Spatial Sorting of Heterogeneous Agents in Stratified Regions

Axel Grøn Roepstorff, 202004982

May 1, 2023

AARHUS UNIVERSITY
Department of Economics and Business Economics
School of Business and Social Sciences
Ba.Oecon

Supervisor: Torben M. Andersen

Number of characters incl. spaces, tables, and graphs: 70,414 (29.3 standard pages)

Danish Title: Spatial Sortering af Heterogene Agenter i Stratificerede Regioner

This thesis may be published

Abstract

Spatial inequality is an empirical fact of modern economies. Viewing it through the traditional lens of consumer goods paints an unclear picture since public amenities have a remarkable effect on location choices. This paper considers the spatial sorting of heterogeneous agents engaged in a CES-type production function. Additionally, it shows a novel method of calculating discrete location choices under idiosyncratic selection criteria. The resulting model presents a reason for agglomeration that combines sorting, selection, and economies of scale with quasi-idiosyncratic utility shocks that covary along education with dynamic replacement of agents. Additionally, the paper evaluates the model through simulation and provides evidence for convergence to a steady state. Finally, the model simulates the effects of a shock to housing endowments and shows how this has a diverging effect on regional education levels.

Contents

1 Introduction

The field of spatial economics has been researching the cross-section of geography and economics for more than fifty years, trying to reconcile the fact that economic activity has a tendency to agglomerate in certain areas. Taken to the extreme, there is not a single aspect of economic decision making that is not permeated by the influence of geography. All markets are at least partially distinguished by their distance from other competing markets to a degree that has been recognised as far back as Marshall's seminal Principals of Economics (Marshall, 1890). To accurately reflect economic decision making, it is imperative to incorporate the implications of the spatial dimension, but the addition of this element has a severe potential to act as a Pandora's Box where the search for increasingly complex analyses results in meaningless results with no robustness. To understand the need for simplicity, one must only read Krugman's 1991 paper "Increasing Returns and Economic Geography". Explaining geographic agglomeration within the framework of classical utility maximisation by introducing trade costs and increasing returns to scale requires very few additional assumptions to tractably predict stable spatial equilibria. Since the 90's, landmark papers in economics have since allowed for models of wide heterogeneity on both the supply and demand side and even the introduction of idiosyncratic preferences following the discrete choice modelling as presented by McFadden (1974). This paper attempts to synthesise these two schools of thought and presents a theoretical model that allows for both a heterogeneous labour market and what might be called quasi-idiosyncratic preferences. Inspired by Behrens et al (2014), the model allows for sorting and heterogeneity in productivity. Labour is the only input factor, but it is separated by productivity types as in Roger and Wasmer (2011). Drawing from the insights of (Bødker and Kaarsen, 2023), the model allows for recognition of the fundamental differences between cities of different sizes. The model allows for stratification of cities into different types to recognise that not all cities compete for labour on the same terms. From here, follows a walkthrough of relevant literature, appropriate previous models, the model and its assumptions, a simulated estimation of model equilibrium, and a simulated policy experiment of a shock to housing.

2 Literature

Economic activity inside countries is not uniformly distributed. This fact is so ubiquitous and uncontroversial that it is the opening statement of five of the articles quoted in this paper. From

Krugman's model of New Economic Geography (Krugman, 1991) to Redding & Rossi-Hansberg (2017), many different economists have tried to reconcile the multiple troubles of identifying which mechanisms drive economic agglomeration. Some, like Krugman (1991), focus on the effects of economies of scale and transport costs. Others, following Roback (1982), construct a spatial equilibrium based on amenities on the demand side. Both schools of thought have had increasingly advanced models win ground, relying on the models pioneered by Melitz (2003)(see Ottaviano, 2011). On the one hand, allowing for heterogeneous input factors opens for a wide variety of extensions and changes the set of equilibrium conditions since equilibrium only has to hold for identical types, and on the other hand, systematic modelling of idiosyncratic preferences allows for models that make less strict assumptions on human behaviour. Behrens et al (2014) and Ahlfeldt et al. (2015) respectively, represent models that incorporate stochastic modelling in spatial economics that ascribe to the two separate schools of thought. Behrens et al (2014) allows for a heterogeneous set of workers sorting according to productivity, and The Danish Secretariat of the Economic Councils (DØRS, 2021, Bødker and Kaarsen, 2023) allows for a home-field advantage type of model that sorts people according to their hometown at age 21. Using these ideas as a jumping off point, it can be relevant to return to the preference for public amenities and the empirically founded heterogeneity in this preference (Koster et al. (2016), Bayer et al. (2007), and Mughan et al. (2022)) along education level. This connection is fairly standard in the literature, being postulated without references in The Handbook of Urban and Regional Economics Moretti (2004). Neglecting a potential educational bias in sorting according to amenity preferences will in this case result in mistaken predictions for sorting and a construction of false causal relations.

Different schools of thought ascribe different reasons for industrial agglomeration. On the supply side, it can in general be separated into three different categories. The 'New Economic Geography' school has economies of scale and input-output linkages result in path-dependent agglomeration. Alternatively, the externalities literature, in general ascribing to either Glaeserian or Marshall-Arrow-Romer externatilities (Glaeser et al., 1992), attributes this agglomeration to knowledge and learning spillovers from the concentration of firms, diversified or otherwise. Finally, the sorting and selection literature has won purchase since Melitz (2003) as has been well exemplified by recent micro-founded models (see Behrens et al., 2014, Behrens and Robert-Nicoud, 2015). These models show that sorting and selection can drive the largest cities to have a higher productivity by having higher urban living costs . The existence of externalities presents an option for a pareto improvement. As a matter of fact, however, oftentimes place-based policies are created to move jobs from the healthy cities to the less well-endowed regions (Neumark and Simpson, 2015). This paper will focus on the causal relationships proposed by different models, construct a new synthesised model, and finally expose this model to a place-based policy shock for a comparison with similar literature.

3 A Walkthrough of Relevant Models

3.1 The Behrens, Duranton and Robert-Nicod Model

In Behrens et al (2014), Behrens, Duranton, and Robert-Nicod present a model that allows for sorting according to productivity in a model that is reminiscent of Melitz (2003). They model a country of individuals that are homogenous in their preferences yet heterogenous in their productivity φ, an interaction of the parameters talent, t, and serendipity, s, that are drawn in a two-stage process. This combination ends up determining a worker's selection into either labour or entrepreneurship. The unique city-wide productivity cutoff determines which individuals innovate and which individuals work. This is determined as a cutoff where entrepreneurship profits must match the worker payoff calculated by the city-wide wage compensated according to the productive equivalent of a unit of work. As a result, workers in general have their type determined by productivity with some overlap between productivities in each type. Costs of agglomeration are modelled with an urban cost factor, the elasticity of which to agglomeration is assumed to be larger than a term ε which determines the elasticity of substitution. As the urban cost elasticity and ε converge, an empirically motivated assertion, the model predicts a unique equilibrium of perfect sorting. The model allows for heterogeneous productivity types that complement each other. There is implicit substitution of these tasks that is determined endogenously. It does, however, only determine a difference in jobs with regards to supplying labour vis-a-vis entrepreneurship and thus only allows for two productivity types in a setting of continuous productivity differences. Since the returns to productivity are increasing in city size, and vice versa, rising urban costs drive higher productivity workers to larger cities. Additionally, the model creates a framework for dynamic relocation very similar to the method set up in this paper.

3.2 The Ahlfeldt, Reddiung, Sturm, and Wolf Model

Ahlfeldt, Redding, Sturm & Wolf's 2015 paper present a modern approach to modelling of idiosyncratic location match preferences. Grounded in urban economics, they use literature-standard Cobb-Douglas production and utility functions interacting labour input and floor-area with an employment-density-driven externality measure, an approach replicated in Redding and Rossi-Hansberg (2017). In choosing these simple fundamentals, the model derives much of its explanatory power from the inclusion of trade costs, travel costs, amenities, and idiosyncratic shocks between city blocks to model urban variation. There are two possible drivers of agglomeration. First of all, high employment density spills over in productivity gains in a block. Additionally, high employment density spills over on the public amenities side. This represents a distinctly Glaeserian view of agglomeration theory. The idiosyncratic shock z_{ijo} is considered to be an independent draw of a Frechét distributed random variable with certain parameters T_i, E_j, ϵ that in a structured fashion affect the values of the shock.

$$F(z_{ijo}) = e^{-T_i\, E_j, z_{ijo}^{\varepsilon}}$$

This enters the indirect utility in such a way that the indirect utility of working in block j and living in block i for an individual o is equivalently Frechét distributed. Finally, the Frechét distribution's properties determine that the share of workers and residents in any given region pairing has a closed form solution. This property is well-suited for analyses of homogenous agents and is a reason for the method's prevalence in discrete choice modelling.

This method and its benefits are however to my knowledge only well documented for homogenous agents. Bødker and Kaarsen (2023) do calibrate their model to specific types of heterogeneity, but that is done manually for a discrete set of cities. Given that one of the objectives of this paper is an enquiry into a spectrum of heterogeneous agents, it is unclear that this result would hold in such a case. Additionally, this model is mainly concerned with the sorting of space in a dense urban area. As a result of this, labour mobility in terms of commuting costs and trade costs are essential components of the model' explanations, two points that will not be the main focus of this paper. This model does, however, provide an advantageous starting point. The model described in this paper has chosen the same simplistic framework for consumption and let the idiosyncratic shock interact with utility in a similar way.

3.3 The Bødker & Kaarsen Model

Bødker & Kaarsen (2023) expounds on the framework given by Ahlfeldt et al. (2015) using the model defined in DØRS (2021) by the same authors. It serves as a point of inspiration for many of the relations created in the forthcoming model and contributes specifically with some intuition on Danish spatial sorting. The stated goal is to create a model that fits the geographical parameters of a small country like Denmark, and its contributions are threefold, of which two cover some of the same problems that this paper attempts to address: the use of spatial general equilibrium models on countries with one uniquely large city and the use of idiosyncratic preference in consumer location preferences. The model is primarily concerned with policy-decisions and the migration of workers based on their geographically predetermined type. Production is assumed to be of a Cobb-Douglas type with labour and land as the input factors. This implies that all labourers find themselves competing on even footing and that firms contribute to housing demand. Additionally, there are no implicitly assumed economies of scale since there is no fixed cost of operation and production is homogenous of degree one. The fraction of consumers settling in a given town i and working in a town j conditional on having type f is given by the following equation.

$$\pi_{i,j|f} = \frac{B_{i|f}E_{j|f}R_{i,j|f} \cdot e^{-\phi \cdot travel\,time} \left(q_i^{1-\beta}\right)^{-\epsilon} \cdot {w_j}^{\epsilon}}{\sum_r \sum_s B_{r|f}E_{s|f}R_{r,s|f} \cdot e^{-\phi \cdot travel\,time} \left(q_r^{1-\beta}\right)^{-\epsilon} \cdot {w_s}^{\epsilon}}$$

In which $B_{i|f}E_{j|f}R_{i,j|f}$ represent preferences for living in city i while working in city j and having any eventual specific relations between the two cities, and ϵ represents the heterogeneity in preferences between cities and types. q_i and w_j are the housing rent and the wage in the respective cities. Obviously, since the housing market enters into both the consumer and producer side of the equation, competition on the housing market increases doubly in denser areas. One of the advantages of this model is that it in itself results in labour flows that take the form of a gravity equation where the size and distance between cities determines flows.

However, the model does not discuss the implications of heterogeneity within education and productivity. While workers might vary in productivity across regions, this is not driven by their innate productivity but rather a serendipitous correlation between their preference for a region and its underlying capital mass. Thus, the wage premium for cities is assumed to be identical for all workers. This means that a worker is not expected to relate their job preference to the labour market and thus takes the citywide wage as a given.

3.4 A Consolidation of Expectations

These models present an overview of some relevant articles within the 'New' New Economic Geography spatial literature as coined by Ottaviano (2011). Many of these considerations will be determining for the forthcoming model. It is, however, important to align some expectations. Behrens et al. (2014) produce a model that attempts to explain perfect sorting in a large and highly diverse country. This model will not be well suited to describe diverse cities in a country with few regions and one unique superstar region. Additionally, it does not assume that amenities are prevalent. Opposingly, the models described by Ahlfeldt et al. (2015) and Bødker and Kaarsen (2023) allow for idiosyncracies but no heterogeneity. The main contribution of the following model is the synthesis of these two methods.

4 Theoretical Model

The model sets up a country split into three types of regions: the metropolis, the medium-sized regions and rural regions. Any given region will be placed into one of these strata which have identical matching preference distributions within themselves but differ across types. This matching captures any excess utility from locating into a region in addition to the consumption of manufacturing goods and land. To capture the variety within a population, the amenity preference for a region is determined globally as a random variable. Each worker will thus experience a vector of different amenity preference matches for each possible region. This added heterogeneity allows for sorting that is driven according to productivity but informed by unobserved characteristics. However, as a simplifying assumption, regions are assumed to exist in autarky; there exists neither trade nor commuting. Finally, the model is based on a dynamic flow of agents inspired by Behrens et al. (2014). Essentially, it is assumed that workers are perfectly mobile at the moment they enter the job market and perfectly immobile in all further periods. To capture this, the model is based in discrete units of time where a generation of size $\bar{n}$ enters the economy just as a generation of equal size disappears. This creates a model that does not take full migration into account while still allowing for a degree of mobility.

4.1 Necessary Assumptions

The essential question of the model is as follows: could a worker decide to remain in a region in spite of better consumption conditions elsewhere? To argue for an affirmative answer to this question, it is necessary to allow for some degree of region-specific bias. While it is possible to let this bias be deterministically set by exogenous factors, it is more reflective of the multiplicity of a population to determine this stochastically. The problem thus becomes one of specification. How exactly should this bias be modelled? Ahlfeldt et al (2015) lets some regions enjoy a homogeneously better disposition for a high amenity, and Kaarsen & Bødker (2023) assume that there is a home-region advantage associated with living situation at the age of 21. Finally, Tabuchi and Thisse (2002) use discrete choice theory to model this same randomness using quasilinear utility functions. All of these can help explain agglomeration, but the articles themselves assume homogeneity in the type of economic agent and their place in the workforce. This is the fundamental distinction of this model from the existing literature. In this model, the preference for a specific region is conditional on their educational attainment which too is drawn from a random distribution. Education is furthermore considered to be a measure of productivity in keeping with literature standards (Moretti, 2004). The preference can be understood as a wide variety of factors of which a number covary with educational attainment. If these effects are strong and heterogeneous enough, there is reason to believe that otherwise identical regions will differ in size and demographic makeup according to this systematically stochastic yet idiosyncratic bias. Setting an initiation point for the model will be covered in further sections. First, it is relevant to analyse the incentives of agents.

4.2 Consumer side

The consumer enters the market with full knowledge of their utility function and of a set of individual-specific parameters. The first, φ, determines the educational attainment of the individual and determines the remaining draws. With the productivity known, the consumer draws a set of preferences equal to the number of regions in the country. For each region, there is one utility function j of which the consumer will choose their utility maximising option.

$$u_j = a_j c_j^{\alpha} h_j^{1-\alpha} \tag{1}$$

$$u^* = max\{u_1, u_2 ... u_n\} \tag{2}$$

where c is a CES-type aggregated good following Krugman (1980), and its expansion in Feenstra (2015) and h is the amount of land consumed. The land endowment in a region, $\bar{H}_j$, is considered to be perfectly inelastic. Following standard literature, land rents are assumed to exit the economy (see Moretti, 2004, Roback, 1982, Behrens et al., 2014). In other words, all housing consumption should be viewed as a contemporaneous rental paid to property owners outside of the economy. This is also necessary for housing consumption in one period to have no direct effect on housing consumption in other periods. Finally, housing being a strict expense puts less strain on the assumption of static expectations. Knowing the utility function, it is clear that marshallian demand for the manufacturing good and housing can be characterised by

$$c_j(\varphi) = \alpha \frac{w_j(\varphi)}{P_j} \tag{3}$$

$$h_j(\varphi) = (1-\alpha) \frac{w_j(\varphi)}{r_j} \tag{4}$$

where P is the price index and r is the land rent.[1]. The location choice of the consumer takes place only once and is taken immediately when they enter the labour market. It is assumed that moving costs are prohibitive after the initial location choice. Each consumer will choose the utility maximising location and supply one unit of their labour. This means that contemporaneous labour supply in a region is perfectly inelastic. The consumer has static expectations; they do not take into account the trend of the region's size and labour distribution but rather expect that they are already in a steady state, an assumption that turns out to be self-fulfilling as shown in section 5. They make their choice recognising only the immediate gain of the contemporaneous labour market. This choice of selection combines principles of Bødker & Kaarsen (2023), Behrens et al (2014) and Ahlfeldt (2015).

4.3 Solving the Labour Market

4.3.1 Labour Supply

The total production of a region j is determined by the availability of effective labour L_j which is characterised by imperfect substitution between tasks differentiated by productivity φ, extending on the aggregates presented in Roger and Wasmer (2011). Each worker produces one unit of labour.

$$L_j = \left(\int_{\underline{\varphi}}^{\overline{\varphi}} \delta(\varphi)\, l_j(\varphi)^{\rho} d\varphi \right)^{\frac{1}{\rho}} \tag{5}$$

$l_j(\varphi)$ is the supply of workers of productivity type φ in the firm, δ is the productivity parameter as a function of the type φ, N is the amount of workers in the region, and ρ is the elasticity of substitution between jobs. Furthermore, $\rho < 1$, implying imperfect subsitutability of tasks. While this is a fairly intuitive measure in the case of discrete distributions of educational attainment, it deserves some further explanation in the continuous case where each education draw is

[1] This is derived in appendix A

technically unique. To reconcile this in a continuous framework, the worker is assumed to have perfect knowledge of a continuous approximation of the educational distribution in the region and will use that multiplied by the size of the region to capture the degree to which they are overrepresented.

$$l_j(\varphi) = p(\varphi|j)N_j, \tag{6}$$

where $p(\varphi|j)$ is the probability of finding a person of a given productivity in the region. This fully specifies the effective labour supply which is exogenous for the individual firms.

This representation of the final labour product is similar to the Ethier (1982) representation of monopolistic competition, but it should not be seen as a direct translation. Most importantly, the Ethier model realises gains from further specialisation of intermediate goods. Larger economies gain from an increase in produced varieties of intermediaries. There is not the same specialisation effect in this model since all regions accept the same spectrum of labour types. Agglomeration does therefor not happen through increased task variety. The aggregate supposes that each worker participates in intermediate tasks. The productivity of a worker is inversely correlated with the amount of equivalent workers in the region. Seen through a task lens, each educational type is uniquely suited to one task on the continuum. Their over(under) representation represents how well they can match their competence with the task demand. All tasks are equally important in production but are arranged in terms of their educational requirements represented by φ. In this paper's specification of the model, and for further calculations, the productivity δ is considered to be monotonically increasing in φ and indeed strictly be φ. This assumption has been chosen for computational reasons, but it is feasible that δ could be indeed both strictly concave, convex or even non-monotonic. Specifiying this relationship could merit its own paper, and as a result, productivity is assumed to simply scale linearly with φ. Intuitively, only a limited amount of workers will be able to actually take care of their specific tasks. Any excess workers will take care of adjacent tasks at the cost of imperfect substitution between tasks. Conversely, an underrepresented type will be compensated for the lack of qualified work within their type. This in itself is a drastic deviation from Behrens et al. (2014). In their model, productive workers will seek the cities since productivity yields increasing returns to density whereas this model results in a sort of 'biggest-fish-in-the-pond'-effect where the highest educated workers have an incentive to locate in a region where they will be far better educated than their counterparts through sheer lack of competition, discouraging the existence of perfectly sorted equilibria. To some extent, this can be viewed as following the externalities described in Glaeser et al. (1992) which find that underrepresentation is the biggest predictor for firm growth. Assuming this type of labour aggregate imposes as a rule that all firms in all regions will have the same structure of tasks that must be completed in order to produce. All regions will need garbage collectors, carpenters, health professionals and CEO's, and this demand is implicitly assumed to scale exactly to region size.

4.3.2 Labour Demand

Firms face a production function using labour as a single input with a fixed labour cost following Krugman (1980). Region-wide production can be viewed as

$$Y_j = \bar{\delta}_j(\varphi)L_j - F \tag{7}$$

where $\bar{\delta}_j(\varphi)$ is the average value creation from education of the region. This measure implies that even though tasks are imperfectly substitutable, there are productivity spillovers that come as a result of having a higher educated workforce. The nationwide equal fixed cost, F, is considered to be a labour cost that must be incurred before production can take place following Krugman (1991). This cost does *not* scale with the average productivity, a fact that increases the diverging effects of sorting. Since profit maximisation and utility maximisation are symmetric across all regions, regional subscripts will be dropped for these derivations to ease notation. Marginal product of labour as a function of education can be determined as

$$\frac{\partial Y}{\partial l(\varphi)} = \frac{\partial Y}{\partial L}\frac{\partial L}{\partial l}$$

Applying this to the production function yields.[2]

$$MPL(\varphi) = \bar{\delta}(\varphi)\delta(\varphi)\left(\frac{L}{l(\varphi)}\right)^{1-\rho} \tag{8}$$

Akin to monopolistic competition (see Feenstra, 2015), workers are assumed to have an infinitesimally small contribution to average education. This simplifying assumption eases calculation considerably since derivatives can be taken without use of the product rule. For a sufficiently large N and M and small φ, the marginal labour contribution of a worker in a firm can be argued to be completely decoupled from the effect it has on the marginal returns to workers in general. While this marginal product of labour has been found for the region as a whole it is applicable for the individual firms as stated by lemma 1.

[2]Proof in appendix A

Lemma 1. *Average marginal productivity of labour is invariant of firm size and only defined by regionwide parameters.*

$$M\bar{P}L = \bar{\delta}(\varphi)\frac{L}{N}$$

This shows that while there are economies of scale through the fixed costs, there are no productivity gains from being a larger firm. A proof of this lemma is shown in appendix A. Furthermore, in the following calculations, δ is assumed to take the form $\delta(\varphi) = \varphi$.

4.3.3 Clearing the Market

The workers take their wage as a given and do not vary their effort with wage. The real wage is thus determined by firm demand with labour supply being contemporaneously inelastic. The labour market will clear when the price is equal to the marginal cost of producing times a markup as is standard in the case of Dixit-Stiglitz-type monopolistic competition. Rearranging yields

$$\frac{w(\varphi)}{P} = \frac{\bar{\varphi}\varphi}{1+\mu}\left(\frac{L}{l(\varphi)}\right)^{1-\rho} \tag{9}$$

where $1+\mu = \frac{\sigma}{\sigma-1} \implies \mu = \frac{1}{\sigma-1}$ following Feenstra (2015).

Following Lemma 1, average real wage will be

$$\frac{\bar{w}_j(\varphi)}{P_j} = \frac{\bar{\varphi}_j}{1+\mu}\frac{L_j}{N_j} \tag{10}$$

The ratio of $\frac{L}{N}$ is homogenous of degree zero. Additionally, in a symmetric equilibrium, average wages must be identical. As such, this ratio, while calculated on a region-wide basis, is applicable on the firm level as well. If there is one uniquely optimal combination of workers, every firm must construct exactly this combination to sell under the same price in a symmetric equilibrium. To reconcile this, workers, and especially those in short supply in a region, must be considered as almost freelancers, selling a continuum of their one unit of labour to the highest bidders. This ensures that firms can have a continuous input of labour.

4.3.4 Pricing and demand

In a symmetric equilibrium where price and consumption are identical across all firms, all monoponistically competing firm set a fixed markup over marginal costs which must be average wages over the average marginal labour productivity. $P = (1+\mu)\frac{\bar{w}}{M\bar{P}L}$ Using price as the numeraire in the economy allows for an expression where wage is simply the marginal productivity divided by the markup.

$$\bar{w} = \frac{M\bar{P}L}{1+\mu} \tag{11}$$

Additionally, free entry implies a zero profit condition which sets prices as

$$\frac{P}{\bar{w}} = \frac{1}{M\bar{P}L} + \frac{F}{N\bar{c}}^{3} \tag{12}$$

where $\bar{c}$ is average consumption of each individual intermediate good.

The number of firms is determined endogenously with the full employment condition. Firms end in a symmetric equilibrium where every firm produces the same $N\bar{c}$ thus implying that production can be given by the equilibrium of (11) and (12).

$$\begin{aligned}
\frac{1+\mu}{M\bar{P}L} &= \frac{1}{M\bar{P}L} + \frac{F}{N\bar{c}} \\
\frac{\mu}{M\bar{P}L} &= \frac{F}{N\bar{c}} \\
N\bar{c} &= \frac{FM\bar{P}L}{\mu}
\end{aligned} \tag{13}$$

Total demand of each firm allows us to finally close the manufacturing sector realising that total labour demand is a scaling of each firm's labour demand which must be equal to the contemporaneous labour supply to fulfill the full employment condition. It is clear in a symmetric equilibrium for firms that $M\,\bar{L}_d = \bar{L}$. Labour demand for a firm is given by $\left(F + \frac{N\bar{c}}{M\bar{P}L}\right) = F(1+\frac{1}{\mu})$. Collecting this allows us to predict the mass of firms in equilibrium.

$$M = \frac{L}{F(1+\frac{1}{\mu})} = \frac{L}{F\sigma} \tag{14}$$

The limit of this equation as σ goes to infinity is 0 which is reflective of the fact that that case would represent perfect competition where no firms could charge any markup in which case there would be no firms in the region.

4.4 Housing and Total Indirect Utility

Knowing that demand for the housing good is a fraction of the real wage over rents r from (4), it is clear that the rent can be found as

$$r = \frac{\bar{\varphi}L(1-\alpha)}{H(1+\mu)} \tag{15}$$

Substituting (9), (14), and (15) into the demand function yields the indirect consumption utility function[4]

[3]Proven in appendix A

[4]For a full derivation, see appendix A

$$\tilde{v} = \left(\left(\frac{\alpha\varphi\bar{\varphi}}{(1+\mu)l(\varphi)^{1-\rho}} \right)^{\frac{\sigma-1}{\sigma}} \left(\frac{1}{F\sigma} \right)^{\frac{1}{\sigma}} L^{\frac{\sigma+\rho(1-\sigma)}{\sigma}} \right)^{\alpha} \left(\frac{\varphi\bar{H}}{L^{\rho}l(\varphi)^{1-\rho}} \right)^{1-\alpha} \tag{16}$$

This expression formalises that there is a push and pull of housing versus the access to varieties. This is analogous to the trade-off described in Behrens et al. (2014) where urban costs - which are caught by the housing term in this model - and the love for variety interact to guarantee that cities do not grow infinitely. It starkly decreases the odds of there being a completely divergent equilibrium for a sufficiently large $\bar{H}_j$ and small α since the final worker in a region would be able to consume the entire housing good. In fact, it is possible to show that agglomeration of labour is strictly centrifugal in the case of identical regions, a proposition that will be shown in a later section. In other words, the model creates a set of bounds for the parameters if density is to drive agglomeration as standard literature assumes. This model also creates a dilemma for the social planner. A purely production maximising planner will encourage the agglomeration in one region to maximise the availability of varieties. Specifically, the increase of housing outside of the capital area will act to actively decrease the availability of varieties to consumers.

4.5 Idiosyncratic Region Preferences

To accurately include region-specific idiosyncratic amenity matches, it is necessary to specify the type of the distribution for region-preference. Various approaches have been used in the literature from Tabuchi & Thissé (2002), drawing from discrete choice theory and modelling using a conditional logit model, to Ahlfeldt (2015) and Kaarsen & Bødker (2023) using McFadden's trick in a Frechét distribution to model the relationship. My approach leans heavily upon the approach chosen in Kaarsen & Bødker (2023) while modifying the choice of conditional parameter. In my model, regional preference is correlated with educational attainment. Specifically, the distribution of a region's match parameter follows a pareto-distribution with parameters ν, a_{min}, φ and ψ that are all strictly positive and can vary between regions of different types. The distribution is given as follows:

$$a_j \sim Pareto\left(a_{min,j}, \frac{\nu_j}{\varphi^{\psi_j}}\right) \tag{17}$$

$$P(a_j < a) = F_{A,j}(a) = 1 - \left(\frac{a_{min,j}}{a}\right)^{\frac{\nu_j}{\varphi_j^{\psi}}} \; \forall\, a \geq a_{min,j} \tag{18}$$

$$f_j(a) = \frac{\frac{\nu_j}{\varphi^{\psi_j}} \cdot a_{min,j}^{\frac{\nu_j}{\varphi^{\psi_j}}}}{a^{\frac{\nu_j}{\varphi^{\psi_j}}+1}} \tag{19}$$

The Pareto distribution has the desirable effect of being a transformation of the exponential distribution and thus following its property of having an identical shape when it is bounded from the left. This allows for a consistent solution to the question of region-worker match draws. It is necessary to note the inclusion of φ in these distributions. As φ grows, the shape parameter decreases for all values of $\psi > 0$. Proposition one proves the existence and convergence of a solution to the problem of preference prediction in the case of three regions.

Proposition 1. *Let there be three regions $a = \{a_1, a_2, a_3\}$ with region preferences according to the distributions described above differing only in the degree of correlation with φ such that $\psi_1 < \psi_2 < \psi_3$ The likelihood of region 1 having the highest amenity draw can be characterised by the following equation which converges.*

$$P(a_1 = max(a)) = \int_{a_{min}}^{\infty} \int_{a_{min}}^{\infty} \frac{a_{min}}{max\{c,k\}^{\frac{\nu}{\varphi^{\psi_1}}}} \frac{\frac{\nu}{\varphi^{\psi_2}} a_{min}^{\frac{\nu}{\varphi^{\psi_2}}}}{c^{\frac{\nu}{\varphi^{\psi_2}}+1}} \cdot \frac{\frac{\nu}{\varphi^{\psi_3}} a_{min}^{\frac{\nu}{\varphi^{\psi_3}}}}{k^{\frac{\nu}{\varphi^{\psi_3}}+1}} dc\, dk \tag{20}$$

Proof. Utilising the law of total probability, it is possible to rewrite a joint probability into an integral of marginal and conditional probabilities.

$$\begin{aligned}
P(a_1 > a_2 \cap a_1 > a_3) &= \int_1^{\infty} P(a_1 > a_2 \cap (a_1 > a_3 | a_3 = k)) \cdot P(a_3 = k) dk \\
&= \int_1^{\infty} P(a_1 > a_2 \cap a_1 > k) \cdot P(a_3 = k) dk \\
&= \int_1^{\infty} \int_1^{\infty} P(a_1 > a_2 \cap a_1 > k | a_2 > c) \cdot P(a_2 = c) \cdot P(a_3 = k)\, dc\, dk \\
&= \int_1^{\infty} \int_1^{\infty} P(a_1 > c \cap a_1 > k) \cdot P(a_2 = c) \cdot P(a_3 = k)\, dc\, dk
\end{aligned}$$

By using the law of total probability twice, it is possible to these joint probabilities into integrals of probabilities conditional on constants. We can finally use marginal probabilities to decompose the joint probability.

$$P(a_1 > c \cap a_1 > k) = P(a_1 > c | a_1 > k) \cdot P(a_1 > k)$$

The characteristics of the pareto distribution, imply that the distribution of a variable conditional on the variable being larger than a constant has the same shape parameter with a new minimum.

It is clear that this can be collected in the joint probability

$$\begin{aligned}P(a_1 > c \cap a_1 > k) &= (1 - P(a_1 < c | a_1 > k)) \cdot (1 - P(a_1 < k)) \\ &= min\left\{ \left(\frac{k}{c}\right)^{\frac{\nu}{\varphi^{\psi_1}}}, 1 \right\} \cdot \left(\frac{a_{min}}{k}\right)^{\frac{\nu}{\varphi^{\psi_1}}} \\ &= min\left\{ \left(\frac{a_{min}}{c}\right)^{\frac{\nu}{\varphi^{\psi_1}}}, \left(\frac{a_{min}}{k}\right)^{\frac{\nu}{\varphi^{\psi_1}}} \right\} \\ &= \frac{a_{min}^{\frac{\nu}{\varphi^{\psi_1}}}}{max\{c, k\}^{\frac{\nu}{\varphi^{\psi_1}}}}\end{aligned}$$

Finally combining all the likelihoods yields

$$P(a_1 > c \cap a_1 > k) \cdot P(a_2 = c) \cdot P(a_3 = k) = \frac{a_{min}^{\frac{\nu}{\varphi^{\psi_1}}}}{max\{c, k\}^{\frac{\nu}{\varphi^{\psi_1}}}} \frac{\frac{\nu}{\varphi^{\psi_2}} a_{min}^{\frac{\nu}{\varphi^{\psi_2}}}}{c^{\frac{\nu}{\varphi^{psi_2}}+1}} \cdot \frac{\frac{\nu}{\varphi^{\psi_3}} a_{min}^{\frac{\nu}{\varphi^{\psi_3}}}}{k^{\frac{\nu}{\varphi^{\psi_3}}+1}}$$

And thus

$$P(a_1 = max(a)) = \int_{a_{min}}^{\infty} \int_{a_{min}}^{\infty} \frac{a_{min}^{\frac{\nu}{\varphi^{\psi_1}}}}{max\{c, k\}^{\frac{\nu}{\varphi^{\psi_1}}}} \frac{\frac{\nu}{\varphi^{\psi_2}} a_{min}^{\frac{\nu}{\varphi^{\psi_2}}}}{c^{\frac{\nu}{\varphi^{\psi_2}}+1}} \cdot \frac{\frac{\nu}{\varphi^{\psi_3}} a_{min}^{\frac{\nu}{\varphi^{\psi_3}}}}{k^{\frac{\nu}{\varphi^{\psi_3}}+1}} dc\, dk$$

To argue conversion, it is sufficient to consider the convergence of the final part of (20)

$$\int_{a_{min}}^{\infty} \int_{a_{min}}^{\infty} \frac{\frac{\nu}{\varphi^{\psi_2}} a_{min}^{\frac{\nu}{\varphi^{\psi_2}}}}{c^{\frac{\nu}{\varphi^{\psi_2}}+1}} \cdot \frac{\frac{\nu}{\varphi^{\psi_3}} a_{min}^{\frac{\nu}{\varphi^{\psi_3}}}}{k^{\frac{\nu}{\varphi^{\psi_3}}+1}} dc\, dk$$

which must converge as it is the product of two density functions. (20) is smaller than this for all relevant values of a_{min}, ψ and ν and hence converges. □

While analytically complex, this integral can easily be computed, and the result is simple and unambiguous; the higher the education, the higher the likelihood of choosing region 3 *ceteris paribus*. For a visual representation of this, figure 1 shows the expected share of workers with a given productivity that will draw their highest match for each region for all given levels of education up to 6. If worker distribution had no impact on productivity and production, and workers had no consumption preferences or all regions had identical consumption characteristics, this would fully specify the final distribution of education types across regions in the model when combined with full information of the global distribution of education. The implication of such a degree of heterogeneity is massive. The typical equilibrium condition of equal wages and consumption is now neither necessary nor sufficient as an unreasonably large draw of productivity would imply a willingness to settle in a region with consumption prospects close to zero for the individual. Additionally, the workers draw their region specific matching preferences *ex ante*

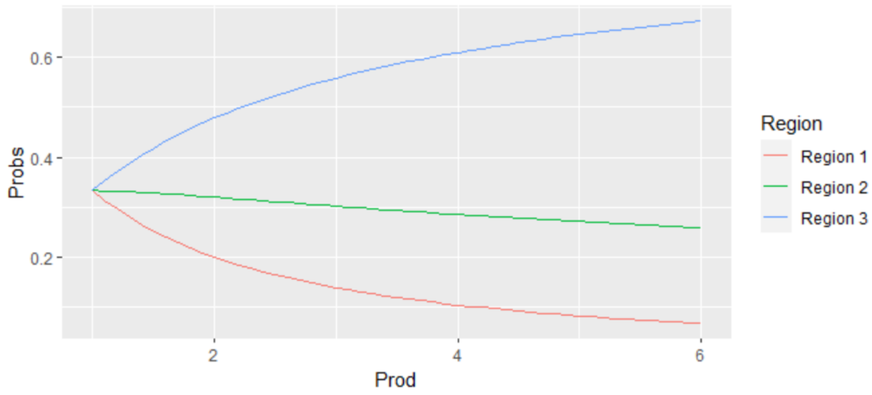


Figure 1: The probabilities of choosing regions 1, 2, and 3 given a certain productivity

in a distinct deviation from Behrens et al. (2014). Workers therefore take their decisions with full information, and their equilibria criteria can therfor not be carried over indiscriminately. However, if the educational distribution is to remain unchanged, it is possible to create a sort of stochastic equilibrium criterion. This will be covered in the numerical section.

4.6 The Model Closed

Finally, it is possible to close the model in some sense. To reiterate, utility is modelled as an interaction between consumption given by an indirect utility function and a region-individual match.

$$u_i = a_i \left(\frac{\alpha \varphi \bar{\varphi}}{(1+\mu) l(\varphi)^{1-\rho}} \left(\frac{L}{F(1+\frac{1}{\mu})} \right)^{\frac{1-\sigma}{\sigma}} \right)^{\alpha} \left(\frac{\bar{H} \varphi}{(1+\mu) L^{\rho} l(\varphi)^{1-\rho}} \right)^{1-\alpha} \tag{21}$$

Assuming that one region has a preference match that scales even slightly better will, all things being equal, result in agglomeration compared to the case of no discrimination. The crux of the time sequences is that there are structural parameters that would determine perfectly the distribution of all productivity types of the population which is perturbed by labour dynamics and inequality which are determined for each period. Combining equations (20) and (21) it is possible to create a more illuminating expression of the location trade-off.

$$u_1 > u_j \implies a_1 > a_j \frac{\left(\left(\frac{\alpha\varphi\bar{\varphi}_j}{(1+\mu)l_j(\varphi)^{1-\rho}}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{1}{F\sigma}\right)^{\frac{1}{\sigma}} L_j^{\frac{\sigma+\rho(1-\sigma)}{\sigma}}\right)^{\alpha} \left(\frac{\varphi\bar{H}_j}{L_j^{\rho}l_j(\varphi)^{1-\rho}}\right)^{1-\alpha}}{\left(\left(\frac{\alpha\varphi\bar{\varphi}_1}{(1+\mu)l_1(\varphi)^{1-\rho}}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{1}{F\sigma}\right)^{\frac{1}{\sigma}} L_1^{\frac{\sigma+\rho(1-\sigma)}{\sigma}}\right)^{\alpha} \left(\frac{\varphi\bar{H}_1}{L_1^{\rho}l_1(\varphi)^{1-\rho}}\right)^{1-\alpha}}$$

$$a_1 > a_j \left(\left(\frac{\bar{\varphi}_j}{\bar{\varphi}_1}\left(\frac{l_1(\varphi)}{l_j(\varphi)}\right)^{1-\rho}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{L_j}{L_1}\right)^{\frac{\sigma+\rho(1-\sigma)}{\sigma}}\right)^{\alpha} \left(\left(\frac{L_1}{L_j}\right)^{\rho}\left(\frac{l_1(\varphi)}{l_j(\varphi)}\right)^{1-\rho}\right)^{1-\alpha} \left(\frac{\bar{H}_j}{\bar{H}_1}\right)^{1-\alpha}$$

$$a_1 > a_j \left(\frac{\bar{\varphi}_j}{\bar{\varphi}_1}\right)^{\alpha\frac{\sigma-1}{\sigma}} \left(\frac{l_1(\varphi)}{l_j(\varphi)}\right)^{(1-\rho)(1-\frac{\alpha}{\sigma})} \left(\frac{L_j}{L_1}\right)^{\frac{\alpha(\rho+\sigma)-\rho\sigma}{\sigma}} \left(\frac{\bar{H}_j}{\bar{H}_1}\right)^{1-\alpha} = a_j\omega_{j,1} \tag{22}$$

This term can be separated into three distinct components which might be considered three keys to agglomeration. Firstly, there is a strict gain from average productivity in a given region; a decrease in the average productivity ratio will strictly decrease the attractiveness of region j. This might be described as attractiveness through a sorting channel. This term is only a part of the manufacturing good as average productivity does not skew one's consumption of exogenous housing. Secondly, all things being equal, being relatively underrepresented in a region increases relative utility. In other words, workers enjoy externalities from diversification. Finally, relative effective labour supply has an initially ambiguous and complex relationship to agglomeration. However, under careful examination, it is possible to understand whether centrifugal or centripetal forces dominate. Let us examine the exponent to relative effective labour supply.

$$\frac{\alpha(\rho+\sigma)-\rho\sigma}{\sigma} > 0$$

$$\alpha(\rho+\sigma) > \rho\sigma$$

$$\alpha > \frac{\rho\sigma}{(\sigma+\rho)} \tag{23}$$

Examining whether this condition can ever hold requires checking if the right-hand-side can be below 1.

$$\frac{\rho\sigma}{(\sigma+\rho)} < 1$$

$$\rho\sigma < \sigma+\rho$$

$$\sigma(\rho-1) < \rho$$

$$\sigma > \frac{\rho}{\rho-1} < 0 \tag{24}$$

(24) implies that any value of α will have an associated upper bound of σ and ρ which will determine whether relative density contributes to agglomeration or detracts from it as shown

in figure 2. More specifically, any value of $\rho < \alpha$ will result in path dependency. That being said, the actual effects of different values of ρ and σ are difficult to predict. As an example, an increase in ρ will decrease the importance of the second term and increase path dependency but simultaneously increase the value of L for all regions.

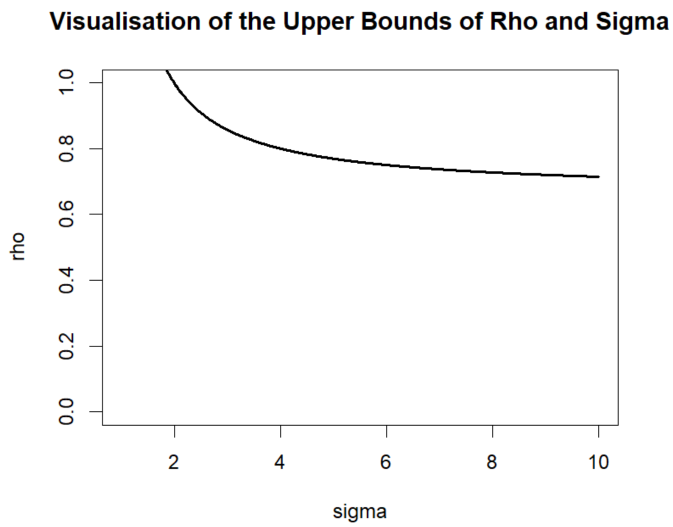


Figure 2

5 A Numerical Example

This section will consist of both simulated and analytical arguments for convergence to some sort of steady state under very few exogenous agglomeration-driving parameters. The example creates a country of three regions, one for each region type. Recalling the indirect utility function, all regions have been given identical endowments of all parameters of importance in the model save one. A full specification of these parameters is given in table 1.

Model simulation takes place as follows. In period $t = 0$, each region is endowed with a sample of 20000 workers whose education is drawn $i.i.d$ from a pareto distribution with parameters $k = 3, x_{min} = 1$, where k is the shape parameter and x_{min} is the the scale parameter. Asymptotically, this results in three indistinguishable regions. In all forthcoming periods, a new sample of $\bar{n}$ is drawn and settles in their utility-maximising region according to the conditions that have been supplied in the preceding section. Just as a new generation localises, a randomly chosen sample

Parameter	Region 1	Region 2	Region 3
$\hat{H}$	20000	20000	20000
$\bar{N}\|_{t=0}$	20000	20000	20000
F	50	50	50
α	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$
μ	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
σ	4	4	4
ρ	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$
a_{min}	1	1	1
ν	5	5	5
ψ	1	1.5	2

Table 1: Fixed Parameters for Regions

of the population equal to $\bar{n}$ perishes. There is no probability of a member of the new generation dying before taking part in the economy. This ensures that regions are not growing in size. To understand the impact of the amenities, it is relevant to first examine the case of settlement in periods $t = 1$ and consider the criteria for a hypothetical equilibrium in period $t = T$.

5.1 Estimating the labour PDF

In a continuous distribution of labour, the workers must be assumed to have a perfect estimate of the density of labour in their region. While theoretically possible, a completely non-parametric estimation of this density often fails to represent the qualities of the distribution of labour. Using commands like `density` and `estimatePDF` in R results in flawed estimates that converge quickly towards zero and do not follow the bounds implicit in the model. Specifically, the slow tail decay that will be relevant for labour whose productivity has been drawn from a pareto-distribution can be impossible to estimate consistently under non-parametric circumstances if the sample does not reach that level. As it turns out, however, a maximum-likelihood estimated pareto-distribution fits the data extremely well. Table 2 shows the mean squared errors of the population moments compared to estimated moments of a pareto-distribution collected from all 100 simulation periods. Additionally, figure 3 plots the empirical CDF of education for region 3 compared to a fitted pareto distribution showing that the two are visually indistinguishable. [5]

[5] A larger version of this graph can be found in appendix B for improved legibility

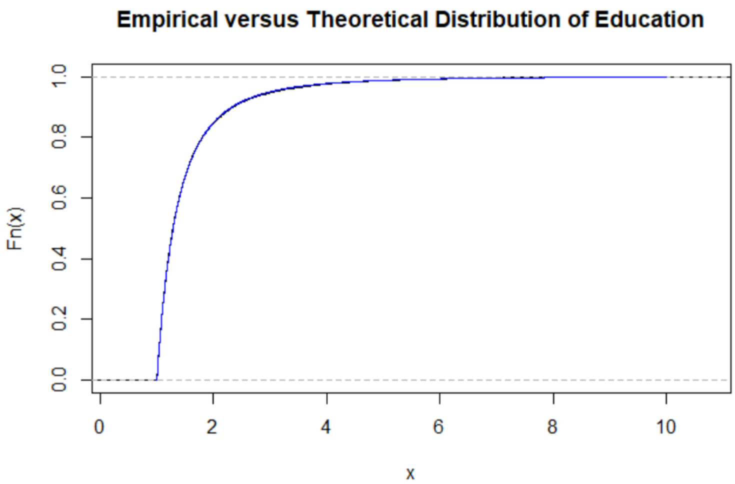


Figure 3: Empirical CDF compared to Parametric Estimation of the CDF for Region 3

The MSE of variance is far larger than the MSE for the means. This likely the result of the actual educational distribution being truncated at 10 for numerical tractability while consumers are assumed to estimate a non-bounded pareto distribution. There is as such a very small mass in the theoretical distribution (at the highest, 0.2%) that lies above 10 and converges towards infinity, contributing to the variance estimate. Seeing that region 3 systematically is estimated to have the lowest scale parameter, it is clear that this region is expected to have the highest deviation of non-parametric variance compared to parametric. Any fears that this approximation does not capture the distribution well, however, should be dispelled by the graphical evidence in figure 3.

Regions	Region 1	Region 2	Region 3
MSE of mean	0.0000	0.0000	0.0000
Maximum squared error from mean	0.0002	0.0002	0.0003
MSE of variance	0.0077	0.02507	0.1730
Maximum squared error of variance	0.0169	0.0465	0.2804

Table 2

After having examined the model under both parametric and non-parametric estimation, the choice of estimation does not seem to have a large effect on the characteristics of the equilibria, but the parametric approach requires far less computational power while smoothing out some apparent irregularities that result from non-parametrical estimation.

5.2 Period 1

In the first period, all regions are identical in terms of consumption parameters. It is thus only necessary to consider the likelihood of one region scoring the highest idiosyncratic preference. Following proposition one, it is is easy to predict the flow of workers for the first period. Multiplying the probability of drawing a given productivity by the probability of choosing the region yields the unconditional expected distribution of the workforce in each of the three regions as visualised in figure 4. This, too presents visual evidence in favour of a parametric estimation of labour productivity. Assuming that this theoretical distribution is pareto-distributed will be-

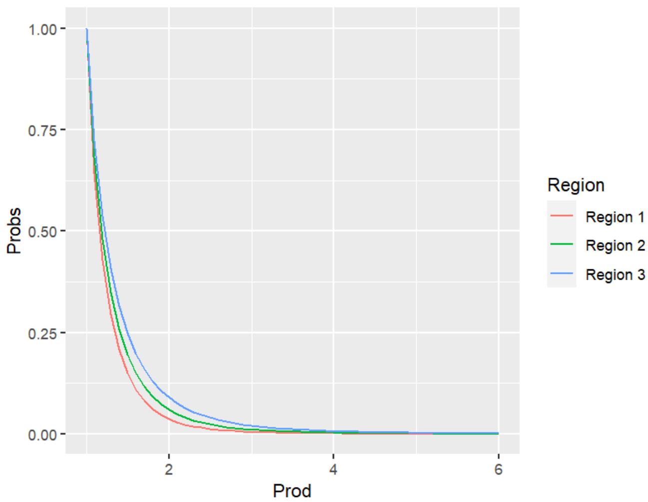


Figure 4: The unconditional distribution of education across regions 1, 2, and 3

come valuable to examine how the consumption side of utility $\tilde{u}$ affects both agglomeration and sorting. After the region choice has been made, a representative sample of the global population perishes. In the first period, all regions are of equal size, meaning that they will each lose $\frac{1}{3}\bar{n}$ of their population. Change in population is thus as follows:

$$\Delta N = \left\{ \begin{array}{lll} \Delta N_1 & = (0.300 - 0.333)\bar{n} & = -0.033\bar{n} \\ \Delta N_2 & = (0.332 - 0.333)\bar{n} & = -0.001\bar{n} \\ \Delta N_3 & = (0.368 - 0.333)\bar{n} & = 0.035\bar{n} \end{array} \right\}$$

That fact that total change in population in this calculation ends up being 0.001 is expected since rounding $\frac{1}{3}$ to decimal points systematically rounds away one unit of the lowest described numeral and does not reflect a change in total population over time. The immigration in each region is calculated from the given parameters in table 1 using equation (22). From the unconditional distributions, it can be seen that in this period, $\bar{\varphi}_{1,NG}$ is lower than the current average, an effect that is mirrored in region 3. Hence, both the effects of average productivity and agglomeration contribute in the same way to increase(decrease) the relative attractiveness of region 3(1). Using Bayes' rule, the contribution of a new generation to average productivity can easily be calculated by integrating over the distributions.

$$\bar{\varphi}_{j,NG} = \int_1^\infty \varphi \frac{P(\tilde{a}_j = max(\tilde{a})|\varphi)\, p(\varphi)}{P(\tilde{a}_j = max(\tilde{a}))} d\varphi \tag{25}$$

This will in turn recalibrate the wages and utilities of every consumer in the economy. The calculation of this would be outside the scope of this paper, but it is a valid conjecture that this model will implicitly feature gentrification in some regions since housing is determined by an individuals wage relative to the average wage which will be decreasing if the average new worker in the economy is more skilled than the former average. Some initial evidence of this is shown in the housing policy shock.

By assumption, in period 1, $\omega_{j,i} = 1$ since all regions are identical. However, for all periods after 1, the labour composition effects must be incorporated to find the changes. From there, the picture becomes significantly more muddled since all elements can be larger than or smaller than 1 depending on the specific labour composition of each region. Thus, I will now move to a characterisation of a hypothetical equilibrium.

5.3 Period T

The change in total population for each region is given by the amount of migrants less the amount of deaths as found in the previous section. Assuming that there is some sort of steady state makes it clear that there must be some condition where inflows and outflows are the same in expectation. The more interesting question, however, is the effect of this change on the labour makeup of the region. It is clear that the likelihood of choosing region 3 is growing in an individual's productivity by assumption, and as such the labour market should unconditionally have a higher average education in region 3 than in the other regions. While this turns out to be true in this example, it is pertinent to consider whether these differences are exacerbated or dampened by the labour market. To accurately predict the location choices of workers, we can imagine the matching preferences as being controlled by the values of ω. The controlled vector of matching preferences will be called $\tilde{a_j}$. The equilibrium condition for regional stability is that all regions in expectation have the same number of migrants as they have lost members of the population. That is:

$$\frac{N_j}{N_1 + N_2 + N_3} = \overline{P(a_j = max(\tilde{a}))} \tag{26}$$

When the probability of settling in a region is identical to the probability of losing a member of the populace, no region will grow beyond their bounds. As mentioned previously, this equilibrium without labour dynamics, what has been called the unconditional equilibrium, is easily specified whereas the disrupted equilibrium is far more difficult to calculate. Following propositions 1 and 2, it is clear that the probability of settling in a region is determined by

$$P(\tilde{a}_1 = max(\tilde{a})|\varphi) = \int_1^\infty \int_1^\infty \frac{1}{max\{c\,\omega_{2,1}, k\,\omega_{3,1}\}^{\frac{2}{\varphi}}} \frac{\frac{2}{\varphi^{1.5}}\frac{2}{\varphi^2}}{(c\,\omega_{2,1})^{\frac{2}{\varphi^{1.5}}+1}(k\,\omega_{3,1})^{\frac{2}{\varphi^2}+1}} dc\, dk \tag{27}$$

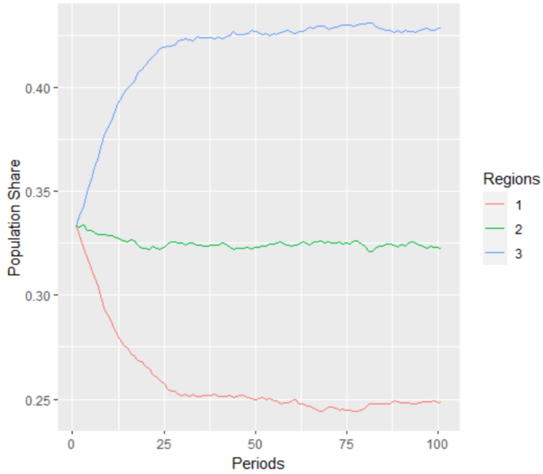
$$\overline{P(\tilde{a}_1 = max(\tilde{a}))} = \int_1^{10} \int_1^\infty \int_1^\infty \frac{1}{max\{c\,\omega_{2,1}, k\,\omega_{3,1}\}^{\frac{2}{\varphi}}} \frac{\frac{2}{\varphi^{1.5}}\frac{2}{\varphi^2}}{(c\,\omega_{2,1})^{\frac{2}{\varphi^{1.5}}+1}(k\,\omega_{3,1})^{\frac{2}{\varphi^2}+1}} p(\varphi) dc\, dk d\varphi \tag{28}$$

The economy will only remain in a steady state when this integral for all regions j equals the fraction of the population living in the region. This condition presents a weak equilibrium, but it is possible to construct a stricter equilibrium condition for a total steady state: that all productivity type shares are asymptotically temporally invariant in equilibrium. That is,

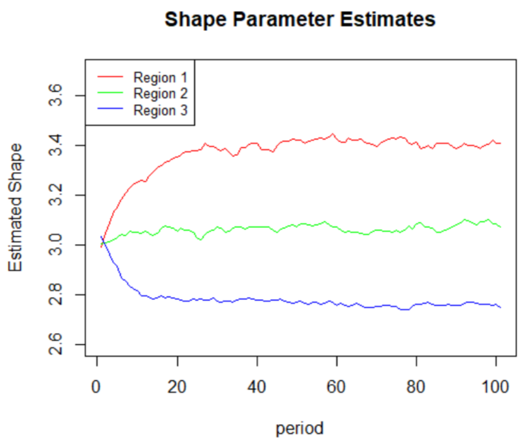
$$\frac{p_j(\varphi)N_j}{p(\varphi)N} = P(a_j = max(\tilde{a})|\varphi) \quad \forall\varphi \tag{29}$$

While not a perfect measure of the empirical distribution, some evidence in favour of convergence satisfying the strict equilibrium condition could be supplied by the time series of estimated shape parameters shown in figure 5. Visually, the estimators seem fairly consistent, and measuring the variance from periods 40 to 100 shows that there is little variance within each region's estimated distribution after period 40 as shown in table 3.

Assuming that there is an equilibrium in model, we can test the estimated scale parameters against the unconditional estimates. Simulating a sample of 60000 workers sorting exclusively under the matching preference requirement yields an estimate of the unconditional shape parameter. A standard t-test finds that all parameters are statistically significant from the unconditional case. Region 1 has a significantly higher shape parameter while region 3 has a significantly lower shape parameter. Additionally, the unconditional prediction of population shares predicts a



(a) Population shares of the three regions



(b) Shape parameter estimates for each region

Figure 5

lower spread across the three regions than is found in the simulation. In summary, under the parameters chosen, the simulated evidence predicts that agglomeration of economic activity will be exacerbated by the circumstances of labour markets and housing whereas the effects of sorting will be diminished by the gains from diversity.

Moment	Region 1	Region 2	Region 3
Mean	3.5275	3.0791	2.7495
Unconditional Mean	3.7868	3.019	2.6000
Standard Deviation	0.0153	0.0155	0.0099
t-statistic	-104.85^{***}	3.4353^{**}	107.23^{***}
p-value	0.000	0.001	0.000

Table 3: Estimates of the shape parameters from periods 40 through 100 compared to unconditional estimates

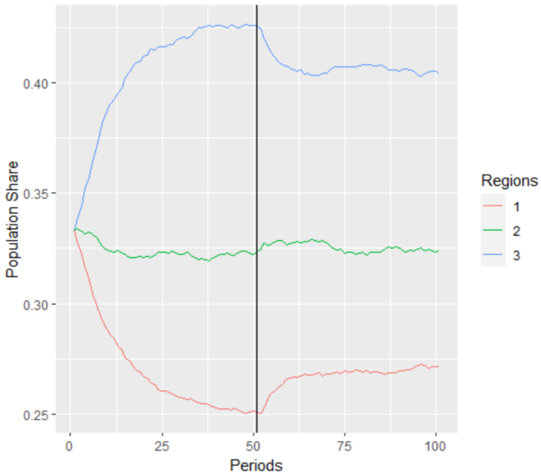
5.4 Shocking the Housing Endowments

Assuming that a social planner wishes to counter trends of agglomeration, the model can be fitted to certain policy simulations. This section has already implicitly shown that a temporary forced relocation of workers will have no effect on the lasting equilibrium. Instead, the social planner could choose to induce spatial diversity by redistributing the housing endowments. Assuming that the fixed housing endowment covers a level of constant reinvestment from the state matching temporal depreciation, it would be possible to redistribute housing by reallocating some of the capital region's housing reinvestment to the periphery to permanently increase the housing endowment in the peripheral region. Leaving the total housing mass intact, this would be a permanent transfer of part of one region's housing endowment to another. In this example this is done by reducing the endowment in region 3 by 10 % and increasing the endowment in region 1 by the same 10 % permanently from period $t = 50$. Figure 6(a) shows how this endowment change would affect the population shares. After having reached a stable equilibrium, region 1 (3) is shocked with a positive (negative) shock of 10% of the housing endowment. Following the full housing condition, rents will decrease in turn, and housing consumption will increase. Table 4 summarises the contemporaneous effects on the various $\omega_{j,i}$'s with the rows representing regions j calculated using equation (22) To take an example, the relative attractiveness $\omega_{3,1}$ will

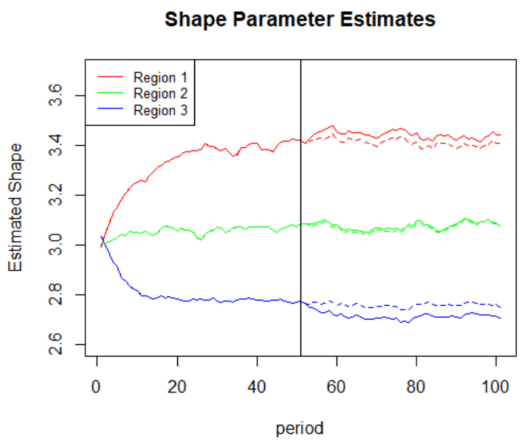
Regions	Region 1	Region 2	Region 3
Region 1	-	1.0322	1.0691
Region 2	0.9687	-	1.0357
Region 3	0.9353	0.9655	-

Table 4

increase compared to the counterfactual; $\omega^P_{3,1}|_{t=51} = 1.0691\omega_{3,1}|_{t=51}$. As the table shows, the model would predict that the likelihood of choosing region 1 over region 3 increases by approximately seven percent, no other effects taken into account. As shown in figure 6(a), the resulting effect is a significant reduction (increase) of the population share of region 3 (1). Taking a mean of periods 30 to 50 and comparing it with the mean of periods 80 to 100 finds that the average population share of region 3 fell by 5.34% relative to the estimated mean. T-testing under a null-hypothesis of shared mean finds a p-value at numerical zero for R. Conversely, a t-test of the means for region 1 in the same period finds that the average population share increases by 7.95% which is equally significant. Finally, the average share of region 2 increases by 0.09% which is statistically significant at a 1 % significance level but practically inconsequential. One of the strengths of a simulated study is that it is possible to compare perfect counterfactuals. To this end, it is possible to compare the estimated shape parameters in section 5.3 to the estimated parameters for the new steady state after the full effects of the housing policy have been incor-



(a) Population shares of the three regions



(b) Shape parameter estimates of the three regions

Figure 6

porated. The model prediction is that region 3 will experience a growth in average education even as the population share decreases as lower educated workers will migrate to the region with a higher housing supply. This can be seen in figure 6(b). This effect counteracts housing redistribution and thus decreases the efficacy of such a policy choice. Such an effect taking place also counteracts any political goals of similar educational distributions across regions. Testing this formally finds that the differences in means for regions 1 and 3 compared to their counterfactuals after the policy shock are significant at any confidence level while the difference for region 2 is not significant at a 5 % significance level. In conclusion, the model predicts that redistributing housing will have a lasting effect on both spatial sorting and the educational distribution across regions after convergence to a steady state Comparing the predictions to the converse experiment played out in the model of Bødker and Kaarsen (2023) serves as a clear illustration of some of the limitations of this paper's model. Given that there is no implicit mobility in the model, it is only possible to construct a crude estimate of the population share effects. Additionally, geographical distance has not been incorporated at all into the model. Finally, the simulation behind this project has not been fitted to create any real predictions for welfare.

6 Discussion

This model is essentially an inversion of the typical assumptions behind international trade theory. Goods cannot be traded but input factors have control over their chosen region. Having all goods be nontradeable is not representative of intraregional patterns of trade, but this simplifying assumption is not without precedent in the literature (Behrens et al., 2014). For relevant policy applications, it would be imperative to understand the implications of trade costs on the model's predictions. Additionally, to accurately use this model for policy purposes, an exceedingly large dataset would be required to find any meaningful differences in the distributional differences. In future work, it would be interesting to consider the inversion of the model for calibration purposes. Additionally, the current measures for matching scaling to education, ψ, are currently set exogenously. It would allow for a far more dynamic model presenting a larger suite of policy options.

Choosing a functional form for utility and production is a contentious topic. Correctly specifying consumption patterns presents a challenge large enough that literature oftentimes returns to the tried and true Cobb-Douglas utility function. In keeping with this tradition, I have made the same choice to maintain some degree of tractability in the model. The addition of any degree of stochastic preference makes the understanding of utility far more complex.

The model has also been created with the express purpose of considering different worker types as imperfectly substitutable variations of the same input factor. In a highly educated and dense city, the model should predict a lower wage premium for highly educated workers. The empirical background for this is questionable at best, and some empirical research could easily invalidate this setup. The model as it stands has not had measures of welfare specified and has only been shown as an example for a specific set of parameters. Welfare optimisation will also be far more politically complicated as the model by assumption assumes larger welfare gains for highly educated workers. Any policy choice constructed to maximise total welfare would thus implicitly favor the needs of the most educated class.

Finally, a dynamic model should optimally create a well-specified and argued framework for expectations. This is not the case in this model. Agents are given static expectations of their regions, making no assumptions as to the trend of consumption in their region. This implies assuming bounded rationality on the consumer side since a fully rational would make reasonable assumptions of the trend of all regions and discount the utility in future periods accordingly. There is no way to incorporate this without complicating the model to a point that is far beyond the scope of this paper. It does turn out, however, to resolve itself as the model predicts a fairly stable equilibrium in which these static expectations are rational.

This model presents a possible explanation for agglomeration driven by sorting and selection on the consumption side in regions with no trade while introducing a centripetal force in the consumption of housing. It condenses location choice into an intuitive trade-off, but it should not be considered a finished work that could be put into use.

7 Conclusion

Spatial economics and its ties to economic geography has resulted in many models of varying complexity. The seminal paper of Krugman (1991) created a model of high tractability under strict restrictions. Among these is the existence of amenities, the consumption of housing and possible heterogeneity of the workforce. The reason for this is obvious: by omitting these variables, the model becomes relatively intuitive and tractable. Recent literature has attempted to reconcile multiple of these omitted factors with landmark gains being made in models of stochastic heterogeneity leading to this model which combines all three. The resulting model has many drawbacks; The defining integral does not as of yet have a closed form solution, there is still no trade and migration incorporated in the model, and the effects of changing parameters have not been fully mapped out. The model does, however, make strong predictions that can be interpreted numerically and does condense into a simple and intuitive choice. It shows that the need for housing can dominate the love for variety in completely homogenous regions, and under given parameters it does exhibit a likeness to convergence to some steady state. Additionally, it is well suited to an evaluation of the effects of labour markets on educational attainment and their interaction with demand-side amenity preferences. Using microdata, it is plausible that the model could be calibrated to create actual predictions in policy evaluation. To assist this, it would be interesting to do further calculations to specify the effects of taking certain policy measures. The next step in further modelling would be an incorporation of trade costs and eventually commuting costs. Finally, the model as of now is completely dependent on exogenous and time-invariant scales of productivity distributions. For further research, it would be relevant to examine potential explanatory variables for ψ to create a model where the scale to the matching parameter could be determined endogenously.

8 Bibliography

References

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., and Wolf, N. (2015). The economics of density: evidence from the berlin wall. *Econometrica*, 83:2127–2189.

Bayer, P., Ferreira, F., and McMillan, R. (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, 115(4):588–638.

Behrens, K., Duranton, G., and Robert-Nicoud, F. (2014). Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, 122(3):507–553.

Behrens, K. and Robert-Nicoud, F. (2015). Chapter 4 - agglomeration theory with heterogeneous agents. In Duranton, G., Henderson, J. V., and Strange, W. C., editors, *Handbook of Regional and Urban Economics*, volume 5 of *Handbook of Regional and Urban Economics*, pages 171–245. Elsevier.

Bødker, J. and Kaarsen, N. (2023). Byer, produktivitet og velfærd – en geografisk ligevægtsmodel for danmark. *Samfundsøkonomen*, 2023(1):51–62.

DØRS (2021). *Produktivitet 2021*. The Danish Secretariat of the Economic Councils.

Ethier, W. J. (1982). National and international returns to scale in the modern theory of international trade. *The American Economic Review*, 72(3):389–405.

Feenstra, R. C. (2015). *Advanced International Trade: Theory and Evidence*. Princeton University Press, Princeton, NJ, 2nd edition.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. (1992). Growth in cities. *Journal of Political Economy*, 100(6):1126–1152.

Koster, H. R. A., van Ommeren, J. N., and Rietveld, P. (2016). Historic amenities, income and sorting of households. *Journal of Economic Geography*, 16(1):203–236.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *American Economic Review*, 70(5):950–959.

Krugman, P. R. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99(3):483–499.

Marshall, A. (1890). *Principles of Economics*. Macmillan.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, pages 105–142. Academic Press, New York.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Moretti, E. (2004). Human capital externalities in cities. In *Handbook of regional and urban economics*, volume 4, pages 2243–2291. Elsevier.

Mughan, S., Sherrod Hale, J., and Woronkowicz, J. (2022). Build it and will they come?: The effect of investing in cultural consumption amenities in higher education on student-level outcomes. *Research in Higher Education*, 63:60–91.

Neumark, D. and Simpson, H. (2015). Chapter 18 - place-based policies. In Duranton, G., Henderson, J. V., and Strange, W. C., editors, *Handbook of Regional and Urban Economics*, volume 5 of *Handbook of Regional and Urban Economics*, pages 1197–1287. Elsevier.

Ottaviano, G. I. P. (2011). 'new' new economic geography: firm heterogeneity and agglomeration economies. *Journal of Economic Geography*, 11:231–240.

Redding, S. J. and Rossi-Hansberg, E. (2017). Quantitative spatial economics. *Annual Review of Economics*, 9:21–58.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy*, 90(6):1257–1278.

Roger, M. and Wasmer, M. (2011). Heterogeneity matters: labour productivity differentiated by age and skills. Number G2011/04.

Tabuchi, T. and Thisse, J.-F. (2002). Taste heterogeneity, labor mobility and economic geography. *Journal of Development Economics*, 69(1):155–177.

9 Appendix A: Proofs

9.1 Utility Maximising Consumption (3) & (4)

Proof. Setting up the Lagrangian:

$$\mathcal{L} = a\, c^{\alpha} h^{1-\alpha} - \lambda(Pc + rh - w) \tag{30}$$

FOC

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial c} &= a\,\alpha \frac{h^{1-\alpha}}{c^{1-\alpha}} - \lambda P \overset{!}{=} 0 \\
\frac{\partial \mathcal{L}}{\partial h} &= a\,(1-\alpha)\frac{c^{\alpha}}{h^{\alpha}} - \lambda r \overset{!}{=} 0 \\
\frac{\partial \mathcal{L}}{\partial \lambda} &= pc + rh - w \overset{!}{=} 0 \\
\frac{\frac{\partial \mathcal{L}}{\partial c}}{\frac{\partial \mathcal{L}}{\partial h}} &= \frac{\alpha h}{(1-\alpha)c} = \frac{\lambda p}{\lambda r} \\
c &= \frac{\alpha}{1-\alpha}\frac{r\,h}{p} \\
hr(1 - \frac{\alpha}{1-\alpha}) &= w \\
h &= \frac{(1-\alpha)w}{r} \\
c &= \frac{\alpha w}{p}
\end{aligned}$$

□

9.2 Marginal Productivity of Labour (9)

Proof. Recognising the chain rule, it is clear that

$$\begin{aligned}
\frac{\partial Y}{\partial l(\varphi)} &= \bar{\delta}(\varphi)\frac{1}{\rho}\left(\int_{\underline{\varphi}}^{\bar{\varphi}} (\delta(\varphi)l(\varphi))^{\rho} d\varphi\right)^{\frac{1}{\rho}-1} \rho\delta(\varphi)^{\rho} l(\varphi)^{\rho-1} \\
&= \bar{\delta}(\varphi)\delta(\varphi)^{\rho}\left(\int_{\underline{\varphi}}^{\bar{\varphi}} (\delta(\varphi)l(\varphi))^{\rho} d\varphi\right)^{\frac{1-\rho}{\rho}} l(\varphi)^{\rho-1} \\
&= \bar{\delta}(\varphi)\delta(\varphi)^{\rho} L^{1-\rho} l(\varphi)^{\rho-1} \\
&= \bar{\delta}(\varphi)\delta(\varphi)^{\rho}\left(\frac{L}{l(\varphi)}\right)^{1-\rho}
\end{aligned}$$

□

9.3 Lemma 1

Proof. Multiplying all input factors by λ finds that

$$
\begin{aligned}
MPL(\lambda l(\varphi)) &= \bar{\varphi}\varphi \left(\frac{\left(\int_{\underline{\varphi}}^{\bar{\varphi}} \lambda^{\rho} (\varphi l_j(\varphi)^{\rho} d\varphi \right)^{\frac{1}{\rho}}}{\lambda l_j(\varphi)} \right)^{1-\rho} \\
&= \bar{\varphi}\Omega\varphi \left(\frac{\left(\int_{\underline{\varphi}}^{\bar{\varphi}} (\varphi l_j(\varphi)^{\rho} d\varphi \right)^{\frac{1}{\rho}}}{l_j(\varphi)} \right)^{1-\rho} \\
MPL(\lambda l(\varphi)) &= MPL(l(\varphi))
\end{aligned}
$$

In a symmetric equilibrium, all prices must be identical, and all firms must therefor have the same average MPL. This can be obtained by integrating marginal productivity of labour over all values of φ

$$
\begin{aligned}
M\bar{P}L &= \bar{\varphi} \int_{\varphi}^{\bar{\varphi}} \varphi \left(\frac{L}{l_j(\varphi)} \right)^{1-\rho} \frac{l_j(\varphi)}{N} d\varphi \\
&= \bar{\varphi} \frac{L^{1-\rho}}{N} \int_{\varphi}^{\bar{\varphi}} \varphi l_j(\varphi)^{\rho} d\varphi \\
&= \bar{\varphi} \frac{L^{1-\rho}}{N} L^{\rho} \\
&= \bar{\varphi} \frac{L}{N}
\end{aligned}
$$

□

9.4 Zero profit condition (13)

Proof. Profits can be described as

$$
\Pi = yp - y\frac{\bar{w}}{M\bar{P}L} - F\bar{w} \stackrel{!}{=} 0
$$

Recognising that $y = N\bar{c}$, that is, the population size times the average consumption

$$
\begin{aligned}
p &= \frac{w}{M\bar{P}L} + \frac{F\bar{w}}{N\bar{c}} \\
\frac{p}{\bar{w}} &= \frac{1}{M\bar{P}L} + \frac{F}{N\bar{c}}
\end{aligned}
$$

□

9.5 Equilibrium rents r (15)

Proof. Letting p be the numeraire allows for an easy setup of housing consumption. Setting up a full housing condition implies

$$\begin{aligned} N\bar{h} &= H \\ \bar{h} &= \frac{(1-\alpha)\bar{w}}{r} = \frac{(1-\alpha)\bar{\varphi}L}{Nr(1+\mu)} \\ N(1-\alpha)\frac{\bar{\varphi}L}{Nr(1+\mu)} &= H \\ r &= \frac{\bar{\varphi}L(1-\alpha)}{H(1+\mu)} \end{aligned}$$

□

9.6 Indirect utility function (19)

Proof. Recognising that a symmetric equilibrium implies that all consumption is equal across all firms results in the budget share being split equally across all firms. Thus $c_i(\varphi) = \frac{\alpha w(\varphi)}{M}$. Letting indirect utility of consumption of the manufacturing good be denoted v_c, and the housing consumption utility be denoted v_h

$$\begin{aligned} v_c &= \left(\left(M \cdot \frac{\alpha w(\varphi)}{M}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\alpha} \\ v_c &= \left((\alpha w(\varphi))^{\frac{\sigma-1}{\sigma}}\left(\frac{1}{M}\right)^{\frac{1}{\sigma}}\right)^{\alpha} \\ v_c &= \left(\frac{\alpha\bar{\varphi}\varphi L^{1-\rho}}{(1+\mu)l(\varphi)^{1-\rho}}\right)^{\frac{\sigma}{\sigma-1}}\left(\frac{L}{F\sigma}\right)^{\frac{1}{\sigma}} \\ v_c &= \left(\frac{\alpha\bar{\varphi}\varphi}{(1+\mu)l(\varphi)^{1-\rho}}\right)^{\frac{\sigma}{\sigma-1}}\left(\frac{1}{F\sigma}\right)^{\frac{1}{\sigma}} L^{\frac{\sigma+\rho\cdot(1-\sigma)}{\sigma}} \\ v_h &= \frac{(1-\alpha)w(\varphi)}{r} \\ v_h &= \frac{(1-\alpha)\bar{\varphi}(1+\mu)\varphi L^{1-\rho}H}{(1-\alpha)\bar{\varphi}(1+\mu)Ll(\varphi)^{1-\rho}} \\ v_h &= \frac{\varphi H}{L^{\rho}l(\varphi)^{1-\rho}} \\ \tilde{v} &= \left(\left(\frac{\alpha\bar{\varphi}\varphi}{(1+\mu)l(\varphi)^{1-\rho}}\right)^{\frac{\sigma}{\sigma-1}}\left(\frac{1}{F\sigma}\right)^{\frac{1}{\sigma}} L^{\frac{\sigma+\rho\cdot(1-\sigma)}{\sigma}}\right)^{\alpha}\left(\frac{\varphi H}{L^{\rho}l(\varphi)^{1-\rho}}\right)^{1-\alpha} \end{aligned}$$

□

10 Appendix B: Figures and Graphs

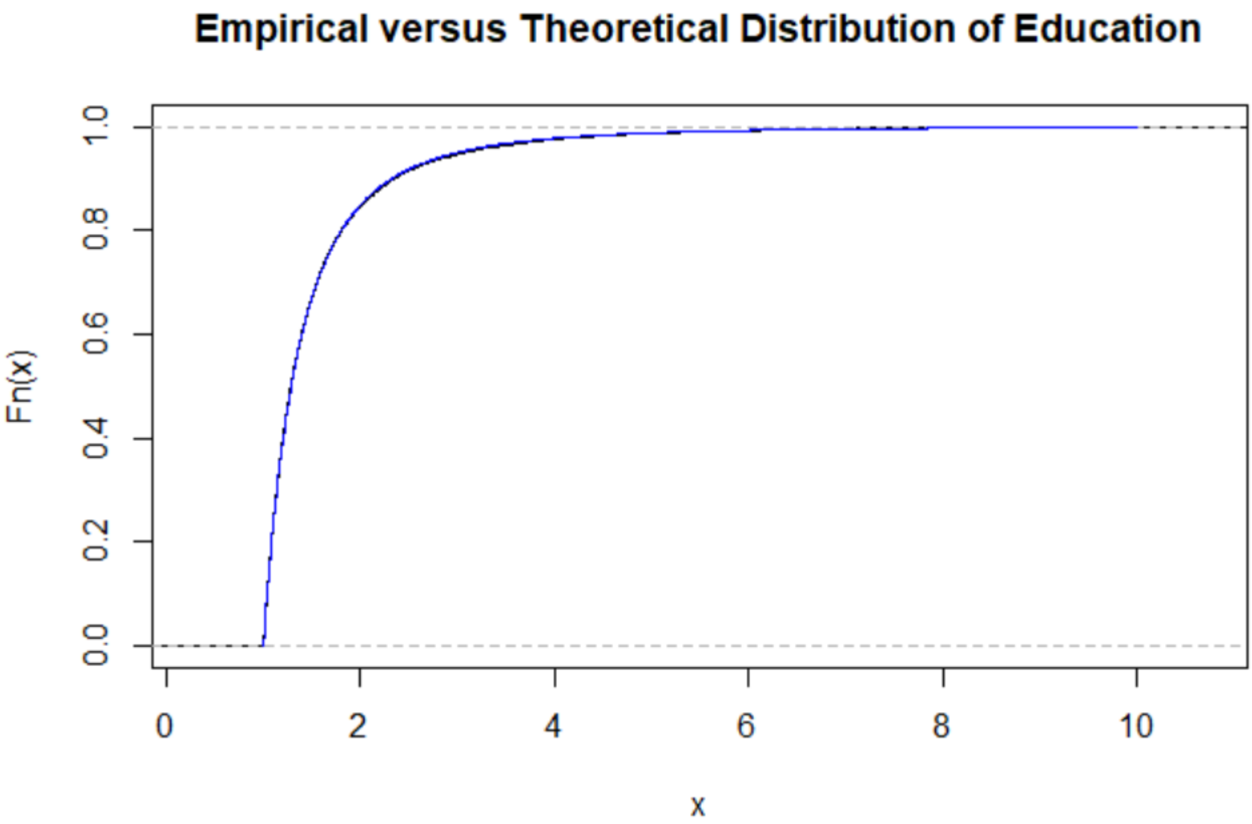


Figure 7: Non-parametric and Parametric Estimation of the CDF for Region 3