(u)F^*(u) < 0. \tag{2.8}$$

Gradient (2.4) is used to infer that the rental price of land is decreasing in distance u to the CBD with $F^*(u)$ being the cost-minimizing housing floorspace per unit of land at u. The decrease in land rents is due to the fact that firms have to be compensated for lower revenues at larger distances to the CBD in order to keep profits at zero. As land becomes cheaper, firms use more of it in the production process. This implies

[4] Formally, this can be shown by differentiating the Hicksian demand for housing $H^h[P(u), \bar{V}]$ with respect to u.

[5] Shephard's Lemma says that $\frac{\partial c[R(u)]}{\partial R(u)} = L^*(u) = \frac{1}{F^*(u)}$, with $\frac{1}{F^*(u)}$ being the cost-minimizing input of land at u per unit of floorspace.

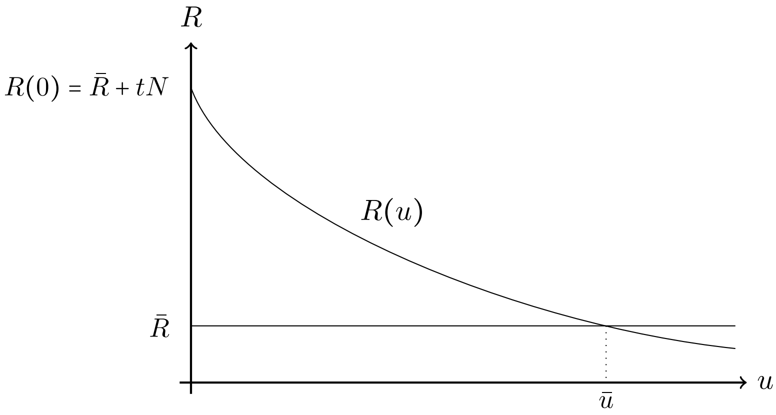


FIGURE 2.1: Land rent gradient in the Monocentric City Model

that housing floorspace per unit of land is decreasing as firms distribute the constructed floorspace over a larger amount of land.[6] The gradient for housing floorspace is thus

$$F'(u) < 0. \tag{2.9}$$

Land on the entire real line is suitable for the alternative of agricultural use at rental price $\bar{R}$. It is only developed for residential use by firms as long as $R(u) > \bar{R}$, implying that the city boundary $\bar{u}$ is determined by the condition

$$R(\bar{u}) = \bar{R}, \tag{2.10}$$

which is also illustrated in figure 2.1.

2.2.3 Urban Equilibrium

Population density $n(u)$ at distance u from the CBD is described by

$$n(u) = \frac{F^*(u)}{H^*(u)} = -\frac{R'(u)}{t}, \tag{2.11}$$

for which the second equality is inferred from (2.4) and (2.8). Equation (2.11) suggests that more housing floorspace and smaller lot sizes lead to an increase in population density at location u.[7] The gradients (2.5) and (2.9) immediately imply that

$$n'(u) < 0, \tag{2.12}$$

[6] Duranton and Puga (2014) use the observation of larger gardens at the city boundary for illustration.
[7] A skyscraper with small apartments obviously offers the highest density.

saying that population density is decreasing in distance u to the CBD.
In equilibrium the city has to offer enough room to accommodate all N residents, which translates into

$$\int_0^{\bar{u}} n(u)\mathrm{d}u = N. \tag{2.13}$$

Substituting $n(u)$ from (2.11) into equation (2.13) and applying condition (2.10) yields

$$\int_0^{\bar{u}} -\frac{R'(u)}{t}\mathrm{d}u = \frac{R(0) - \bar{R}}{t} = N. \tag{2.14}$$

Equation (2.14) pins down the rental price of land at the CBD ($u = 0$), given by

$$R(0) = \bar{R} + tN, \tag{2.15}$$

which is illustrated by the intercept in figure 2.1. Equation (2.15) in turn determines the price of housing at the CBD, as the zero-profit condition for firms has to be satisfied for every $u \in [0, \bar{u}]$. Applying identity (2.7) yields

$$P(0) = c[R(0)] = c(\bar{R} + tN). \tag{2.16}$$

Equation (2.16) can be used to evaluate the city residents' level of utility. A residential equilibrium within the city implies that all city residents must achieve the same utility level, as otherwise relocation incentives are prevalent. Thus the utility of an individual at location $u = 0$ has to equal the utility level of individuals at any other location $u \in [0, \bar{u}]$, i.e. $V(u) = V(0) \;\; \forall \;\; u \in [0, \bar{u}]$, implying that

$$V(u) = v[P(0), w] = v[c(\bar{R} + tN), w] = \bar{V} \qquad \forall \;\; u \in [0, \bar{u}]. \tag{2.17}$$

Identity (2.17) delivers the model's crucial insight for the purpose of my thesis, as (2.17) can be solved for city population N as a function of the model's parameters $t, w, \bar{V}$ and $\bar{R}$[8]:

$$N = N(t, w, \bar{V}, \bar{R}). \tag{2.18}$$

According to the Monocentric City Model, any change to one of the model's parameters ($t, w, \bar{V}$ and $\bar{R}$) induces urban growth or urban decline.

[8]For the purposes of my thesis an open-city urban equilibrium is examined, in which utility level $\bar{V}$ is exogenous and city population N is endogenous. In a closed-city urban equilibrium, city population N is exogenous with utility $\bar{V}$ being determined by the model.

2.2.4 System of Cities

So far, the model only permits to investigate how parameter changes affect the population of one particular city. In order to explain differences across cities, a system of more than one city has to be introduced.
Suppose there exists an urban system of $i = 1, ..., C$ cities. Changes to the model's parameters $t_i, w_i, \bar{V_i}$ in one particular city i do not affect the parameters of other cities within the urban system. Agricultural land rent $\bar{R}$ is common to all cities.[9]
A residential equilibrium across cities requires that utility levels are equalized across all C cities, as otherwise, people have an incentive to move between them[10], implying that

$$\bar{V_i} = \bar{V} \quad \forall \ i = 1, ..., C. \tag{2.19}$$

Together with identity (2.17), the condition for a residential equilibrium across and within cities can be stated as

$$V_i(u) = v[c(\bar{R} + t_i N_i), w_i] = \bar{V} \quad \forall \ u_i \in [0, \bar{u}_i], \ \ \forall \ i = 1, ..., C. \tag{2.20}$$

Identity (2.20) can be solved for city i's population N_i as a function of the city-specific parameters t_i and w_i and the parameters $\bar{V}$ and $\bar{R}$ common to all cities, implying that

$$N_i = N(t_i, w_i, \bar{V}, \bar{R}). \tag{2.21}$$

Equation (2.21) permits to evaluate why cities differ in size.

2.3 Transportation Infrastructure

Improvements to the city's transportation infrastructure can contribute to urban growth. If commuting within a particular city becomes cheaper, effects are twofold. First, people have an incentive to migrate to that city. And second, residents within the city have an incentive to move away from the center as commuting to the workplace becomes less expensive. In this section, I present the theoretical implications of infrastructure improvements implied by the Monocentric City Model and review the empirical literature on the connection between transportation and urban growth.

[9] Land on the entire real line is equally productive for agricultural use.
[10] Note that a residential equilibrium across cities is not the same as a residential equilibrium within the city.

2.3.1 Implications of the Monocentric City Model

Assume that cities within the urban system only differ in the quality of their provided transportation infrastructure with t_i representing transportation costs in city i. A higher quality of provided transportation infrastructure is represented by lower transportation costs as commuting within the city is facilitated.[11] Referring to identity (2.20), a residential equilibrium requires that utility levels are equalized within and across all cities of the urban system, implying that

$$V_i(u) = v[c(\bar{R} + t_i N_i), w] = \bar{V} \quad \forall\ u_i \in [0, \bar{u}_i], \ \ \forall\ i = 1, ..., C. \tag{2.22}$$

Identity (2.22) implicates that

$$\frac{\mathrm{d}N_i}{\mathrm{d}t_i} < 0, \tag{2.23}$$

as otherwise a residential equilibrium cannot be achieved. Thus cities offering a superior transportation infrastructure exhibit a larger population.

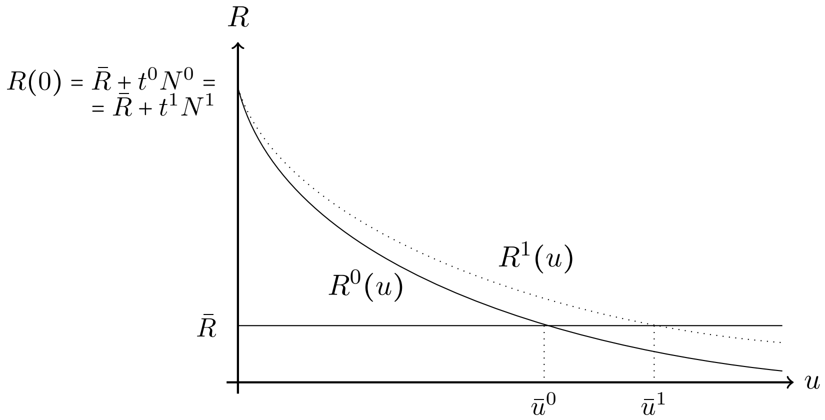


FIGURE 2.2: Impact on land rents due to improvements to the city's transportation infrastructure

Figure 2.2 illustrates how improvements to the transportation infrastructure affect the city's land prices.[12] t^0 and N^0 denote transportation costs and population before the improvement, t^1 and N^1 transportation costs and population afterwards, with $t^0 > t^1$. The change in transportation costs affects all individuals in the city, except those living at the CBD, as they do not commute to work and thus bear no transportation costs. Since the city's utility level $\bar{V}$ is exogenous, land rents at the CBD, denoted by $R(0)$,

[11]The improvement can result in reductions in costs or commuting time.

[12]Local infrastructure can be improved by the provision of additional highways or an improved public transportation system among other actions.

stay unaffected. Knowing that

$$\bar{R} + t^0 N^0 = \bar{R} + t^1 N^1, \tag{2.24}$$

one can infer that $N^1 > N^0$.[13] The solid curve in figure 2.2 depicts land rents before the change in t, the dotted curve depicts the relation between distance and land rents afterwards. As all city residents, beside those at the CBD, benefit from the reduction in transportation costs, there has to exist a mechanism restoring a residential equilibrium. This is guaranteed by the immigration of residents from other cities, which drives up the rental price of land, i.e. $R^1(u) > R^0(u) \; \forall \; u \in (0, \bar{u}^1]$. The resulting increase in house prices, implied by identity (2.17), compensates for the savings in transportation expenditures. Figure 2.2 also shows that the city boundary shifts rightwards ($\bar{u}^1 > \bar{u}^0$), i.e. the city becomes larger in spatial dimension. This implication can be interpreted as suburbanization, as the city's fraction of people living in the city center is decreasing, whereas the fraction living close to the boundary is increasing.[14] The shift is simply due to the fact that a larger population requires more room to be accommodated by the city. As noted by Duranton and Puga (2015), the increase in population has two effects: First, it shifts the city boundary $\bar{u}$ rightwards. Second, it increases density n throughout the city, as residents choose smaller apartment. The increase in density can never be large enough to induce a leftwards shift of the city boundary.

In summary, the Monocentric City Model implies that cities offering a well-functioning transportation infrastructure exceed cities with a poor infrastructure in population, density and spatial extent. Any improvements to a city's transportation infrastructure lead to subsequent growth in population accompanied by suburbanization.

2.3.2 Empirical Findings

While the impact of the city's transportation infrastructure on urban growth is of main interest for my thesis, the empirical literature addressing it is relatively scarce. According to Duranton and Turner (2012) this scarcity is due to "...a lack of clear predictions regarding the effect of the level of transportation infrastructure on subsequent city growth, and the difficulty of dealing with the simultaneous determination of employment or population growth and transportation infrastructure in cities".

The majority of empirical papers on the topic focus on the infrastructure's impact on suburbanization. Baum-Snow (2007a) suggests the following regression model to estimate the impact of changes to the city's transportation infrastructure on the city's

[13] More precisely, $N^1 = \frac{t^0}{t^1} N^0$.

[14] See Baum-Snow (2007b) for an extensive discussion.

center population:

$$\Delta_{t,t+1}\log(N_{i,t}^{center}) = \beta_0 + \beta_1\Delta_{t,t+1}T_{i,t} + \delta_{i,t}'\beta_2 + u_{i,t}. \tag{2.25}$$

i indexes for cities, t indexes for time. $N_{i,t}^{center}$ denotes city i's population living within the city center at time period t. $\Delta_{t,t+1}$ represents the change in variables between periods t and $t+1$. $T_{i,t}$ describes the infrastructure's quality provided by city i, which the literature predominantly approximates by the number of highways or railroads that proceed within the city. $\delta_{i,t}$ is a vector of relevant control variables that accounts for various city characteristics. $u_{i,t}$ is the model's error term. The coefficient of interest in regression model (2.25) is β_1, which indicates how changes to the infrastructure's quality affect subsequent suburbanization. Estimates for β_1 cannot be interpreted causally as regression model (2.25) exhibits reverse causality. Transportation infrastructure is not randomly assigned to cities. It heavily depends on the model's outcome variable $N_{i,t}^{center}$, as local governments react to inner-city population movements. Thus any empirical analysis intending to measure the causal impact of the city's infrastructure needs to find the infrastructure's exogenous source of variation, i.e. an instrument that predetermines changes in the infrastructure's quality, but has no causal effect on subsequent suburbanization.

Burchfield et al. (2006) examine differences in urban sprawl across US metropolitan areas. They use satellite data to track the evolution of land use across the US on 8.7 billion 30 x 30 meter grid cells.[15] Although the paper's focus is not on the infrastructure's impact on urban sprawl, they find that those cities promoting the commute by car are significantly more scattered.[16] Interpreting "car-friendliness" as higher quality of the city's infrastructure, the authors find that any improvements to it induce city growth in spatial dimension as city centers become less densely populated. Baum-Snow (2007a) examines whether the provision of highways in US cities promotes the leaving of residents from the center to the suburbs, as is predicted by the Monocentric City Model. To account for potential endogeneity, Baum-Snow (2007a) uses the number of highways in a 1947 national interstate highway plan as exogenous source of variation for the actual number of highways across US cities. IV estimates indicate that the allocation of one additional highway reduces the central city's population by 18 percent. He claims that the provision of new highways represents about one third of the overall decline in central city population between the years 1950 and 1990. Similar results are presented by Garcia-López (2012), who investigates how local improvements to the transportation infrastructure have affected the suburbanization process in Barcelona, Spain. Contrary

[15]The data set allows them to determine which grid cells of land in the US are developed and which are not.

[16]I.e. those cities have a higher percentage of undeveloped land.

to Baum-Snow (2007a), the paper concentrates on the infrastructure's impact on population patterns across districts within one particular city. Garcia-López (2012) tackles the inherent endogeneity by using Barcelona's 19th century railroad and main road system to instrument for the city's modern infrastructure provision. Results indicate that the local provision of additional highways and railroads significantly fostered Barcelona's suburbanization process. Garcia-López et al. (2013) replicate the paper by Baum-Snow (2007a) for Spanish cities. They apply Spain's historical road networks to instrument for the provision of highways and railroads in Spanish cities observed today. In line with Baum-Snow (2007a), they find that improvements to the city's transportation infrastructure (measured by additional highways and railroads) cause decentralization of population within cities. Baum-Snow et al. (2012) remark that the city's transportation infrastructure does not only affect the distribution of residents within the city but also the distribution of firms. Using data on Chinese counties and applying historic route maps from 1962 as exogenous source of variation for modern infrastructure provisions, they show that the city's infrastructure has massive influence on the decentralization of production. This decentralization has the potential to further extend the occurrence of urban sprawl in the future.[17]

The impact of the transportation infrastructure's quality on subsequent city growth in population (or employment) is far less intensely investigated in the literature. Hymel (2009) examines the impact of road congestion on subsequent employment growth in US cities. Road congestion can be interpreted as poor quality of a city's provided transportation infrastructure. Using data on the 85 largest US metropolitan areas between 1983 and 2002, he finds that a 10 percent increase in congestion levels reduces subsequent employment growth by 4 percent. Similar to Baum-Snow (2007a), Hymel (2009) uses the city's count of highways in a 1947 national interstate highway plan to instrument for modern congestion levels across US cities. Most prominently, Duranton and Turner (2012) estimate the impact of the city's initial provision of highways $T_{i,t}$ on the city's subsequent log employment growth $\Delta_{t,t+1}\log(N_{i,t})$ between the years 1983 and 2002 by estimating the regression model

$$\Delta_{t,t+1}\log(N_{i,t}) = \beta_0 + \beta_1 T_{i,t} + \delta'_{i,t}\beta_2 + u_{i,t}. \tag{2.26}$$

They use the 1947 plan of the interstate highway system, a railroad map from 1898 and a map showing major expeditions of exploration between 1835 and 1850 as exogenous source of variation for US highway provision in 1983. IV estimates suggest that a 10 percent increase in the city's initial stock of highways is associated with a 1.5 percent increase in employment over the next twenty years.

[17] Note that the empirical finding of Baum-Snow et al. (2012) is not an implication of the Monocentric City Model, which relies on the assumption that jobs are exogenously located at the city center.

In summary, all theoretical implications of the Monocentric City Model are confirmed by the empirical literature. Provision of superior transportation infrastructure causes city growth in spatial dimension accompanied by the process of suburbanization. Furthermore, improvements to the city's transportation infrastructure promote subsequent population (or employment) growth within that city.

2.4 Supply of Amenities

Roback (1982) is the first to point out that there have to exist substantial differences in the quality of life across cities in order to explain the observed wage and rent differentials. As noted by Glaeser et al. (2001), the quality of life provided by cities will get even more important for people's location decision as long as income levels continue to rise. In this section, I survey the literature on how the city's endowment with amenities can be responsible for urban growth.

2.4.1 Implications of the Monocentric City Model

The standard Monocentric City Model presented in section 2.1 implies that each city within the urban system is endowed with the same level of amenities. The individuals' utility only depends on consumption levels of the composite good $x(u)$ and housing $H(u)$, but not on the city's characteristics. In this section, I relax the assumption of equal amenity provision across cities within the urban system. Each city $i = 1, ..., C$ has a specific quality of life A_i, which is exogenously given.[18]
Utility function $V_i(\cdot)$ of an individual located at city i now additionally depends on the city's amenity provision A_i:

$$V_i(u) = V[A_i, x(u), H(u)], \tag{2.27}$$

Let $U[x(u), H(u)]$ be the sub-utility from consuming $x(u)$ units of the composite good and $H(u)$ units of housing, with $v[P(u), w - tu]$ being the corresponding indirect utility function. The utility representation of equation (2.27) can then be rewritten as

$$V_i(u) = V\left\{A_i, U[x(u), H(u)]\right\} = V\left\{A_i, v[P(u), w - tu]\right\}, \tag{2.28}$$

[18]The model is silent about the drivers of the cities' amenity provision. Some cities might be located in an appealing surrounding, others might have an extensive supply of leisure activities.

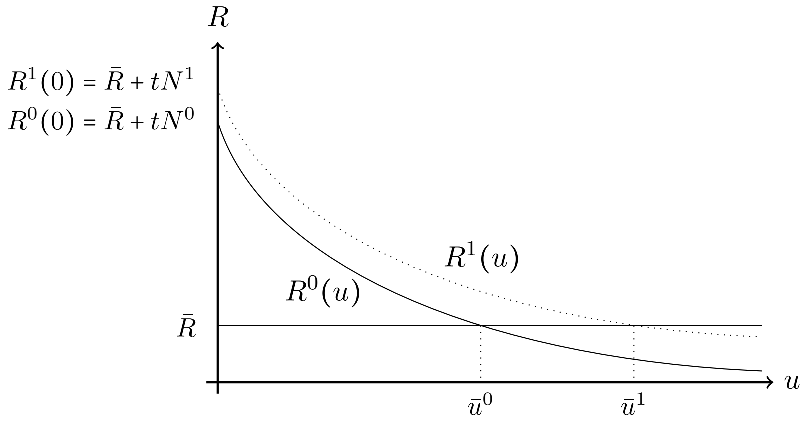


FIGURE 2.3: Impact on land rents due to an increase in the city's amenity supply, an increase in the city's wage level or an inflow of entrepreneurs

with $\frac{\partial V}{\partial A_i} > 0$ and $\frac{\partial V}{\partial v} > 0$. The residential equilibrium condition still requires that utility levels are equalized across and within all C cities within the urban system, independent of the cities' provision of amenities. Together with identity (2.17), this implies that

$$V_i(u) = V\left\{A_i, v[c(\bar{R} + tN_i), w]\right\} = \bar{V} \quad \forall\ u_i \in [0, \bar{u}_i],\ \ \forall\ i = 1, ..., C. \tag{2.29}$$

Totally differentiating identity (2.29) with respect to A_i and N_i and noting that $\frac{\partial v}{\partial N_i} < 0$[19] yields

$$\frac{\mathrm{d}N_i}{\mathrm{d}A_i} = -\frac{\frac{\partial V}{\partial A_i}}{\frac{\partial V}{\partial v}\frac{\partial v}{\partial N_i}} > 0. \tag{2.30}$$

The key message of gradient (2.30) is the positive dependence of city i's population on its amenity endowment, implying that more attractive places are larger than those with a low provision of local amenities.

Figure 2.3 illustrates how an exogenous change to the city's endowment with amenities affects the city's land rent gradient. A^0 denotes the amenity level before the change, A^1 the local amenity level afterwards, with $A^1 > A^0$. Identity (2.28) assures that

$$V\left\{A^0, v[c(\bar{R} + tN^0), w]\right\} = V\left\{A^1, v[c(\bar{R} + tN^1), w]\right\}, \tag{2.31}$$

which implies that $N^1 > N^0$. The solid curve in figure 2.3 depicts land rents before the change in local amenities, the dotted curve depicts land prices after the change. As all residents benefit from the higher amenity provision, an inflow of population ensures that the conditions for a residential equilibrium are fulfilled.[20] This inflow drives up rental

[19]This is due to the fact that land rents and thus housing prices increase in the city's population N_i.

[20]Note that contrary to the case of decreasing transportation costs, individuals at the CBD also benefit from the provision of additional amenities.

prices of land throughout the entire city, i.e. $R^1(u) > R^0(u) \; \forall \; u \in [0, \bar{u}^1]$. The resulting increase in housing prices then counterbalances the individuals' benefits from the higher amenity level. Figure 2.3 also shows that the city boundary shifts rightwards ($\bar{u}^1 > \bar{u}^0$), i.e. the city becomes larger in spatial extent.

To sum up, the Monocentric City Model predicts that more attractive places are larger in population and spatial extent. Stated differently, increasing the provision of local amenities is followed by an inflow of population that causes city growth in spatial dimension.

2.4.2 Empirical Findings

The empirical findings I summarize in this section concentrate on the impact of the city's exogenous amenity supply[21] on the city's population and subsequent population growth. Duranton and Puga (2014) propose the following regression model to estimate the effect of the city's amenity endowment on the city's subsequent growth:

$$\Delta_{t,t+1} \log(N_{i,t}) = \beta_0 + \beta_1 A_{i,t} + \delta'_{i,t}\beta_2 + u_{i,t}. \tag{2.32}$$

$N_{i,t}$ denotes city i's population at time t. $A_{i,t}$ represents city i's amenity endowment. $\delta_{i,t}$ is a vector of relevant control variables. The parameter β_1 indicates how the city's log population growth is affected by the city's quality of life.[22]

One major difficulty in estimating regression model (2.32) is the adequate approximation of the city's amenity endowment. Glaeser et al. (2001) identify mainly four types of amenities, which are decisive for urban growth. First, a substantial supply of services and consumer goods like restaurants, theaters and museums as well as the opportunity of having social interactions with like-minded people. Second, a rich supply of architectual aesthetics, a beautiful landscape and pleasant climatical conditions. Third, well-functioning public services, which prevent urban crime and permit a promising formal education for children. And fourth, the provision of transportation infrastructure, facilitating fast commutes within the city.

Glaeser et al. (2001) examine how the presence of leisure amenities, measured by per capita restaurants, per capita live performance venues, per capita art museums, per capita bowling alleys and per capita movie theaters affect urban growth. They show that only the impact of restaurants and live performance venues is statistically significant. While the positive relationship between the city's stock of human capital and

[21] E.g. the region's climatical conditions or the city's surrounding landscape.

[22] As the focus of this section is on the impact of exogenous amenities, estimates on β_1 should not be biased due to endogeneity.

subsequent growth is mostly attributed to the productivity enhancing influence of human capital (see Lucas (1988) and Moretti (2004) among others), Glaeser et al. (2001) reinterpret the relation as being amenity-driven. They argue that people value living in cities with high human capital endowment, since those places have on average less social problems, offer superior formal education and induce more pleasant social interactions. Costa and Kahn (2000) interpret cities as a natural marriage market, which can be understood as an amenity for young single-living people. The disproportionate supply of potential marriage partners causes an inflow of further young single people.

Brueckner et al. (1999) investigate how the affection for aesthetic architecture can reshape location patterns of residents within cities. Rappaport and Sachs (2003) show that coastal proximity has a significant effect on population density, implying that people migrate to places with an appealing landscape. Specific climatical conditions are found to be one of the strongest predictors of urban growth. Cragg and Kahn (1999) and Costa and Kahn (2003) demonstrate how the influence of local weather conditions has emerged in the household's valuation of US regions since 1950. Rappaport (2007) identifies the county's mean January temperature to be the "climate" variable with the most substantial impact on population growth. Increasing mean January temperature by one standard-deviation causes an increase of 0.6 standard-deviations in population growth at the US county level during the period from 1970 to 2000. He also illustrates the significant growth enhancing impact of the county's mean July temperature. Increasing July temperature by one standard-deviation decreases county population growth between 1970 and 2000 by 0.2 standard-deviations. Thus people move to places with mild winters and cool summers. These results are confirmed by Glaeser et al. (2001) and Cheshire and Magrini (2006) for US and European cities.

Cullen and Levitt (1999) find that reducing the city's crime rate[23] by increasing the police occurrence is associated with subsequent urban population growth, whereas each additional crime reported to the police is associated with almost an one-person decline in population. To address the issue of endogeneity they apply the severity of the state criminal justice system to instrument for crime rates across cities. Contrary to Cullen and Levitt (1999), Ellen and O'Regan (2010) do not detect a statistically significant relation between crime and population growth during the 1990s.[24] Rappaport (1999) finds that population growth of US counties significantly depends on local expenditures on elementary and secondary school education, implying that the provision of promising school education has the potential to cause subsequent growth.

The impact of local transportation infrastructure on subsequent population growth is discussed in detail in section 2.3.2 and thus not further addressed here.

[23] A low crime rate can be interpreted as amenity.

[24] Ellen and O'Regan (2010) also apply the severity of the state criminal justice system as an exogenous source of variation for local crime rates.

An alternative approach in measuring the impact of a city's quality of life on urban growth is proposed by Carlino and Saiz (2008). They use the number of leisure visits to US cities as a proxy for a city's attractiveness. Regressing population growth between 1990 and 2000 on the number of leisure visits yields highly significant results, which imply that more attractive places are associated with higher population growth. Carlino and Saiz (2008) apply the city's coastal proximity and the city's number of historic places as exogenous source of variation for the number of leisure visits to eliminate concerns of endogeneity.
All empirical results confirm the theoretical implications of the Monocentric City Model. Cities endowed with more amenities are on average larger and grow faster than those cities with a poor amenity endowment.

2.5 Agglomeration Economies

Burchfield et al. (2006) point out that the entire US population is crowded within only 2 percent of its available land area. O'Flaherty (2005) remarks that Mexico City accomodates ten percent of Mexico's overall population by only covering 0.1 percent of its land area. There has to exist some form of increasing returns at the city-level (i.e. agglomeration economies) in order to explain the observed extent of people living in crowded urban environments.[25] In this section, I discuss theoretical implications of the Monocentric City Model concerning agglomeration economies, survey the theoretical literature on potential microfoundations and summarize previous empirical findings.

2.5.1 Implications of the Monocentric City Model

The prevalence of agglomeration economies raises people's productivity in the surrounding of other people. This productivity enhancement translates into rising wages. I now relax the assumption of exogenous wages and instead assume that wages w_i positively depend on city i's population N_i, i.e. $w_i = w_i(N_i)$ with $w_i'(N_i) > 0$ and $w_i''(N_i) < 0$.[26] Cities within the urban system exhibit differences in the ability to benefit from agglomeration economies, implying that the exact functional form $w_i(N_i)$ varies across cities. Suppose that the potential of city $i = A$ to exploit agglomeration economies is superior to the potential of city $i = B$. Mathematically, this advantage can be represented by $w_A'(N) > w_B'(N) \ \ \forall N \geq 0$.

[25] In the following, the expressions "increasing returns" and "agglomeration economies" are used interchangeably.

[26] The assumption $w_i''(N_i) < 0$ is introduced for technical reasons.

A residential equilibrium within and across cities requires that

$$V_i(u) = v[c(\bar{R} + tN_i), w_i(N_i)] = \bar{V} \quad \forall\ u_i \in [0, \bar{u}_i], \ \ \forall\ i = 1, ..., C. \tag{2.33}$$

Total differentiation of identity (2.33) with respect to population N_i yields

$$\frac{\partial v}{\partial c}\frac{\partial c}{\partial N_i} = \frac{\partial c}{\partial w_i} w_i'(N_i). \tag{2.34}$$

The left-hand side of equation (2.34) represents the marginal utility loss due to crowding. A growing population N_i drives up land rents and housing prices within the city. The associated decline in housing consumption reduces the residents' utility. The marginal utility gain from exploiting agglomeration economies is represented on the right-hand side of equation (2.34). An increase in population N_i makes residents more productive, which induces higher wages. As the additional income can be spent for consumption of the composite good and housing, the residents' utility level increases. Equation (2.34) implies that the more a city benefits from agglomeration economies, the larger its population N_i.

Figure 2.3 illustrates the model's implications for an exogenous wage increase in one particular city.[27] w^0 represents wages before the exogenous change and $w^1 > w^0$ afterwards. Everyone living in the city benefits from higher salary. A residential equilibrium requires that

$$v[c(\bar{R} + tN^0), w^0] = v[c(\bar{R} + tN^1), w^1] = \bar{V} \tag{2.35}$$

Equation (2.35) implies that benefits from higher wages have to be compensated by an inflow of population, i.e. $N^1 > N^0$. As illustrated in figure 2.3, land rents increase throughout the city, which enhances the cities spatial extent ($\bar{u}^1 > \bar{u}^0$). The existence of agglomeration economies, i.e. the fact that wages positively depend on the city's population, causes an additional rise in wages, followed by further migration to the city. The described cycle ends just as the wage increase cannot compensate for higher house prices anymore.

The key message of the Monocentric City Model with respect to agglomeration economies is that they have a decisive effect on the city's population and vice-versa. The more a city can benefit from agglomeration economies, the larger it is and the faster it grows. As certain conditions in the city change, such that agglomeration economies can evolve, this will result in subsequent urban population growth.

[27] For simplification, I assume that the exogenous wage increase is not due to a rising city population.

2.5.2 Theoretical Microfoundations of Agglomeration Economies

Since the microfoundations of knowledge spillovers constitute a central issue of this dissertation and are further investigated in chapter 3, I survey the existent micro-founded theories of agglomeration economies prevalent in the literature. Duranton and Puga (2004) identify three sources able to justify the positive relation between population and productivity at the city-level. Those sources encompass sharing, matching and learning. Duranton and Puga (2004) propose a model in which agglomeration economies arise from sharing a greater variety of intermediate goods in the production sector. They argue that a greater local labor supply enables more intermediate producers to enter the city. By applying the CES production function à la Stiglitz (1977), a greater variety of inputs causes aggregate increasing returns to scale at the city-level.[28] Duranton (1998) examines, how agglomeration economies can result from sharing the gains from individual specialization. The model is based on the idea that a greater local labor supply enables individuals to focus on specific tasks in the production chain. Following Smith (1776), this specialization enhances individual productivity leading to city-wide increasing returns scale. Krugman (1991) picks up the idea of Marshall (1890) and claims that agglomeration economies can result from risk reduction when sharing a large urban labor force. His model shows that in the presence of idiosyncratic firm-specific productivity shocks, each firm benefits from labor market pooling.[29] This improvement is due to the fact that the covariance between firm-specific shocks and market wages is reduced.

Helsley and Strange (1990) set up a model in which increasing returns to scale at the city-level arise from an improved quality of matches between workers and firms. Assuming a fixed number of firms in the city, increasing the urban labor force causes an improvement of each match. The improved match quality enhances each worker's productivity, leading to increasing returns to scale from matching. A larger labor force also enhances each individual's probability of being matched. This idea stems from Pissarides (1979), who introduces an aggregate matching function relating the number of matches to the number of workers and firms searching.[30] The concept of aggregate matching functions was transferred into an urban environment by Glaeser (1999), Berliant et al. (2006) and Assmann and Stiller (2015) among others. The mentioned papers have in common that agglomeration economies arise from the fact that every individual moving to the city raises the chances of being matched for every other individual in the city.

The third possible source of agglomeration economies suggested by Duranton and Puga

[28]Increasing returns to scale are equivalent to agglomeration economies, i.e. output changes overproportionally as inputs are changed.

[29]It turns out that each firm has higher expected profits.

[30]The model of Pissarides (1979) can be applied for several topics related to the city, not only in the context of the local labor market; e.g. matching of singles for relationships or in the context of any social interaction taking place in the city.

(2004) are knowledge spillovers. Marshall (1890) asserts that most intellectual knowledge flows are achieved in a dense urban environment. Jacobs (1969) argues that cities serve as a natural meeting point for diverse individuals. She claims that the combination of diverse knowledge during face-to-face interactions leads to the creation of new ideas, which raises the city's productivity. Lucas (1988) picks up Jacobs' view and states that human capital externalities, predominantly prevalent in cities, are the main engine for aggregate growth. Glaeser (1999) sets up a model, in which individuals in the city learn from each other via random face-to-face interactions. He shows that the transmitted knowledge increases in the city's density. A related model from Peri (2002) shows that especially young people move to cities, as they benefit most from the provided learning opportunities. Berliant et al. (2006) use a random meeting framework à la Pissarides (1979) to model social interactions in cities. In their model meeting partners have the choice, whether the meeting turns into a face-to-face interaction in which knowledge is transmitted. Heterogeneity in the meeting partners' knowledge types determines the interaction's extent of knowledge transmission.

The model presented in chapter 3 builds on the work of Berliant et al. (2006).[31] Contrary to Berliant et al. (2006), two types of knowledge spillovers are exchanged during urban face-to-face interactions. The transmission of knowledge enhances the involved meeting partners' productivity, while the creation of knowledge influences the city's overall productivity level. The extent of both types of knowledge spillovers depends on the heterogeneity of the meeting partners' knowledge types. It is shown that individuals choose an excessively narrow range of partners to interact with, leading to a lower than socially optimal creation of new ideas. Davis and Dingel (2012) are the first to examine knowledge spillovers in a spatial equilibrium framework for a system of cities. They find that highly skilled individuals self-select into larger cities, as they overproportionally benefit from superior learning opportunities.

2.5.3 Empirical Findings

Previous empirical literature concentrates on the quantification of agglomeration economies, i.e. on the measurement of the population's (or population density's) impact on workers' individual productivity. Ciccone and Hall (1996) propose the following regression model to estimate the extent of prevalent agglomeration economies:

$$\log(w_{i,t}) = \beta_0 + \beta_1 \log(N_{i,t}) + \delta_{i,t}'\beta_2 + u_{i,t}, \tag{2.36}$$

$w_{i,t}$ denotes the workers' average productivity at time t in city i, approximated by the city's average output per worker or the city's average wage level. $N_{i,t}$ depicts the city's

[31]The content of chapter 3 is based on Assmann and Stiller (2015).

population or population density. $\delta_{c,t}$ is a vector of relevant control variables. The estimate for β_1 is biased as long as the inherent reverse causality in regression model (2.36) is not accounted for. Beside the population's effect on local productivity, the city's productivity in return influences city size. In order to derive unbiased estimates for β_1, one needs to find exogenous sources of variation for the city's population or population density.

Ciccone and Hall (1996) detect a significant impact of population density on productivity at the US state-level. As exogenous source of variation for modern levels of population density across US states they use data on population across US states during the 19th century, data on the presence of railroads in 1860 and data on the states' proximity to the eastern seaboard. Ciccone and Hall (1996) show that increasing the state's density by one percent raises average productivity by around 5 percent. Applying the same IV approach, Ciccone (2002) identifies results of similar magnitude for regions across Germany, Italy, France, Spain and the UK. Increasing density by one percent enhances average productivity across European regions by 4.5 percent. Combes et al. (2010) estimate the effect of local density on productivity for French cities. Additional to historical instruments like population patterns during the 19th century, they use a set of geological characteristics that have predetermined the distribution of early settlements. They estimate the elasticity of average wages to density to be around 2 percent. Glaeser and Resseger (2010) find that the significant connection between population density and productivity is also prevalent at the US city-level and that the connection is significantly stronger for cities with higher skill levels. This indicates that knowledge spillovers among high-skilled workers are particularly important for agglomeration economies. Glaeser and Maré (2001) present further evidence that agglomeration economies result from local knowledge spillovers. Controlling for individual levels of education and experience, they show that individual wages are increasing more quickly in dense urban places with a highly skilled workforce. Similar results for Spanish cities are suggested by De la Roca and Puga (2013), who find that workers experience steeper wage increases during the time they work in larger and denser cities. The findings by Glaeser and Maré (2001) and De la Roca and Puga (2013) are only consistent with the hypothesis of learning by knowledge spillovers in cities, as other attempts of an explanation only suggest immediate wage increases at the time an individual enters the city.

Audretsch and Feldmann (2004) note that the urban environment[32] is decisive for the extent of local knowledge spillovers and thus for agglomeration economies to evolve. The intractable nature of knowledge spillovers renders their direct measurement impossible. But one can use information on employment growth, wage growth and innovative output of cities to make conclusions about the industry structure's impact on local knowledge spillovers. Mainly two attitudes towards the optimal urban environment are existent

[32]I.e. specific characteristics of the local industry structure.

in the literature. Marshall (1890) argues that the prevalence of firms from the same industry (specialization) brings about the most effective technological spillovers in the city. Jacobs (1969), however, advocates the idea that the combination of knowledge from different industries (diversity) generates the highest subsequent city growth.
Most prominently, Glaeser et al. (1992) and Henderson et al. (1995) estimate the impact of the city's industry structure on subsequent employment and wage growth. Glaeser et al. (1992) find that employment growth within city-industries increases in the city's diversity of its industry structure. Henderson et al. (1995) find that traditional manufacturing industries that do not heavily rely on innovation benefit from specialization, while diversity enhances employment growth among innovative high-technology firms. Feldmann and Audretsch (1999) show that a diversied city-industry structure is significantly positively connected with the emergence of innovative output measured by new product introductions. Thus the literature predominantly suggests that cities with a diversified environment induce innovative output (or knowledge creation).
Glaeser et al. (1992) show that while employment growth is positively related to diversity, the city's wage growth is, however, positively related to the city's extent of specialization. This result can be interpreted as the outcome of better learning opportunities (i.e. more knowledge transmission) among workers with similar knowledge backgrounds.
Concerning urban density, empirical findings suggest that it accelerates both the creation and transmission of knowledge. Carlino et al. (2007) show that doubling urban employment density raises individual innovative output measured by per capita patents by about 20 percent. Glaeser et al. (1992), Henderson et al. (1995) and Glaeser and Resseger (2010) find that wage growth is strongly correlated with urban density. Interpreting those empirical observations as driven by knowledge spillovers, they suggest that denser cities accelerate both the creation and transmission of knowledge.
I will use the empirical observations summarized above to justify the assumptions I make about the extent of knowledge spillovers in face-to-face interactions in the theoretical model presented in chapter 3.
In summary, there exists unambiguous empirical evidence for a positive connection between the city's population (or population density) and productivity. The positive relation can most adequately be explained by the process of learning from other highly skilled individuals in the city. There is further evidence that the city's industry structure has a decisive impact on the extent of local knowledge spillovers and thus on the city's extent of agglomeration economies.

2.6 Entrepreneurial Activity

The emergence of innovative start-up firms in cities is supposed to be an important driver of subsequent urban growth. Those firms might have the potential to make ideas created through local knowledge spillovers available to the market. If this is true then the fastest growing cities are those that have the ability to attract entrepreneurs.

2.6.1 Implications of the Monocentric City Model

Now, I assume that there exist two types of workers in the model: Entrepreneurs (E_i) and industrial workers (I_i) with $E_i + I_i = N_i$. For simplicity, I assume that both types have identical preferences and earn the same wage w_i, which positively depends on the number of entrepreneurs E_i in city i, i.e. $w_i = w(E_i)$ with $w'(E_i) > 0$ and $w''(E_i) < 0$. The wages' positive dependence on the number of entrepreneurs reflects the general sentiment that entrepreneurs bring their ideas to the city and thus enhance city i's technology level.

A residential equilibrium within and across cities requires that

$$V_i(u) = v[c(\bar{R} + tN_i), w_i(E_i)] = \bar{V} \quad \forall\ u_i \in [0, \bar{u}_i],\ \ \forall\ i = 1, ..., C, \tag{2.37}$$

with $N_i = E_i + I_i$. Totally differentiating identity (2.37) with respect to E_i and N_i yields

$$\frac{\mathrm{d}N_i}{\mathrm{d}E_i} = -\frac{\frac{\partial v}{\partial w_i} w_i'(E_i)}{\frac{\partial v}{\partial c}\frac{\partial c}{\partial N_i}} > 0. \tag{2.38}$$

Gradient (2.38) implies that an inflow of entrepreneurs is followed by an inflow of further population to the city, as both worker types benefit from higher wages.[33]

Figure 2.3 illustrates the effect on land rents throughout the city as the city experiences an exogenous inflow of entrepreneurs. Let E^0 be the number of entrepreneurs before the exogenous change and $E^1 > E^0$ the number afterwards. The inflow of entrepreneurs translates into higher wages, i.e. $w(E^1) > w(E^0)$. A residential equilibrium requires that

$$v[c(\bar{R} + tN^0), w(E^0)] = v[c(\bar{R} + tN^1), w(E^1)] = \bar{V}, \tag{2.39}$$

which implies that the increase in wages has to be compensated by an inflow of population[34], i.e. $N^1 > N^0$. This population increase raises land rents and housing prices and shifts the city boundary to the right, i.e. $\bar{u}^1 > \bar{u}^0$. As described in section 2.5.1, the inflow of further population ends just as the wage increase, induced by the additional

[33]Mathematically, gradient (2.38) is positive since $\frac{\partial v}{\partial w_i} > 0$, $w_i'(E_i) > 0$ and $\frac{\partial v}{\partial c}\frac{\partial c}{\partial N_i} < 0$.

[34]The population inflow incorporates entrepreneurs and industrial workers.

entrepreneurial activity, cannot compensate for higher house prices anymore.
In summary, the Monocentric City Model predicts that cities with a greater supply of entrepreneurs are larger in population size. Increasing the city's number of entrepreneurs is followed by an inflow of additional population. It has to be noted that model's implications crucially depend on the assumption that the prevalence of entrepreneurs raises the city's technology level.

2.6.2 Empirical Findings

Glaeser et al. (1992) propose the following regression model to identify the entrepreneurship's impact on urban growth:

$$\Delta_{t,t+1} \log(N_{i,t}) = \beta_0 + \beta_1 \log(E_{i,t}) + \delta'_{i,t}\beta_2 + u_{i,t}, \tag{2.40}$$

City i's entrepreneurial activity at period t is denoted by $E_{i,t}$. There are mainly two proxies for urban entrepreneurship prevalent in the literature. First, the number of newly founded start-up firms.[35] And second, the city's average establishment size. Higher levels of local entrepreneurship are associated with the emergence of more start-up businesses and a smaller average establishment size. Previous empirical literature focuses on the city's level of employment as outcome variable of interest[36], thus $N_{i,t}$ denotes city i's employment at period t. $\delta_{i,t}$ represents a vector of relevant control variables. The coefficient β_1 indicates how urban employment growth depends on the city's initial supply of entrepreneurs.
Vernon (1960) and Chinitz (1961) supply non-formal evidence on the relationship between entrepreneurship and its urban environment. Further influential non-formal articles stem from Jacobs (1969) and Saxenian (1994). Glaeser et al. (1992) are the first to detect a significant positive connection between local entrepreneurial activity and subsequent economic growth.[37] Similar results for the US and several European countries are provided by Miracky (1993), Rosenthal and Strange (2003), Ghani et al. (2011) and Samila and Sorenson (2011) among many others.
Only recently, the literature starts to address the issue of reverse causality inherent in regression model (2.40). Entrepreneurship might have a positive effect on subsequent employment growth, but in the same way does a prosperous economic environment attract potential entrepreneurs. In order to give estimates of β_1 a causal interpretation, one has to detect exogenous sources of variation for modern rates of entrepreneurship across

[35] Usually only start-up firms having less than a certain amount of employees are counted.
[36] This is due to the fact that the creation of new jobs is the most important task that is associated with entrepreneurship.
[37] The authors measure the local level of entrepreneurship by the city's average establishment size and urban growth by the growth of city employment.

cities. So far, the literature provides only three approaches trying to circumvent potential endogeneity by applying an Instrumental Variable approach. Most prominently, Glaeser et al. (2012) propose the proximity of US metropolitan statistical areas (MSAs) to historical coal mines as a potential source of exogenous variation for modern levels of entrepreneurship. They argue that the city's industry structure close to coal mines is dominated by large enterprises operating in the steel industry. The prevalence of those large steel companies discourages potential entrepreneurs from entering the city.[38] Lee (2014) uses homestead exemption levels set by state bankruptcy law in 1975 to instrument for local entrepreneurship. The approach's idea is that cities in states with higher exemption levels attract more potential founders. As entrepreneurs have to take high risks, they might respond to loss mitigating policies. Both approaches find a persistent link between entrepreneurship and subsequent employment growth after instrumenting for modern rates of local entrepreneurship.[39] In chapter 4, I propose the distribution of 19th century patent activity across the US as an alternative instrument.[40] I argue that cities with more patents during the 19th century exhibit a higher probability of becoming a place dominated by large companies, which deters potential entrepreneurs from entering the city. Contrary to Glaeser et al. (2012), cities dominated by large enterprises do not necessarily depend on the steel industry. This disentanglement from the city's industry structure reduces possible problems at the second stage of IV regressions, as modern growth rates are less correlated with the instrument. IV results suggest that local entrepreneurship loses almost its entire impact on subsequent employment growth and becomes insignificant once entrepreneurship is instrumented for by 19th century patent activity.

While there exists a general sentiment in the literature that entrepreneurship enhances subsequent urban growth, there is some reason to believe that at least part of the relation is due to the entrepreneurs' self-sorting into fast-growing cities.

2.7 Alternative Sources of Urban Growth

Beside the presented determinants of urban growth, Duranton and Puga (2014) address two additional factors that, taken by itself, may not be able to explain urban growth but certainly have an impact on the observed heterogeneity. Those factors encompass the city's housing supply and the theory of random urban growth.

The Monocentric City Model presented in section 2.2 does not incorporate supply side

[38]See section 4.2 for an extensive discussion of why that is the case.

[39]The approaches proposed by Glaeser et al. (2012) and Lee (2014) are presented in detail in section 4.2.

[40]The content of chapter 4 relies on Assmann (2015).

restrictions. In reality, the city's supply of housing is affected by exogenous geographical constraints, like water or mountains, and endogenous land-use regulations.[41] These restrictions are not able to explain the observed extent of city growth, but they can contribute to the comprehension of the observed growth heterogeneity across cities. Suppose that city i exhibits an exogenous increase in wages. The wage enhancement generates an incentive for population inflow.[42] The incentive's consequences crucially depend on the city's housing supply elasticity. Cities with restrictive surroundings or strict land-use regulations like San Francisco, CA or Miami, FL are expected to mainly react with rising housing prices and no or minor population growth as the additional demand cannot be compensated with extra housing space. Cities with less restrictive surroundings and lax land-use regulations like Orlando, FL or Phoenix, AZ are expected to react with a mitigated increase in housing prices and with population growth. This is due to the fact that these cities are able to adjust their housing supply to the extra demand. Glaeser et al. (2008) confirm this conjecture by showing that US cities with an inelastic housing supply experienced significantly larger price increases and less population growth during the 1980s boom and during the post-1996 boom. To measure housing supply elasticities across US cities, Glaeser et al. (2008) apply the proxy for supply side restrictions proposed by Saiz (2010).[43] He uses satellite-generated data on the presence of water bodies and steep-sloped terrain to estimate the amount of developable land across US cities. Thus there exists evidence that differences in housing supply restrictions across cities are partially responsible for the observed city growth heterogeneity.

A broad range of recent urban growth literature concentrates on the detection of pure statistical processes to explain the observed pattern of city size distributions across countries. This strand of literature imputes that city sizes are mainly predetermined and only marginally influenceable by urban policies.[44] Many authors[45] have confirmed that the size distribution of US cities resembles a Zipf Distribution (Zipf (1949)), a special case of the Pareto Distribution.[46] In the context of cities, Zipf's Law states that city sizes within countries are inversely proportional to their rank in the city size distribution.[47] Gabaix (1999) illustrates how Zipf's Law can be inferred from the application of Gibrat's Law of proportional growth (Gibrat (1931)). Gibrat's Law says that all cities within a

[41] Saiz (2010) finds that the two restrictions are highly correlated. Cities with more geographical constraints also exhibit stricter land-use regulations.

[42] See section 2.5.1 for an extensive discussion.

[43] Glaeser et al. (2008) cite an earlier version from 2008, which uses the same methodology to estimate housing supply elasticity across cities.

[44] See Gabaix and Ioannides (2004) for an extensive literature review.

[45] See Gabaix (1999) and Krugman (1996) among others.

[46] The original text of Zipf (1949) deals with the frequency distribution of words in specific languages.

[47] Suppose that N_r denotes the population of the city occupying rank r in a country's city size distribution. The population of the country's largest city is denoted by N_1 as it occupies the distribution's first rank. Zipf's Law states that city r's expected size is determined by $N_r = \frac{N_1}{r}$. Thus the second largest city is expected to have half the population of the largest city and the third largest city is expected to exhibit a third of the largest city's population.

country exhibit time-variant annual growth rates which are independent from city size. Assuming that the country's initial population and initial number of cities are fixed and introducing a lower bound on city sizes, Gabaix (1999) shows that the country's steady state city size distribution follows Zipf's Law. Eeckhout (2004) and Rossi-Hansberg and Wright (2007) attempt to impose economic rationalizations for Zipf's Law. They set up models in which industry- and city-wide productivity shocks generate urban growth rates coinciding with Gibrat's Law.

2.8 Discussion

I summarize the literature on potential drivers of urban growth, which encompass the provision of transportation infrastructure, the supply of local amenities, the existence of agglomeration economies and the prevalence of entrepreneurs in cities. I also address the topics of housing supply restrictions and random growth theory.

There are two things to note here: First, during the last years the literature turns its focus on micro-based foundations of agglomeration economies. As noted by Glaeser (1999), the observed urban wage evolution is best explained by the cities' ability to induce knowledge spillovers.[48] Chapter 3 of my thesis further investigates the microfoundations of agglomeration economies. I present theoretical work that examines whether knowledge spillovers in urban face-to-face interactions reach their optimal extent. Second, measuring the impact of city characteristics on city growth is associated with empirical difficulties. Regression models relating city growth to one of the described determinants suffer from reverse causality, which causes OLS estimates to be biased. Ciccone and Hall (1996) are the first to address the endogeneity issue in a regional context by applying Instrumental Variable regressions. This strategy requires the identification of exogenous sources of variation for today's city characteristics. Those instruments encompass historical population patterns or historical characteristics that are assumed to have influenced them. The key assumption, required for consistency of IV estimates, is that instruments which have affected population patterns a long time ago in the past are mostly unrelated to today's productivity. While the strategy of identifying exogenous sources of variation is well-established in the empirical literature on the link between urban growth and city characteristics, it is only recently applied in measuring the causal impact of local entrepreneurship. The work presented in chapter 4 intends to remove this shortcoming. I present IV regression results using the regional distribution of 19th

[48]Wages increase gradually during the time workers are employed in cities and do not drop after leaving the city.

century patents across the US as exogenous source of variation for modern rates of local entrepreneurship. I find that the general sentiment of the local entrepreneurship's growth enhancing impact is only restrictedly supported by the data.

Chapter 3

Knowledge Spillovers in Cities

3.1 Introduction

Cairncross (2001) argues that new communication technologies induce the "death of distance", rendering cities in their role as meeting points irrelevant because individuals can avoid the cost of urban living and still benefit from information flows. However, empirical observation shows that cities are becoming even more important as the increasing concentration of economic activity in space continues.[1] New communication methods may indeed facilitate the transmission of codified information with a stable meaning expressed in a standardized system of symbols, but the exchange of complex uncodified knowledge brings along distinct challenges. Only real-life face-to-face interactions give people the opportunity to use all relevant forms of communication techniques at the same time. As these interactions are local, spillovers of complex knowledge are also spatially restricted. Audretsch and Feldmann (1996) show that more knowledge intensive industries exhibit a significantly stronger geographic concentration. They also find evidence for the great importance of cities in the process of creating ideas. Less than 5 percent of product innovations occur outside from metropolitan areas and more than 45 percent of these innovations come from the four metropolitan areas New York City, Los Angeles, San Francisco or Boston. Jaffe et al. (1993) quantify the extent of knowledge spillovers and find that new patents disproportionately often cite patents that were invented in the same city or region. While the importance of local knowledge spillovers as agglomeration force is empirically well established, the literature on the underlying mechanisms is less developed. This further stresses the importance of a more detailed theoretical understanding of local spillovers.

[1] See chapter 1 for illustrative data.

The objective of this chapter is to investigate the effects of two different types of knowledge spillovers (transmission and creation) on urban productivity, innovation and city size. I present co-authored theoretical work[2], which applies a search-theoretic spatial equilibrium framework to analyze both types of knowledge spillovers from urban face-to-face interactions. First, cities give individuals the opportunity to increase their productivity through knowledge transmission (learning). And second, knowledge creation (innovation) increases the city's rate of technological change, which raises the productivity of each worker affected by the innovation. The transmission of knowledge can be thought of as the result of workers' observation and imitation of each other's techniques. This process is facilitated when interacting workers have a similar knowledge background. Knowledge creation results in the form of new ideas from the combination of interacting workers' existing knowledge. We adopt the view of Jacobs (1969) and assume that every interaction among workers, independently of their knowledge background, has the potential to bring about innovations. One major difference between these two types of knowledge spillovers is apparent: Workers benefit individually from the process of imitating other workers. The increased productivity directly causes higher wages. On the contrary, innovations are treated as a non-excludable local public good. This assumption can be justified as the contribution to the emergence of innovations is often not directly credited to the inventors and thus not fully compensated. Our model economy consists of two asymmetric locations, the city and the periphery. Only the city provides the opportunity to exchange knowledge via local face-to-face interactions. Urban workers choose the range of other workers they are willing to interact with. Since they do not consider the interactions' impact on the economy's rate of technological change, they only accept a smaller than socially optimal range of matches.[3] The resulting suboptimal extent of knowledge spillovers generally leads to socially inefficient city sizes. To the best of my knowledge there exists no theoretical model that incorporates both types of knowledge spillovers in an urban context.

This chapter is organized as follows: Section 3.2 summarizes empirical observations from section 2.5.3 to infer predictions about the extent of knowledge spillovers in face-to-face interactions. In section 3.3, the model environment is introduced. Section 3.4 presents the model economy's Steady-State Equilibrium. In section 3.5, the market outcome is compared to the outcome that results from the Social Planner's Problem and the different types of inefficiencies that can emerge are analyzed. Section 3.6 summarizes and concludes.

[2]The content of this chapter is based on Assmann and Stiller (2015).

[3]There exists a very restricted parameter space resulting in a range of matches that is larger than socially optimal. More on that finding can be found in section 3.5.2.

3.2 Spillovers in Face-to-Face Interactions

In section 2.5.3, I summarize empirical observations on the connection between the city's industry structure and the city's subsequent employment growth, wage growth and innovative output. These observations can be used to infer predictions about the extent of knowledge transmission and knowledge creation in urban face-to-face interactions. Such an approximation to knowledge spillovers is inevitable as their source and extent are not directly observable from the data. The inferred predictions are only valid if the city's industry structure is responsible for the city's composition of face-to-face interactions. Thus one has to assume that cities with a diversified industry structure bring about interactions between individuals with diverse knowledge types, whereas specialized cities cause face-to-face interactions between individuals with a similar knowledge background. Table 3.1 summarizes the empirical observations made in section 2.5.3 and illustrates the respective predictions for the outcome of knowledge spillovers in face-to-face interactions. These predictions are used to justify the model's assumptions about the extent of knowledge spillovers made in section 3.3.5.

Empirical Observation	**Prediction for f-2-f interactions**
Both, the creation and transmission of knowledge, are positively affected by urban density.	Dense urban areas bring about more face-to-face interactions.
The creation of knowledge (or innovative output) is not harmed by the city's diversity of industries.	The creation of knowledge (or innovative output) is not harmed by the city's diversity of knowledge types.
The transmission of knowledge decreases in the city's diversity of industries.	The transmission of knowledge decreases in the city's diversity of knowledge types.

TABLE 3.1: Stylized facts about the types of knowledge spillovers in f-2-f interactions

3.3 Economic Environment

This section presents the search-theoretic model of a spatial economy incorporating two different types of knowledge spillovers (creation and transmission of knowledge) in the city. The model is related to the work of Berliant et al. (2006) and Glaeser (1999), but additionally incorporates the creation of knowledge.[4] The basic idea is the following: The dense environment of cities provides workers with the opportunity to get into contact via face-to-face meetings, whereas the rural area does not.[5] Urban workers are brought together by a random meeting-technology with the interaction's outcome of knowledge

[4] Knowledge creation is also referred to as innovative output or innovation in the following.

[5] E.g. because the area is too spacious, meeting points like public squares are not prevalent, etc.

transmission and creation being influenced by the combination of the meeting partners' unobservable knowledge types. The partners' knowledge type is unknown before the meeting, but revealed after a first contact. This framework is adopted from Pissarides (2000), who uses it in the context of stochastic job matchings. In this type of model it is crucial to distinguish between a meeting and a match. Whether a meeting between two workers becomes a match depends on the realized productivity.[6] Meetings with low realizations are canceled after a very first contact because it is worthwhile to wait for a partner with a more adequate knowledge type to be matched with. Neoclassical Theory implies that innovative output is available at no charge to everyone in the city such that workers are not fully compensated for their created knowledge, making innovative output a local public good. Its existence gives rise to social inefficiencies as social benefits of generated innovations exceed private benefits. Therefore workers accept only the matches maximizing their expected personal outcome, not taking into account that each accepted match contributes to publicly available innovations in the city.

3.3.1 Basic structure

The model economy is populated by infinitely-lived workers and consists of two heterogeneous regions: The city and the periphery. Time is continuous and in each point of time workers decide in which region to locate. All action takes place in the city, whereas the periphery is modeled as simple as possible. In the city, individuals have the possibility to interact face-to-face. Living in a crowded urban environment is associated with economic cost.[7] Each worker in the city generates congestion costs of $t > 0$. N denotes the number of individuals living in the city, so the total congestion costs each worker faces upon entering are tN.[8] There is no crowding in the periphery, so workers living there do not face any congestion costs.

3.3.2 Economic Agents

Workers are heterogeneous in their horizontally differentiated knowledge background. The variety of the economy's knowledge base is displayed by a unit circle, represented in figure 3.1. The approach of using a unit circle to illustrate the economy's knowledge base is adopted from Helsley and Strange (1990) and was used by Berliant et al. (2006) and Brueckner et al. (2002) among others. Each worker is endowed with a specific

[6] In the following the label "contact" is tantamount to "meeting" and the label "face-to-face-interaction" is tantamount to "match".

[7] Pollution, road congestion and high house prices are only a few examples for the burden of urban living.

[8] The results of the model analysis are robust to well-established transformations of the congestion cost function. E.g. the results stay unaltered when quadratic congestion costs tN^2 are used.

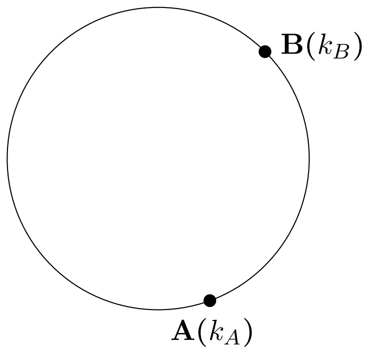


FIGURE 3.1: Knowledge Space

knowledge type k represented by its position on the circle's circumference K.[9] The circumference K can be interpreted as the economy's knowledge space representing all types of knowledge in the economy, e.g. economics, mathematics, physics, etc. Location $k \in K$ is drawn from a uniform distribution and exogenously assigned to each worker. In figure 3.1, knowledge type k_A is assigned to worker A, whereas knowledge type k_B is assigned to worker B. The distance between k_A and k_B on the circumference measures the horizontal difference between the two types of knowledge. The model incorporates no vertical differentiation, i.e. all workers have an equal level of education. Furthermore, position k on the unit circle is only of relevance for workers in the city and irrelevant for workers in the periphery since only the city facilitates the exchange of knowledge via face-to-face interactions.

Workers are heterogeneous in knowledge background, but homogeneous in preferences. Flow output (or flow income) y is spent on a homogeneous consumption good. The exact determination of y is discussed in section 3.3.4. Flow utility U is linear in y, yielding

$$U = U(y) = y. \tag{3.1}$$

The utility function's linearity implies that maximizing lifetime utility is equivalent to maximizing lifetime income.

3.3.3 Meeting Technology

The reason for workers to enter the city is the opportunity to increase their productivity by the exchange of knowledge. Before introducing the exact modeling strategy of knowledge spillovers, the emergence of meetings (contacts) in the city is clarified. The model is based on the framework of stochastic job matching used in Pissarides (2000) and relies on the assumption that there exists a well-behaved meeting function relating

[9] In the following, an individual with that characteristic is labeled as worker k.

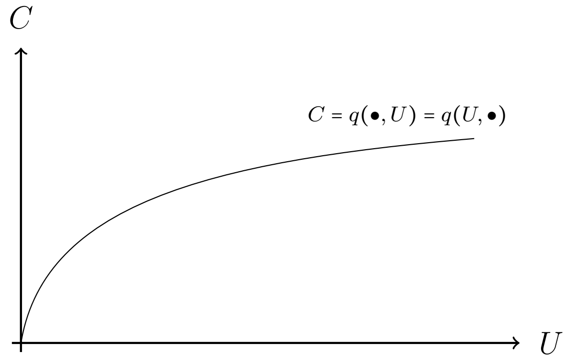


Figure 3.2: Meeting Technology

the city's number of contacts to the city's number of workers searching for face-to-face interactions. This framework generates a connection between the city's density and its number of meetings. Suppose the city is populated by N individuals. A fraction $m \in (0,1)$ of those N individuals is matched, i.e. are currently involved in face-to-face interaction. The number of matched individuals is denoted by M. Thus the fraction of individuals unmatched is $u = 1 - m$ with the number of unmatched individuals denoted by U, implying $N = M + U$. It is important to be clear about the difference between a meeting (contact) and a match (face-to-face interaction). Whether a meeting turns into a match is the decision of the involved meeting partners and depends on the potential productivity gains. Matched workers are not able to search for new partners, so that only unmatched individuals are engaged in the search process. A meeting always requires two parties, one in the first and one in the second position. In Pissarides (2000), the number of job contacts per unit of time depends on the number of firms in the first and the number of unemployed in the second position. The presented model of this chapter contains no firms and no unemployed. The number of job contacts is thus replaced by the number of meetings and firms and unemployed are replaced by the city's number of unmatched workers. Since individuals meet symmetrically, all unmatched workers can either be in the first or second position, thus the meeting function can be described by

$$C = q(U,U). \tag{3.2}$$

C denotes the number of meetings per unit of time. The meeting function $q(\cdot)$ is increasing and concave in both arguments and homogeneous of degree $\gamma > 1$. The last assumption ensures that the probability of having a meeting increases with the city's density of unmatched workers. $q(\cdot)$ is assumed to fulfill the Inada conditions. Figure 3.2 illustrates the meeting function's behavior. The meeting technology randomly selects unmatched workers from the pool of possible meeting partners U. The meeting rate, the rate at which unmatched worker have contacts per unit of time, is given by

$$\mu(U) = \frac{C}{U} = \frac{q(U,U)}{U}. \tag{3.3}$$

Using the assumption of homogeneity of degree $\gamma > 1$, the meeting rate $\mu(U)$ can be expressed as

$$\mu(U) = \frac{q(U,U)}{U} = \frac{U^{\gamma} q(1,1)}{U} = U^{\gamma-1} q(1,1). \tag{3.4}$$

In order to derive a meeting rate which is linear in the number of unmatched workers, γ is set to 2.[10] The expression $q(1,1)$ determines how many contacts an individual is able to have per unit of time and denoted as meeting intensity α in the following.[11] Therefore the meeting rate of an unmatched worker in the city can be written as

$$\mu(U) = q(1,1)U = \alpha U. \tag{3.5}$$

3.3.4 Production Technology

Production in the periphery is modeled as simple as possible in order to focus on the city's production process. Therefore flow output $\bar{y}$ of each individual living in the periphery is assumed to be $\bar{y} = 0$.

All structure is put on the city's production technology. Suppose that an individual of knowledge type k is currently matched with an individual of knowledge type k'. The flow output of worker k crucially depends on the partner's type k' and is represented by

$$y(k,k') = A + e(k,k'). \tag{3.6}$$

A denotes the city's total factor productivity (TFP)[12] common to all individuals in the city. It crucially depends on the city's overall level of knowledge creation (see section 3.3.5.2). $e(k,k')$, denotes the personal effectiveness of individual k currently matched with individual k' and is determined by interaction's knowledge transmission (see section 3.3.5.1). If individual k is currently unmatched, it has personal effectiveness of zero and flow production y is solely determined by the TFP A, i.e. $y = A$. Thus both the TFP A and the personal effectiveness $e(k,k')$, are influenced by knowledge spillovers resulting from face-to-face interactions.

[10] This simplifying assumption is also used by Berliant et al. (2006).

[11] A value of $\alpha = 0.1$ indicates that during one period of time unmatched agents can meet 10 percent of all unmatched individuals in the city. A value of $\alpha = 2$ means that during the same time unmatched individuals can talk to each unmatched individual twice.

[12] For tractability the total factor productivity is assumed to be a flow value.

3.3.5 Knowledge Spillovers

This section clarifies how the extent of knowledge transmission and creation depends on the combination of knowledge types.

3.3.5.1 Knowledge Transmission

The transmission of knowledge is equivalent to the intellectual exchange described in Berliant et al. (2006). The heterogeneity of workers in terms of their position in the knowledge space K plays a crucial role for the personal effectiveness. Consider two workers brought together by the city's meeting technology: one endowed with knowledge type $k \in K$, the other endowed with knowledge type $k' \in K$. Suppose that both accept to be matched. The matching partners' personal effectiveness crucially depends on the distance between knowledge types k and k' in the knowledge space K, measured by the Euclidean metric $d(k, k')$.[13] The highest degree of knowledge transmission and thus the highest personal effectiveness is attained when the two meeting partners are exactly similar in knowledge type ($k = k'$) as illustrated in figure 3.3. In this case it is straightforward to communicate and exchange information. They already use the same vocabulary and techniques, so they can start exchanging knowledge and applying it to attain a higher personal effectiveness right away. The assumption of decreasing knowledge transmission with increasing diversity of knowledge types is motivated by the stylized empirical facts about the outcome of face-to-face interactions presented in section 3.2. For better illustration one can think of two urban economists. Both use the same terminology and techniques and once matched (e.g. have a research collaboration), immediately start to learn from each other and become more productive. As knowledge distance $d(k, k')$ increases, knowledge transmission becomes more cumbersome. Since workers do not have a lot in common, they will have problems understanding each other's terminology and will not be able to just imitate each other's techniques. Thus, individuals with very heterogeneous knowledge backgrounds have to put in a lot of effort before the transmission of knowledge can begin. Once able to communicate they will find it difficult to apply the gained knowledge in their respective fields. Here one can think of a match between an economist and a dentist. Both will have major problems in understanding the vocabulary and imitating each other's techniques. Once they have managed to communicate, it remains questionable whether they can apply the gained knowledge in their respective occupation. The personal effectiveness $e(k, k')$ is described by:

$$e(k, k') = e_0 - e_1 d(k, k'). \tag{3.7}$$

[13]Distance in the knowledge space is used as a measure for knowledge diversity and has nothing to do with physical distance.

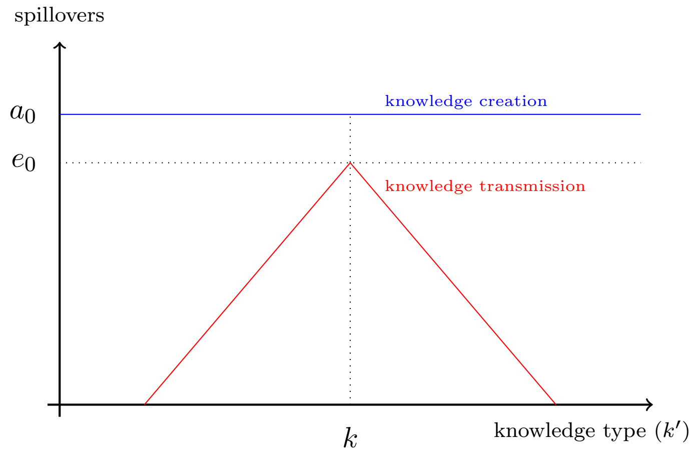


FIGURE 3.3: Knowledge Transmission and Knowledge Creation in f-2-f interactions

Parameter $e_0 > 0$ denotes the maximum level of personal effectiveness, which is achieved when the two meeting partners coincide in knowledge background ($k = k'$). Parameter $e_1 > 0$ describes the knowledge transmission's sensitivity to heterogeneity of knowledge types.[14]

3.3.5.2 Knowledge Creation

Matches in the city not only facilitate the transmission of existing knowledge but also cause the creation of new knowledge that in turn raises the city's technology level. How does the distance between knowledge types $d(k, k')$ affect the creation of new knowledge? For knowledge creation it is unclear whether the outcome should decrease in the distance $d(k, k')$. One can revisit the two examples from above. Two economists working in the same field personally benefit from a research collaboration. They can additionally use the collaboration to write papers that contribute to the creation of new knowledge. An economist and a dentist will find it very hard to apply the gained information to increase their individual productivity. The techniques are too different to benefit from in their

[14] $e_1 = 0$ indicates that heterogeneity of knowledge types among individuals is irrelevant for the transmission of knowledge and thus for the personal effectiveness. Each match between two workers, independently of the knowledge types they are endowed with, generates the same amount of knowledge transmission. $e_1 \to \infty$ in turn indicates that only workers with the same knowledge background have a chance to increase productivity through the transmission of knowledge. As soon as the knowledge types differ to a minimal extent they are not able to communicate. Berliant et al. (2006) assume that there exists an optimal distance $d^* > 0$ between knowledge types that creates the highest productivity. Their justification for this assumption is that "when individuals are too alike, they cannot accomplish much and little knowledge will be obtained". This assumption is not adopted in the model presented in this chapter as it distinguishes between the transmission and creation of knowledge. Berliant et al. (2006) combine these spillovers into one effect. Once the transmission and creation of knowledge are analyzed separately, it makes sense to assume a maximum effectiveness when agents are alike ($k = k'$), because the pure transmission of knowledge (in the absence of knowledge creation) is easiest when agents do not have to overcome any knowledge barrier.

respective occupation and the extent of knowledge transmission is thus expected to be rather low. But it is possible that the combination of their knowledge causes the creation of something new, e.g. they could find a way to create a more cost-efficient treatment plan or accounting system. We are content with the weak assumption that each match, independent of the diversity in knowledge types, contributes in equal measure to the city's technology level. This is in line with the stylized empirical facts on the extent of knowledge creation presented in section 3.2. Accordingly, the creation of new knowledge $a(k,k')$ by a currently matched agent of knowledge type k is independent of his partner's knowledge type k' and always equal to

$$a(k,k') = a_0. \tag{3.8}$$

Adopting the Neoclassical view, created knowledge is assumed to be a local public good which is distributed among all N individuals living in the city. The overall knowledge creation is directly translated into the city's total factor productivity A. In each point of time, the TFP A is determined by the created knowledge of a matched worker (a_0) times the city's number of matched individuals (M) divided by the overall number of individuals living in the city (N):

$$A = \frac{Ma_0}{N}. \tag{3.9}$$

Since created knowledge is a local public good, the model contains potential sources for social inefficiencies. Workers choose the range of acceptable matches in order to maximize their personal effectiveness. They do not take into account their impact on knowledge creation. Section 3.4 provides an extensive discussion concerning these inefficiencies. Figure 3.3 summarizes the impact of the knowledge distance $d(k,k')$ on the extent of knowledge transmission and creation in face-to-face interactions.[15]

3.3.6 Choice of the Knowledge Spread

The meeting rate $\mu(U) = \alpha U$, derived in section 3.3.3, indicates that the number of contacts increases linearly at rate α in the city's number of unmatched individuals U. So what determines which meeting turns into a match and which meeting is canceled after a first contact? Consider a meeting between two individuals of knowledge types k and k'. Given position k in the knowledge space, each individual chooses a knowledge spread $\delta_k > 0$, which determines the range of workers individual k is able to have a face-to-face interaction with. The knowledge spread δ_k is geometrically represented by the arc around knowledge type k leading to a knowledge horizon $[k-\delta_k/2, k+\delta_k/2]$. This horizon can be interpreted as the set of disciplines, individual k has at least elementary knowledge

[15] In figure 3.3, a_0 is chosen to be larger than e_0. However, since the relation between a_0 and e_0 is an empirical question, the model allows for all other potential coherences between a_0 and e_0.

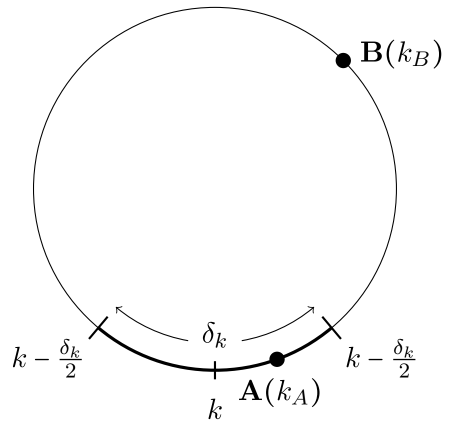


FIGURE 3.4: Knowledge Horizon

about. This elementary knowledge is indispensable to exchange knowledge in face-to-face interactions with an individual of knowledge type k'. Only if k' is located within the knowledge horizon of k, i.e. $k' \in [k - \delta_k/2, k + \delta_k/2]$, a face-to-face interaction is possible. Extending the knowledge horizon by increasing knowledge spread δ_k enables individual k to interact with a wider range of workers. Figure 3.4 displays knowledge space K. Since worker A with knowledge type k_A is located within the knowledge horizon of individual k, a face-to-face interaction is possible. Individual B with knowledge type k_B is situated outside the knowledge horizon of individual k. Consequently no transmission and creation of knowledge can occur, since worker k has no elementary understanding of B's field of knowledge. As workers are ex ante symmetric[16], only symmetric equilibria have to be considered. Symmetry implies that all workers choose the same knowledge spread ($\delta_k = \delta \;\; \forall k$). Either both individuals accept or both reject to be matched.[17] The next section clarifies how the exact choice of δ comes about in the market solution.

3.3.7 Expected Lifetime Utility

The expected lifetime utility V_m of a matched individual of knowledge type k depends on the interacting individuals' knowledge types, k and k', and on the city's number of unmatched workers U, i.e. $V_m = V_m(k, k', U)$. An unmatched individual of knowledge type k has currently no face-to-face interaction, thus his expected lifetime utility V_u only depends on its own knowledge type k and the number of unmatched individuals U, yielding $V_u = V_u(k, U)$. There exists an exogenous separation rate $\lambda > 0$, at which ongoing face-to-face interactions are split up. Assets can be traded over time at the

[16]Knowledge types are revealed only after two workers have met, i.e. workers are heterogeneous ex post, but indistinguishable ex ante.

[17]If worker k is not located within the knowledge horizon of k', then worker k' is automatically not located within the knowledge horizon of k.

exogenous interest rate $r > 0$.[18]
The value of a match between workers of type k and k', given a number of unmatched individuals U, satisfies the following Bellman Equation:

$$rV_m(k,k',U) = \underbrace{A + e(k,k')}_{y(k,k')} + \lambda\Big[V_u(k,U) - V_m(k,k',U)\Big] \tag{3.10}$$

Being matched with a worker of type k' can be interpreted as an asset held by worker k. Given perfect capital markets, the left hand side (capital cost of the asset) has to equal the right hand side (rate of return on the asset). A currently matched worker of knowledge type k has flow income $y(k,k')$ for the time being matched. The state changes from matched to unmatched at the exogenous separation rate λ with an associated net return of $V_u(k,U) - V_m(k,k',U)$.
The value of worker k being currently unmatched at a given number of unmatched individuals U, satisfies the following Bellman Equation:

$$rV_u(k,U) = A + \alpha U \int_{k-\frac{\delta}{2}}^{k+\frac{\delta}{2}} \Big[V_m(k,k',U) - V_u(k,U)\Big] dk' \tag{3.11}$$

Being unmatched can also be interpreted as an asset held by worker k. A is the worker's flow income due to the city's total factor productivity. Meetings occur at a rate $\mu(U) = \alpha U$. Only individuals of knowledge type $k' \in [k - \delta/2, k + \delta/2]$ are accepted for a face-to-face interaction.[19] The term inside the integral can be interpreted as the net return from changing the state from being unmatched to being matched with a worker of type k'.
(3.11) implies that a fraction δ of meeting partners are accepted for face-to-face interactions. Thus, the number of matches $p(U)$ per unit of time is given by

$$p(U) = \mu(U)\delta = \alpha\delta U, \tag{3.12}$$

with $\mu(U)$ being the rate of meetings per unit of time and δ being the probability of a meeting to turn into a match. $\alpha\delta$ can be interpreted as matching intensity, i.e. the rate at which individuals have face-to-face interactions per unit of time. It linearly increases in the city's density of unmatched individuals U. The uniform distribution of knowledge types implies that the expected distance in knowledge types of matched individuals is $\delta/4$. Thus (3.10) and (3.11) can be expressed as

$$V_m(k,k',U) = \frac{A}{r} + \frac{e_0 - e_1 d(k,k')}{r+\lambda} + \frac{\lambda}{r+\lambda}\frac{\alpha\delta U}{r+\lambda+\alpha\delta U}\frac{e_0 - e_1\frac{\delta}{4}}{r}. \tag{3.13}$$

[18]This assumption requires perfect capital markets.
[19]Other partners are not accepted, because worker k is not able to exchange knowledge with them.

$$V_u(\delta, U) = \frac{A}{r} + \frac{\alpha\delta U}{r + \lambda + \alpha\delta U}\frac{e_0 - e_1\frac{\delta}{4}}{r}. \tag{3.14}$$

For the analysis only (3.14) is of relevance as workers choose their location and knowledge spread δ in an unmatched state. This is due to the fact that a worker enters the city in that state. A/r is lifetime income generated by the city's total factor productivity. Since A is a local public good, worker k makes use of it independent of being matched or not. $\frac{e_0 - e_1\frac{\delta}{4}}{r}$ is expected lifetime income that would have been generated if individual k was matched at every point in time. As individual k is not consistently matched, it is discounted by the rate $\frac{\alpha\delta U}{r+\lambda+\alpha\delta U}$.[20]

Worker k chooses knowledge spread δ to maximize expected lifetime utility (3.14). Two opposing effects are taken into account. First, increasing the knowledge spread δ extends the knowledge horizon. Thus worker k is able to have face-to-face interactions with a wider range of individuals which increases the probability to become matched. But second, increasing knowledge spread δ also raises the expected distance in knowledge types for accepted matches.[21] The resulting decrease in expected knowledge transmission reduces the expected flow income of individual k.

3.4 Equilibrium Analysis

In this section, the symmetric Steady State Nash Equilibrium is established. The equilibrium is defined by the workers' choice of knowledge spread δ, the city's number of unmatched individuals U and the resulting city size N.

3.4.1 Steady State Population

The Steady State Equilibrium requires the city's number of matched and unmatched workers to be constant over time. In the symmetric case, this relationship implies that flows into and out of the pool of unmatched workers are equal:

$$\underbrace{\alpha\delta U}_{\substack{\text{flow out of the pool of}\\ \text{unmatched individuals}}} = \underbrace{\lambda M = \lambda(N - U).}_{\substack{\text{flow into the pool of}\\ \text{unmatched individuals}}} \tag{3.15}$$

[20]The discount rate can be interpreted as discounted matching rate, i.e. it is the matching rate $\frac{M}{N}$ with the interest rate r in the denominator.

[21]I.e. individual k also enters matches with individuals that have not much in common with him.

Identity (3.15) can be solved to derive the city's number of unmatched individuals as an implicit function of the city's total population N in steady state:[22]

$$U = \frac{\lambda}{\lambda + \alpha\delta U} N \tag{3.16}$$

3.4.2 Steady State Equilibrium (Market Solution)

The optimal knowledge spread δ^* maximizes the city's expected lifetime utility (3.14). It is determined by the tradeoff between increasing the probability of having face-to-face interactions and increasing the interactions' expected extent of knowledge transmission. Given the optimal choice of δ^*, individuals enter the city until lifetime utility in the city and the periphery are equalized, such that no relocation incentives exist.[23] Together with equation (3.16) the equilibrium values N^* and U^* are determined. The Steady State Equilibrium, which is also referred to as Market Solution in the remainder of the text, is defined as follows:

Definition 1: Steady State Equilibrium (Market Solution)
The Steady State Equilibrium is an allocation $\{\delta^*, U^*, N^*\}$ that satisfies the conditions:

(1) Workers maximize expected lifetime utility by choosing knowledge spread δ^*: $\delta^* = \text{argmax}_\delta V_u(\delta, U^*)$.

(2) Lifetime utility in the city equals lifetime utility in the periphery:
$V_u(\delta^*, U^*) - tN^* = 0$.

(3) The condition for Steady State Population is satisfied: $U^* = \frac{\lambda}{\lambda + \alpha\delta^* U^*} N^*$.[24]

Condition (1) says that the workers' choice of δ satisfies the following first order condition:[25]

$$\frac{\partial V_u}{\partial \delta} = \delta^2 + \frac{2(r+\lambda)}{\alpha U}\left(\delta - \frac{2e_0}{e_1}\right) = 0 \tag{3.17}$$

(3.17) implicitly defines the equilibrium knowledge spread as a function $\delta(U)$. In the following, this condition is referred to as the knowledge spread condition KS. One can see from KS that the larger the city's number of unmatched individuals U, the narrower the chosen knowledge spread δ. Intuitively, this relation is due to the population density's positive impact on the matching rate. It allows workers to be more picky regarding

[22] This expression is analogous to the Beveridge Curve in Labor Market Theory (see Pissarides (2000)).
[23] This resembles the residential equilibrium condition across cities in the Monocentric City Model presented in chapter 2.
[24] Conditions (1) and (2) determine $\{\delta^*, U^*\}$, condition (3) then automatically determines N^*.
[25] In the following, δ^* is referred to as δ.

their interaction partners. This relationship leads to a downward sloping KS-locus as is represented by the blue curve in figure 3.5.

The Market Solution gives rise to several inefficiencies. First, workers do not consider the impact of their choice of knowledge spread δ on the Steady State Population of unmatched individuals, which gives rise to an inefficiency referred to as matching externality in the following. Second, workers do not consider their impact on the local flow of knowledge creation and take total factor productivity A as an exogenous variable. Instead they choose knowledge spread δ in order to maximize their expected lifetime utility from knowledge transmission, i.e. they choose δ to maximize their personal effectiveness. This inefficiency is referred to as innovation externality.[26]

A residential equilibrium (defined by condition (2)) requires lifetime utility to be equalized across locations, such that relocation incentives disappear. This definition implies that workers migrate to the city until the city's attainable expected lifetime utility, $V_u - tN$, equals the periphery's attainable lifetime utility, 0. Accordingly, the city's steady state number of unmatched workers is determined by the following condition:[27]

$$\underbrace{\underbrace{\frac{\alpha\delta U}{\lambda + \alpha\delta U}\frac{a_0}{r}}_{A/r} + \frac{\alpha\delta U}{r + \lambda + \alpha\delta U}\frac{e_0 - e_1\frac{\delta}{4}}{r}}_{V_u(\delta,U)} - tN = 0. \tag{3.18}$$

Using the steady state population condition (3.16), the city's equilibrium number of unmatched individuals is defined as a function $U(\delta)$, referred to as equilibrium entry condition EE.

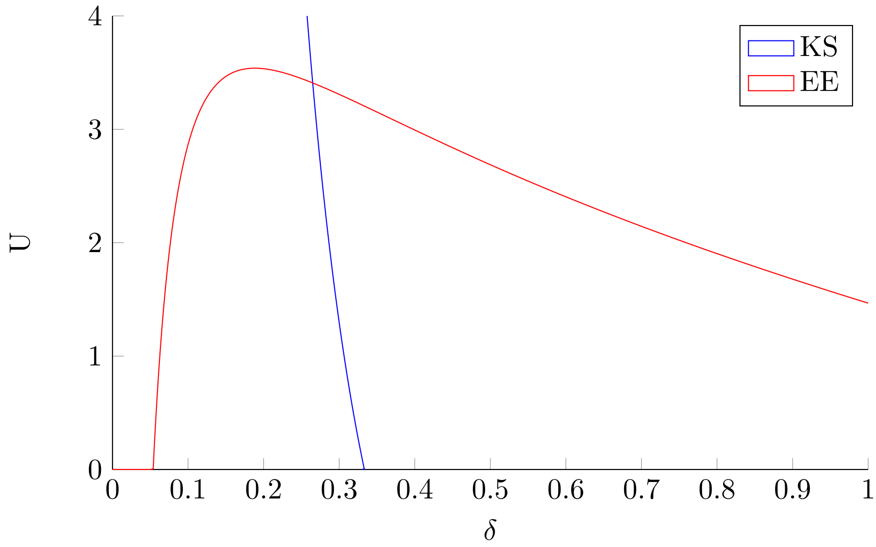
The knowledge spread's influence on equilibrium population N is twofold. An increase in δ raises the matching rate, but also diminishes the average extent of knowledge transmission. It turns out that this interaction makes the EE-locus hump shaped as is represented by the red curve in figure 3.5.[28]

The model's market solution is determined by the intersection point of the KS- and EE-locus.

[26] Both, the matching and innovation externality are discussed extensively in section 3.5.

[27] The expression for A in equation (3.18) is derived by making use of the condition for Steady-State Population, equation (3.16), and the fact that $M = N - U$. Furthermore, since workers in the city get the TFP in every period, independent of being matched or unmatched, the knowledge creation a_0 is divided by the interest rate r.

[28] Figure 3.5 shows a hump shaped relationship between δ and U. This coherence automatically implies a hump shaped connection between δ and N.



Chosen parameter values: $\alpha = 0.2, r = 0.2, \lambda = 0.15, e_0 = 1, e_1 = 6, a_0 = 1, t = 0.5$.
$\{\delta^*, U^*, N^*\} = \{0.265, 3.409, 7.516\}$

FIGURE 3.5: Steady State Equilibrium (Market Solution)

3.5 Social Inefficiencies

3.5.1 Social Planner's Solution

The Steady State Equilibrium gives rise to various inefficiencies. In order to explore their extent and interactions, equilibrium conditions are compared to the Social Planner's optimal choice. The Planner sets the knowledge spread δ and the population allocation N simultaneously.

The model contains three sources of externalities: Congestion externalities arise because individuals do not consider the impact of their location decision on congestion costs. Entering the city bears costs $t > 0$ for every worker living in the city. Matching externalities arise because individuals do not consider the impact of the chosen knowledge spread δ on the mass of unmatched workers. Increasing the own probability of being matched reduces the chances for everyone else. Most prominently, innovation externalities arise, because individuals do not consider the impact of the chosen knowledge spread δ on innovations and thus on the city's TFP A. The Definition of the Social Planner's Solution is as follows:

Definition 2: Social Planner's Solution

The Social Planner's Solution is an allocation $\{\hat{\delta}, \hat{U}, \hat{N}\}$ that satisfies the conditions:

(1) The Social Planner chooses the knowledge spread and the city's number of unmatched workers simultaneously in order to maximize the expected lifetime utility of a worker living in the city: $\{\hat{\delta}, \hat{U}\} \in \text{argmax}_{\delta,U} V_u(\delta, U)$.

(2) The condition for Steady State population is satisfied: $\hat{U} = \dfrac{\lambda}{\lambda + \alpha\hat{\delta}\hat{U}}\hat{N}$.

The exact optimality conditions are stated in appendix A. Here, we focus on the effects, the Social Planner takes into account when choosing knowledge spread δ and city population N. First, the optimality condition for the knowledge spread:

$$\frac{\partial V_u}{\partial \delta} = \underbrace{\frac{\partial A}{\partial \delta}}_{\substack{\text{innovation extern.}\\ (>0)}} + \underbrace{\frac{\partial A}{\partial U}\frac{\partial U}{\partial \delta}}_{\substack{\text{matching extern.}\\ (<0)}} + \frac{\partial e}{\partial \delta} + \underbrace{\frac{\partial e}{\partial U}\frac{\partial U}{\partial \delta}}_{\substack{\text{matching extern.}\\ (<0)}} = 0 \tag{3.19}$$

Matching externalities cause the equilibrium knowledge spread to be larger than socially optimal. Individuals are too broad in their acceptance of partners, because they do not consider that they prevent potentially more productive matches by lowering the mass of unmatched workers.
The innovation externality causes the equilibrium knowledge spread to be smaller than socially optimal. The chosen knowledge spread is too narrow as individuals do not consider their innovative output in interactions with diverse knowledge types.
Depending on the externalities' relative importance, it is possible that workers are either too picky or too generous in their choice of partners. The following analysis focuses on the case of a sufficiently important role of innovations such that the innovation externality outweighs the matching externalities and the equilibrium knowledge spread is smaller than socially optimal.

The inefficiencies' impact on location decisions can be elaborated by the Social Planner's optimality condition for the population allocation. The Planner chooses U such that net lifetime utility for the representative worker in the city is maximized, implying the first order condition:

$$\frac{\partial V_u}{\partial U} = \frac{\partial A}{\partial U} + \frac{\partial e}{\partial U} - t\frac{\partial N}{\partial U} = 0 \tag{3.20}$$

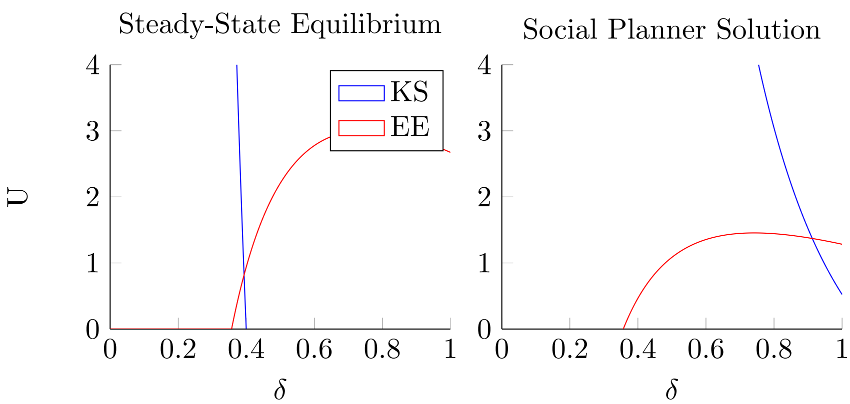
For the case of a smaller than socially optimal equilibrium knowledge spread, the inefficiencies concerning the location decision are again twofold: First, the congestion externality leads to a larger than socially optimal city size as workers do not consider their impact on city-wide congestion costs. Second, the inefficient choice of δ leads to restricted agglomeration forces as knowledge spillovers do not reach their optimal extent. This inefficiency causes smaller than socially optimal cities.

In summary, the presence of various externalities causes the equilibrium choices of δ and N to be inefficient. Direction and extent of inefficiencies crucially depends on the relative importance of innovation for the choice of δ and on the relative importance of knowledge spillovers for the choice of N.

3.5.2 Inefficiency Patterns

Depending on the parameter configuration, the market solution can exhibit three distinct inefficiency patterns in the workers' choices of knowledge spread and location: Overselectivity and Underpopulation, Overselectivity and Overpopulation as well as Underselectivity and Overpopulation. The existence of these inefficiency patterns is verified by construction.

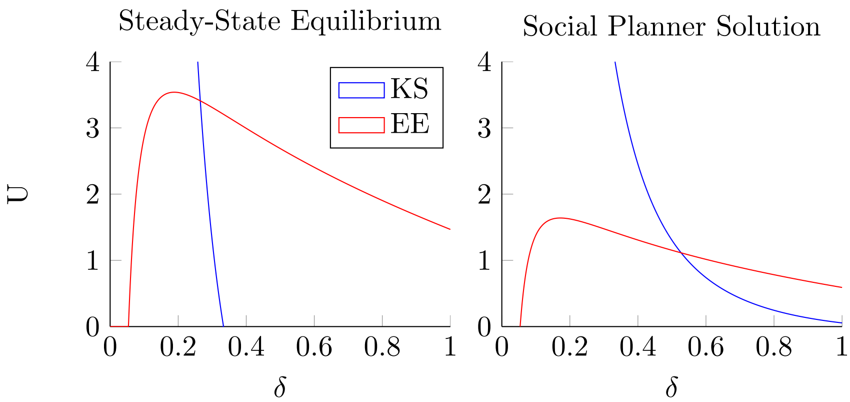
Case 1: Overselectivity and Underpopulation:
The case is marked by excessively narrow knowledge spreads (overselectivity) and smaller than socially optimal cities (underpopulation) in the market equilibrium. Overselectivity occurs if the relative importance of innovation is sufficiently high. The Social Planner realizes that workers refuse too many innovative face-to-face interactions. The city's underpopulation directly follows from the significant overselectivity, which does not allow the agglomeration force of innovation to develop its full extent. An example for the resulting Market Equilibrium and Social Planner Solution is depicted in figure 3.6.



$\alpha = 0.1, r = 0.2, \lambda = 0.8, e_0 = 2, e_1 = 10, a_0 = 2.7, t = 0.8.$
$\{\delta^*, U^*, N^*\} = \{0.3937, 0.8192, 0.8522\}, \{\hat{\delta}, \hat{U}, \hat{N}\} = \{0.9125, 1.3709, 1.5852\}$

FIGURE 3.6: Overselectivity and Underpopulation

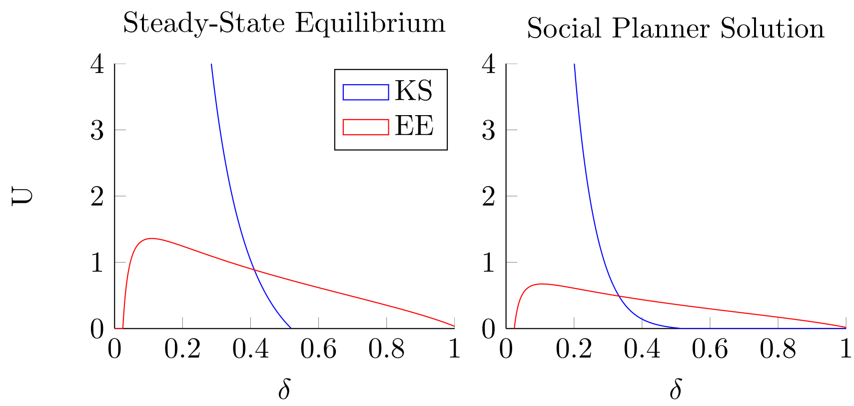
Case 2: Overselectivity and Overpopulation:
In this situation, workers are too picky in their choice of partners (overselectivity) and the resulting city size is larger than socially optimal (overpopulation). As in case 1, overselectivity is driven by the relatively high importance of innovation, which would be recognized by a Social Planner. However, in this case, overselectivity is not as pronounced such that the congestion externality outweighs the innovation externality. The Social Planner therefore chooses smaller city sizes than in the market equilibrium. The described Market Equilibrium and Social Planner Solution is illustrated in figure 3.7.



$\alpha = 0.2, r = 0.2, \lambda = 0.15, e_0 = 1, e_1 = 6, a_0 = 1, t = 0.5.$
$\{\delta^*, U^*, N^*\} = \{0.265, 3.409, 7.516\}, \{\hat{\delta}, \hat{U}, \hat{N}\} = \{0.527, 1.114, 1.985\}$

FIGURE 3.7: Overselectivity and Overpopulation

Case 3: Underselectivity and Overpopulation:
The third case is marked by underselectivity and overpopulation in the absence of innovation ($a_0 = 0$). As the innovation externality is zero in this example, the matching externality leads to chosen knowledge spreads that are larger than socially optimal. The interplay with congestion externalities again causes overpopulation in the market solution, i.e. cities are larger than socially optimal. The described scenario is illustrated in figure 3.8.



$\alpha = 0.16, r = 0.1, \lambda = 0.01, e_0 = 2.6, e_1 = 10, a_0 = 0, t = 0.9.$
$\{\delta^*, U^*, N^*\} = \{0.411, 0.886, 6.051\}, \{\hat{\delta}, \hat{U}, \hat{N}\} = \{0.332, 0.490, 1.767\}$

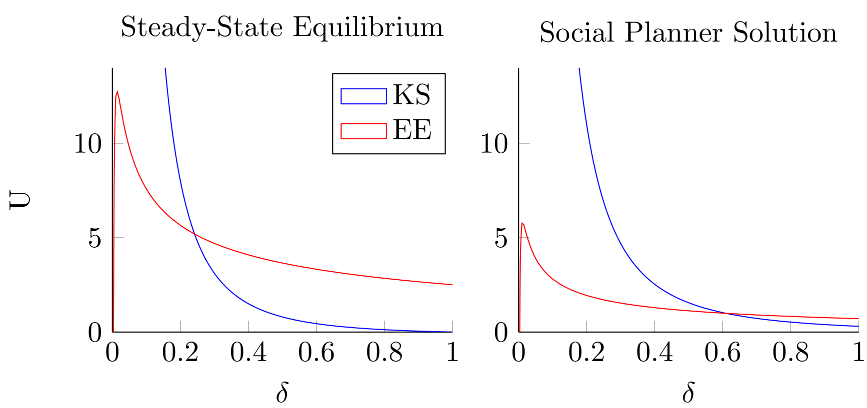
Figure 3.8: Underselectivity and Overpopulation

In summary, these results highlight the relevance of the innovation's agglomeration force. If the role of innovation is sufficiently important, workers are too picky in their choice of interaction partners. Depending on the externality's magnitude, cities can be smaller or larger than socially optimal. If innovation is irrelevant, however, excessively large knowledge spreads due to the matching externality cause larger than socially optimal cities.

3.5.3 Predicted Inefficiency Pattern

While the preceding section establishes the existence of three distinct inefficiency patterns, it is now determined which of these patterns is the empirically most relevant. In order to do so, the model's key parameters are calibrated using parameter values suggested by existing empirical work. The goal of this exercise is to establish the model's qualitative predictions.

The model's key parameters are the ones that govern the relative importance of knowledge creation and transmission, i.e. e_0, e_1 and a_0. It is impossible to measure the parameters directly as the effect of knowledge transmission on productivity is not observable. However, observed urban wage premiums can be used as an approximation to their respective impact. Measurements of static urban wage premiums are interpreted as an approximation for the role of the technology level and measurements of dynamic wage premiums as an approximation for the role of learning. This interpretation follows the logic that workers immediately benefit from the higher technology level upon moving to the city, while the buildup of know-how happens over time. For the model's qualitative predictions, the relative importance of innovation and learning is crucial. Recent studies from Carlsen et al. (2013), D'Costa and Overman (2014), De la Roca and Puga (2013) all find that the static wage premium is more pronounced than the dynamic wage premium. Carlsen et al. (2013) quantify the share of the static premium in urban lifetime earnings premiums as 2/3 while the remaining 1/3 stems from dynamic premiums. Controlling for observable and unobservable individual characteristics, they find the static premium for Norwegian urban workers to be 3.3%. Accordingly, the parameter a_0 is set to 0.033. The value of the dynamic premium from learning is consequently set to 0.0165. This value corresponds to $e_0 - e_1\delta/2$ in the model. We normalize $e_0 - e_1 1/2 = 0$. As δ is an endogenous variable, e_0 and e_1 cannot be uniquely determined, a reasonable parameter configuration in line with the previous findings is to set $e_0 = 0.02$ and $e_1 = 0.04$.



$\alpha = 0.3, r = 0.01, \lambda = 0.05, e_0 = 0.02, e_1 = 0.04, a_0 = 0.033, t = 0.1.$
$\{\delta^*, U^*, N^*\} = \{0.242, 5.193, 44.301\}, \{\hat{\delta}, \hat{U}, \hat{N}\} = \{0.610, 0.991, 4.587\}$

FIGURE 3.9: Predicted Inefficiency Pattern

The model's qualitative predictions are less sensitive to the remaining parameters as long as their magnitude is broadly in line with reality. The parameters of the arrival and separation rate of matches are chosen in line with the literature on matching in the labor market. Following Hobijn and Sahin (2009), the arrival rate α is set to 0.3 and the separation rate λ to 0.05. The interest rate r is 0.01 and congestion costs t are set to 0.1. For this parameter configuration, the resulting equilibrium pattern, as depicted in figure 3.9, is marked by overpopulation and overselectivity. The model predicts that workers choose a range of interaction partners that is too narrow and that city sizes are larger than socially optimal. This qualitative pattern is robust to any parameter configuration that is in line with the existing literature on urban wage premiums and matching rates.

3.6 Discussion

Knowledge Spillovers are assumed to be one of the major reasons for the existence of urban growth. The aim of this chapter is to further investigate the microfoundations of agglomeration economies in a search-theoretic model of knowledge spillovers in urban face-to-face interactions. The presented model explicitly incorporates two types of knowledge spillovers, the creation and transmission of knowledge and analyzes how they affect the migration decision of individuals. It consists of two asymmetric locations: The city and the periphery. Only the city provides individuals with the opportunity to exchange knowledge via face-to-face interactions. In each point of time, workers decide where to locate. Furthermore, workers in the city decide over the range of individuals they are willing to interact with. The intensities of knowledge creation and knowledge transmission in those interactions crucially depend on the heterogeneity of knowledge background of the interacting individuals: First, the individual buildup of skills through knowledge transmission increases in the similarity of knowledge backgrounds. And second, the creation of knowledge is independent of knowledge backgrounds.

The model's market solution exhibits three sources of inefficiencies. Since the city's created knowledge is a local public good, workers only focus on the buildup of personal skills when deciding about the range of individuals they accept to be matched with (innovation externality). Congestion externalities arise because individuals do not consider the impact of their location decision on city-wide congestion costs. Matching externalities arise because individuals do not consider the impact of their choice of the knowledge spread on the mass of unmatched individuals.

Depending on the parameter values, workers choose a range of matching partners that can be smaller or larger than socially optimal. The more important the role of knowledge creation in face-to-face interactions, the more likely it is that the chosen range of interaction partners is smaller than socially optimal. This means that people overestimate the importance of interacting with other individuals having a relatively similar knowledge background.

The interplay of agglomeration and dispersion forces determines the city size in the model. Moving to the city provides the chance to benefit from local knowledge spillovers. However, these face-to-face interactions come at the price of urban congestion costs. The model analysis shows that the inefficient decision on the range of individuals to interact with also causes socially inefficient city sizes. For parameter values based on previous empirical literature on urban wage premiums and matching, we find that workers are too picky in their choice of partners and that the resulting city size is larger than socially optimal. This result suggests that agglomeration forces do not reach their optimal extent, such that cities grow excessively fast.

Chapter 4

Entrepreneurship and City Growth

4.1 Introduction

Many urban economists have argued that entrepreneurial activity at the city-level plays a crucial role in generating subsequent economic growth.[1] The idea behind this presumption is the following: New start-up firms make ideas that are created through local knowledge spillovers available to the market. These ideas are then taken up and transformed by other firms in the surrounding. This dynamic environment in turn enhances the city's productivity and employment growth. But the location of the entrepreneur's start-up firm is certainly not exogenously predetermined. Entrepreneurs will choose the city in which they expect the highest potential profit. On average this will be the place where the highest overall productivity and employment growth are expected over the next years. Urban growth thus also has a positive effect on the rate of local entrepreneurship. Furthermore, there exist many potential city- or region-specific characteristics such as climatical conditions or certain properties of the city's population that influence local entrepreneurship and urban growth simultaneously.

For those reasons, the positive impact of local entrepreneurship on urban growth observed in the data[2] could be interpreted as a pure correlation between the two variables. Due to the inherent reverse causality, one should use core and caution when interpreting simple linear regression results. In this chapter[3], I address the endogeneity issue by applying an Instrumental Variable approach. I use the US patent activity during the 19th

[1]See Glaeser et al. (1992), Johnson and Parker (1996), Fritsch and Mueller (2004) and Acs and Armington (2004) among many others.

[2]The positive connection between the two variables becomes obvious in figure 4.1, in which the measure of entrepreneurship is on the x-axis and the measure of urban growth on the y-axis.

[3]The content of this chapter is based on Assmann (2015).

century as my instrument. More precisely, the regional distribution of registered patents between the years 1836 and 1873 across the United States is used as an exogenous source of variation for current rates of entrepreneurship at the level of US Metropolitan Statistical Areas (MSAs). Regions that exhibit a high number of patents during the 19th century had a higher probability of generating an environment of large companies that prevents potential entrepreneurs from entering the city[4] and also prohibits the intergenerational transmission of entrepreneurial skills. My empirical strategy is motivated by Chinitz (1961), who addresses the point by comparing the origins for various industries that were located in New York and Pittsburgh. One of the implications of Chinitz's work is that current rates of local entrepreneurial activity across the US are to some extent predetermined by the patent activity in a distant past.

I use publicly available patent data for the years between 1836 and 1873 provided by the United States Patent and Trademark Office to measure the regional distribution of patents at that time. This variable is my exogenous source of variation for the rate of entrepreneurship in US MSAs in 1993, which is measured by the city's number of newly founded start-up firms in 1993 with less than 20 employees.[5] I estimate the local entrepreneurship's impact on subsequent urban growth between the years 1993 and 2002, measured by the city's employment growth and payroll growth. The timeframe between 1993 and 2002 is applied for the empirical analysis to circumvent redefinitions of MSA borders and to rule out the impact of the financial crises in the US.

The empirical results are as follows. First, OLS regression results confirm previous findings that local entrepreneurial activity significantly accelerates subsequent urban growth. Increasing the city's number of small start-up businesses by one percent, raises subsequent employment growth by .1582 percent.[6] Second, using the IV approach, however, I find that the local entrepreneurship's impact on subsequent urban growth is substantially diminished.[7] Given that the instrument's orthogonality condition is not violated[8], results indicate that the predominant sentiment of the local entrepreneurship's importance for subsequent urban growth cannot be maintained. According to the results, a steady well-functioning urban industry structure without the birth of many start-ups

[4] The expressions "Metropolitan Statistical Areas (MSAs)" and "Cities" are used interchangeably in the following.

[5] I also present results for local entrepreneurship being proxied by the city's average establishment size.

[6] In this specification I include Census Division fixed effects and well-known city growth covariates like log population 1990, log density 1990, log share of city residents with a bachelor degree 1990, log house prices 1993, log mean January temperature and log mean July temperature. The results are qualitatively unaltered, if I use payroll growth instead of employment growth and the city's average establishment size instead of the births of small start-ups.

[7] The impact even turns into insignificantly negative, depending on how I define the sphere of influence of historical patents. The result does not depend on the proxies used for local entrepreneurship and urban growth and are robust to various robustness checks.

[8] I.e. the regional distribution of patent activity in the 19th century does not predetermine the regional distribution of economic growth between 1993 and 2002.

can as well be related to growth. Nevertheless, the results should be interpreted with a caution, especially in regards to the connection between historical patent activity and current urban growth, which is discussed extensively in section 4.5.
The structure of the chapter is as follows. Section 4.2 presents an overview of existing literature on the connection between local entrepreneurship and urban growth. Section 4.3 starts with a theoretical model that links local entrepreneurship with city growth. The model can be used as starting point for the empirical analysis of this paper. Subsequently, I describe the used data, OLS regression results are presented and the prevalent endogeneity issue is discussed. In section 4.4 I introduce my instrument for modern rates of local entrepreneurship. I examine how patent activity in 19th century has predetermined later firm size distributions across US cities. Conclusively I explain the concept of spatial rings that determines by which patents a city is assumed to be influenced. Section 4.5 describes in detail the results of the proposed instrument at the first and second stage of the IV regression. I also conduct several robustness checks to test the sensitivity of the results. Section 4.6 summarizes and concludes.

4.2 Empirical Strategy

In this chapter, I use an Instrumental Variable approach to tackle the potential issue of reverse causality. I propose the regional distribution of US patent activity during the 19th century[9] as exogenous source of variation for modern levels of entrepreneurship across US MSAs. I argue that cities with more registered patents during that time exhibit a higher probability of generating an environment of large companies that prevents potential entrepreneurs from entering the city. This line of thought is motivated by Chinitz (1961) and related to Saxenian (1994), who claim that the entrepreneurs' scepticism towards large companies arises from the "aura of second-class citizenship"[10] that surrounds entrepreneurs in cities dominated by those large enterprises. Chinitz (1961) further ascertains that entrepreneurs face more financial constraints in "large company" cities, since financial institutions are more restrained towards small-scale businesses. In equal measure "large company" cities cause potential entrepreneurs difficulties in the access to intermediate goods as small start-up businesses heavily depend on outsourced tasks that are supplied by the local economy. As entrepreneurs presumably have information about those financial constraints and potential difficulties in the access to local inputs, they have reasons to decide against those cities.

[9]The US patent activity is measured by the number of registered inventions between the years 1836 and 1873. See section 4.4 for details.
[10]The quote stems from Chinitz (1961).

To the best of my knowledge there exist only two approaches which try to circumvent the inherent endogeneity by identifying exogenous sources of variation for modern levels of local entrepreneurship. Most prominently, Glaeser et al. (2012) propose the proximity of US MSAs to historical coal mines as an instrument for modern levels of local entrepreneurship. They argue, that today's industry structure of cities close to coal mines is dominated by large enterprises operating in the steel industry. The prevalence of those large steel companies discourages potential entrepreneurs to enter the city. IV regression results presented by Glaeser et al. (2012) indicate that local entrepreneurship is even more important for subsequent growth than respective OLS estimates suggest. Lee (2014) instead uses homestead exemption levels set by state bankruptcy law in 1975 to instrument for local entrepreneurship. His approach relies on the idea that cities in US states with higher exemption levels attract more potential entrepreneurs, since they enable entrepreneurs to avoid paying creditors back part of the loan. Lee (2014) finds that using IV regression does not significantly alter OLS estimates on the impact of local entrepreneurship on subsequent urban growth. Both approaches leave room for criticism. Cities that are overproportionally assembled with companies from the steel industry might not only deter potential entrepreneurs, but also face mature disadvantages in the current growth process, causing the IV estimates presented by Glaeser et al. (2012) to be potentially upward biased. Cities in states with high exemption levels might additionally benefit from other business-friendly legislations as business policies within states are not independent from each other, causing the presented IV results by Lee (2014) to be potentially upward biased as well. Furthermore, the instrument proposed by Lee (2014) provides no variation at the city-level, but only at the US state-level.

As for Glaeser et al. (2012) and Lee (2014), the validity of my proposed instrument crucially depends on the assumption that it is not related to today's urban growth. For my approach, this requires that drivers of patent activity in the past have no causal impact on the drivers of today's city growth. I argue that my instrument has the advantage of predetermining the city's size distribution of establishments without predetermining the city's industry structure. This is more of a concern for the strategy proposed by Glaeser et al. (2012), as cities close to coal mines are dominated by companies from the declining steel industry. The work of Boldrin and Levine (2013) supports my argument that patent activity in the distant past is suitable as an exogenous source of variation for modern rates of local entrepreneurship. They find that there exists no empirical evidence that the number of officially registered patents has a positive effect on current or subsequent productivity and argue that the emergence of patents is not equivalent to the emergence of innovations. Boldrin and Levine (2002) provide theoretical support for the empirical observations presented by Boldrin and Levine (2013). Their model illustrates how the positive aspects of patents (the right to own and sell ideas) can be compensated by their negative aspects (the right to control the use of ideas).

4.3 Relationship between Local Entrepreneurship and Urban Growth

4.3.1 A Simple Theoretical Model

Following Glaeser et al. (1992), Henderson et al. (1995) and Lee (2014), I use a simplified model of local entrepreneurship and subsequent urban economic growth that guides the empirical work.

Consider an economy consisting of C cities. Each city $i = 1, ..., C$ accommodates one representative firm that produces output $Y_{i,t}$ at time period t. Each of these representative firms uses a decreasing-returns-to-scale production technology $F(\cdot)$ to produce a homogeneous good.

$$Y_{i,t} = F(L_{i,t}) = A_{i,t} L_{i,t}^{\alpha}, \qquad \alpha \in (0,1) \tag{4.1}$$

For simplicity, I assume that each firm uses only one input: Labor $L_{i,t}$. Productivity of labor is determined by the city-wide technology level $A_{i,t}$. Assuming perfect labor mobility leads to equal market wages across all cities $i = 1, ..., C$, i.e. $w_{i,t} = w_t$. Market wages grow at the exogenous rate of $1 + g$ over time, with the growth rate g being determined by the national economy. City i's number of workers during time periods t and $t + 1$ is then determined by the Neoclassical first order expressions

$$\begin{aligned} w_t &= \alpha A_{i,t} L_{i,t}^{\alpha-1} \\ w_{t+1} &= (1+g) w_t = \alpha A_{i,t+1} L_{i,t+1}^{\alpha-1}, \end{aligned} \tag{4.2}$$

with the wage being on the lefthand side and the worker's marginal product being on the righthand side.

Taking logs and subtracting the first expression from the second in (4.2) yields

$$\ln(1+g) = \Delta \ln(A_{i,t}) + (\alpha - 1) \Delta \ln(L_{i,t}), \tag{4.3}$$

where $\Delta \ln(A_{i,t}) = \ln(w_{A,t+1}) - \ln(A_{i,t})$ and $\Delta \ln(L_{i,t}) = \ln(L_{i,t+1}) - \ln(L_{i,t})$.

Equation (4.3) then can be rewritten as

$$\Delta \ln(L_{i,t}) = \frac{\Delta \ln(A_{i,t}) - \ln(1+g)}{1-\alpha}, \tag{4.4}$$

thus employment changes within cities are entirely driven by city-wide changes in technology level $A_{i,t}$. Assuming that the level of local entrepreneurship has a decisive impact on subsequent urban economic growth, I specify the technology growth process as follows:

$$\Delta \ln(A_{i,t}) = G(e_{i,t}, X_{i,t}, X_{r,t}). \tag{4.5}$$

The level of local entrepreneurship in period t is denoted by $e_{i,t}$. The vector $X_{i,t}$ summarizes all city-wide characteristics and the vector $X_{r,t}$ all region-wide characteristics that have an impact on the growth of the urban technology level.[11] From equation (4.4) it immediately follows that urban employment growth $\Delta \ln(L_{i,t})$ is as well determined and approximated by these variables, i.e.

$$\Delta \ln(L_{i,t}) \approx G(e_{i,t}, X_{i,t}, X_{r,t}). \tag{4.6}$$

The main goal of the empirical analysis of this chapter is to detect how the initial level of local entrepreneurship $e_{i,t}$ affects subsequent employment growth $\Delta \ln(L_{i,t})$ within a city, controlling for differences in all other potential factors of influence $X_{i,t}$ and $X_{r,t}$, i.e. the empirical goal is to measure $\frac{\partial \Delta \ln(L_{i,t})}{\partial e_{i,t}}$.

4.3.2 Data Description

This section introduces the data used to measure the two main variables of interest: Local entrepreneurship and subsequent economic growth across US MSAs. For the analysis, I use the period from 1993 to 2002 in order to avoid redefinitions of MSA borders and to rule out the impact of the financial crises, as the question whether local entrepreneurship is able to mitigate the cities' consequences of country-wide shocks (e.g a shock to the financial system) is beyond the scope of this chapter.[12] A timeframe of 10 years should be sufficient to identify economic consequences that are due to differences in the initial level of local entrepreneurship.

My proxy for entrepreneurship in 1993 is the city's number of newly founded establishments (start-up businesses) with less than 20 employees in 1993. Focusing on small start-up businesses captures the public notion of entrepreneurship most adequately: The more innovative start-up firms are emerging during the near past, the higher the city's entrepreneurial activity. Concentrating on establishments with less than 20 employees rules out the inclusion of expansions by large already existing establishments, as those are not considered as entrepreneurial activity. In the following, I call the city's number

[11] City i is located within region r. For the empirical analysis one can think of a city's superordinated region r as US State or US Census Division.

[12] This is an issue that could be analyzed in a stand-alone paper.

of those newly founded establishments "Start-up Births in 1993". The required information needed to construct the variable are taken from the publicly available Statistics of U.S. Business (SUSB) Employment Change Data. For each year, the data set indicates the number of newly founded establishments in each MSA across the United States. It further contains information about the cities' establishment size distributions by classifying establishments into three groups: less than 20, 20-499 and 500 or more employees. My alternative proxy for local entrepreneurship in 1993 is the city's average number of employees per establishment. The higher the city's number of start-up firms, the lower the mean number of employees per establishment. Thus more local entrepreneurial activity is associated with smaller average establishment size. This negative connection is discussed in detail by Glaeser et al. (2010). In order to construct the variable "Average Establishment Size in 1993" I need information on total employment at the MSA-level. The required data are derived from the publicly available Statistics of U.S. Business (SUSB) Annual Data. Information on the number of establishments within a city can be drawn from the publicly available Statistics of U.S. Business (SUSB) Employment Change Data. I present OLS and IV results, using average establishment size as proxy for entrepreneurship, in the Robustness Checks in section 4.5.4.3. However, I focus on the city's number of start-up births as my main proxy for local entrepreneurship as it measures entrepreneurial activity more directly.

Consistent with the literature on local entrepreneurship and subsequent urban growth and also consistent with the theoretical model presented in section 4.3.1, I use the variable "Employment Growth 1993-2002" as my main proxy for urban economic growth. Implicitly it is assumed that a city's positive economic development is primarily reflected in the creation of new jobs. An alternative proxy, for which results are presented in the following, is the city's "Payroll Growth 1993-2002". This variable is less frequently used in the urban growth literature but has the advantage of incorporating productivity changes of workers into my measure of local economic growth and thus provides additional information. All required information to construct the proxies for urban economic growth between 1993 and 2002 are contained in the SUSB Annual Data. Table 4.1 presents summary statistics for all described variables. My data set consists of 321 Primary Metropolitan Statistical Areas (PMSAs)[13] according to the US Census definition in 2000. The only MSAs excluded from the sample are Anchorage, AK and Honolulu, HI, as I want to concentrate on cities with a homogeneous surrounding, i.e. I do not want to include extensive outliers due to their geographical location. Unless otherwise stated, I refer to this sample of 321 MSAs located on the US mainland.

US MSAs had an average employment of 248,840 in 1993 with Enid, OK having the lowest and Los Angeles-Long Beach, CA having the highest level of employment. These

[13]Whenever I use the expression MSA I refer to Primary Metropolitan Statistical Areas (PMSAs) and not to Combined Metropolitan Statistical Areas (CMSAs).

Summary Statistics	**Min**	**Max**	**Mean**	**S.D.**
Employment 1993	18,764	3,495,140	248,840.2	436,996.4
Employment 2002	20,452	3,791,362	295,415.2	506,870.3
Payroll 1993	335,607	$1.23 \cdot 10^8$	6,464,595	$1.32 \cdot 10^7$
Payroll 2002	501,290	$2.13 \cdot 10^8$	$1.09 \cdot 10^7$	$2.22 \cdot 10^7$
Log Emp. Growth 93-02	-.2605	.5505	.1609	.1006
Log Payroll Growth 93-02	.0456	.9667	.4699	.1408
Start-up Births 1993	105	20,602	1,368.06	2,376.71
Emp.Share in Start-ups 93	.0260	.1235	.0575	.0178
Av.Est. Size 1993	8.29	24.13	14.82	2.53

Note: The sample contains 321 US Metropolitan Statistical Areas. *Log Employment Growth 1993-2002* is derived by $Log(Employment_{2002}) - Log(Employment_{1993})$. *Log Payroll Growth 1993-2002* is derived by $Log(Payroll_{2002}) - Log(Payroll_{1993})$. The *Employment Share in Start-ups 1993* is the number of residents working in start-ups in 1993 divided by the level of employment in 1993. The average number of workers per start-up is assumed to be 10.

Table 4.1: Summary Statistics on Entrepreneurship 1993 and Urban Growth 93-02

two MSAs also keep the extreme values in the year 2002. Overall, the average employment level across MSAs increased to 295,415 employees in 2002. Between 1993 and 2002 the average MSA payroll increased from about 6.46 million US dollars to 10.09 million US dollars with New York, NY and Enid, OK having the highest and lowest payroll in both periods. Average employment growth over the ten years was roughly 16 percent across MSAs with Kankakee, IL having the lowest and Las Vegas, NV having the highest employment growth rate during that timeframe. The highest payroll growth was observed in Austin-San Marcos, TX, where the payroll almost doubled. The lowest payroll growth rate was measured in Flint, MI. My two measures of urban economic growth, employment growth and payroll growth, have the desired property to be highly correlated with a correlation coefficient of .9173. The same is true for the two proxies of local entrepreneurship, Start-up Births in 1993 and Average Establishment Size in 1993. Here, the correlation coefficient is equal to -.8182.[14] Start-up Births in 1993 indicate that Elmira, NY was the city with the lowest level of local entrepreneurial activity in 1993 with only 105 new start-ups, whereas New York, NY was the city with the highest entrepreneurial activity with 20,602 new start-up businesses. For Start-up Births one has to be careful in interpreting the results, as they have no validity without controlling for the city's population.[15] The city's share of employees working in small start-up businesses, however, contains more meaningfulness by its own.[16] Focusing on that variable, it turns out that Rochester, MN is the least entrepreneurial MSA with only 2.6 percent and Naples, FL is the most entrepreneurial MSA with over 12 percent being employed in newly founded start-ups. The city's Average Establishment Size indicates

[14]The negative correlation is desired here, as more entrepreneurial activity is associated with low average establishment size.

[15]Larger cities generate more start-ups just by construction.

[16]To construct this variable, I assume that the average number of employees per start-up is 10. The ranking of MSAs according to this variable does not depend on the guess about the average number of employees per start-up. It is sufficient to relate the number of start-ups to the number of employees in the city. Therefore the assumption of 10 employees is not critical to the presented results

that Rochester, MN has the lowest entrepreneurial activity with over 24 workers per establishment and Barnstable-Yarmouth, MA has the highest entrepreneurial activity with only around 8 workers per establishment.

4.3.3 Correlation between Local Entrepreneurship and Urban Growth

Figure 4.1 illustrates the correlation between the rate of local entrepreneurship in 1993[17] and subsequent employment payroll growth between 1993 and 2002. The sample includes 321 US MSAs as described in section 4.3.2. In both pictures, the slopes of the fitted OLS regression lines have the expected sign. Increasing the 1993 employment share in newly founded start-ups by one percent increases employment growth and payroll growth by .1917 percent and .2555 percent, respectively. Using standardized beta-coefficients, the results imply that increasing the 1993 share of employment in small start-up firms by one standard-deviation raises subsequent employment and payroll growth by .55 and .52 standard-deviations, respectively.

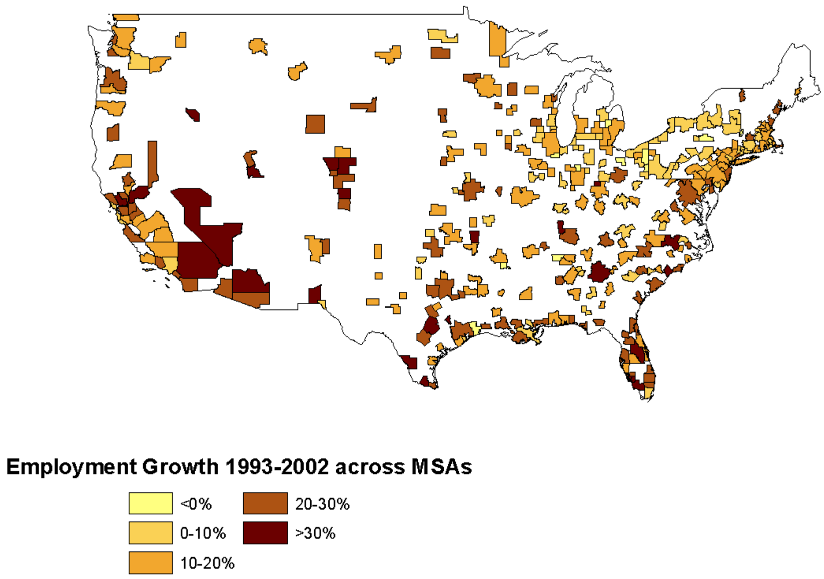
Note: The sample contains 321 US Metropolitan Statistical Areas. For data construction see the notes to table 4.1.

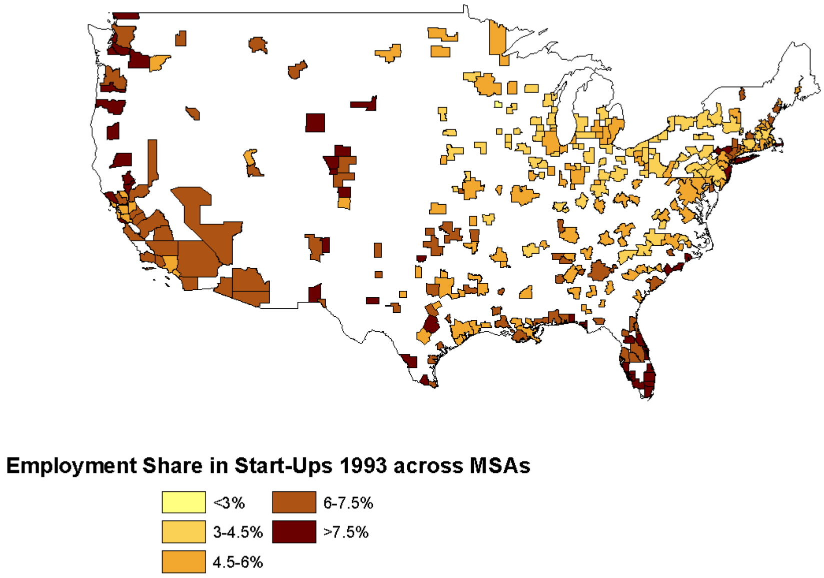
FIGURE 4.1: Entrepreneurship 1993 and Urban Growth 93-02

Figure 4.2 illustrates the spatial coincidence of entrepreneurial activity in 1993[18] and employment growth 1993-2002 in all 321 MSAs across the US mainland. The upper and lower panel informally show that cities which exhibited considerable entrepreneurial activity in 1993 had on average higher employment growth during the next ten years. This relation explains the positive connection between the two variables from figure 4.1.

[17]Here, entrepreneurship is proxied by the share of employment in newly founded establishments with less than 20 employees. This is done as the absolute number of start-ups has no useful interpretation by its own. In an OLS regression I can as well use the absolute number of newly founded start-ups as I control for the city population.

[18]Local Entrepreneurship is measured by the share of employees in small newly founded establishments with less than 20 employees





Note: MSA borders are chosen according to the US Census definitions of 2000.

FIGURE 4.2: (a) Employment Growth 93-02 and (b) Employment Share in Start-Ups 1993

4.3.4 OLS Relationship between Local Entrepreneurship and Urban Growth

In this section, I control for various factors that potentially influence local entrepreneurial activity and subsequent urban economic growth simultaneously. The sample used remains unchanged with 321 MSAs on the US mainland. The econometric model follows from the theoretical model developed in section 4.3.1:

$$Growth_i^{93-02} = \alpha + \beta\, ln(Entrepreneurship_i^{1993}) + \gamma\, ln(Employment_i^{1993}) + \\ + \delta\, CensusDivision_i + OtherControls_i{}'\eta + \epsilon_i \tag{4.7}$$

Subscript i indexes for MSAs. The variable $Growth_i^{93-02}$ represents the measure for urban economic growth during the years 1993-2002, which is proxied by log employment growth (as my main proxy) or log payroll growth (as an alternative measure for urban growth). Local entrepreneurial activity is measured by the MSA's number of newly founded establishments with less than 20 employees.[19]

A: Log Emp.Growth	**(1)**	**(2)**	**(3)**
Log Start-up Births 93	.1988***	.1530***	**.1582*****
	(.0254)	(.0230)	(.0298)
p-value	.000	.000	.001
initial empl.	Yes	Yes	Yes
C.D. fixed effects	No	Yes	Yes
City Growth covariates	No	No	Yes
N	321	321	321

B: Log Payroll Growth	**(1)**	**(2)**	**(3)**
Log Start-up Births 93	.2734***	.2223***	**.2268*****
	(.0335)	(.0311)	(.0415)
p-value	.000	.000	.001
initial empl.	Yes	Yes	Yes
C.D. fixed effects	No	Yes	Yes
City Growth covariates	No	No	Yes
N	321	321	321

Note: The sample contains 321 Metropolitan Statistical Areas in the US. Cluster Robust Standard Errors are reported in parentheses. ***Significance at 1 percent level, **Significance at 5 percent level, *Significance at 10 percent level. The nine U.S. census divisions are Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic and New England. City Growth Covariates include log mean January temperature, log mean July temperature, log city population 1990, log density 1990, log 1990 share of residents with bachelor degree or higher, log house prices in 1993.

TABLE 4.2: OLS Regression - Entrepreneurship 1993 and Urban Growth 93-02

Table 4.2 illustrates the results of OLS regressions with log employment growth and log payroll growth as dependent variables and the log number of start-up births in 1993 as

[19] I focus on the absolute number of start-ups, as I control for city size. An OLS regression with the share of employees in start-up firms as independent variable delivers exactly the same results.

independent variable of interest. Due to space limitation, only the estimated coefficients for that variable are presented here. In specifications (1)-(3), I include different sets of control variables to test the sensitivity of my OLS results to their inclusion.

Panel A shows the results when measuring urban growth by log employment growth, whereas Panel B uses log payroll growth as proxy for city growth. Contrary to Glaeser et al. (2012), I continuously use clustered standard errors with observations being clustered according to the nine US Census Divisions.

Specification (1) in column 1 only controls for the MSA's initial log employment level in 1993. By controlling for the level of employment in 1993, possible concerns regarding correlations being driven by mean reversion can be excluded. Increasing the city's number of newly founded start-up firms by one percent increases subsequent employment and payroll growth by .1988 and .2734 percent, respectively.

In specification (2) I additionally include fixed effects for the nine US Census Divisions.[20] Thereby I control for differences in economic growth due to varieties in regional characteristics. Using US Census Division fixed effects slightly lowers the effects of local entrepreneurial activity on economic growth, meaning that at least part of the observed correlation turns out to be a result of differences in regional characteristics. But the impact of entrepreneurial activity remains highly significant, independent of the proxy used for urban economic growth. Increasing the number of newly founded start-ups with less than 20 employees by one percent raises subsequent employment and payroll growth by .1530 and .2223 percent, respectively.

The results for core specification (3) are presented in column 3. This specification additionally includes standard control variables from the literature on urban growth. Those include the log mean January and log mean July temperature (to control for local (dis)advantages due to climatical conditions), the log share of people with a bachelor degree or higher in 1990 (to control for highly skilled individuals selecting fast growing cities), the log city population in 1990, the log population density in 1990 and log housing prices from the year 1993 (to control for core attributes of metropolitan areas). Data on average temperatures in January and July are derived from the website www.usclimatedata.com.[21] Information on overall population, population density and share of high skilled individuals are drawn from the US Census in 1990. Data on housing prices in 1993 are derived from the publically available HPI-Index.[22] The inclusion of those control variables has no major effect on the local entrepreneurial activity's impact on subsequent urban growth. All coefficients keep their algebraic sign and remain almost unaltered in magnitude. Increasing the 1993 number of newly founded start-ups in a

[20]The nine US census divisions are Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic and New England.

[21]Whenever climate data for a specific city were not accessible, I chose available data from the city closest to it.

[22]If house prices for a MSA were not accessible, I chose available data for the city closest to it.

city by one percent raises subsequent employment and payroll growth in that city by .1582 and .2268 percent, respectively.

For an alternative economic interpretation of the measured results, one can also look at the estimated beta-coefficients, which are normalized by standard-deviations. These standardized beta-coefficients can be found in Table B.1 in appendix B. Increasing the city's number of start-up births in 1993 by one standard-deviation raises subsequent employment and payroll growth by .45 and .47 standard-deviations, respectively. Here I refer to core specification (3), in which all relevant control variables are included. Those results illustrate that the magnitude of the entrepreneurship's impact on subsequent urban growth measured in an OLS regression is economically meaningful and substantial. Overall, OLS results indicate that local entrepreneurship has a significant impact on the rate of subsequent urban economic growth, even after controlling for factors that potentially influence both variables simultaneously. The results also indicate that the measured impact is robust to the usage of various proxies for urban economic growth between the years 1993 and 2002.

4.3.5 Potential Reverse Causality

The inclusion of control variables dissolves at least part of the concerns when trying to measure the causal relationship between local entrepreneurship and subsequent urban growth. But one major issue that is not resolved yet is the potential reverse causality. For OLS to deliver unbiased results, one has to assume that the city choice of potential entrepreneurs is not in the least affected by the city's growth. Moreover, one has to control for all sources that affect local entrepreneurship and urban economic growth simultaneously. But knowing that the economic development inside a city can be decisive for the success or failure of newly founded businesses makes the assumption of exogenous city choice unrealistic. Entrepreneurs will predominantly choose cities in which they anticipate the highest profit for their firm. On average this is the city where they expect the highest productivity and employment growth during the next years.

$$\begin{tikzcd} Entrepreneurship_c^{1993} \arrow[r, bend left] & Growth_c^{1993-2002} \arrow[l, bend left] \end{tikzcd}$$

FIGURE 4.3: Potential Reverse Causality between Entrepreneurship 1993 and Urban Growth 93-02

That is, the presence of newly founded firms stimulates the local economy and generates new jobs but a rapid growing city also attracts more entrepreneurs with innovative business ideas. If the relationship between the two variables depicted in figure 4.3 is true, then the OLS regression results from table 4.2 are all biased and partially driven by selection bias.[23] The only way to give the coefficients a causal interpretation is to find an exogenous source of variation for the rate of local entrepreneurship in 1993 that neither has a causal effect on, nor any other statistical relationship with subsequent urban economic growth, as is represented in figure 4.4.[24] The described IV approach would resolve the issue of reverse causality and deliver unbiased coefficient estimates.

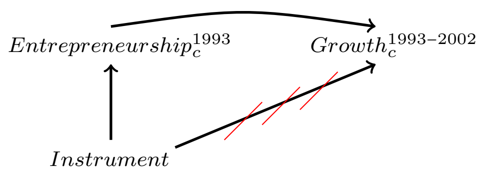


FIGURE 4.4: Exogenous Source of Variation

4.4 US Patent Activity 1836-1873

I propose the regional distribution of patent activity across the United States during the 19th century as exogenous source of variation for modern rates of local entrepreneurship. Following Chinitz (1961) and Saxenian (1994), I argue that the city's industrial history predetermines subsequent entrepreneurial activity.

4.4.1 Patent Data

The United States Patent and Trademark Office stores the entire US patent activity from 1790 till present. For the examined period from 1836 to 1873 the data set additionally contains publicly available information on the name of the invention, the exact registration date and the inventor's name and place of residence. Altogether the data set encompasses 146,119 utility patents that were officially registered between the years 1836 and 1873. Unfortunately no information on the exact ZIP code for the inventor's place of residence are available. The only information existing is the inventor's home town and home state. This inaccuracy brings about several consequences. In some cases the entire name or notation of towns in the US has changed since the 19th century. In these cases the towns' names were adjusted to today's name or notation whenever they

[23]If we presume that selection bias is the main problem, i.e. entrepreneurs sort into fast growing cities, then the OLS estimates should be upward biased.

[24]The graphical representation depicted in figure 4.4 is borrowed from Cameron and Trivedi (2005).

Summary Statistics	**Min**	**Max**	**Mean**	**S.D.**
County	0	14686	45.33	359.41
MSA	0	20149	353.93	1421.92

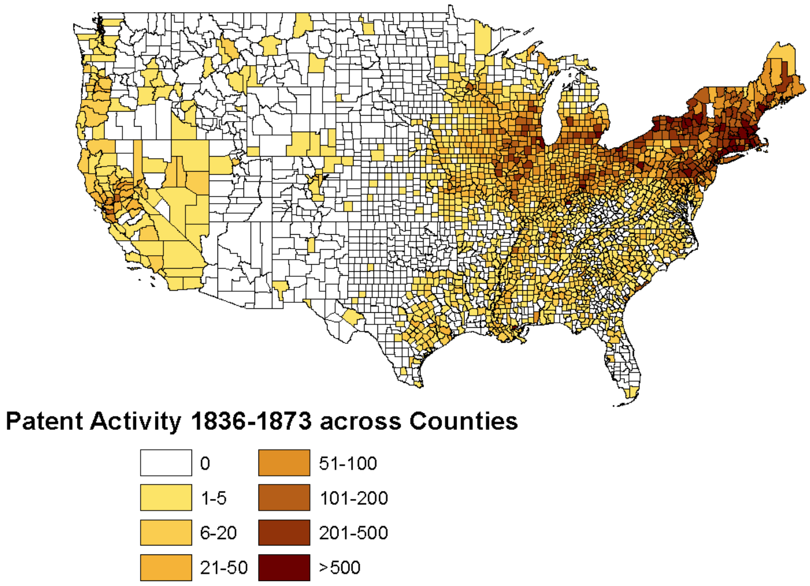
Table 4.3: Summary Statistics on US Patent Activity 1836-1873

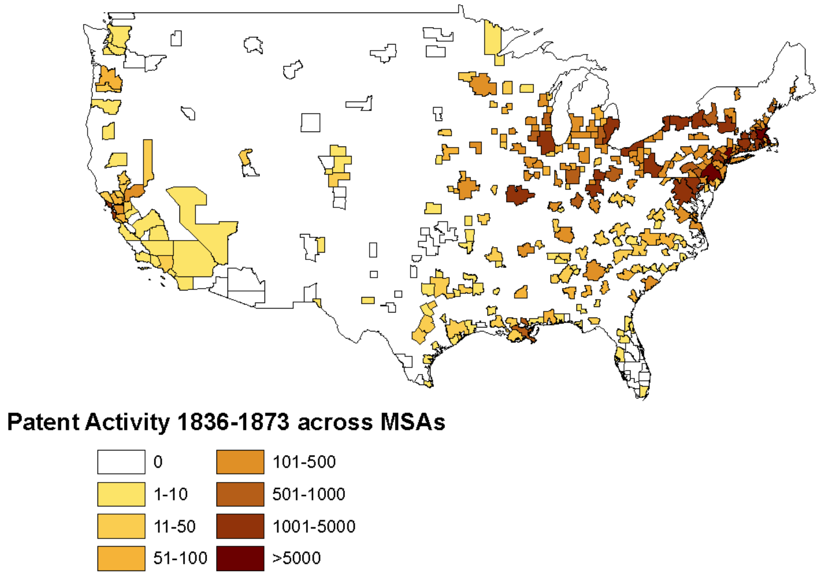
Patent Activity 1836-1873	**County**	Σ_{County}
1.	New York County, NY	14686
2.	Philadelphia County, PA)	6869
3.	Suffolk County, MA	6413
4.	Kings County, NY	4176
5.	New Haven County, CT	2804
	MSA	Σ_{MSA}
1.	New York, NY	20149
2.	Boston, MA	10437
3.	Philadelphia, PA	8131
4.	Chicago, IL	3594
5.	Albany-Schenectady-Troy, NY	3121

Table 4.4: Counties and MSAs with highest number of patents 1836-1873

were unambiguously reconstructable. Furthermore, some towns that existed between 1836 and 1873 do not exist anymore. Nevertheless most of these "ghost towns" can be matched to today's US counties.[25] With only few exceptions the provided information in the data set are sufficient to match nearly all registered patents between the years 1836 and 1873 to US counties by today's definition. Overall, I lose 415 patents from inventors living in the US, because they could not be matched to US counties. Furthermore, I drop 4,350 patents registered by inventors living outside the United States as those patents are not relevant for my analysis. This reduction leaves me with 140,939 observations on patents that were registered between the years 1836 and 1873 and can be unambiguously matched to US counties by today's definition. Table 4.3 displays summary statistics for the patent activity in 3,109 counties and 321 MSAs by today's definition. The average number of registered patents per county was about 45 and about 354 per MSA, respectively. The large standard deviation for patents per county and MSA display the high variation in the regional distribution of patents. Table 4.4 lists the five counties and MSAs with the highest number of registered patents in the US between 1836 and 1873. Figure 4.5 illustrates the heterogeneity in patent activity among (a) US counties and (b) US MSAs.

[25] County borders are defined according to the US Census 2000. Whenever I use the expression counties by today's definition I refer to these county borders.





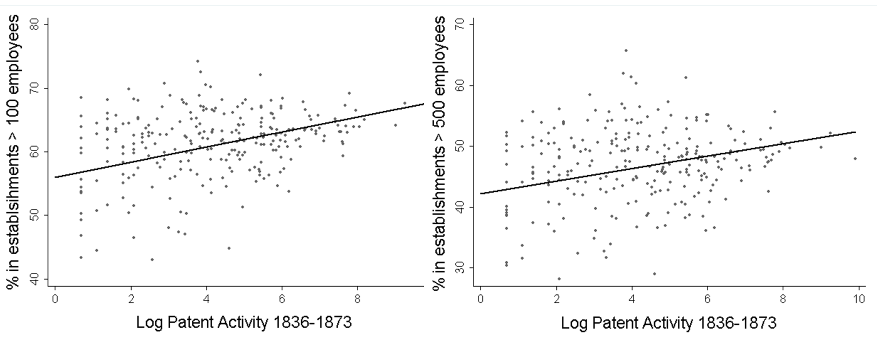
Note: County and MSA borders are chosen according to the US Census definitions of 2000.

FIGURE 4.5: (a) Number of patents by county 1836-1873 and (b) Number of patents by MSA 1836-1873

4.4.2 Implications for Today's Industry Structure

Above, I pick up the argument of Chinitz (1961) and Saxenian (1994), who claim that a city's industrial history predetermines subsequent levels of local entrepreneurship. I suggest that regions with a high occurrence of registered patents during the time period between 1836 and 1873 were exposed to a higher probability of becoming a deterring "big company" city for entrepreneurs, which is the starting point of my empirical strategy.

Obviously, only very few registered patents have the potential to result in successful business concepts. And even though the registered invention does, it is not clear whether the resulting company persists and grows over time and still exists as a large establishment today. But the probability of having generated large establishments is higher for those regions that had a high patent activity between the years 1836 to 1873. The emergence of patents in the 19th century can also be interpreted as an indicator for the state of economic development of US regions at that time. According to Lamoreaux (2010), the end of the 19th century was exactly the time when many US cities started to become dominated by large-scale enterprises. The economic development at the beginning and middle of the 19th century was thus decisive for the question which cities experienced a rising of large establishments. Information from the Statistics of US Business (SUSB) Annual Data lend a support for the presumed relationship. The data set contains information on all US MSAs' shares of employees working in establishments with more than 100 or 500 employees respectively.



Note: The sample contains 321 US Metropolitan Statistical Areas. Patents are counted within today's MSA borders.

FIGURE 4.6: US Patent Activity 1836-1873 and Establishment Size 1993

As figure 4.6 illustrates, patent activity in the 19th century still has an impact on the relative importance of large-scale companies in 1993. Cities that exhibited more registered patents between 1836 and 1873 have a significantly higher share of residents working in large-scale establishments. Increasing the number of registered patents in the 19th century by one percent increases the city's share of residents in companies with more than 100 employees by .120 percent. For establishments with more than 500 employees the corresponding point estimate is .102 percent. Both estimates are significant at a 1 percent significance level. The results are robust to the inclusion of relevant control variables.

4.4.3 Assumptions about the patents' sphere of influence

Taking the number of historical patents within today's MSA borders is a first order approximation at best for the status of that city's economic development during the 19th century. The borders that we observe today are the result of a long adjustment process which had absolutely no relevance at that time. In fact, the growth of most MSAs was starting on the basis of very narrow geographical borders. In the course of years people from the surrounding were then attracted to cities which led them to grow in population and spatial dimension. Furthermore, there is no reason to believe that the influence of patents should stop at the city border, especially as these borders had no relevance in the 19th century. As an alternative and preferred definition of historical MSA borders, I thus take the central points of MSAs by today's definition and construct spatial rings around those points. Every patent that lies within such a spatial ring[26] is assumed to have influenced the city's subsequent firm size distribution. These alternative definitions of city borders also determine the presumed sphere of influence of 19th century patents on surrounding cities. This approach, nevertheless, has some drawbacks. First, it is not clear whether the central points of MSAs by today's definition were the starting point for urban sprawl beginning in the 19th century, which is implicitly presupposed when I construct spatial rings around today's MSAs' central points. On average, however, the arising inaccuracy is quite trivial. Second, some patents lie within multiple spatial rings and thus count for more than one MSA. This problem is not as severe as it first seems since I do not rule out that newly created patents could have an effect on more than one surrounding MSA at the same time. Third, I take a stand on the radius of spatial rings, which defines the patents' sphere of influence. One can never know for sure which cities benefited from a specific invention in the surrounding. In the following, I present results for spatial rings of three distance horizons, i.e. three different values for the radii of the spatial rings: 50, 75 and 100 miles. When considering a radius of 100 miles, I face the problem that a wide range of patents is counted for more than one MSA at the same time. Nevertheless, it sounds plausible that patents within a spatial ring of 100 miles radius around a city's central point might still have had the potential to have

[26]As described earlier almost all patents could be matched with U.S. counties by today's definition. A registered patent is counted for a specific city if the central point of the county, where the inventor's place of residence is located, lies within the spatial ring around that city.

an impact on that city.[27] Which radius of spatial rings to prefer is also a question of how well the instrument is performing. As it turns out in the following, especially for a radius of 75 miles the proposed instrument performs technically very well at the first stage (patent activity in the 19th century predetermines today's local entrepreneurship significantly) and also delivers accurate results at the second stage (patent activity in the 19th century is not able to predict a city's current economic growth rate).

4.5 IV Regression: Historical Patent Activity, Local Entrepreneurship and Urban Growth

4.5.1 First Stage Results

IV estimation provides unbiased and consistent estimates if the instrumental variable (the log count of patents between 1836 and 1873 in the surrounding of a specific city) is associated with changes in the initial level of local entrepreneurship in 1993, but does not cause changes in the city's subsequent economic growth. In this section, I check whether the proposed instrument is valid at the first stage by testing whether there exists a significant relationship between the regional distribution of historical patent activity and local entrepreneurship in 1993. In order for the instrument to be valid at the first stage, that connection needs to hold after controlling for relevant factors that can influence the variables of interest simultaneously. Thus I conduct OLS regressions with the city's log number of start-up births in 1993 as dependent variable and the log count of historical patents in predefined borders as explanatory variables of interest. I add one patent to each MSA, as the log count of patents would otherwise not be defined for those MSAs that had no registered patents during the 19th century. In each regression I control for the initial level of employment[28], US Census Division fixed effects[29] and include standard control variables from the urban growth literature.[30] Results of the OLS regressions are presented in table 4.5, in which I only state the estimated coefficients for the independent variables of interest. I use four definitions of historical MSA borders. First, I use MSA

[27] I stop at a radius of 100 miles as larger distances were very cumbersome to cover in the 19th century and thus it is unlikely that patents at a larger distance had a significant impact on cities (see Lamoreaux (2010))

[28] I.e. the log level of employment in 1993.

[29] I.e. I control for the nine US census divisions: Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic and New England.

[30] As described in section 4.4.3 these variables include the mean January and mean July temperature (in order to control for local (dis)advantages due to climate conditions), the share of people with bachelor degree in 1990 (to control for highly skilled individuals selecting fast growing cities), the overall population in 1990, the population density in 1990 and housing prices from the year 1993 (to control for core attributes of metropolitan areas).

Log Count of registered patents + 1 within...	**Log Start-ups 93**
Today's MSA borders	-.0585 ***
	(.0080)
p-value	.000
F-statistic	24.670
p-value (F-statistic)	.001
Partial R^2	.179
Circle of 50 miles radius	-.0478 ***
	(.0075)
p-value	.000
F-statistic	13.413
p-value (F-statistic)	.006
Partial R^2	.151
Circle of 75 miles radius	-.0433 ***
	(.0065)
p-value	.000
F-statistic	16.3858
p-value (F-statistic)	.004
Partial R^2	.146
Circle of 100 miles radius	-.0443 ***
	(.0066)
p-value	.000
F-statistic	15.974
p-value (F-statistic)	.004
Partial R^2	.147

Note: The sample contains 321 Metropolitan Statistical Areas in the US. Cluster Robust Standard Errors are reported in parentheses. ***Significance at 1 percent level,**Significance at 5 percent level, *Significance at 10 percent level. It is controlled for the nine U.S. census divisions (Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic and New England). City Growth Covariates that are controlled for include log mean January temperature, log mean July temperature, log city population 1990, log density 1990, log 1990 share of residents with bachelor degree or higher, log house prices in 1993.

Table 4.5: First Stage - US Patent Activity 1836-1873 and Entrepreneurship 1993

borders by today's definition as historical city borders. In the following, I construct spatial rings of radius 50, 75 and 100 miles around the central points of today's MSAs. Every patent within those borders is assumed to have affected the city and is thus counted for that specific city.[31] As the results from table 4.5 show, the relationship between US patent activity in the 19th century and local entrepreneurship in 1993 is highly significant after the inclusion of relevant control variables. Using today's MSA borders, I find that increasing the city's number of patents in the 19th century by one percent reduces the city's number of start-up births in 1993 by .0585 percent. The connection is also highly significant and similar in magnitude for every other proposed definition of historical city borders, i.e. for spatial rings of 50, 75 and 100 miles. Also postestimation diagnostics for the first stage indicate that the proposed instrument is valid. The observed F-statistics (to test for significance of the instrument) have p-values close to 0 for every definition of city borders. The values of the partial R^2 range from 14.6

[31] I.e. the patent is counted for a MSA if the central point of the inventor's county is within the spatial ring around the MSA's central point.

percent to 17.9 percent, which are relatively small in magnitude but not worrisome.[32] Therefore, all results indicate that the log count of patents registered between 1836 and 1873 passes the first stage validity test for every proposed definition of MSA borders.

4.5.2 Second Stage Results

Validity of the proposed instrument at the second stage is more worrisome than validity at the first stage. It is well conceivable that historic patent activity influences the current economic development of cities through other ways than just entrepreneurship. It is possible that conditions that have an effect on economic growth today might have had an impact on the inventors' locational choice in the 19th century, even though over 150 years have passed. The crucial question is, whether a connection between the two variables exists after controlling for relevant city- and regional-fixed effects.
Table 4.6 illustrates the results of OLS regressions of log employment growth and log payroll growth as dependent variables on the log count of registered patents between the years 1836 and 1873 within predefined borders as independent variables of interest. All relevant control variables are included in the regressions. As before, I use four different definitions of the patents' sphere of influence: today's MSA borders, and spatial rings of 50, 75 and 100 miles radius around the central points of today's MSAs. All patents within the presupposed borders of a city are assumed to have had an impact on that specific city.[33] Again, only estimated coefficients on the explanatory variable of interest are presented due to limitations of space. As the results from table 4.6 show, the null hypothesis of no causal relationship between historic US patent activity and urban growth between the years 1993 and 2002 cannot be rejected after including relevant control variables, independent of the proxy used for urban economic growth. Especially when the patents' sphere of influence is defined by spatial rings of radius 50 miles and larger, the data show almost no causal relationship between the two variables of interest, with p-values of 50 percent and higher. This result constitutes a major advantage compared to the instrument proposed by Glaeser et al. (2012). While the proximity to coal mines predetermines the cities' industrial structure and thus at least part of the city's growth path[34], the 19th century patent activity only seems to predetermine the firm size distribution of cities without having a major influence on the city's future growth.

[32]See Cameron and Trivedi (2005) for a discussion of Partial R^2.
[33]I.e. those patents are assumed to have influenced the city.
[34]Cities closer to coal mines are dominated by firms from the declining steel industry.

Log Count of registered patents +1 within...	**Log Emp.Growth 93-02**	**Log PR.Growth 93-02**
Today's MSA borders	-.0050	-.0076
	(.0032)	(.0067)
p-value	.150	.285
Circle of 50 miles radius	-.0030	-.0050
	(.0042)	(.0073)
p-value	.492	.516
Circle of 75 miles radius	.0005	-.0006
	(.0036)	(.0068)
p-value	.903	.928
Circle of 100 miles radius	.0006	-.0008
	(.0038)	(.0072)
p-value	.880	.913

Note: The sample contains 321 Metropolitan Statistical Areas in the US. Cluster Robust Standard Errors are reported in parentheses. ***Significance at 1 percent level,**Significance at 5 percent level, *Significance at 10 percent level. It is controlled for the nine U.S. census divisions (Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic and New England). City Growth Covariates that are controlled for include log mean January temperature, log mean July temperature, log city population 1990, log density 1990, log 1990 share of residents with bachelor degree or higher, log house prices in 1993.

Table 4.6: Second Stage - US Patent Activity 1836-1873 and Urban Growth 1993-2002

4.5.3 IV Regression Results

After having discussed the technical validity of the instrument at the first and second stage, I present IV regression results with the log count of patents registered between the years 1836 and 1873 as exogenous source of variation for rates of local entrepreneurship in 1993. Corresponding results are illustrated in table 4.7. Ignoring the issue of reverse causality yields elasticity point estimates of .1582 when urban growth is measured by log employment growth and .2208 when urban growth is instead measured by log payroll growth (see table 4.2 for OLS results). Defining the patents' sphere of influence by today's MSA borders yields IV estimates of .0862 and .1306, respectively. The entrepreneurship's impact on subsequent employment growth remains significant at a 5 percent level, but both estimates are substantially smaller than the respective OLS estimates. This reduction reflects a severe loss in the importance of local entrepreneurship, once potential endogeneity is taken into account. Using spatial rings of 50, 75 and 100 miles radius around the MSAs' central points as the patents' sphere of influence further lowers the elasticity point estimates. For a radius of 50 miles, IV elasticity estimates are .0633 when urban growth is measured by employment growth and .1043 when urban growth is instead measured by payroll growth. Using a radius of 75 miles yields elasticity estimates of -.0106 and .0147, respectively. A radius of 100 miles further reduces the elasticity point estimates to values of -.0135 and .0185, respectively. The most striking observation from table 4.7 is the fact, that the local entrepreneurship's impact on subsequent urban growth shrinks substantially compared to OLS results and is no longer statistically significant. The estimated impact of local entrepreneurship on subsequent

Log Start-up Births 93 instrumented by Log Count of registered patents + 1 within...	**Log Emp.Growth**	**Log PR.Growth**
Today's MSA borders	.0862 **	.1306
	(.0379)	(.0834)
p-value	.023	.117
Circle of 50 miles radius	.0633	.1043
	(.0698)	(.1208)
p-value	.364	.388
Circle of 75 miles radius	-.0106	.0147
	(.0792)	(.1414)
p-value	.894	.917
Circle of 100 miles radius	-.0135	.0185
	(.0819)	(.1462)
p-value	.869	.900

Note: The sample contains 321 Metropolitan Statistical Areas in the US. Cluster Robust Standard Errors are reported in parentheses. ***Significance at 1 percent level,**Significance at 5 percent level, *Significance at 10 percent level. It is controlled for the nine U.S. census divisions (Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic and New England). City Growth Covariates that are controlled for include log mean January temperature, log mean July temperature, log city population 1990, log density 1990, log 1990 share of residents with bachelor degree or higher, log house prices in 1993.

Table 4.7: IV Regression - Entrepreneurship 1993 and Urban Growth 93-02

employment growth even turns into insignificantly negative once the patents' sphere of influence is assumed to be 75 miles or larger. These results have to be taken serious, as the proposed instrument yields technically the best results for radii of 75 and 100 miles. Overall, IV results with the log count of historical patents in the surrounding of cities as exogenous source of variation suggest that the entrepreneurship's role in influencing subsequent growth is generally overestimated.

4.5.4 Robustness Checks

4.5.4.1 Weighted Least Squares Estimation

One potential concern when interpreting IV results from table 4.7 is that the estimates are distorted due to a relationship between entrepreneurship and urban growth that is especially prevalent in smaller MSAs. Thus I conduct re-estimations of OLS and IV regressions in which I weight observations with the cities' population in 1990 to diminish the impact of small and increase the influence of large cities on the results. Respective results are presented in tables 4.8 and 4.9.

Weighted OLS estimation suggests a significant positive impact of local entrepreneurship on subsequent economic growth measured by log employment growth and log payroll growth. Increasing the number of newly founded start-ups by one percent raises subsequent employment growth by .1346 percent and subsequent payroll growth by .1995 percent, respectively (see table 4.8). These elasticities are slightly below the estimated

A: Log Emp.Growth	**(1)**	**(2)**	**(3)**
Log Start-up Births 93	.1859***	.1150***	**.1346*****
	(.0385)	(.0404)	(.0213)
p-value	.001	.002	.000
initial empl.	Yes	Yes	Yes
C.D. fixed effects	No	Yes	Yes
City Growth covariates	No	No	Yes
N	321	321	321

B: Log Payroll Growth	**(1)**	**(2)**	**(3)**
Log Start-up Births 93	.2581***	.1888***	**.1995*****
	(.0502)	(.0517)	(.0271)
p-value	.001	.006	.000
initial empl.	Yes	Yes	Yes
C.D. fixed effects	No	Yes	Yes
City Growth covariates	No	No	Yes
N	321	321	321

Note: The sample contains 321 Metropolitan Statistical Areas in the US. Cluster Robust Standard Errors are reported in parentheses. Observations are weighted by the city population in 1990. ***Significance at 1 percent level,**Significance at 5 percent level, *Significance at 10 percent level. The nine U.S. census divisions are Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic and New England. City Growth Covariates include log mean January temperature, log mean July temperature, log city population 1990, log density 1990, log 1990 share of residents with bachelor degree or higher, log house prices in 1993.

Table 4.8: Weighted OLS Regression - Entrepreneurship 1993 and Urban Growth 93-02

Log Start-up Births 1993 instrumented by Log Count of registered patents + 1 within...	**Log Emp.Growth**	**Log PR.Growth**
Today's MSA borders	.0706	-.0162
	(.0490)	(.1362)
p-value	.150	.905
Circle of 50 miles radius	.0317	-.0235
	(.0974)	(.1885)
p-value	.745	.901
Circle of 75 miles radius	-.0205	-.0738
	(.1021)	(.2013)
p-value	.841	.714
Circle of 100 miles radius	-.0146	-.0811
	(.0904)	(.1974)
p-value	.871	.681

Note: The sample contains 321 Metropolitan Statistical Areas in the US. Cluster Robust Standard Errors are reported in parentheses. Observations are weighted by the city population in 1990. ***Significance at 1 percent level,**Significance at 5 percent level, *Significance at 10 percent level. It is controlled for the nine U.S. census divisions (Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic and New England). City Growth Covariates that are controlled for include log mean January temperature, log mean July temperature, log city population 1990, log density 1990, log 1990 share of residents with bachelor degree or higher, log house prices in 1993.

Table 4.9: Weighted IV Regression - Entrepreneurship 1993 and Urban Growth 93-02

elasticities of an unweighted OLS regression (see table 4.2). The elasticities estimated by weighted OLS regression become substantially reduced, once log start-up births 1993 is instrumented by the log count of historical patents in the surrounding of cities (see table 4.9). The local entrepreneurship's impact on subsequent employment growth even turns into insignificantly negative for spatial rings of 75 and 100 miles. The impact on subsequent payroll growth is insignificantly negative for all definitions of spatial rings. Overall, weighting observations by the city's population does not alter the results of section 4.5.4.3 qualitatively. Once I use the regional distribution of 19th century patents as exogenous source of variation, differences in entrepreneurial activity across US MSAs are no longer able to explain regional differences in urban growth.

4.5.4.2 Estimation without 1870 US Territories

Not all of the 321 US cities that are included in my sample were an official part of the United States in the 19th century. Some of today's official US states were organized as so called "US territories" at that time. Maybe cities that were located in former US territories only had a lower number of registered patents in the 19th century because a lack of existing institutions made it cumbersome to register inventions. Thus I conduct re-estimations that only include those 282 US MSAs that were located in an official US state in 1870.[35] Results of OLS and IV regressions are reported in tables 4.10 and 4.11. Using the reduced sample of 282 US MSAs leads to higher OLS elasticity estimates for entrepreneurship compared to elasticity estimates for the whole sample of 321 MSAs (see table 4.2). Raising the city's number of start-up births by one percent increases employment growth during the next ten years by .1865 percent and payroll growth during the same timeframe by .2599 percent, respectively (see table 4.10). To address the issue of reverse causality, I instrument the log number of start-up births in 1993 with the log count of historical patents in the surrounding of cities. Counting patents within today's MSA borders slightly reduces the entrepreneurship's impact on subsequent growth, but the impact remains significant at a 1 percent significance level. If the definition is changed to spatial rings of 50 miles, the impact of entrepreneurship on subsequent growth remains significant at a 5 percent level for employment growth and significant at a 10 percent level for payroll growth. For spatial rings of a radius of 75 miles and larger, entrepreneurship loses its significant impact and p-values above 20 percent are observed.

Thus using the sample of 282 MSAs slightly reduces the instrument's impact on OLS results. But the observed pattern remains unaltered: The entrepreneurship's impact on city growth is diminished once the regional distribution of patents from the 19th

[35] 39 cities from the following US states (and former US territories) are excluded: Washington, Montana, Idaho, Utah, Arizona, New Mexico, Colorado, Wyoming, South Dakota, North Dakota.

A: Log Emp.Growth	**(1)**	**(2)**	**(3)**
Log Start-up Births 93	.2005***	.1692***	**.1865*****
	(.0236)	(.0253)	(.0359)
p-value	.000	.000	.001
initial empl.	Yes	Yes	Yes
C.D. fixed effects	No	Yes	Yes
City Growth covariates	No	No	Yes
N	282	282	282

B: Log Payroll Growth	**(1)**	**(2)**	**(3)**
Log Start-up Births 93	.2734***	.2384***	**.2599*****
	(.0319)	(.0366)	(.0421)
p-value	.000	.000	.000
initial empl.	Yes	Yes	Yes
C.D. fixed effects	No	Yes	Yes
City Growth covariates	No	No	Yes
N	282	282	282

Note: The sample contains 282 Metropolitan Statistical Areas in the US. Cluster Robust Standard Errors are reported in parentheses. Observations are weighted by the city population in 1990. ***Significance at 1 percent level,**Significance at 5 percent level, *Significance at 10 percent level. The nine U.S. census divisions are Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic and New England. City Growth Covariates include log mean January temperature, log mean July temperature, log city population 1990, log density 1990, log 1990 share of residents with bachelor degree or higher, log house prices in 1993.

Table 4.10: OLS Regression without 1870 Territories - Entrepreneurship 1993 and Urban Growth 93-02

Log Start-up Births 1993 instrumented by Log Count of registered patents + 1 within...	**Log Emp.Growth**	**Log PR.Growth**
Today's MSA borders	.1749 ***	.2018 ***
	(.0317)	(.0630)
p-value	.000	.001
Circle of 50 miles radius	.1767 **	.2018 *
	(.0714)	(.1197)
p-value	.013	.092
Circle of 75 miles radius	.1160	.1317
	(.0829)	(.1374)
p-value	.162	.338
Circle of 100 miles radius	.1090	.1297
	(.0875)	(.1494)
p-value	.213	.385

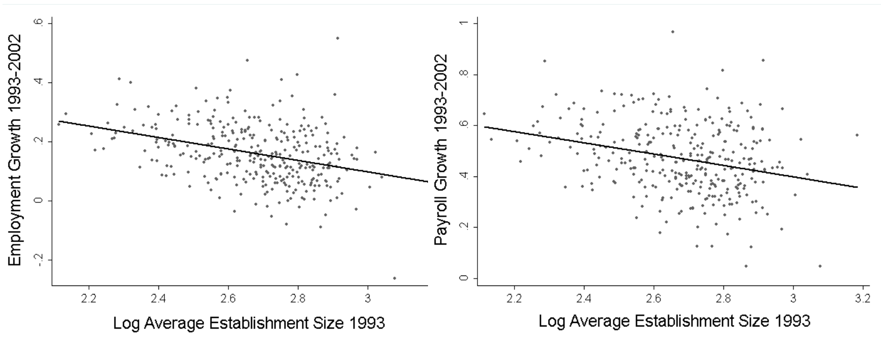
Note: The sample contains 282 Metropolitan Statistical Areas in the US. Cluster Robust Standard Errors are reported in parentheses. ***Significance at 1 percent level,**Significance at 5 percent level, *Significance at 10 percent level. It is controlled for the nine U.S. census divisions (Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic and New England). City Growth Covariates that are controlled for include log mean January temperature, log mean July temperature, log city population 1990, log density 1990, log 1990 share of residents with bachelor degree or higher, log house prices in 1993.

Table 4.11: IV Regression without 1870 Territories - Entrepreneurship 1993 and Urban Growth 93-02

century is used as exogenous source of variation for local entrepreneurship. The impact also becomes insignificant once I assume the patents' sphere of influence to be 75 miles and larger.

4.5.4.3 Average Establishment Size as proxy for Local Entrepreneurship

Using the city's average establishment size as proxy for the city's entrepreneurial activity is widely used in the literature. The higher the city's number of small start-up firms, the lower the city's mean number of employees per establishment. Thus more local entrepreneurial activity is associated with smaller average establishment size.



Note: The sample contains 321 US Metropolitan Statistical Areas.

FIGURE 4.7: Entrepreneurship 1993 (proxied by average establishment size) and Urban Growth 1993-2002

Figure 4.7 illustrates the perceived negative correlation between my alternative proxy for local entrepreneurship and subsequent employment growth and payroll growth. Note that local entrepreneurial activity decreases with average establishment size, which explains the negative connection in the scatterplot. Table 4.12 presents the results of OLS regressions with log employment growth and log payroll growth between 1993 and 2002 as dependent variables and log average establishment size in 1993 as independent variable of interest. Specification (3) of table 4.12 includes all relevant control variables. Decreasing average establishment size by one percent leads to an increase in subsequent employment growth by .1808 percent and to an increase in subsequent payroll growth by .2354 percent, respectively. Both estimated elasticities are significant at the 5 percent level. Validity of IV results requires that the instrument succeeds at the first stage, i.e. there has to exist a significant connection between the number of registered patents during the 19th century in the surrounding of a city and the city's average establishment size in 1993 after controlling for relevant factors that could influence both variables simultaneously.[36] Table 4.13 presents IV regression results with local entrepreneurship

[36]Initial Employment, US Census Division fixed effects and City Growth covariates.

A: Log Emp.Growth	**(1)**	**(2)**	**(3)**
Log Average Est.Size 93	-.2478***	-.1689**	**-.1808****
	(.0453)	(.0551)	(.0722)
p-value	.001	.015	.037
initial empl.	Yes	Yes	Yes
C.D. fixed effects	No	Yes	Yes
City Growth covariates	No	No	Yes
N	321	321	321

B: Log Payroll Growth	**(1)**	**(2)**	**(3)**
Log Average Est.Size 93	-.3360***	-.2181***	**-.2354****
	(.0541)	(.0720)	(.1003)
p-value	.000	.001	.047
initial empl.	Yes	Yes	Yes
C.D. fixed effects	No	Yes	Yes
City Growth covariates	No	No	Yes
N	321	321	321

Note: The sample contains 321 Metropolitan Statistical Areas in the US. Cluster Robust Standard Errors are reported in parentheses. ***Significance at 1 percent level,**Significance at 5 percent level, *Significance at 10 percent level. The nine U.S. census divisions are Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic and New England. City Growth Covariates include log mean January temperature, log mean July temperature, log city population 1990, log density 1990, log 1990 share of residents with bachelor degree or higher, log house prices in 1993.

Table 4.12: OLS Regression - Entrepreneurship 1993 (proxied by average establishment size) and Urban Growth 93-02

Log Average Est. Size 1993 instrumented by Log Count of registered patents + 1 within...	**Log Emp.Growth**	**Log PR.Growth**
Today's MSA borders	-.2152**	-.3261*
	(.0852)	(.1908)
p-value	.012	.087
Circle of 50 miles radius	-.1290	-.2123
	(.1429)	(.2504)
p-value	.367	.396
Circle of 75 miles radius	.0220	-.0304
	(.1632)	(.2949)
p-value	.893	.918
Circle of 100 miles radius	.0286	-.0392
	(.1722)	(.3127)
p-value	.868	.900

Note: The sample contains 321 Metropolitan Statistical Areas in the US. Cluster Robust Standard Errors are reported in parentheses. ***Significance at 1 percent level,**Significance at 5 percent level, *Significance at 10 percent level. It is controlled for the nine U.S. census divisions (Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic and New England). City Growth Covariates that are controlled for include log mean January temperature, log mean July temperature, log city population 1990, log density 1990, log 1990 share of residents with bachelor degree or higher, log house prices in 1993.

Table 4.13: IV Regression - Entrepreneurship 1993 (proxied by average establishment size) and Urban Growth 93-02

Log Count of registered patents + 1 within...	**Log Average Est.Size 93**
Today's MSA borders	.0234 ***
	(.0049)
p-value	.000
F-statisic	9.865
p-value (F-statistic)	.014
Partial R^2	.085
Circle of 50 miles radius	.0235 ***
	(.0043)
p-value	.000
F-statisic	13.491
p-value (F-statistic)	.006
Partial R^2	.107
Circle of 75 miles radius	.0209 ***
	(.0038)
p-value	.000
F-statisic	16.748
p-value (F-statistic)	.004
Partial R^2	.100
Circle of 100 miles radius	.0209 ***
	(.0039)
p-value	.000
F-statisic	15.084
p-value (F-statistic)	.000
Partial R^2	.096

Note: The sample contains 321 Metropolitan Statistical Areas in the US. Cluster Robust Standard Errors are reported in parentheses. ***Significance at 1 percent level,**Significance at 5 percent level, *Significance at 10 percent level. It is controlled for the nine U.S. census divisions (Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic and New England). City Growth Covariates that are controlled for include log mean January temperature, log mean July temperature, log city population 1990, log density 1990, log 1990 share of residents with bachelor degree or higher, log house prices in 1993.

TABLE 4.14: First Stage - US Patent Activity 1836-1873 and Entrepreneurship (proxied by average establishment size) 1993

being proxied by average establishment size. IV results are qualitatively identical to the results from table 4.7 (in which entrepreneurship is proxied by start-up births). If I count historical patents within today's MSA borders, the estimated impact of local entrepreneurship on subsequent growth becomes slightly larger compared to the OLS regression results.[37] Once I use the preferred definitions of spatial rings around today's MSA central points, the estimated impact of local entrepreneurship on subsequent city growth diminishes substantially and becomes insignificant.[38] For spatial rings of 75 and 100 miles, the estimated elasticity even turns into insignificantly positive when urban growth is measured by employment growth. Thus local entrepreneurship has an insignificantly negative impact on subsequent employment growth.

Table 4.14 illustrates that my proposed instrument remains valid at the first stage as the relationship between patent activity and average establishment size in 1993 is highly

[37]If entrepreneurship is measured by average establishment size, a larger impact means that the elasticity estimate becomes more negative.

[38]I.e. the elasticity estimates become less negative or even positive.

significant for each of the four definitions of the patents' sphere of influence. All p-value of the observed F-statistics suggest that the instrument remains valid at the first stage. Overall, the results of this section confirm the results of section 4.5.3. Once I use the regional distribution of 19th century patents as exogenous source of variation for local entrepreneurship across US MSAs, entrepreneurial activity loses its significant impact on subsequent urban growth. This finding persists if local entrepreneurship is proxied by the city's average establishment size.

4.5.5 Critical Résumé of IV Results

The overall pattern of table 4.7 suggests that the local entrepreneurship's impact on subsequent urban growth is diminished or, depending on the patents' assumed sphere of influence, disappears completely once local entrepreneurship is instrumented by the log count of registered patents between 1836 and 1873. This observation is robust to the use of different proxies for local entrepreneurship (log number of newly founded start-ups and log average establishment size), robust to the use of different proxies for urban growth (employment growth and payroll growth) and also robust to the use of various historical city borders (today's borders, spatial rings of 50, 75 and 100 miles around MSAs' center points). Furthermore, the results remain qualitatively unchanged when MSAs are weighted by population or MSAs from former US territories are excluded from the sample.

These results mirror the findings by Glaeser et al. (2012), who show that estimated IV elasticities are even larger than the respective OLS estimates, i.e. they argue that local entrepreneurship plays an even more important role for city growth once endogeneity is taken into account. I suggest that my instrument performs better at the second stage of the IV regression than the instrument proposed by Glaeser et al. (2012). The proximity to coal mines partially predetermines the city's industry structure, while the count of patents in the 19th century only predetermines the city's firm size distribution. This constitutes a major advantage compared to Glaeser et al. (2012) as the predetermination of the city's industry structure has a substantial impact on today's growth rates.

So why do I observe IV estimates that indicate a less important role for local entrepreneurship in determining the rate of subsequent urban growth? A less positive interpretation for my results is the following: Cities with many registered patents during the 19th century in the surrounding developed an environment of large enterprises that repels potential entrepreneurs. The same cities might exhibit various characteristics that once attracted inventors and are now positively correlated with urban growth. If that is true the orthogonality condition needed for consistent IV estimation is violated,

leading to downward biased coefficient estimates in my analysis. A more positive interpretation of my results focuses on the selection bias. Potential entrepreneurs choose the city, in which their expected profits are highest. On average they will choose cities with the highest growth rates, meaning that entrepreneurs self-select into fast growing places. The usage of my instrument then resolves part of the reverse causality and delivers results for which the self-selection bias is subtracted out.[39]
Even though I showed that my instrument performs well at the second stage[40], there always remains the concern that the instrument is correlated with characteristics promoting subsequent urban growth. Nevertheless, I think that the observed decline of the local entrepreneurship's importance related to subsequent urban growth cannot solely be explained by a violation of the orthogonality condition. My analysis suggests that there is reason to believe that the declining impact of local entrepreneurship is mainly explained by resolving the self-selection problem.

4.6 Discussion

The main goal of this chapter is to measure the causal impact of local entrepreneurial activity on subsequent urban growth. The general sentiment in the urban economics literature is that this connection is positive and causal. Starting point for my analysis is the concern about potential reverse causality inherent to OLS regressions on the relationship between local entrepreneurship and urban growth. The issue is tackled by identifying an exogenous source of variation for modern rates of local entrepreneurship. I propose the regional distribution of patents between the years 1836 and 1873 as an instrument, arguing that regions that exhibit a high number of patents during the 19th century had a higher probability of generating an environment of large companies that prevents potential entrepreneurs from entering the city (see Chinitz (1961)). Thus historical patent activity predetermines a city's firm size distribution and in turn present rates of local entrepreneurship. IV regression results suggest that the positive connection between local entrepreneurship and urban growth is substantially diminished or disappears completely, depending on how I define the patents' sphere of influence. Results have to be interpreted with a caution as there remain some issues regarding historical patent activity being correlated with current urban growth rates across US cities. The general sentiment sees local entrepreneurial activity as one of the key drivers of urban growth and claims to promote the emergence of local start-ups as much as possible. The

[39]Of course the observed results can be driven by both forces at the same time, i.e. while the self-selection bias is resolved, results are nevertheless downward biased due to a violation of the orthogonality condition.

[40]Especially for spatial rings of 75 and 100 miles.

results of this chapter, however, indicate that this might be a misguided effort, as the positive effect measured in the data is partially driven by entrepreneurs self-sorting into fast growing cities.

Chapter 5

Conclusion

The main goal of my thesis is to examine potential determinants responsible for the observed extent and heterogeneity of city growth. More precisely, my work concentrates on the impact of local knowledge spillovers and on the influence of local entrepreneurial activity. In chapter 2, I present a summary on previous empirical findings regarding the impact of potential key drivers on subsequent urban growth. I focus on the city's provision of transportation infrastructure, the city's supply of amenities, the role of agglomeration economies and the city's endowment with entrepreneurial activity. I introduce a simplified version of the Monocentric City Model to derive theoretical implications of how the described factors are expected to influence city growth. For their empirical examination it turns out that potential reverse causality is the key problem in quantifying the factors' influences. Only by the identification of exogenous sources of variation for modern city characteristics it is possible to derive causal estimates of the factors' importance. Chapter 3 addresses the topic of agglomeration economies in detail. Empirical observations on the development of urban workers' wages suggest that increasing returns at the city-level are best explained by learning through knowledge spillovers in cities. I present a search-theoretic model à la Pissarides (2000) of urban face-to-face interactions to examine the role of local knowledge spillovers for the extent of agglomeration economies. The model explicitly incorporates two types of spillovers, the creation and transmission of knowledge, which are exchanged in urban face-to-face interactions. The extent of those spillovers crucially depends on the heterogeneity of the interacting individuals' knowledge types. Assuming that created knowledge is a public good, the model's key insight is that individuals have face-to-face interactions with an inefficient range of people in the city. This is mainly due to the fact that individuals do not take into account their personal impact on the city's overall creation of knowledge. The described inefficiency in knowledge creation causes socially inefficient city sizes as agglomeration forces do not reach their optimal extent. Chapter 4 focuses on the relationship between the city's initial level of local entrepreneurship and subsequent city growth. The prevalence of entrepreneurial activity is generally assumed to enhance urban growth. As entrepreneurs have an incentive to self-select into fast growing cities, I argue that the positive relation might be driven by reverse causality. While the identification of exogenous sources of variation is well established for other potential drivers of urban growth, those sources are only recently detected for current levels of local entrepreneurship. In the empirical analysis presented in chapter 4, I propose the regional patent distribution during the 19th century across the US as an instrument for modern rates of local entrepreneurship. I find that the entrepreneurship's impact on city growth becomes substantially diminished and insignificant once I conduct Instrumental Variable regressions to take the inherent endogeneity into account. This result puts popular policy advices for cities concerning the recruitment of start-up businesses into question. In summary, my thesis provides three important research contributions. First, I present the theoretical implications of potential key drivers for subsequent urban growth in a unified framework and deliver a summary on previous empirical findings for each of them. Second, the theoretical model presented in chapter 3 is the first to explicitly incorporate the transmission and creation of knowledge in an urban environment. It contributes to the comprehension of the knowledge spillovers' role for agglomeration economies. Third, I am among the first to detect potential instruments for modern rates of entrepreneurship. Using the 19th century patent activity as instrument for modern rates of entrepreneurship, I show that the perceived positive relation between entrepreneurship and urban growth might be driven by reverse causality.

Appendix A

Knowledge Spillovers in Cities

The optimality conditions for the knowledge spread $\frac{\partial V_u}{\partial \delta} = 0$ and for the number of unmatched agents in the city $\frac{\partial V_u}{\partial U} = 0$ can be expressed as follows:

Optimality condition for the knowledge spread:

$$\begin{aligned}
\frac{\partial V_u}{\partial \delta} &= \underbrace{\frac{\lambda\alpha U}{(\lambda+\alpha\delta U)^2}\frac{a_0}{r}}_{\frac{\partial A}{\partial \delta}} + \\
&+ \underbrace{\frac{\lambda\alpha\delta}{(\lambda+\alpha\delta U)^2}\frac{a_0}{r}}_{\frac{\partial A}{\partial U}} \underbrace{\left(-\frac{\alpha U^2}{\lambda+2\alpha\delta U}\right)}_{\frac{\partial U}{\partial \delta}} + \\
&+ \underbrace{\frac{(r+\lambda)\alpha U(e_0 - e_1\frac{\delta}{2}) - (\alpha\delta U)^2\frac{e_1}{4}}{r(r+\lambda+\alpha\delta U)^2}}_{\frac{\partial e}{\partial \delta}} + \\
&+ \underbrace{\frac{(r+\lambda)\alpha\delta(e_0 - e_1\frac{\delta}{4})}{r(r+\lambda+\alpha\delta U)^2}}_{\frac{\partial e}{\partial U}} \underbrace{\left(-\frac{\alpha U^2}{\lambda+2\alpha\delta U}\right)}_{\frac{\partial U}{\partial \delta}} = 0
\end{aligned}$$

Optimality condition for the number of unmatched in the city:

$$
\begin{aligned}
\frac{\partial V_u}{\partial U} \quad = \quad & \underbrace{\frac{\lambda\alpha\delta}{(\lambda+\alpha\delta U)^2}\frac{a_0}{r}}_{\frac{\partial A}{\partial U}} + \\
& + \underbrace{\frac{(r+\lambda)\alpha\delta(e_0 - e_1\frac{\delta}{4})}{r(r+\lambda+\alpha\delta U)^2}}_{\frac{\partial e}{\partial U}} - \\
& -t \underbrace{\frac{\lambda + 2\alpha\delta U}{\lambda}}_{\frac{\partial N}{\partial U}} = 0
\end{aligned}
$$

Appendix B

Entrepreneurship and Urban Growth

A: Log Emp.Growth	**(1)**	**(2)**	**(3)**
Log Start-up Share	.5711***	.4395***	**.4545*****
	(.0254)	(.0230)	(.0298)
p-value	.000	.000	.001
initial empl.	Yes	Yes	Yes
C.D. fixed effects	No	Yes	Yes
City Growth covariates	No	No	Yes
N	321	321	321

B: Log Payroll Growth	**(1)**	**(2)**	**(3)**
Log Start-up Share	.5611***	.4563***	**.4654*****
	(.0335)	(.0311)	(.0415)
p-value	.000	.000	.001
initial empl.	Yes	Yes	Yes
C.D. fixed effects	No	Yes	Yes
City Growth covariates	No	No	Yes
N	321	321	321

Note: The sample contains 321 Metropolitan Statistical Areas in the US. Cluster Robust Standard Errors are reported in parantheses. ***Significance at 1 percent level,**Significance at 5 percent level, *Significance at 10 percent level. The nine U.S. census divisions are Pacific, Mountain, West North Central, West South Central, East North Central, East South Central, Middle Atlantic, South Atlantic and New England. City Growth Covariates include log mean January temperature, log mean July temperature, log city population 1990, log density 1990, log 1990 share of residents with bachelor degree or higher, log house prices in 1993.

Table B.1: Standardized OLS Regression - Entrepreneurship 1993 and Urban Growth 93-02

Bibliography

Acs, Zoltan and Catherine Armington (2004) “Employment Growth and Entrepreneurial Activity in Cities,” *Regional Studies*, Vol. 38(8), pp. 911–927.

Alonso, William (1964) *Location and Land Use: Toward a General Theory of Land Rent*, Cambridge, MA: Harvard University Press.

Assmann, Dirk (2015) “Entrepreneurship and City Growth: An Empirical Investigation with US Patent Data from the 19th Century,” *Unpublished Working Paper*.

Assmann, Dirk and Johannes Stiller (2015) “Knowledge Spillovers in Cities: The Creation and Transmission of Knowledge,” *Unpublished Working Paper*.

Audretsch, David B. and Maryann P. Feldmann (1996) “R&D Spillovers and the Geography of Innovation and Production,” *American Economic Review*, Vol. 86(3), pp. 630–640.

Audretsch, David and Maryann P. Feldmann (2004) “Knowledge spillovers and the geography of innovation,” in *Handbook of Regional and Urban Economics*, Vol. 4: Elsevier, Chap. 61, pp. 2713–2739.

Baum-Snow, Nathaniel (2007a) “Did Highways Cause Suburbanization?,” *The Quarterly Journal of Economics*, Vol. 122(2), pp. 775–805.

——— (2007b) “Suburbanization and transportation in the monocentric model,” *Journal of Urban Economics*, Vol. 62(3), pp. 405–423.

Baum-Snow, Nathaniel, Loren Brandt, Vernon J. Henderson, Matthew A. Turner, and Qinghua Zhang (2012) “Roads, Railways and Decentralization of Chinese Cities,” *Unpublished Working Paper*.

Berliant, Marcus, Robert R. Reed III, and Ping Wang (2006) “Knowledge exchange, matching, and agglomeration,” *Journal of Urban Economics*, Vol. 60(1), pp. 69–95.

Boldrin, Michele and David K. Levine (2002) “The Case Against Intellectual Property,” *American Economic Review*, Vol. 92(2), pp. 209–212.

——— (2013) "The Case Against Patents," *Journal of Economic Perspectives*, Vol. 27(1), pp. 3–22.

Brueckner, Jan K. (1987) "The structure of urban equilibria: A unified treatment of the muth-mills model," in *Handbook of Regional and Urban Economics*, Vol. 2: Elsevier, Chap. 20, pp. 821–845.

Brueckner, Jan K., Jaques-Francois Thisse, and Yves Zenou (2002) "Local Labor Markets, Job Matching, and Urban Location," *International Economic Review*, Vol. 43(1), pp. 155–171.

Brueckner, Jan K., Jaques-Frnacois Thisse, and Yves Zenou (1999) "Why is central Paris rich and downtown Detroit poor?: An amenity-based theory," *European Economic Review*, Vol. 43(1), pp. 91–107.

Burchfield, Marcy, Henry G. Overman, Diego Puga, and Matthew A. Turner (2006) "Causes of Sprawl: A Portrait from Space," *The Quarterly Journal of Economics*, Vol. 121(2), pp. 587–633.

Cairncross, Frances (2001) *The Death of Distance 2.0; How the Communications Revolution will Change our Lives*, Cambridge, MA: Havard Business School Press.

Cameron, Colin and Pravin K. Trivedi (2005) *Microeconometrics using STATA*, Cambridge, NY: Cambridge University Press.

Carlino, Gerald A., Satyajit Chatterjee, and Robert M. Hunt (2007) "Urban density and the rate of invention," *Journal of Urban Economics*, Vol. 61(3), pp. 389–419.

Carlino, Gerald A. and Albert Saiz (2008) "City Beautiful," *Working Papers 08-22, Federal Reserve Bank of Philadelphia.*

Carlsen, Fredrik, Jorn Rattso, and Hildegunn E. Stokke (2013) "Education, experience and dynamic urban wage premium," *Working Paper Series Norwegian University of Science and Technology.*

Cheshire, Paul C. and Stefano Magrini (2006) "Population growth in European cities: Weather matters but only nationally," *Regional Studies*, Vol. 40(1), pp. 23–37.

Chinitz, Benjamin (1961) "Contrasts in Agglomeration: New York and Pittsburgh," *American Economic Review*, Vol. 51(2), pp. 279–289.

Ciccone, Antonio (2002) "Agglomeration effects in Europe," *European Economic Review*, Vol. 46(2), pp. 213–227.

Ciccone, Antonio and Robert E. Hall (1996) "Productivity and the Density of Economic Activity," *American Economic Review*, Vol. 86(1), pp. 54–70.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastian Roux (2010) "Estimating Agglomeration Economies with History, Geology, and Worker Effects," in *Agglomeration Economics*: NBER Books, National Bureau of Economic Research, Chap. 1, pp. 15–66.

Costa, Dora L. and Matthew E. Kahn (2000) "Power Couples: Changes In The Locational Choice Of The College Educated," *The Quarterly Journal of Economics*, Vol. 115(4), pp. 1287–1315.

——— (2003) "Understanding the American Decline in Social Capital, 1952–1998," *Kyklos*, Vol. 56(1), pp. 17–46.

Cragg, Michael I. and Matthew E. Kahn (1999) "Climate consumption and climate pricing from 1940 to 1990," *Regional Science and Urban Economics*, Vol. 29(4), pp. 519–539.

Cullen, Julie B. and Stephen D. Levitt (1999) "Crime, Urban Flight, And The Consequences For Cities," *The Review of Economics and Statistics*, Vol. 81(2), pp. 159–169.

Davis, Donald R. and Jonathan I. Dingel (2012) "A Spatial Knowledge Economy," *NBER Working Papers 18188, National Bureau of Economic Research, Inc.*

D'Costa, Sabine and Henry Overman (2014) "The urban wage growth premium: Sorting or learning?" *Regional Science and Urban Economics*, Vol. 48, pp. 168–179.

Duranton, Gilles (1998) "Labor specialization, transport costs, and city size," *Journal of Regional Science*, Vol. 38(4), pp. 553–573.

Duranton, Gilles and Diego Puga (2004) "Micro-foundations of urban agglomeration economies," in *Handbook of Regional and Urban Economics*, Vol. 4: Elsevier, Chap. 48, pp. 2063–2117.

——— (2014) "The Growth of Cities," in *Handbook ofEconomic Growth*, Vol. 2: Elsevier, Chap. 5, pp. 781–853.

——— (2015) "Urban Land Use," in *Handbook of Regional and Urban Economics*, Vol. 5: Elsevier, Chap. 8, pp. 467–560.

Duranton, Gilles and Matthew A. Turner (2012) "Urban growth and transportation," *The Review of Economic Studies*, Vol. 79(4), pp. 1407–1440.

Eeckhout, Janl (2004) "Gibrat's law for (All) cities," *American Economic Review*, Vol. 94(5), pp. 1429–1451.

Ellen, Ingrid G. and Katherine O'Regan (2010) "Crime and urban flight revisited: The effect of the 1990s drop in crime on cities," *Journal of Urban Economics*, Vol. 68(3), pp. 247–259.

Feldmann, Maryann P. and David Audretsch (1999) "Innovation in cities:: Science-based diversity, specialization and localized competition," *European Economic Review*, Vol. 43(2), pp. 409–429.

Fritsch, Michael and Pamela Mueller (2004) "Regional Growth Regimes Revisited - The Case of West Germany," *Papers on Entrepreneurship, Growth and Public Policy, Max Planck Institute of Economics.*

Gabaix, Xavier (1999) "Zipf's law for cities: An explanation," *The Ouaterly Journal of Economics*, Vol. 114(3), pp. 739–767.

Gabaix, Xavier and Yannis M. Ioannides (2004) "The evolution of city size distributions," in *Handbook of Regional and Urban Economics*, Vol. 4: Elsevier, Chap. 53, pp. 2341–2378.

Garcia-López, Miquel-Àngel (2012) "Urban spatial structure, suburbanization and transportation in Barcelona," *Journal of Urban Economics*, Vol. 72(2), pp. 176–190.

Garcia-López, Miquel-Àngel, Adelheid Holl, and Elisabet Viladecans-Marsal (2013) "Suburbanization and highways: when the romans, the bourbons and the first cars still shape Spanish cities," *Working Papers 2013/5, Institut d'Economia de Barcelona (IEB).*

Ghani, Ejaz, William R. Kerr, and Stephen O'Connell (2011) *South Asia - Entrepreneurship, Growth, and Job Creation*, Oxford: Oxford University Press.

Gibrat, Robert (1931) *Les inégalités économiques; applications: aux inégalités des richesses, à la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc., d'une loi nouvelle, la loi de l'effet proportionnel.*, Paris: Librairie du Recueil Sirey.

Glaeser, Edward L. (1999) "Learning in Cities," *Journal of Urban Economics*, Vol. 46(2), pp. 254–277.

Glaeser, Edward L., Hedi D. Kallal, Jose A. Sheinkam, and Andrei Shleifer (1992) "Growth in Cities," *Journal of Political Economy*, Vol. 100(6), pp. 1126–1152.

Glaeser, Edward L., Joseph Kerr, and Albert Saiz (2008) "Housing supply and housing bubbles," *Journal of Urban Economics*, Vol. 64(2), pp. 198–217.

Glaeser, Edward L., Sari Pekkala Kerr, and William R. Kerr (2012) "Entrepreneurship and Urban Growth: An Empirical Assessment with Historical Mines," *Harvard Business School Working Papers.*

Glaeser, Edward L., William R. Kerr, and Giacomo A.M. Ponzetto (2010) "Clusters of Entrepreneurship," *Journal of Urban Economics*, Vol. 67(1), pp. 150–168.

Glaeser, Edward L., Jed Kolko, and Albert Saiz (2001) "Consumer City," *Journal of Economic Geography*, Vol. 1(1), pp. 27–50.

Glaeser, Edward L. and David C. Maré (2001) "Cities and Skills," *Journal of Labor Economics*, Vol. 19(2), pp. 316–342.

Glaeser, Edward L. and Matthew G. Resseger (2010) "The Complementarity between Cities and Skills," *Journal of Regional Science*, Vol. 50(1), pp. 221–244.

Helsley, Robert W. and William C. Strange (1990) "Matching and agglomeration economies in a system of cities," *Regional Science and Urban Economics*, Vol. 20(2), pp. 189–212.

Henderson, Vernon, Ari Kuncoro, and Matthew A. Turner (1995) "Industrial Development in Cities," *Journal of Political Economy*, Vol. 103(5), pp. 1126–1152.

Hobijn, Bart and Aysegül Sahin (2009) "Job-finding and separation rates in the OECD," *Economic Letters*, Vol. 104(3), pp. 107–111.

Hymel, Kent (2009) "Does traffic congestion reduce employment growth?," *Journal of Urban Economics*, Vol. 65(2), pp. 127–135.

Jacobs, Jane (1969) *The Economy of Cities*, New York: Random House.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson (1993) "Geographic Localization of Knowledge Spillovers as Evidence by Patent Citations," *The Quarterly Journal of Economics*, Vol. 108(3), pp. 577–598.

Johnson, Peter and Simon Charles Parker (1996) "Spatial Variation in the Determinants and Effects of Firm Births and Deaths," *Regional Studies*, Vol. 30(7), pp. 679–688.

Krugman, Paul (1991) *Geography and Trade*, Cambridge, MA: MIT Press.

——— (1996) "Confronting the mystery of urban hierarchy," *Journal of the Japanese and International Economies*, Vol. 10(4), pp. 1120–1171.

Lamoreaux, N. (2010) "Entrepreneurship in the United States, 1865-1920," in *The Invention of Enterprise: Entrepreneurship from Ancient Mesopotania to Modern Times*: Princeton University Press.

Lee, Yong Suk (2014) "Entrepreneurship, Small Businesses, and Economic Growth in Cities," *Department of Economics Working Papers, Williams College.*

Lucas, Robert Jr. (1988) "On the mechanics of economic development," *Journal of Monetary Economics*, Vol. 22(1), pp. 3–42.

Marshall, Alfred (1890) *Principals of Economics*, London: Macmillan.

Mills, Edwin S. (1967) "An aggregate modelof resource allocation in a metropolitan area," *American Economic Review Papers and Proceedings*, Vol. 57(2), pp. 197–210.

Miracky, William (1993) "Economic Growth and Business Cycles in Cities: The Role of Locational Externalities," *Thesis, Massachusetts Institute of Technology.*

Moretti, Enrico (2004) "Human capital externalities in cities," in *Handbook of Regional and Urban Economics*, Vol. 4: Elsevier, Chap. 51, pp. 2243–2291.

Muth, Richard F. (1969) *Cities and Housing*, Chicago: University of Chicago Press.

O'Flaherty, Brendan (2005) *City Economics*, Cambridge, MA: Harvard University Press.

Peri, Giovanni (2002) "Young workers, learning, and agglomerations," *Journal of Urban Economics*, Vol. 52(3), pp. 582–607.

Pissarides, Christopher A. (1979) "Job matchings with state employment agencies and random search," *Economic Journal*, Vol. 89(356), pp. 818–833.

——— (2000) *Equilibrium Unemployment Theory*, Cambridge, MA: MIT Press Books.

Rappaport, Jordan (1999) "Local Growth Empirics," *Development Working Paper n.23, Harvard University, Center for International Development.*

——— (2007) "Moving to nice weather," *Regional Science and Urban Economics*, Vol. 37(3), pp. 375–398.

Rappaport, Jordan and Jeffrey D. Sachs (2003) "The United States as a Coastal Nation," *Journal of Economic Growth*, Vol. 8(1), pp. 5–46.

Roback, Jennifer (1982) "Wages, Rents and the Quality of Life," *The Journal of Political Economy*, Vol. 90(6), pp. 1257–1278.

De la Roca, Jorge and Diego Puga (2013) "Learning By Working In Big Cities," *Working Papers wp2013 1301, CEMFI.*

Rosenthal, Stuart S. and William C. Strange (2003) "Evidence on the Nature and Sources of Agglomeration Economies," *Handbook of Regional and Urban Economics*, Vol. 4, pp. 2119–2171.

Rossi-Hansberg, Esteban and Mark L. J. Wright (2007) "Urban structure and growth," *Review of Economic Studies*, Vol. 74(2), pp. 597–624.

Saiz, Albert (2010) "The Geographic Determinants of Housing Supply," *The Quaterly Journal of Economics*, Vol. 125(3), pp. 1253–1296.

Samila, Sampsa and Olav Sorenson (2011) "Venture Capital, Entrepreneurship and Economic Growth," *The Review of Economic Studies and Statistics*, Vol. 93(1), pp. 338–349.

Saxenian, AnnaLee (1994) *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*, Cambridge, MA: Havard University Press.

Smith, Adam (1776) *An Inquiry into the Nature and Causes of the Wealth of Nations*, London: Printed for W. Strahan, and T. Cadell.

Stiglitz, Joseph E. (1977) "The theory of local public goods," in *The Economics of Public Services*: MacMillan Press, pp. 274–333.

Vernon, Raymond (1960) *Metropolis 1985*, Cambridge, MA: Havard University Press.

Zipf, George Kingsley (1949) *Human Behaviour and the Principle of Least Effort: An Introduction to Human Ecology*, Cambridge: Addison Wesley.