Search and the City*

C.N. Teulings and P.A. Gautier†

July 3, 2003

Abstract

Can search frictions be the source of increasing returns to scale (IRS) that brings people together in big cities? The empirical studies on this issue are not encouraging, since they favor constant returns in search. However, this research is troubled by all sorts of selection bias. We develop a model of an economy with several regions, which differ in scale. Within each region, workers have to search for a job-type that matches their specific skills well. They face a trade-off between match quality and the cost of extended search. This trade off differs between regions, because search in larger regions is more efficient. Then, interregional mobility and trade lead to a pattern of specialization where large regions undertake more search intensive activities. Empirical evidence for the United States corroborates the implications of the model. Search can explain about 75% of the wage differentials between large metropoles and small cities.

*We are grateful to Daron Acemoglu, Peter Diamond, participants at Princeton, the MIT macro lunch, Meteor Maastricht, IRES Louvain le-Neuve, the KNAW-Tinbergen Institute conference on search and assignment models in Amsterdam, the ESPE in Athens, the SED meeting in Stockholm, the Theil memorial conference in Amsterdam for useful comments and suggestions on earlier versions of this paper.

†Tinbergen Institute, Erasmus University Rotterdam, Burg. Oudlaan 50, 3062 PA Rotterdam, The Netherlands, teulings@few.eur.nl, gautier@tinbergen.nl

1 Introduction

Why does a large fraction of scientists, rocket engineers, top managers and brain surgeons on the one hand, and high school drop outs and low skilled immigrants on the other hand live in large metropolitan areas? According to the CPS, the standard deviation of worker's human capital is about 20% higher in a large Metropolitan area like Boston than in a regular city like Austin (Texas). Differences in average human capital between Boston and Austin account for a 12 % higher wage in Boston. The dispersion of occupations, ranked in a similar way as human capital gives similar figures.

This paper seeks to explain these stylized facts by increasing returns to scale (IRS) in job search. We think of cities as areas with a high density of job seekers and vacancies. Under IRS, search is more efficient in large metropoles than in small towns or rural areas. Consequently, large metropoles specialize in activities that require a lot of search, for example because specialized inputs are involved.

The question why people live in cities has received a lot of attention in the literature, Lucas (1988, 1999) and Glaeser et al. (1992). All explanations require some form of IRS, like knowledge transfers in Jacobs (1969), "love for variety" and fixed production cost in Krugman (1991), Ciccone and Hall (1996) and Fujita et al. (1999). However, the literature has not been able to pin down the actual source of IRS in cities empirically. At first sight, search is not an obvious candidate. Though IRS in search is a logical assumption from a theoretical point of view, see Diamond (1982), most empirical studies based on aggregate data find CRS, see the overview of Petrongola and Pissarides (2000). We argue that studies on aggregate data severely underestimate returns to scale because of selection biases. In Teulings and Gautier (2003) we argue that heterogeneous workers and firms face a trade off between match quality and the cost of search. The lower the cost of search, the more selective agents are when they have to decide on whether or not to match. Part of the IRS in search shows up, not in a lower equilibrium rate of unemployment, but in a better match quality. Empirical studies based on aggregate unemployment and vacancy data miss this effect. However, this effect can never explain why we do not observe any IRS at all: some reduction in search cost (that is: lower unemployment) has to be left over to justify the greater selectivity.

The specialization of large metropoles on search intensive activities provides another reason why IRS in search does not necessarily translates into a lower unemployment rate. Though there are much more vacancies available in cities like New York and Boston, they are all highly specialized. Hence, from the point of view of an individual job seeker, only a small fraction of these vacancies are relevant. Given the wide dispersion of jobs that is available, these cities also attract workers with a wide variety of human capital. From the firm's point of view, only a small fraction of the job seekers are relevant. Contrary to the greater selectivity in accepting matches, this mechanism can explain why we would not observe any returns to scale at all, or even negative returns to scale. When there is a small supply of metropolitan areas and when there are some extremely search intensive activities, the initial advantage in search efficiency of metropolitan areas might be more than offset by specialization in search intensive activities. The mechanism is related to Shimer's (1999) analysis of why regions with a lot of youngsters, who are unemployment more often than elderly, have a lower unemployment rate. This paper offers a formal model of this type of regional specialization in production based on regional differences in scale, and therefore in search efficiency. The paper provides empirical evidence for the relevance of this model, based on the spatial distribution of workers and jobs in the United States.

The model that we propose is an extension of Shimer and Smith (2000) and Teulings and Gautier (2003), who analyze search frictions in matching models with heterogeneous workers and tasks. These models start from the hedonic pricing and assignment models of Rosen (1974), Sattinger (1975, 1995), and Teulings (1995, 2002). In a frictionless Walrasian market, workers and tasks end up in the optimal match. In a market with search frictions, workers and firms are prepared to also accept suboptimal matches in order to save on search cost.[1] The higher the cost of search, the more job seekers are willing to accept suboptimal matches. Job seekers and firms face a trade-off between the cost of prolonged search and the quality of the match. For the sake of simplicity, we rule out on the job search.

Regions only differ by the scale of the market, which is treated as exogenous, i.e. we do not explain why Boston is larger than Austin. The regions are linked to each other by both labor mobility and trade. Labor mobility equalizes real reservation wages between regions and thereby determines the regional skill distribution. Trade offers each region the opportunity to produce that commodity that yields the maximum value added, given its labor market scale relative to that of the other regions. Obviously, labor mobility and trade are directly related: a region that produces a very complex commodity will have a highly skilled workforce. We analyze the interaction between regional differences in the efficiency of the search process and the pattern of regional specialization, both in the composition of the workforce and the type of commodities that a region produces.

[1] If all firms are single worker firms we can use "tasks" and firms interchangeably.

In Teulings and Gautier (2003) we characterized the cost of search (defined as the relative difference between value added in the optimal assignment and the reservation wage), taking value added in the optimal assignment as given. Since there is no analytical solution of the equilibrium, we applied Taylor expansions. However, search frictions do not only drive a wedge between optimal value added and the reservation wage, but they also affect the value added itself, by a general equilibrium effect on task prices. There are essentially two effects, *skill compression* and *skill spoiling.* This paper provides a characterization of these effects, also by means of Taylor expansions. The skill compression effect is directly related to our IRS assumption. Since the skill distribution is less dense in the tails than around the median, the cost of search are larger in the tails. Hence, search frictions compress the "effective" relative to the actual skill distribution. The skill compression effect is crucial for the analysis, since it is the force that drives small scale regions with large search frictions to reduce their labor demand in the tails of the task distribution.

It is less easy to provide an intuition for the skill spoiling effect at this stage. We show that part of the skills of the workforce remain underutilized, by the interaction of the shape of the supply distribution and the productivity differentials between high and low skilled workers.[2] This yields the prediction that the mean of the skill distribution is larger in non-metropolitan areas than in metropolitan areas, keeping the task distribution constant, to offset the skill spoiling effect. Our empirical evidence provides support for this surprising implication.

Apart from the implications for the distributions of human capital and job complexity, our model yields also implications for regional wage differentials. Since the scale of a region is the fixed factor, for example due to the limited stock of housing, the owners of this fixed factor will capture the rents. These rents raise the cost of living in metropolitan areas, and therefore raise nominal reservation wages. Glaeser and Maré (2001) show indeed that migrants entering a city receive an immediate wage gain suggesting that the observed city wage premium is not due to unobserved ability differences but reflects a true nominal wage increase. We will relate the measured nominal wage gains for working in a metropolitan area to the predictions of our model.

Some other papers have looked at the relation between search frictions and scale of the markets. In Kim Sunwoong (1989) firms offer a larger variety of job requirements in cities while workers specialize more. Glaeser (1999) argued that for labor market pooling to work, workers

[2] More specifically, a particular type of task is done by a range of skill types in a search equilibrium, while it is done by a unique skill type in a Walrasian assignment. For tasks done by sub-modal skill types this implies that more high than low skill types come in (since the skill density is rising within this interval), while the opposite holds for tasks done by supra-modal skill types. If high skill types are more productive at all jobs, this yields an increase in the output of tasks done by sub-modal skill types relative to those done by supra-modal skill types.

must be able to change employers without changing residencies. Dumais et al. (1997) find that industry location is far more driven by labor mix than by any other explanatory variable.

The plan of the paper is the following. Section 2 presents some stylized facts on the distribution of workers and jobs between cities in the United States. Section 3 discusses the basic structure of the model. Section 4 solves the assignment problem for a Walrasian world. In Section 5, we introduce search frictions. Section 6 analyzes patterns of specialization between regions, by labor mobility and trade. In Section 7, we reevaluate our empirical evidence in the light of our model and test its main implications. Section 8 concludes.

2 Some stylized facts

Regions differ in the density of their economic activity. On the one extreme, there are the large metropolitan areas like Washington, New York and Chicago. The activities of their inhabitants are highly integrated. On the other extreme there are sparsely populated areas like Montana and Wyoming where people live almost in autarky. Figure 1 plots the value of the Gautier and Teulings (2003) labor market density index for the US. The darker areas reflect the densest labor markets. The map shows that the most dense areas are located at the East-coast and that the least dense areas are in the Mid-West. This index is based on the following idea. Consider a job at a particular location, and consider the likelihood that it will be occupied by a particular worker living in the neighborhood of that job. In a small, low density local labor market, only a small number of workers are available for this job, and hence the probability for each of these workers to actually occupy it is high. Alternatively, in a dense metropolitan labor market, many workers are potential candidates to fill the job and hence the probability for each individual worker to actually occupy it is low. Let $\hat{\chi}$ denote this index. It can take values between zero and unity, the denser the market, the lower $\hat{\chi}$. Washington D.C. and New Jersey/New York-North are the densest regions with $\hat{\chi} = 0.18$ while rural Montana is the least dense region with $\hat{\chi} = 0.95$. The West-coast cities are intermediately dense. This fits the intuition. East-coast cities typically have a better infrastructure than the more spread out cities on the West coast.

These regional differences in the density of economic activity are correlated to the distribution of human capital and jobs in each of these regions. Table 1 provides some evidence on this relation. We construct indexes for human capital and the "complexity" of jobs based on the following simple methodology. Human capital has multiple components, for example education and experience. We combine these components in a single statistic s by running a regression

$$w = x'\beta + u$$

where w denotes the individual's log wage, x is a vector of human capital components,[3] β is a vector of parameters, and u is an error term. We pool CPS data for the period 1988-92. Our human capital or skill index s is defined as: $s \equiv x'\beta$. This method implies that the unit of measurement of s is log wages. Similarly, we construct an index for job complexity c by running a regression of wages on a vector of job characteristics z and using the regression coefficients γ for aggregation: $c \equiv z'\gamma$.[4] Again, the unit of measurement of the complexity index is log wages. Let $\widehat{\mu}^s$ and $\widehat{\mu}^c$ be the regional mean and $\widehat{\sigma}^s$ and $\widehat{\sigma}^c$ the regional standard deviation of this skill and complexity index respectively. We do these calculations for 150 regions, see Gautier and Teulings (2003) for details of this classification.

Table 1 reports the coefficients of an OLS regression of these moments for each region on the density index $\hat{\chi}$. Dense regions have a better skilled work force and on average more complex jobs. Just based on their higher average human capital, workers in Washington DC earn on average 5% higher wages than workers in Montana. Just looking at job complexity, the difference is even 6%. Moreover, the dispersion of both worker skills and job complexities is 5 % higher in dense regions. The evidence on these issues presented in section 7 will be even stronger. We also present the coefficient of a regression of wages on $\hat{\chi}$, where we control for all available human capital variables. Labor market density has a large effect on wages, Washington DC workers have a 30 % higher wage than workers with similar skills, employed at similar jobs in Montana.

Figures 2 and 3 plot the mean and the standard deviations of s and c for each region. The larger circles reflect denser areas. Boston and Washington DC are examples of regions having both a high $\widehat{\mu}^c$ and $\widehat{\sigma}^c$, while Houston and Seattle have a high $\widehat{\mu}^c$ but a low $\widehat{\sigma}^c$. Los Angeles and Miami have a low $\widehat{\mu}^c$ and a high $\widehat{\sigma}^c$, while in the rural areas of Arkansas both $\widehat{\mu}^c$ and $\widehat{\sigma}^c$ have a low value. However, the larger circles are predominantly located in the North-East quadrant. Summing up, large metropoles: (1) offer more complex jobs and attract better skilled workers, (2) have a larger variety of job and skill levels, and (3) pay higher wages. The first fact has drawn a lot of attention in the literature, i.e. Lucas (1988) but the second fact has largely been ignored. Our model provides an explanation for these stylized facts.

[3] We include: years of schooling, a third order polynomial in experience, highest completed education, race, sex, being married, having a full or part time contract and various cross terms of experience, education and being married ($R^2 = 0.35$). The exact specification does not matter much for the results in Table 1.

[4] The vector z includes 520 occupation and 242 industry dummies ($R^2 = 0.41$). Various commentators asked why we estimated separate regressions for the skill and the complexity index. What we do is what theory dictates: s and c are highly correlated. Hence, we estimate them simultaneously, the one will be a proxy for the unobserved component of the other and vice versa, see Gautier and Teulings (2003) for a detailed discussion.

3 Structure of the model

The economy considered in this paper is made up of a large number of regions, two of which are shown in Figure 4. We start with the discussion of the economy in a single region.

Each region produces a single composite tradable commodity by means of the input of non-tradable intermediate tasks, indexed by their complexity c, using a Leontieff technology. These intermediate tasks are produced by risk neutral firms, where each firm produces a single task type. The production technology of tasks exhibits constant returns to scale (CRS) and is equal across regions. Firms of each c type can enter freely and there is perfect competition on the task markets. Hence, a zero profit condition applies for each c type in each region and there exists a unique set of task prices that clears regional task markets.[5] To do these tasks, firms hire workers with various skill types s. Worker types are the only factors of production in the economy. Workers are risk neutral and supply a fixed amount of labor. Hence, each regional economy faces an assignment problem, see Sattinger (1975) and Teulings (1995): what worker type s should be assigned to what task type c, and vice versa? In the absence of search frictions, a Walrasian auctioneer would implement an efficient assignment that maximizes value added in the regional economy, where each worker type s would be assigned to her "optimal" c type task.

Since the production technology is CRS and equal across regions, regions would face the same production frontier per worker in a world without search frictions. However, when there are search frictions, workers do not wait forever till they find their optimal task. Instead, they also accept tasks that yield a lower value added than their optimal task type. Since the contact technology underlying the search process is characterized by increasing returns to scale (IRS) and since regions differ by their scale (or, better: density), the production frontier (conditional on search frictions) per worker differs between regions. Large regions are more efficient. In a region of infinite size, the cost of search would drop to zero. Hence, this region would be equivalent to the Walrasian economy. We use this region of infinite size, or equivalently, the Walrasian benchmark as the point of reference for our analysis.

The regions are related to each other by: (1) labor mobility (the upper part of Figure 4) and (2) interregional trade in the region-specific tradable commodities (the lower part of figure 4). Both labor mobility and transportation of commodities are costless. Each region is small relative to the economy as a whole, so that it takes economy wide reservation wages and commodity prices as given. Both conditions give rise to a *no arbitrage condition* so that real wages for worker types and tradable commodity prices are equalized between regions. The number of workers (or:

[5] We think of all non-tradables, like for example retail trade, as just intermediate c-type tasks that are required for the production of the tradable commodity.

scale, or: density) of a region is determined by the available stock of real estate. This is the fixed factor in a region that we treat exogenous here. Hence, real estate owners capture any region specific rents via the rental price of their property. The workers utility function is homothetic, so that all workers within a region face the same cost of living index, irrespective of their level of income. The rental price of real estate adjusts until the cost of living index has fully absorbed potential nominal wage differences between regions.

Regions can "choose" what tradable composite commodity they want to produce, taking the price of each type of tradable commodity as given. The "choice" of the tradable economy is equivalent to the choice of a set of Leontieff coefficients that specify how much of each c-type task is required for the production of one unit of that tradable commodity. In equilibrium, each region specializes in that tradable which makes best use of its comparative advantage. Large regions with a comparative advantage in search specialize in search intensive production, small regions specialize in search extensive production.

The exposition of the model proceeds as follows. In Section 4, we discuss in greater detail the equilibrium assignment in a Walrasian equilibrium. We focus on the general equilibrium effect of shifts in the skill and complexity distributions on value added in the optimal assignment of each skill type s. In Section 5, we introduce search frictions. Interregional interactions by labor mobility and trade are discussed in Section 6.

4 The Walrasian assignment in a single region

4.1 Assumptions

On the supply side of the labor market we assume that the skill distribution in a region is normal:

$$s\tilde{\ }N\left(\mu^{s},\sigma^{s}\right)$$

Throughout the paper we adopt the conventions that non-underlined Greek letters are region specific parameters and that underlined Greek letters are economy wide parameters. It is convenient for future use to define the parameter vector: $\pi^{s}\equiv[\mu^{s},\sigma^{s}]'$ consisting of the mean and standard deviation of the skill distribution. On the demand side of the labor market, job or tasks differ by their complexity level c. Tasks are combined in a single composite commodity by a Leontieff technology. Tasks play no other role than being input in the production of this composite commodity. The distribution of c type tasks required for production of one unit of the composite commodity (the Leontieff coefficients) is also normal, with mean μ^{c} and standard deviation σ^{c}. These parameters are the counterparts of μ^{s} and σ^{s}, but now for the demand side of the regional labor markets. Analogous to π^s, we define: $\pi^c \equiv [\mu^c, \sigma^c]'$. For future reference we define $\Delta\pi \equiv \pi^s - \pi^c$.

Tasks are traded on the regional market for intermediate inputs. Let $p(c)$ be the log price for a task that clears that regional market and let $p^\#$ be the log price of the composite commodity. We take the latter price as a numeraire, $p^\# = 0$. Then, the Leontieff technology implies that the weighted sum of task prices must be equal to unity, or in logs:

$$p^\# = \ln \left[\int_{-\infty}^{\infty} \frac{1}{\sigma^c} \phi \left(\frac{c - \mu^c}{\sigma^c} \right) \exp \left[p(c) \right] dc \right] = 0 \tag{1}$$

where $\phi(\cdot)$ is the standard normal density function. Since all markets are assumed to be perfectly competitive and since there are no other factors of production than labor, a zero profit condition applies and wages are equal to value added. Let $y^0(s,c)$ be log value added of an s-type worker employed at a c-type job. By definition it satisfies:

$$y^0(s,c) \equiv p(c) + \underline{f}(s,c) \tag{2}$$

where $\underline{f}(s,c)$ is the log productivity of a type s worker in a type c task.

Assumption: the log productivity of worker type s in a c-type job satisfies:[6]

$$\underline{f}(s,c) \equiv \underline{\xi} \left(1 - e^{c-s} \right)$$

where $\underline{\xi}$ is an economy wide technology parameter. This specification of production technology of tasks implies that the level of productivity is log supermodular: $\underline{f}_{sc} > 0$. This log supermodularity captures the idea that highly skilled workers have a *comparative advantage* at complex job types. The larger c, the larger the relative productivity gain of an additional unit of s. Furthermore, this specification implies *absolute advantage* for better skilled workers: $\underline{f}_s > 0$ for any combination s and c.

4.2 The Walrasian equilibrium for $\Delta\sigma = 0$

We start with the characterization of the equilibrium for the special case that $\sigma^s = \sigma^c$, or $\Delta\sigma = 0$. Since wages are equal to value added, worker types are assigned to the task-type where they produce the highest value added in market equilibrium. Let $c(s)$ be this value of c that maximizes $y^0(s,c)$ for a particular s. Assuming $p(c)$ to be differentiable,[7] $c(s)$ satisfies:

$$y_c^0 [s, c(s)] = p_c [c(s)] + \underline{f}_c [s, c(s)] = 0 \tag{3}$$

[6] Teulings and Gautier (2003) apply the technology $f(\bar{s}, \bar{c}) = \bar{s}\bar{c}$. Since we have not yet defined the units of measurement of s and c, we can apply any rising transformation. When we take: $\overline{s} = -\exp(-s)$ and $\overline{c} = \underline{\xi} \exp(c)$, both technologies are equivalent (up to a general efficiency differential $\underline{\xi}$).

[7] This can in fact be proven, see Teulings (1995) for the relevant conditions.

The function $c(s)$ describes the assignment of workers to jobs. Comparative advantage can be shown to imply $c_s(s) > 0$: better skilled workers are assigned to more complex tasks. Let $y(s)$ denote log value added of a type s worker in equilibrium. By definition:

$$\begin{aligned} y(s) &\equiv y^0[s, c(s)] \\ y_s(s) &= \underline{f}_s[s, c(s)] \end{aligned} \tag{4}$$

The first equation applies identically for all s, and hence its first derivative applies, which is the second equation. The situation is depicted in y, c-space for a particular s-type in panel A of Figure 5. The curve $y^0(s, c)$ denotes log value added of this particular s-type for various c-types. By definition, its maximum is $y(s)$, which is attained at $c = c(s)$. The difference $y(s) - y^0(s, c)$ measures the relative loss of value added due to suboptimal assignment. In the Walrasian equilibrium, suboptimal assignments are irrelevant, since they do not occur in equilibrium. However, they will be relevant in the presence of search frictions, since then job seekers and firms trade off the output loss due to suboptimal assignment against the cost of further job search, see section 5. The optimal assignment must clear the market for tasks, in logs:

$$-\ln\sigma^s - \frac{1}{2}\left(\frac{s-\mu^s}{\sigma^s}\right)^2 + \underline{f}[s, c(s)] = -\ln\sigma^c - \frac{1}{2}\left(\frac{c(s)-\mu^c}{\sigma^c}\right)^2 + y^{\#} + \ln c_s(s) \tag{5}$$

where $y^{\#}$ is log output of the composite commodity. The left hand side is the log supply of labor of type s (the log normal density function) plus its log productivity in task type $c(s)$. The first two terms on the right hand side are the log demand for task type $c(s)$ (the log of the Leontieff coefficient plus log output) and the last term is the log Jacobian $\frac{dc}{ds} = c_s(s)$. In general, the differential equations (4) and (5) determining $y(s)$ and $c(s)$ do not have an analytical solution. However, as can be checked easily, the solution is simple in the special case where $\Delta\sigma = 0$.

$$\begin{aligned} c(s) &= s - \Delta\mu \\ y_s(s) &= \underline{\xi} e^{-\Delta\mu} \end{aligned} \tag{6}$$

The returns to skill depend on the difference between the means of the skill and complexity distribution, $\Delta\mu$. The higher the mean of the skill distribution, the higher the supply of highly skilled workers and the lower therefore their wage surplus compared to less skilled workers, see Figure 5, Panel A.[8] The return to skill $y_s(s)$ is inversely related to μ^s by imperfect substitutability between worker types across jobs: the increase in the supply of high skilled workers causes

[8] For $\Delta\pi = 0$, $\underline{\xi}$ is equal to the inverse of the *compression elasticity* or the *complexity dispersion parameter*, which is defined as the percentage decrease of the return to skill per percent increase of the skill level of workers evaluated at the going rate of return to skill, see Teulings (2003). The lower $\underline{\xi}$, the lower is the substitutability of worker types.

their wages to go down, the decrease in the supply of low skilled workers causes their wages to go up. An increase in μ^s causes the wage for the median s-type to go up, since the skill levels above the median earn more than half of the value of output. Hence, their wage reductions carry more weight than the wage increase in the lower half of the distribution. The wage of the median worker therefore has to go up since the substitution effects sum to zero. The break even point can be shown to be at $s = \mu^s + \sigma^{s^2}$. Note that $y_s(s)$ does not depend on s. Hence, log wages are a linear function of s in the special case $\Delta\sigma = 0$.

We are now in position to make further assumptions on the distribution of skill supply and complexity demand in the Walrasian benchmark region:

Benchmark parameters: $\Delta\pi = 0, \pi^c = \underline{\pi} \equiv [0, \underline{\sigma}]'$[9]

Equation (6) shows that using $\Delta\mu = \mu^c = 0$ as the benchmark case involves no loss of generality. Any other value of these parameters could be undone by a proper rescaling of s and c and a redefinition of $\underline{\xi}$ to undo the effect on $y_s(s)$. Setting $\Delta\sigma = 0$ is a restrictive assumption. However, this assumption greatly simplifies the analysis, since it is the only value for $\Delta\sigma$ for which the assignment problem in the Walrasian economy has an analytical solution and for which log value added $y(s)$ is linear in s. The normality of the skill distribution and the linearity $y(s)$ imply that the wage distribution is log normal, which is a reasonable description of empirical wage distributions. Its variance is:

$$\text{Var}[w] = \underline{\sigma}^2 y_s^2 = \underline{\sigma}^2 \underline{\xi}^2 \equiv \underline{\Sigma}^2 \tag{7}$$

Since contrary to $\underline{\sigma}^2$, the variance of wages, $\underline{\Sigma}^2$ can be observed directly form the data. Hence, we will cast most of our analysis in terms of the latter.

4.3 The Walrasian equilibrium for $\Delta\sigma \neq 0$

For the general case of $\Delta\sigma \neq 0$, no analytical solution is available, but we can provide a second order Taylor expansion of the solution around $\Delta\sigma = 0$. An approximate solution for the general case is presented in Appendix A.1. In Table 2 we present this solution in the form of a set of derivatives with respect to s, $\Delta\pi$, and π^c for the benchmark parameters and the median value of $s = \mu^s = 0$. The log wage function $y(s)$ is convex for $\Delta\sigma < 0$ and it is concave for $\Delta\sigma > 0$. Figure 6, panel B provides an intuition for the derivatives for the case where $\Delta\sigma > 0$. An increase in σ^s raises the relative supply of high and low-skilled workers at the expense of the supply of workers with an intermediate skill level. Consequently, wages go down in both

[9] Throughout the paper we just write 0 for the null vector.

extremes of the wage distribution and rise for intermediate skill levels: $y_{ss\Delta\sigma}(s) < 0$. Since the upper tail has a greater weight in value added than the lower tail, an equal relative increase in both tails at the expense of the median skill types is per saldo equivalent to an increase in the mean skill level, leading to a decrease in the return to skill. The opposite applies for $\Delta\sigma < 0$. Hence: $y_{s\Delta\sigma}(s) < 0$. All this shows that choosing $\Delta\sigma = 0$ as the benchmark is a serious restriction which however simplifies the analysis considerably since an analytical solution for $y(s)$ is available only for that case.

4.4 Choosing the right composite commodity

Till sofar, we treated the parameters π^c as exogenous. In this section they will be endogenized. Suppose that each value of π^c corresponds to a particular variety of a composite commodity. This idea of a region choosing to produce a particular type of composite commodity is directly related to the empirical evidence presented in Section 2. Take for example, Detroit, a region producing cars. The production of cars requires the input of different c-type tasks which are used in fixed proportions (wheels are no substitute for the window or the motor). These tasks are on average close to the nation wide mean level of complexity, μ^c. However, producing cars requires both very simple tasks (assemblage) and very complex tasks (design). Hence, the standard deviation of task types, σ^c, is above the nation wide average. As an opposite example, consider Arkansas, a region that specializes in agriculture. Agricultural production requires only relatively simple and homogeneous tasks so that both μ^c and σ^c are below their nation-wide average

Suppose that there is a market price for each variety of the composite commodities. Let this market price be denoted by $\underline{p}(\pi^c)$; $\underline{p}(\pi^c)$ is a hedonic price index in the sense of Rosen (1974). It measures the price of a commodity as a function of its characteristics. Since regions are assumed to be small relative to the total market, a region has to take $\underline{p}(\pi^c)$ as given. Then, competition forces regions to produce that commodity that maximizes its value added net of production cost, $\underline{p}(\pi^c) - p^{\#}$, taking task prices $p(c)$ as given:

$$\begin{aligned} \underline{p}(\pi^c) - p^{\#} &= 0 \\ \underline{p}_{\pi^c}(\pi^c) - \frac{dp^{\#}}{d\pi^c} &= 0 \\ \underline{P}_{\pi^c\pi^{c\prime}}(\pi^c) - \frac{d^2p^{\#}}{d\pi^c d\pi^{c\prime}} &< 0 \end{aligned} \tag{8}$$

where we use an upper case to denote the matrix of second derivatives. The first two equations are the first and the second order condition, while the third equation is due to the zero profit condition: wages will adjust to absorb all value added until the production cost of a unit of the composite commodity is equal to its market price.

Is is convenient to normalize $\underline{p}(\pi^c)$ and its derivatives for the benchmark parameters to zero without loss of generality:

Normalization: $\underline{p}(\underline{\pi}) = \underline{p}_{\pi^c}(\underline{\pi}) = 0$

The normalization $\underline{p}(\underline{\pi}) = 0$ is equivalent to using the benchmark commodity as the numeraire; the normalization $\underline{p}_{\pi^c}(\underline{\pi}) = 0$ is equivalent to a definition of the relative units of measurement of commodities of various types. Furthermore:

Assumption: $|\underline{P}_{\pi^c\pi^{c\prime}}(\pi^c)| \leq 0$

This assumption will assure an interior maximum for π^c, since in our model $\frac{d^2p^\#}{d\pi^c d\pi^{c\prime}}$ is of higher order. $\underline{p}(\pi^c)$ can be interpreted as a utility function, measuring the value of a unit of consumption with characteristics π^c. Then, the matrix of second derivatives, $\underline{P}_{\pi^c\pi^{c\prime}}(\pi^c)$, measures the elasticities of substitution between these characteristics.

5 Search Frictions

5.1 Assumptions

When there are search frictions, s-type workers meet only a limited amount of c-type firms per unit of time. In that case, it does not make sense for an s-type worker to wait for ever till the optimal task $c(s)$ comes along. A worker's matching set will be larger than just this unique optimal job type. Mutatis mutandis, the same applies for firms. Figure 6 depicts the matching sets of both the Walrasian equilibrium and the search equilibrium in s, c-space: the Walrasian equilibrium is represented by the diagonal, the search equilibrium by a band around it. Hence, after a contact between a worker and a firm takes place, both sides have to decide whether to match or to continue search. The larger the density of the labor market in a region, the more contacts per unit of time a worker expects to have and hence the more choosy she will be when deciding whether or not to match, since the expected revenues of continued search are high.

In this section, we analyze these issues more formally. For this purpose, we make the following assumptions. Only firms with vacancies and unemployed workers search. Central to our analysis is the notion that there is IRS in the contact technology. Let Λ be the size of the labor force. The contact rates $\lambda_{i\to j}$ for worker (job) type i to run into job (worker) type j are:[10]

$$\begin{aligned} \lambda_{s\to c} &\equiv v(c)\Lambda \\ \lambda_{c\to s} &\equiv u(s)\Lambda \end{aligned} \tag{9}$$

[10] These matching rates are consistent with a quadratic contact technology: $m = \lambda uv$, where m denotes contacts, u unemployment and v vacancies.

where $v(c)$ and $u(s)$ are the densities of vacancies of type c and of unemployed of type s respectively, both per unit of total labor supply. The contact rates are increasing in Λ. Search frictions are therefore smaller in dense cities than in rural areas. When the size of the regional labor force becomes infinitely large, $\Lambda = \infty$, the market outcome converges to the hypothetical Walrasian equilibrium. For simplicity, we assume that unemployed workers receive neither unemployment benefits nor a value of leisure. When a worker and a firm decide to match, the surplus is shared by Nash bargaining as in Pissarides (2000), with $\underline{\beta}$ being workers' bargaining power, $0 < \underline{\beta} < 1$. Matches are destroyed at a Poisson rate $\underline{\delta}$ and the future is discounted at rate $\underline{\rho}$. We study the economy while it is on a golden growth path, where the discount rate is equal to the growth rate of the labor force.[11] There is free entry of firms so that the asset value of a vacancy is zero in equilibrium.

5.2 Characterization of the equilibrium

Free entry makes the expected pay off of a firm with a vacancy equal to the cost of maintaining a vacancy. When a contact between a worker and a firm occurs, both sides have to decide whether or not they want to match. A match takes place when the reservation wage of the worker is less than the value of output. Let $R(s)$ be the reservation wage of worker type s. Then, the reservation wage of the worker and the free entry condition for vacancies read:

$$R(s) = \underline{\beta}\kappa \int_{m_c(s)} v(c) \left[Y^0(s,c) - R(s) \right] dc \tag{10}$$

$$P(c)\,\underline{\varphi} = (1-\underline{\beta})\kappa \int_{m_s(c)} u(s) \left[Y^0(s,c) - R(s) \right] ds \tag{11}$$

where $\kappa \equiv \frac{\Lambda}{\underline{\delta}+\underline{\rho}}$, captures the scale of a labor market, $\underline{\delta}+\underline{\rho}$ is the unemployment inflow rate into job seeking, $\underline{\delta}$ for destroyed matches and $\underline{\rho}$ for the growth of the work force, and where $\underline{\varphi}$ is the flow cost of maintaining a vacancy of type c in terms of output of this task type. Remember that $\underline{\delta}$ and $\underline{\rho}$ are the same in all regions so that all the variation in κ comes from the size of the labor force, Λ. The left hand side of (11) is the monetary flow cost of maintaining a vacancy of type c, that is $\underline{\varphi}$ units of task type c times the task price $P(c)$ (lower cases are the logs of the corresponding upper cases). The sets $m_s(c)$ (or: $m_c(s)$) are the subsets of c (or: s) with whom s (or: c) is willing to match. These subsets are determined by the condition that value added exceeds the reservation wage of the worker:

$$y^0(s,c) \geq r(s) \Leftrightarrow s \in m_s(c) \Leftrightarrow c \in m_s(s)$$

[11] This assumption is not critical for our results, but simplifies the analysis, since it implies that the net discounted cost of unemployment as a fraction of the reservation wage is equal to the unemployment rate.

The steady state flow equilibrium condition reads:

$$\frac{1}{\sigma^s}\phi\left(z^s\right) - u\left(s\right) = \kappa u\left(s\right)\int_{m_c(s)} v\left(c\right) dc \tag{12}$$

where: $z^s \equiv \frac{s-\mu^s}{\sigma^\sigma}$. The left hand side is equal to total employment for type s, being equal to total labor supply minus unemployment. The right hand side measures the new hires of type s, divided by unemployment inflow. Both must be equal in equilibrium. Finally, output of task type c is defined by:

$$\kappa v\left(c\right)\int_{m_s(c)} u\left(s\right)\underline{F}\left(s,c\right) ds - \underline{\varphi} v\left(c\right) = \frac{1}{\sigma^c}\phi\left(z^c\right)Y^{\#} \tag{13}$$

where z^c is defined analogously to z^s. This equation is equivalent to equation (5), but now for search equilibria. The left hand side is the total expected output of today's inflow of filled vacancies minus the cost of maintaining vacancies. The right hand side is task demand for type c, the Leontieff coefficient times output of the composite commodity.

The notation introduced previously for the Walrasian case can be extended for the characterization of the effect of search frictions of size κ on the reservation wage of the worker. For this purpose, we define the cost of search $x\left(s\right)$ as the difference between the maximum of log value added in the optimal assignment $y\left(s\right)$ and the log reservation wage $r\left(s\right)$:

$$r\left(s\right) \equiv y\left(s\right) - x\left(s\right) \tag{14}$$

The cost of search $x\left(s\right)$ of a type s worker is the log of the ratio of her value added in the optimal assignment and her reservation wage. The situation is depicted in the lower panel of Figure 6. Panel A represents the Walrasian case, where the log reservation wage $r\left(s\right)$ is equal to $y\left(s\right)$; hence, $x\left(s\right) = 0$. All workers of type s will be assigned to task type $c\left(s\right)$. Panel B represents the case with search frictions. The log reservation wage $r\left(s\right)$ is less than the maximum log value added $y\left(s\right)$, the difference being $x\left(s\right)$. An s-type job seeker accepts all c-type tasks for which $y^0\left(s,c\right) > r\left(s\right)$. In the Walrasian equilibrium, the worker is assigned to the unique job type $c\left(s\right)$ that maximizes $y^0\left(s,c\right)$ but under search frictions, she accepts all c-types satisfying the constraint $y^0\left(s,c\right) > r\left(s\right)$. Since wages are set by Nash bargaining, a fraction β of the difference $y^0\left(s,c\right) - r\left(s\right)$ translates into wages. The effect of search frictions on reservation wages can now be decomposed into two components, the direct effect on the cost of search $x\left(s\right)$ and the indirect, general equilibrium effect of search frictions on value added in the optimal assignment $y\left(s\right)$. These components will be discussed in the next two subsections.

5.3 The cost of search

This section relies heavily upon Teulings and Gautier (2003), to which we refer for details. An approximation of $x(s)$ can be derived by applying a second order Taylor expansion to the integrands of the system (10)-(12). For example, the integral in equation (10) is approximately equal to the surface enveloped by the functions $y^0(s,c)$ and $r(s)$ in Figure 6. A Taylor expansion of $y^0(s,c)$ allows us to use the formula for the integration of a parabola for the calculation of the integral. Hence, the second derivative $y^0_{cc}[s, c(s)]$ is of crucial importance for the size of search frictions. There is a simple economic intuition for this: the larger this second derivative, the larger the decrease in value added for suboptimal assignments, $c \neq c(s)$. The expansion of equation (11) uses the log production cost per unit of output of task c, denoted by $p^0(s,c)$. Here, $p^0_{ss}[s, c(s)]$ plays a similar role as $y^0_{cc}[s, c(s)]$ in equation (10). We refer to this second derivative as the *mismatch coefficient* $g(s)$:

$$g(s) \equiv p^0_{ss}[s, c(s)] = y_{ss}(s) + y_s(s)$$

The derivation of the final equality is delegated to Appendix A.2. The larger this mismatch coefficient, the larger the cost of search $x(s)$. For the benchmark parameters $\Delta\pi = 0, \pi^c = \underline{\pi}$, we have $y_{ss}(s) = 0, y_s(s) = \underline{\xi}$, and hence $g(s) = \underline{\xi}$. In general, $g(s)$ depends negatively on $\Delta\mu$ as well as on $\Delta\sigma$, as can be seen from the derivatives of y_s and y_{ss} with respect to these parameters in Table 2. An increase in μ^c while keeping μ^s constant raises the return to skill, and makes it therefore more important to assign the better skilled workers to the more complex jobs, which raises the mismatch coefficient. An increase in σ^c keeping σ^s constant raises the heterogeneity of job complexity, with a similar effect. Combining the Taylor expansions of equation (10)-(12) yields an expression for $x(s)$, see Appendix A.4 for its derivation:

$$\begin{aligned} x(s) &\cong \chi \left[\frac{\sigma^s}{\underline{\sigma}} \frac{\phi(0)}{\phi(z^s)} \sqrt{\frac{g(s)}{\underline{\xi}}} \right]^{2/5} \\ \chi &\equiv \underline{\chi}_0 \kappa^{-2/5} \end{aligned} \tag{15}$$

The scale of the labor force κ comes in with an elasticity of $-2/5$ via the new parameter χ. This elasticity follows from the order of the first non-vanishing terms in the Taylor expansions for the equations (10)-(12), see Appendix A.4: $x(s)$ enters both Bellman equations (10) and (11) with order 3/2 (adding up to 6/2), and it enters the flow equilibrium condition (12) with order 1/2, yielding in total an order of 6/2 - 1/2 = 5/2. Whether this increase comes in by a general increase in the size of the labor force, κ, or by the density of the skill distribution for that s-type, $\frac{1}{\sigma^s}\phi(z^s)$, is irrelevant for the effect on $x(s)$. This trade off between scale and density will be crucial in our analysis of interregional specialization as non-dense regions specialize in the production of homogeneous commodities which reduces σ^s-and offsets the initial negative effect on $x(s)$ of having a low κ.

The parameter $\underline{\chi}_0$ does not play a major role in our analysis, so we do not discuss it any further, see Appendix A.4 for its specification. However, it is conveniently specified such that for the benchmark parameters $\Delta\pi = 0, \pi^s = \underline{\pi}$ and for the median skill type, $s = \mu^s = 0$:

$$x(0) \cong \chi$$

We therefore refer to χ as the baseline cost of search, which is inversely related to the size of the labor market κ. In a hypothetical Walrasian region, $\kappa \to \infty$ and the baseline cost of search are zero: $\chi = 0$. For the sake of notational convenience and since there is a one-to-one relation between χ and κ, see equation (15), we cast the rest of our analysis in terms the baseline cost of search χ instead of in terms of the size of the labor market κ. The factor between square brackets in equation (15) reflects the differences in cost of search due to either (i) $s \neq \mu^s$, or (ii) $\pi^s \neq \underline{\pi}$, or (iii) $g(s) \neq \underline{\xi}$, where the latter depends on $\Delta\pi \neq 0$. We discuss all three factors in turn.

First, the density $\frac{1}{\sigma^s}\phi(z^s)$ reaches a maximum at the median skill type $s = \mu^s$, and hence $x(s)$ reaches a minimum at that point. The cost of search is higher in the tails of the distribution, which is a direct result of the IRS assumption. The situation is depicted in Figure 7, where the locus of $x(s)$ is drawn for two different values of χ. Let χ_1 denote the baseline cost of search in New York and χ_2 the cost of search in Montana. So: $\chi_2 > \chi_1$, due to the scale advantage of New York. For the median worker type $s = 0$, the log cost of search are somewhat higher in Montana. If we move towards the tails of the distribution where the market becomes thinner and thinner, differences in search frictions between New York and Montana get larger and larger, since χ and $\phi(z^s)^{-2/5}$ enter multiplicatively. Cost of search are therefore higher in small regions and higher in the tails of the distribution, and in particular high in the tails of the distribution in small regions.

Second, shifts in μ^s, while keeping $\Delta\mu$ fixed, just shift the mean of s. The effect of variations in σ^s, again keeping $\Delta\sigma$ fixed, is crucial for our analysis. The more dispersed the skill distribution (σ^s is high), the lower the density of the skill distribution at the median and the higher therefore the cost of search.

Finally, the mismatch coefficient $g(s)$ captures the effect of the difference between the skill and the complexity distribution on the cost of search. The more complex or the more heterogeneous jobs are relative to the skill distribution, the higher is the cost of suboptimal assignment as shown by a higher mismatch coefficient. A greater mismatch coefficient raises therefore the cost of search.

The multiplicative structure of equation (15) implies that all derivatives of $x(s)$ to other arguments than χ are of order $O[\chi]$. Since the effect of the baseline cost of search χ on π^s and π^c will be shown to be of $O[\chi]$, the indirect effect of χ via π^s and π^c on $x(s)$ is of order $O\left[\chi^2\right]$. Hence, we only have to account for the direct effect of χ in a first order approximation. The derivatives of $x(s)$ are presented in Table 3.

The cost of search can be decomposed into three factors: the net discounted cost of unemployment and of vacancies, and the efficiency loss due to suboptimal assignment. Generically, workers are not assigned to the job type $c(s)$ that maximizes their log value added $y^0[s, c(s)]$ in the presence of search frictions, since it is costly to wait for a job of this type to come along. The average productivity loss due to suboptimal assignment accounts for one third of the cost of search $x(s)$, while unemployment and the cost of vacancies account for the other two thirds, see Teulings and Gautier (2003). The distribution of the latter two thirds is proportional to the bargaining power of workers and firms.

$$\begin{aligned} \text{unemployment rate} &\cong \frac{2}{3}\underline{\beta} x(s) \\ \text{vacancy rate} &\cong \frac{2}{3}\left(1-\underline{\beta}\right) x(s) \\ \text{sub-optimal assignment} &\cong \frac{1}{3} x(s) \end{aligned} \tag{16}$$

The aggregate outcome looks like Figure 8. In the frictionless equilibrium, all matches are on the diagonal while in the search equilibrium, all matches are within two bounds around the diagonal.

5.4 The general equilibrium effects of search

Having derived an approximation for $x(s)$, we now turn our attention to the general equilibrium effect on $y(s)$. Where the characterization of $x(s)$ is just a restatement of Teulings and Gautier (2003), the characterization of the general equilibrium effects is new. Consider equation (13) that characterizes the equilibrium of the market for task type c in the presence of search frictions. The integral on the right hand side is approximated by a Taylor expansion. Taking logs and using the benchmark assumption $\Delta\pi = 0, \pi^c = \underline{\pi}$ yields:

$$-\frac{1}{2}\left(\frac{s}{\underline{\sigma}}\right)^2 + \underline{f}[s, c(s)] - x(s) - x_s(s) \cong -\frac{1}{2}\left(\frac{c(s)}{\underline{\sigma}}\right)^2 + y^{\#} + \ln c_s(s) \tag{17}$$

The derivation is in Appendix A.4, see in particular equation (44). Equation (17) is the equivalent of equation (13) in the presence of search frictions. The final two terms on the left hand side capture the two types of effects of search frictions on $y(s)$. First, the term $-x(s)$ accounts for the *skill compression effect.* For each skill type s, a fraction $x(s)$ of the actual skill distribution is "lost" due to search frictions, partly due to unemployment, partly due to vacancies, and partly due to the cost of suboptimal assignment. Since $x_{ss}(s) > 0$ and $x_s(\mu^s) = 0$, this loss is larger in the tails of the distribution. Hence, a larger fraction of labor supply is "lost" in the tails than around the median. Search frictions compress therefore the "effective" skill distribution. Second, the term $-x_s(s)$ accounts for the *skill spoiling effect*, which is somewhat more difficult to understand. In a Walrasian world, the task type $c(s)$ is produced by workers of skill type s only. In the presence of search frictions, both better and worse skilled worker types than s are part of the matching set $m_s[c(s)]$ of task type $c(s)$, which is approximately symmetric around $c(s)$. Since $y_s(s) > 0$, the better skilled types have an above average productivity, and the worse skilled types a below average productivity. Next, recall that by assumption, the skill distribution is unimodal. Consider what happens above the mode, where the skill density is falling. The number of low skilled workers becoming part of a symmetric matching set of a relatively complex job is larger than the number of high skilled workers. Hence, average productivity falls. Exactly the opposite argument applies below the modus. Hence, productivity falls for the upper part of the skill distribution and rises for the lower part. This is equivalent to a fall in the average skill level of the labor force. Since $x(s)^{-5/2}$ is proportional to the density of labor supply, $\frac{1}{\sigma^s}\phi(z^s)$, see equation (15), $x_s(s)$ is proporitional to $d\frac{1}{\sigma^s}\phi(z^s)/ds$.

Equation (17) holds identically for all s. Hence, its first two derivatives with respect to s evaluated at $s = 0$ have to apply as well, see Appendix A.5, equation (45). From this we derive the general equilibrium effect of search frictions χ on $y_s(0)$ and $y_{ss}(0)$, the derivatives of $y(0)$. For the calculation of the general equilibrium effects on the level of $y(0)$, it is useful to introduce some matrix notation. Let a "$\sim$" above a function denote a vector containing its first two derivatives at the median skill type $s = 0$, so $\widetilde{y} \equiv [y_s(0), y_{ss}(0)]'$. The relation between $y(0)$ and $\widetilde{y}$ on the one hand and $\Delta\pi$ and π^c on the other hand is given in Appendix A.1. Table 2 presents the derivatives of this function. These derivatives can be used for a first order Taylor expansion of $y(0)$ and $\widetilde{y}$ for $\Delta\pi \neq 0$. Let the superscipt f denote the difference between the benchmark value of $y(s)$ for $\Delta\pi = 0$ and its value for $\Delta\pi \neq 0$, let $y_{\Delta\pi}$ be the vector of derivatives of $y(0)$ and let $\widetilde{Y}_{\Delta\pi'}$ be the matrix of derivatives of $\widetilde{y}$, both with respect to $\Delta\pi$ (all these derivatives are listed in table 2), then:

$$\begin{aligned} y^f(0) &= y_{\Delta\pi}\Delta\pi \\ \widetilde{y}^f &= \widetilde{Y}_{\Delta\pi'}\Delta\pi \end{aligned}$$

Elimination of $\Delta\pi$ from the first equation by matrix inversion yields:

$$y^f(0) = y'_{\Delta\pi}\widetilde{Y}^{-1}_{\Delta\pi'}\tilde{y}^f$$

Now, let $\tilde{y}^f$ stand for the general equilibrium effect of search frictions on $\tilde{y}$ as derived in Appendix A.5, equation (45). Then, we can apply the above equation to calculate the general equilibrium effect on $y(0)$. Since this equation has to apply for all χ, its first derivative with respect to χ has to apply. Hence:

$$y^f_\chi(0) = y'_{\Delta\pi}\widetilde{Y}^{-1}_{\Delta\pi'}\tilde{y}^f_\chi \tag{18}$$

Table 4 summarizes the direct effect on $x(s)$ and the indirect, general equilibrium effect on $y(s)$ of search frictions of size χ. The effect on reservation wages can be calculated from $r(s) = y(s) - x(s)$, by substracting the direct effect from the general equilibrium effect. The general equilibrium effects raise both the derivatives of $y(s)$ (that is: both elements of $\widetilde{y}^f_\chi$ are positive). The first derivative is pushed up both by the *skill spoiling effect* and by the *skill compression effect*. The skill spoiling effect makes skills more scarce, thereby raising return to skill y_s. The negative effect of skill compression in the tails of the distribution has a greater weight in the upper tail than in the lower tail since productivity is higher in the upper tail, thereby making skills more scarce overall and hence raising y_s. The rise in the return to skill causes y for the median worker to go down. Since substitution effects sum to zero, the mean value added E$[Y(s)]$ must be constant. Since there is more value above than below the median of s, the wage of the median worker has to go down to make this happen. The skill compression effect yields a wage distribution which is skewed to the right.

The second derivative is pushed up by the *skill compression effect*: skill types in the tails of the distribution become more scarce, thereby raising task prices in the tails, which in turn translates in higher wages. Relative wages in both tales of the distribution are pushed up. Since y_{ss} goes up, y must go down to keep E$[Y(s)]$ constant. The direct effect of search frictions is illustrated in Figure 7: search frictions are higher in the tails of the distribution, $x_{ss} > 0$.

6 Interregional labor mobility and trade

6.1 Labor mobility

We are now able to analyze the effect of the interaction of search frictions and interregional labor mobility and trade on the skill and complexity distribution and on the level of nominal wages in regions of various size. This subsection focusses on labor mobility, assuming that all regions produce the same commodity. In the next subsection, we analyze the implications of interregional specialization in the production of commodities.

We assume that the long run cost of labor mobility are zero. Hence, workers continue to migrate till real reservation wages for each s-type are equal across regions. By the homotheticity of the utility function, all worker types deflate their nominal reservation wage by the same cost of living index. Hence, log nominal reservation wages net of cost of living are equal across regions. Owners of real estate in a region push up the rents for their property till the cost of living satisfy the no-arbitrage condition. Let $l(\chi)$ denote this log cost of living index for a region with baseline search cost χ. Hence:

$$r(s) - l(\chi) = y(s) - x(s) - l(\chi) = \underline{r}(s) \tag{19}$$

where $\underline{r}(s) = \underline{\xi} s$ are the log reservation wages that apply in the Walrasian benchmark, $\chi = 0$, $\Delta\pi = 0, \pi^c = \underline{\pi}$. Equation (19) is a straightforward extension of equation (14). Since this equation applies identically for all χ, it first derivative with respect to χ has to apply. Similarly, (19) applies for all s, its first two derivatives evaluated at $s = 0$ have to apply. Hence, interregional labor mobility imposes the following no-arbitrage conditions:

$$\begin{aligned} y_\chi(0) - x_\chi(0) &= l_\chi(0) \\ \tilde{y}_\chi - \tilde{x}_\chi &= 0 \end{aligned} \tag{20}$$

where $\tilde{x} \equiv [x_s, x_{ss}]'$ and $\tilde{r} \equiv [r_s, r_{ss}]'$. From Table 4, we know that variations in the baseline cost of search χ have a direct effect, $\tilde{x}_\chi$, and a general equilibrium effect $\tilde{y}_\chi$ (the latter denoted by $\tilde{y}^f_\chi$). For the no-arbitrage condition to be satisfied, these effects have to be offset by shifts in the skill distribution by interregional labor mobility. We approximate these shifts by their effect on the first two moments of the skill distribution, $\pi^s \equiv [\mu^s, \sigma^c]$.[12] Since we keep constant π^c, a shift in π^s is equivalent to a shift in $\Delta\pi$. Their derivatives with respect to the baseline cost of search χ can be derived straightforwardly from (20):

$$\widetilde{Y}_{\Delta\pi}\Delta\pi_\chi + \tilde{y}^f_\chi - \tilde{x}_\chi = 0 \tag{21}$$

The first term is the effect of labor mobility, the second term reflects the indirect general equilibrium effects of search, and the third term the direct effects of search. Solving (21) for $\Delta\pi_\chi$ yields:

$$\Delta\pi_\chi = \widetilde{Y}^{-1}_{\Delta\pi}\left(\tilde{x}_\chi - \tilde{y}^f_\chi\right) = \frac{1}{5\underline{\xi}}\begin{bmatrix} \underline{\Sigma}^4 + \underline{\Sigma}^2\left(2\underline{\xi} - 1\right) + 2\underline{\xi} \\ \frac{1}{\underline{\Sigma}}\left(-2\underline{\xi} + \underline{\Sigma}^2\right) \end{bmatrix} \tag{22}$$

The mean of the skill distribution is an increasing function of the cost of search, $\Delta\mu_\chi > 0$. This is due to the *skill spoiling effect*, which raises the return to skill y_s. The *skill compression*

[12] Expressions for higher order moments of the skill distribution can be obtained by considering higher order derivatives of equation (19).

effect reinforces this effect. This invokes immigration of highly skilled workers, until the market clearing condition $r_s = \underline{\xi}$ is satisfied. The effect of search frictions on the standard deviation of the skill distribution, $\Delta\sigma_\chi$ depends on two effects which are of opposite sign. The direct effect of larger search frictions on reservation wages, $x_{ss\chi}$, is negative, leading to an emigration of skill types in the tail, $\Delta\sigma$ going down. The indirect, general equilibrium effect, $y_{ss\chi}$, works in the opposite direction. The reduction in effective supply in the tails of the distribution pushes up the wages for these skill types. The direct effect dominates when worker types are good substitutes, that is for high values of $\underline{\xi}$:

$$2\underline{\xi} > \underline{\Sigma}^2 \tag{23}$$

Then, smaller regions with larger χ export worker types from the tails of the skill distribution: $\Delta\sigma_\chi < 0$.

Using the same methodology as in equation (18) yields the following expression for $y_\chi(0)$:

$$y_\chi(0) = y'_{\Delta\pi}\Delta\pi_\chi + y^f_\chi(0) = y'_{\Delta\pi}\widetilde{Y}^{-1}_{\Delta\pi}\tilde{x}_\chi = -\frac{1}{5} - \underline{\Sigma}^2 \tag{24}$$

In the final equality, the expression is evaluated, using the derivatives listed in Tables 2 and 4. The indirect general equilibrium effects drop out here. The reason is obviously that the general equilibrium effects on $\tilde{y}$ are exactly offset by labor mobility. On top of that, labor mobility sets $\tilde{y}$ equal to $\tilde{x}$, to compensate the direct effect of search frictions. Hence, only the latter effect enters into $y_\chi(0)$. This effect is negative. The intuition is that search frictions are higher in the tails. To offset this negative effect on relative reservation wages, $y(s)$ has to go up in the tails relative to the Walrasian benchmark. Since substitution effects sum to zero, this inevitably implies that wages around the median have to go down.

The effect of search frictions on the cost of living is larger than unity, that is, greater than the direct effect of the baseline cost of search χ. A simple way to understand why this is the case is to realize that χ is the minimal value of $x(s)$ attained for the median skill type. Since the no arbitrage condition (19) dictates that all log reservation wages have to be reduced by the same amount, the higher value of $x(s)$ in the tails has to be offset by substitution effects that increase $y(s)$ in the tails. Since the value weighted sum of substitution effects equals zero, wages have to go down in the median. The greater $\underline{\Sigma}$, the greater this effect, for the usual reason that there is more value above than below the median.

The effect of search frictions on reservation wages is $y_\chi(0)$, see equation (24), minus $x_\chi(0) = 1$. The effect of cost of living is the same as that on log reservation wages, by the no-arbitrage condition. However, the effect on expected log wages (keeping s constant) is smaller in absolute value. The reason for this difference is that higher expected wages have to compensate for the longer expected duration of unemployment. By the mirror image of the argument that the average log productivity loss due to sub-optimal assignment is one third of the cost of search, $\mathrm{E}\left[y(s) - y^0(s,c)\,|s\right] = \frac{1}{3}x(s)$. Hence, the average log productivity surplus above the log reservation wage equals $\mathrm{E}\left[y^0(s,c) - r(s)\right] = \frac{2}{3}x(s)$. Workers get a share $\underline{\beta}$ of this.

6.2 Interregional commodity trade

The analysis of commodity trade extends the case of labor mobility by the endogenization of the parameters π^c of the tradable commodity produced by a region. The results for $\Delta\pi_\chi$ in Table 5 apply unchanged close to the Walrasian benchmark.[13] Hence, the effect of search frictions on the parameters of the skill distribution satisfies:

$$\pi^s_\chi = \pi^c_\chi + \Delta\pi_\chi \tag{25}$$

Any shift in the complexity distribution due to regional specialization translates one-for-one in a shift in the skill distribution by additional labor mobility.

Owners of real estate raise the rents as far as possible. This forces firms to produce that tradable commodity that maximizes their profits, taking reservations wages as given, which yields a first order condition in the form of (8). Free entry of firms then pushes up wages till this maximum profit is equal to zero. However, equation (8) is not very convenient, since it requires us to calculate the log price of the composite commodity, $p^\#$, by means of equation (1). A simpler alternative is to realize that maximizing profits and imposing a zero profit constraint is equivalent to maximizing the nominal reservation wage of one skill type, for example the median, $s = 0$. Since labor mobility fixes relative reservation wages at their value in the Walrasian benchmark, maximizing the nominal reservation wage for one worker type simultaneously maximizes the nominal reservation wages for all other worker types. The nominal log reservation wage of the median worker satisfies:

$$r(0) = y(0) - x(0) + \underline{p}(\pi^c)$$

Hence, the first order condition reads:

$$\frac{dy(0)}{d\pi^c} - \frac{dx(0)}{d\pi^c} + \underline{p}_{\pi^c}(\pi^c) = 0$$

This condition must apply identically for all χ and hence its first derivative in the Walrasian

[13] $\Delta\pi_\chi$ does not depend on π^c_χ in a first order approximation, since the cross effect is of higher order.

benchmark applies as well:[14]

$$-\underline{P}_{\pi^c\pi^{c\prime}}\pi^c_\chi = \begin{bmatrix} 0 \\ -\frac{\xi}{\underline{\Sigma}}\frac{12+22\underline{\Sigma}^2}{25} \end{bmatrix} \tag{26}$$

If the matrix $\underline{P}_{\pi^c\pi^{c\prime}}$ were negative definite, equation (26) would define a set π^c of dimension zero for each value of χ. The equilibrium set of all π^c would be of the same dimension as χ, that is, only single dimensional, while the characteristics space spanned by π^c is two dimensional. This violates the assumption that there is positive demand for all types π^c.[15] Hence, $\underline{P}_{\pi^c\pi^{c\prime}}$ can only be a market equilibrium if it is negative semi-definite. Competition will force it to be that way. The only matrix $\underline{P}_{\pi^c\pi^{c\prime}}$ that is consistent with this requirement reads:

$$\underline{P}_{\pi^c\pi^{c\prime}} = \begin{bmatrix} 0 & 0 \\ 0 & \underline{p}_{\sigma^c\sigma^c} \end{bmatrix} \tag{27}$$

Combining (26) and (27) yields:

$$\sigma^c_\chi = -\frac{2}{5}\frac{6\underline{\Sigma}^{-2}+11}{5}\frac{\xi\underline{\Sigma}}{\left|\underline{p}_{\sigma^c\sigma^c}\right|} < 0 \tag{28}$$

The dispersion of the complexity distribution σ^c depends negatively on the baseline search friction χ. As discussed in section 3.3, the scale effect in the cost of search for the median skill type $s = 0$ is made up of the level of total labor supply and the density function of the skill distribution for that particular skill type, $\frac{1}{\sigma^s}\phi(0)$. Part of the negative scale effect in low density regions is offset by these regions producing more homogeneous products. This automatically translates into a more homogeneous skill distribution, see equation (25), which reduces σ^s and therefore increases the density. In other words, a region with large search frictions specializes in the production of tradable commodities with a low value of σ^c because this reduces the need for

[14] Substition of equation (24) and applying the formula for the derivative of an inverse matrix

$$\frac{dA^{-1}}{dx} = A^{-1}\frac{dA}{dx}A^{-1}$$

yields:

$$y'_{\Delta\pi\mu^c}\widetilde{Y}^{-1}_{\Delta\pi}\tilde{x}_\chi - y'_{\Delta\pi}\widetilde{Y}^{-1}_{\Delta\pi}Y_{\Delta\pi\mu^c}\widetilde{Y}^{-1}_{\Delta\pi}\tilde{x}_\chi + Y'_{\Delta\pi}\widetilde{Y}^{-1}_{\Delta\pi}\tilde{x}_{\chi\mu^c} - x_{\chi\mu^c} = -\frac{d\underline{p}_{\mu^c}\left(\underline{\pi}+\pi^c_\chi\chi\right)}{d\chi}_{|\chi=0}$$

and mutatis mutandis the same for σ^c. These expressions can be evaluated using table 2. Furthermore:

$$\frac{d\underline{p}_{\pi^c}\left(\underline{\pi}+\pi^c_\chi\chi\right)}{d\chi}_{|\chi=0} = \underline{P}_{\pi^c\pi^{c\prime}}\pi^c_\chi$$

[15] This assumption is not critical for our conclusion. If the assumption were not satisfied for particular combinations of π^c, then we can adjust these prices to make $\underline{P}_{\pi^c\pi^{c\prime}}$ negative semi-definite, since these commodities are not produced in equilibrium anyway. Hence, this case is embedded in the case where the constraint of negative semi-definiteness is imposed.

job search. High density regions do the opposite: they produce tradables with heterogeneous inputs, in order to exploit their comparative advantage in search intensive production.

The model does not yield any prediction regarding μ_{χ}^{c}. Any value can be the optimal choice for a region of particular scale, since scale does not yield comparative advantage in the production of on average more or less complex tradables. Hence, regional specialization does not pay off. This conclusion is contingent on our choice of setting $\Delta\sigma = 0$ for the Walrasian benchmark. This choice implies that the mismatch coefficient $g(s)$ is independent of s. Hence, the cost of suboptimal assignment does not vary with s, so that search frictions are equally important for low and high skilled workers.

Figure 10 offers a graphical illustration. The parabola is the function $\underline{p}(\sigma^c)$. By construction, it reaches a maximum for the benchmark value, $\sigma^c = \underline{\sigma}$. The function $\underline{p}(\sigma^c, \chi)$ is plotted for two values of χ, $\chi_2 > \chi_1$. The optimal choice of σ^c sets marginal revenue $\underline{p}_{\sigma}(\sigma^c)$ equal to marginal cost $p_{\sigma}(\sigma^c, \chi)$. Cost of living adjusts such that it makes this value of σ^c a point of tangency between both curves. Since $p_{\sigma\chi}(\sigma^c, \chi) < 0$ and since $\underline{p}_{\sigma\sigma} < 0$, regions with larger χ choose lower values of σ^c. The negativity of this cross-derivative is therefore crucial for our conclusion. The mechanism relates directly to larger cost of search in the tails of the distribution, see Figure 7. To offset these larger cost in the tails, any increase in σ^c has to be compensated by an increase in σ^s, which in turn increases the cost of search by reducing the density of the skill distribution. Since the difference between the cost of search at the median and in the tails is proportional to χ, metropoles with a low χ have a comparative advantage in producing tradables with a dispersed complexity distribution, resulting in a high σ^c.

6.3 Do the selection effects offset IRS?

The results in this paper depend strongly on IRS in the contact technology. However, most of the empirical evidence seems to suggest that returns to scale are mildly increasing at best. In Teulings and Gautier (2003) we argue that part of the returns to scale are absorbed by workers and firms becoming more choosy as the scale of the market increases. It is optimal to respond to the higher contact rate of match offers by turning down a larger share of the offers. This mechanism accounts for one third of the returns to scale. Empirical research on aggregate data for unemployment and vacancies would therefore underestimate the returns to scale. However, this mechanism can never explain why we do not observe any returns to scale at all, since a greater choosiness can never completely offset the initial gain in the contact rate.

The model in this paper introduces an alternative mechanism which dampens IRS in contacts. Metropoles exploit their comparative advantage by specializing in the production of the heterogeneous commodities. A greater heterogeneity in task demand puts a burden on the search process, by reducing the share of match offers that is acceptable. This specialization mechanism can more than offset the initial advantage in contact technology of large metropoles. In order to analyze this issue more formally, we extend the approximation of x_χ by second order terms:

$$\begin{aligned} x_\chi &= 1 + x'_{\Delta\pi}\widetilde{Y}_{\Delta\pi}^{-1}\tilde{x}_\chi + x'_{\sigma^s}\left(\Delta\sigma_\chi + \sigma^c_\chi\right) \\ &= 1 + \frac{4}{25}\left(\frac{1}{2} + \underline{\xi}\left(1 - \underline{\xi}\underline{\Sigma}^{-2}\right) - \frac{6\underline{\Sigma}^{-2} + 11}{5}\frac{\underline{\xi}^2}{\left|\underline{p}_{\sigma^c\sigma^c}\right|}\right)\chi \end{aligned} \tag{29}$$

where we use $x_{\mu^s} = 0$ in the first equality. The first term captures the direct effect of search frictions. The second term captures the indirect effects on $\tilde{y}_\chi$, via the mismatch coefficient $g(s)$, due to the net effect of labor mobility and the general equilibrium effects of search frictions. The general equilibrium effect of search frictions, $\tilde{y}^f_\chi$, is offset by labor mobility, so that only the labor mobility offsetting $\tilde{x}_\chi$ remains, analogous to equation (24). The third term captures the effect of labor mobility on the skill distribution, partly offsetting relative wage effects, partly offsetting the regional specialization in commodity production. Whether or not an increase in the baseline cost of search χ reduces $x(0)$ depends on the sign of x_χ. We focus the discussion on the second order effect, the second term in the final line of equation (29). For a reasonable value of the variance of log wages $\underline{\Sigma}^2$, the factor $1 - \underline{\xi}\underline{\Sigma}^{-2}$ cancels. Hence, altogether the indirect effects of labor mobility via the mismatch coefficient and the skill distribution push up $x(0)$.[16] If this effect would remain unaccounted, empirical research on aggregate data would over- and not underestimate the returns to scale in the contact technology. However, specialization in commodity production as measured by $\left|\underline{p}_{\sigma^c\sigma^c}\right|$ offers a counteracting effect: the larger this specialization, the closer is x_χ to zero, and hence the more difficult it is to establish IRS from aggregate data alone. In the next section, we shall try to assess whether $\left|\underline{p}_{\sigma^c\sigma^c}\right|$ is sufficiently small to have a substantial effect on x_χ.

7 Empirical evidence

7.1 Introduction

Our model yields a number of testable implications for the composition of worker and job types, the wage of the median worker, and the cost of living in regions of various size. Those testable implications are summarized below:

1. from Table 5: $\Delta\mu_\chi = \frac{4}{5} > 0$, due to the skill spoiling effect;

[16] The intuition is that the no arbitrage condition for labor mobility requires $y_{ss} - x_{ss} = 0$ and hence $y_{ss} = x_{ss} > 0$. Since the mismatch coefficient satisfies $g = y_s + y_{ss}$, an increase in y_{ss} raises g, which in turn raises x.

2. also from Table 5: $\Delta\sigma_\chi = \frac{1}{5}\left(\underline{\xi}^{-1}\underline{\Sigma} - 2\underline{\Sigma}^{-1}\right) \gtrless 0$, since the direct and the indirect effect of the skill compression effect are of opposite sign;

3. from equation (28) $\sigma^c_\chi = -\frac{2}{5}\frac{6\underline{\Sigma}^{-2}+11}{5}\underline{\xi}\underline{\Sigma}\left|\underline{p}_{\sigma^c\sigma^c}\right|^{-1} < 0$, due to regional specialization and commodity trade;

4. from Table 5: $w_\chi = -\frac{6}{5} - \underline{\Sigma}^2 < 0$, since large areas are more efficient in matching workers to jobs;

5. $l_\chi = -\frac{6}{5} - \underline{\Sigma}^2 + \frac{2}{3}\underline{\beta} < w_\chi$, since the higher log real wage in high search cost regions, $w_\chi - l_\chi > 0$, must offset the longer duration of unemployment.

We have a reasonable idea about the empirical value of some of the parameters. For example, the variance of log wages, $\underline{\Sigma}$, can be directly observed from the data. For the US, $\underline{\Sigma}$ is about 0.60, but for most European countries it is substantially lower. We apply $\underline{\Sigma} = 0.50$ as a compromise between both sides of the Atlantic. We apply the standard value for the worker's bargaining power $\underline{\beta} \cong 0.50$. In Teulings and Gautier (2001), we show that this value leads to the smallest efficiency loss. Evidence on the mismatch coefficient $\underline{\xi}$ is much more scattered. The mismatch coefficient happens to be the inverse of complexity dispersion parameter, see Teulings and Vierra (1999). It is also related to Katz and Murphy's (1992) elasticity of substitution between low and high skilled workers, see Teulings (2002). These pieces of evidence suggest $\underline{\xi} \simeq 0.25$.[17] For the average baseline cost of search, $\mathrm{E}[\chi]$, there are two ways to obtain a guesstimate, which however give completely different answers. One alternative is to use equation (16), which says that unemployment makes up for one third of the cost of search. Applying a natural unemployment rate of 5%, this yields a $\mathrm{E}[\chi] = 15\%$. A second alternative invokes the concavity of log wages in s and c, see Figure 6. Gautier and Teulings (2003) provide evidence for this concavity. This evidence suggests values of $\mathrm{E}[\chi]$ up to 75%. Again, we compromise and apply a value for $\mathrm{E}[\chi]$ of 30%.[18] Table 6 summarizes those benchmark values of the model's parameters. We have no sensible estimate for the price elasticity of σ^c, $\left|\underline{p}_{\sigma^c\sigma^c}\right|$. Here, we have to rely on the empirical results that are present below.

Our empirical test applies the empirical measures for skill, complexity, and baseline cost of search, $\hat{s}$, $\hat{c}$, and $\hat{\chi}$ respectively, which were derived in section 2. The skill and complexity

[17] Katz and Murphy's (1992) estimate imply the mismatch parameter to be about 2, which would imply $\underline{\xi} \simeq 0.50$. However, Katz and Murphy's estimate includes both between-occupation and within-occupation-between-skill substitution, while the mismatch coefficient refers only to the latter type of imperfect substitution.

[18] Both guestimates can be reconciled by allowing for on the job search. Ignoring on the job search yields an underestimate of $\mathrm{E}[\chi]$ by the first method and an overestimate by the second method, see ongoing work by Gautier, Teulings, and Van Vuuren (2003).

measures $\hat{s}$ and $\hat{c}$ have been constructed such that $dw/d\hat{s} = dw/d\hat{c} = 1$. Since wages are equal to value added in the Walrasian benchmark, we have $w_s = y_s$. Hence, the empirical skill measure is equal to $\hat{s} = y_s s = \underline{\xi} s$ and hence $\hat{\pi}^s = \underline{\xi}\pi^s$, and the same for $\hat{c}$, since $c'(s) = 1$ for the Walrasian benchmark. We can therefore use our reference value of $\underline{\xi} = 0.25$ to translate from π^s and π^c to $\hat{\pi}^s$ and $\hat{\pi}^c$. The situation is more complicated for our empirical measure for the baseline cost of search, $\hat{\chi}$; $\hat{\chi}$ is constructed such that it is good proxy for the scale of the regional labor market. Hence, it is highly correlated with χ. However, we do not know the transformation rule from the one to the other, $d\chi/d\hat{\chi} \equiv \chi_{\hat{\chi}}$. Hence, like the price elasticity $\left|\underline{p}_{\sigma^c\sigma^c}\right|$, this transformation parameter $\chi_{\hat{\chi}}$ has to be established empirically.

We use the Dumond et al.(1999) regional cost of living indices to test for the relation between search frictions and the cost of living. Finally, we calculate and estimate w_χ by estimating an earnings function for each region with s, c and dummies for each year as the explanatory variables. The region specific constants are our proxy for the wage of the median worker (for which $s = 0$) in that region.

7.2 Estimation results

Using Table 6 and (22) we obtain, $\Delta\sigma_\chi = \frac{1}{5\underline{\xi}\underline{\Sigma}}\left(-2\underline{\xi} + \underline{\Sigma}^2\right) = -0.4$ and from Table 6 and (28), we get: $\sigma^c_\chi < 0 = -\frac{2}{5}\frac{6\underline{\Sigma}^{-2}+11}{5}\frac{\underline{\xi}\underline{\Sigma}}{\left|\underline{p}_{\sigma^c\sigma^c}\right|} - 1.17|\underline{p}^{-1}_{\sigma^c\sigma^c}| = -\frac{7}{20}|\underline{p}^{-1}_{\sigma^c\sigma^c}|$.

Table 7 and 8 give the estimation results for $\Delta\hat{\pi}_{\hat{\chi}}$, $\hat{\pi}^s_{\hat{\chi}}$, $\hat{\pi}^c_{\hat{\chi}}$, $l_{\hat{\chi}}$, and $w_{\hat{\chi}}$. Since the variables to explain are sample statistics, we weight all estimates by the inverse of their standard error. The regression results reported in column 1 are based on the 82 regions for which we have at least 800 observations.[19] Implications 1 till 4 are all supported by the data. All coefficients have the predicted sign and are significant at the 95% level (except $\Delta\hat{\mu}_{\hat{\chi}}$, which is only significant at the 90% level). We were particularly surprised to find support for implication 1, $\Delta\hat{\mu}_{\hat{\chi}} < 0$ due to the skill spoiling effect. One would expect metropolitan areas to have excess supply of high skilled workers, if alone because the location of universities might bias the distribution of highly skilled workers towards these regions. However, our model predicts unambiguously the opposite, and the data confirm this prediction: controlling for the composition of the complexity of labor demand, the mean skill level is higher in low density regions. The greater dispersion of the complexity distribution in dense areas, see implication 3 is also confirmed by the data. In addition, we find a strong effect for the mean of the complexity distribution $\hat{\mu}^c_{\hat{\chi}} < 0$. This result provides evidence neither against nor in favor of our model. The model does not yield

[19] We have also experienced with leaving the "immigrant ports" like El Paso, LA, NY, and Miami out. This did not change the results.

any specific prediction, so any outcome is consistent. However, the fact that the effect is so strong suggests that there is a systematic force that drives metropolitan areas to produce the more complex tradables. One explanation, which is consistent with the flavor of our model, is that the mismatch parameter $g(s)$ is not constant but increasing in s, implying that the impact of search frictions is larger in the upper than in the lower tail of the complexity distribution. If this is the case, then metropolitan areas would specialize in complex tradables, by a similar mechanism as why they specialize in dispersed tradables in the model as it stands. To avoid even more complexity, we did not investigate this route any further.

Most theories of search in the labor market suggest that search frictions are more relevant for young workers, see Topel (1991). In the beginning of their career workers have not yet settled down in a good match. Whenever a better job comes along, they switch. Later on in their career, most workers have found a reasonable match, and jobs therefore tend to last longer. Consequently, search frictions are less relevant for older workers. In column 2, we therefore restrict the analysis to workers younger than 30. This restriction increases all coefficients and their t-values, even though we use fewer observations per region. To make sure that those results are not due to the limitation of the analysis of the set of regions with a sufficient number of observations, we repeat our estimates for exactly the same restricted set of regions, but now for the entire work force in column 3. Column 3 gives similar results as column 1, suggesting that the age restriction, not the smaller set of regions drives the increase of the coefficients from column 1 to column 2. In Table 8 we only include CMSA's because we are worried that other factors might explain the characteristics of agricultural regions. In particular, the large input of land in agricultural production might offer an alternative explanation for the special characteristics of rural areas. Another reason for restricting the sample to CMSA's only is that we have cost of living information for those areas, which allows us to test implication 5. Obviously by restricting the sample to CMSA's only, we decrease the mean and standard deviation of $\hat{\chi}$, see Table 7 and 8. Again, the estimates are consistent with implications 1 till 4 and part of implication 5. We find that $l_\chi < 0$, but not $l_\chi < w_\chi$.[20] We also restrict the sample to CMSA's and young workers only. Again, this makes all coefficients larger and more significant.

We use the coefficients in column 2 to get an impression of the magnitude of search frictions. Theoretically, an estimate of $\chi_{\hat{\chi}}$ can be derived both from implications 1 and 2. However, since the sign of $\Delta\sigma_\chi$ is theoretically ambiguous and since $\Delta\hat{\sigma}_{\hat{\chi}}$ is indeed not significantly different from zero, we focus on the estimate for $\Delta\mu_\chi$. The benchmark value for the transformation parameter

[20] Glaeser and Mare (2001) show that the city wage premium cannot be fully attributed to unobserved ability bias because migrants moving into the city experience an immediate wage gain.

$\chi_{\hat{\chi}}$ listed in Table 6 is derived from this estimate.[21] This transformation parameter can be used to calculate the standard deviation of χ, which is the basis for the coefficient of variation of χ that is also listed in Table 6.[22] The definition of $\chi(\kappa)$ in equation (15), $\ln\chi = \frac{2}{5}\ln\kappa_0 - \frac{2}{5}\ln\kappa$, allows to transform this into a coefficient of variation of the scale of the labor market, κ, which is 28%. Our estimation results therefore suggest that a four standard deviation interval of the log of the scale of the market has a width of $4 \times 0.28 = 1.1$. Search is therefore $e^{1.1} > 3$ times as effective in the top 5% metropolitan areas than in the 5% least densely populated regions.

Similarly, we can calculate the coefficient of variation for σ^c, which is $\hat{\sigma}^c_\chi \times$ stdev$(\hat{\chi})/\hat{\sigma}^c = 0.037(0.22/0.293) = 2.3\%$. Hence, regional specialization in tradable commodities offsets about $2.3/28 = 8\%$ of the initial differences in the scale of the labor market by making dense areas more heterogeneous. Implication 3 yields an estimate for $\underline{p}^{-1}_{\sigma^c\sigma^c}$

$$\begin{aligned} 0.037 &= -\hat{\sigma}^c_{\hat{\chi}} = -\frac{\xi}{\chi_{\hat{\chi}}}\sigma^c_\chi \cong \frac{\xi}{\chi_{\hat{\chi}}}\frac{7}{20}\underline{p}^{-1}_{\sigma\sigma} \Rightarrow \underline{p}^{-1}_{\sigma\sigma} \cong 3 \\ \underline{p}^{-1}_{\sigma^c\sigma^c} &= 0.037\left(\frac{6.7}{0.25}\frac{20}{7}\right) \cong 3 \end{aligned}$$

This number is a factor seventeen lower than the number $\underline{p}^{-1}_{\sigma\sigma} \cong 49$ which comes from (29) in Section 6.3 as the value for which the initial advantage of large regions in search frictions is fully offset by their specialization in the production of search intensive tradables. We can see this directly, from (29). Substituting $\underline{p}^{-1}_{\sigma\sigma} \cong 3$ in (29) yields:

$$\frac{dx(0)}{d\chi} = 1 + \frac{4}{25}\left(\frac{1}{2} - \frac{24+11}{5}\frac{1}{16}3\right)0.3 = 0.96$$

Hence, the comparative advantage of large areas in search is absorbed for 4% by a specialization in the production of search intensive tradables.

Finally, we use the estimates for CMSA's only to compare $l_{\hat{\chi}}$ and $w_{\hat{\chi}}$ in the first column of Table 8 with the theoretically predicted values that attribute all cost of living and wage differences to differences in search frictions. This gives:

$$\begin{aligned} 0.27 &= -l_{\hat{\chi}} \overset{?}{=} -\frac{l_\chi}{\chi_{\hat{\chi}}} = \frac{1.53}{6.7} \cong 0.23 \\ 0.29 &= -w_{\hat{\chi}} \overset{?}{=} -\frac{w_\chi}{\chi_{\hat{\chi}}} = \frac{1.20}{6.7} \cong 0.18 \end{aligned}$$

Hence, search frictions account for $62 - 85\%$ of the cross regional earnings differentials.

[21] $0.030 = \Delta\hat{\mu}_{\hat{\chi}} = \frac{\xi}{\chi_{\hat{\chi}}}\Delta\mu_\chi \cong \frac{\xi}{\chi_{\hat{\chi}}}\frac{4}{5}$

[22]

$$\text{stdev}(\chi) = \frac{\text{stdev}(\hat{\chi})}{\chi_{\hat{\chi}}} \Rightarrow \text{coef.of var}(\chi) \equiv \frac{\text{stdev}(\chi)}{\text{E}(\chi)} = \frac{\text{stdev}(\hat{\chi})}{\chi_{\hat{\chi}}\text{E}(\chi)}$$

To sum up, the empirical results show that the advantages that dense areas have over non-dense areas in terms of low search frictions are exploited in equilibrium by labor mobility and trade. Dense areas produce more complex and or diverse (in terms of task inputs) goods and demand therefore a wider variety of worker skill levels. Consequently, cities are more heterogeneous than the country side. The effects are substantial from a quantitative point of view. However, the degree of specialization is insufficient to fully offset the initial advantage of large regions in search efficiency.

8 Final remarks

We have specified a search model with ex ante heterogeneous workers and jobs and IRS in the contact technology. The model is made analytically tractable by applying Taylor expansions. In previous work, Teulings and Gautier (2001), we were able to characterize the "cost of search" (the relative difference between the maximum value added and the reservation wage) for each worker type. In this paper, we also managed to characterize the general equilibrium effect of search frictions. That is, how does the maximum value added for a particular worker type change due to frictions. It turns out that search frictions give rise to both *skill compression* (reducing the dispersion of the effective skill distribution) and *skill spoiling* (reducing its mean).

When there is interregional variation in scale, interregional labor mobility and trade cause specialization of regions. Labor mobility offsets the general equilibrium effects of search. In particular, given the task distribution, small scale regions with large frictions have on average a better skilled work force. This is an equilibrium response to offset the skill spoiling effect. Empirical evidence corroborates this surprising implication. Furthermore, large scale regions use their comparative advantage in search to specialize in search intensive production, that is, the production of commodities with a highly dispersed distribution of inputs. This implication is also supported by the data. Our evidence suggest that search frictions can explain about one third of the interregional wage differentials.

An important objection one could raise against our analysis is that most of the empirical evidence suggests CRS instead of IRS in matching technology, see Petrongolo and Pissarides (1999). As discussed in Section 1, there are several reasons why this evidence might be misleading. It does not make sense to repeat these arguments here. However, one argument relates directly to the evidence presented in this paper. This argument claims that large regions specialize in search intensive production, which offsets their initial advantage in search effectiveness. This standard selectivity problem biases any analysis based on aggregate statistics, like the number of unemployed and vacancies. The crude calculations presented in Section 7 suggest however that this mechanism absorbs only a small part of the initial advantage of large scale regions. Hence, this selectivity effect seems to be too small to explain the empirical findings in favor of CRS. We offer two suggestions why this number might be too low an estimate of the true effect.

First, we apply a quadratic contact technology which has extremely large returns to scale. This was done purely for reasons of tractability. Smaller returns to scale would imply that the offsetting effect accounts for a larger share in the total returns. Second, our empirical evidence suggests that large scale regions specialize in the production of commodities, not only with greater dispersion of the complexity distribution, but also with a higher mean. In fact, the effect on the mean is much stronger than the effect on the dispersion. As it stands, our model does not offer an explanation for that phenomenon. A story that is consistent with the framework of our model is that the *mismatch coefficient* is increasing in the skill level (we assumed it to be constant). In other words, when an engineer is sub-optimally matched, the output loss is larger than when a high school drop-out is imperfectly matched. Then, large scale regions would have a comparative advantage in the production of commodities with a high mean complexity level. This specialization pattern implies that the average mismatch coefficient is pushed up in large regions, undoing part of the initial advantage in the efficiency of the search process.

9 References

Acemoglu D. and J. Angrist (2000) How large are human capital externalities? Evidence from compulsory schooling laws, Mimeo MIT.

Arrow K.J. (1962) The economic implications of learning by doing, *Review of Economic Studies*, 29, 155-73.

Burdett K. and D. Mortensen (1998) Equilibrium wage differentials and employer size, *International Economic Review*, 39, 257-274.

Ciccone A. and R.E. Hall (1996) Productivity and the density of economic activity, *American Economic Review*, 86, 54-70.

Diamond C. and C. Simon (1990) Industrial specialization and the returns to labor, *Journal of Labor Economics*, 8, 175-201

Diamond P.A. (1981) Mobility costs, frictional unemployment, and efficiency, *Journal of Political Economy*, 89, 798-812.

Diamond P.A. (1982a) Wage determination and efficiency in search equilibrium, *Review of Economic Studies*, 49, 217-227.

Diamond P.A. (1982b) Aggregate demand management in search equilibrium, *Journal of Political Economy*, 90, 881-894

Dumais G., G. Ellison E.L. Glaeser (1997) Geographic concentration as a dynamic process, Mimeo Harvard university.

Dumond, M., B.T. Hirsch and . MacPherson (1999), Wage differentials across labor markets and workers: does cost of living matter? *Economic Inquiry*, 37, 577-598.

Ellison G., E.L. Glaeser (1997) Geographic concentration in U.S. manufacturing industries: A dartboard approach, *Journal of Political Economy,* 105, 889-926.

Fujita M., P. Krugman, A.J. Venables (1999) *The Spatial Economy*, MIT Press, Cambridge MA.

Gautier P.A. and C.N. Teulings (2003) An empirical measure for labor market density, *Review of Economics and Statistics*, Forthcoming.

Gautier P.A. and C.N. Teulings (2002) How large are search frictions?, Mimeo, Tinbergen Institute.

Glaeser E.L. (1999) Learning in cities, *Journal of Urban Economics*, 46, 254-277.

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

Glaeser E.L. and D.C. Mare.(2001) Cities and skills, *Journal of Labor Economics*, 19, 316-342.

Henderson, V., A. Kuncoro, and M. Turner (1995) Industrial development in cities, *Journal of Political Economy*, 103, 1067-1090.

Jacobs J. (1968) *The Economy of cities*, New York, Vintage Books.

Katz, L.F. and K.M. Murphy (1992), Changes in relative wages, 1963-1987: supply and demand factors, *Quarterly Journal of Economics*, 107, 35-78.

Kim Sunwoong (1989) Labor specialization and the extent of the market, *Journal of Political Economy,* 97, 692-705.

Krugman, P. (1991) *Geography and trade,* Cambridge MA: MIT Press.

Lucas R.E. Jr. (1988) On the mechanics of economic development, *Journal of Monetary Economics,* 22, 3-42.

Lucas R.E. Jr. (2001), Externalities and Cities, *Review of Economic Dynamics*, 2001, vol. 4, 245-274.

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

Petrongolo B. and C.A. Pissarides (1999) Looking into the blackbox: A survey of the matching function, working paper LSE.

Pissarides C.A. (1990) *Equilibrium unemployment theory*, Basil Blackwell, London.

PORTELA, M. (2001) Measuring skill: a multi-dimensional index *Economics Letters*, 72, 27-32.

RAUCH J.E. (1993) Productivity gains from geographic concentration of human capital: evidence from the cities, *Journal of Urban Economics* 34, 380-400.

ROMER, P. (1990) Increasing returns and long run growth *Journal of Political Economy,* 94, 1002-37.

ROSEN, S. (1974) Hedonic prices and implicit markets: product differentiation in pure competition", *Journal of Political Economy* , 82, 34-55.

SATTINGER M. (1975) Comparative advantage and the distribution of earnings and abilities, *Econometrica*, 43, 455-68.

SATTINGER M. (1995) Search and the efficient assignment of workers to jobs, *International Economic Review*, 36, 283-302.

SHIMER R.(2001) The impact of young workers on the aggregate labor market, *Quarterly Journal of Economics*, 116, 969-1007.

SHIMER R. AND L. SMITH (2000) Assortative matching and search, *Econometrica*, 68, 343-370.

TEULINGS C.N. (1995) The wage distribution in a model of the assignment of skills to jobs, *Journal of Political Economy*, 280-315.

TEULINGS C.N. (2002) Comparative advantage, relative wages, and the accumulation of human capital, *Tinbergen Institute discussion paper,* 2002-081/3.

TEULINGS C.N. AND P.A. GAUTIER (2003) The right man for the job, *Review of Economic Studies,* forthcoming.

TEULINGS C.N. AND J. VIERRA (1999) Urban versus rural return to human capital in Portugal: A Cook-Book Recipe for applying assignment models, *Tinbergen Discussion Paper*, 98-095/3.

TOPEL R. (1991) Specific capital, mobility, and wages: Why wages rise with job seniority, *Journal of Political Economy*, 145-176.

Appendix

A Derivations

A.1 The full wage equation

The equivalent of wage equation (6) for the case $\Delta\sigma \neq 0$ reads, see Teulings (2002):

$$\begin{aligned} y(s) &= \underline{\xi} e^{-\Delta\mu} \left[(s - \Delta\mu - \mu^c) - \frac{1}{2} \frac{\Delta\sigma}{\sigma^c} (s - \Delta\mu - \mu^c)^2 \right] \\ &\quad - \underline{\xi}^2 e^{-2\Delta\mu} \sigma^c \Delta\sigma (s - \Delta\mu - \mu^c) - \frac{1}{2} \underline{\xi}^2 \left(e^{-2\Delta\mu} - 1 \right) \sigma^{c^2} + \underline{\xi} (1 + \mu^c - e^{-\Delta\mu}) \\ &\quad + \frac{1}{2} \underline{\xi} e^{-\Delta\mu} \sigma^c \left(1 + 3\underline{\xi}^2 e^{-2\Delta\mu} \sigma^{c^2} \right) \Delta\sigma + O\left[\Delta\sigma^2\right] \end{aligned}$$

A.2 The mismatch coefficient

Equivalent to the definition of log value added $y^0(s,c)$ of worker type s in job type c in equation (2), we can define the log cost per unit of output of employing an s-type worker in a c-type job:

$$p^0(s,c) \equiv y(s) - \underline{f}(s,c) \tag{30}$$

Partially differentiating equation (30) with respect to s twice yields:

$$\begin{aligned} p^0_s[s,c(s)] &= y_s(s) - \underline{f}_s[s,c(s)] = 0 \\ p^0_{ss}[s,c(s)] &\equiv g(s) = y_{ss}(s) + y_s(s) > 0 \end{aligned} \tag{31}$$

where the first equation is the first order condition for cost minimization by the firm in the Walrasian equilibrium and where we apply the first order condition and hence $\underline{f}_{ss}[s,c(s)] = -\underline{f}_s[s,c(s)] = -y_s(s)$ in the second equation. The inequality in the second equation is the second order condition of the cost minimization problem. Partially differentiating (3) yields:

$$y^0_{cc}[s,c(s)] = p_{cc}[c(s)] - y_s(s)$$

where we apply $\underline{f}_{cc}[s,c(s)] = -y_s(s)$. Totally differentiating (3) and (2):

$$\begin{aligned} \left\{p_{cc}[c(s)] + \underline{f}_{cc}[s,c(s)]\right\} c_s(s) + \underline{f}_{sc}[s,c(s)] &= 0 \\ y_{ss}(s) - \underline{f}_{ss}[s,c(s)] - \underline{f}_{sc}[s,c(s)]\, c_s(s) &= 0 \end{aligned}$$

Substitution, using $\underline{f}_{cc}[s,c(s)] = -\underline{f}_{sc}[s,c(s)] = \underline{f}_{ss}[s,c(s)] = -y_s(s)$ yields:

$$\begin{aligned} g(s) &= y_s(s)\, c_s(s) \\ p_{cc}[c(s)] &= \frac{y_s^2(s)}{y_s(s) + y_{ss}(s)} + y_s(s) \\ y^0_{cc}[s,c(s)] &= \frac{y_s^2(s)}{y_s(s) + y_{ss}(s)} = \frac{y_s^2(s)}{g(s)} = \frac{y_s(s)}{c_s(s)} \end{aligned}$$

For $\Delta\pi = 0$, we have:

$$g(s) = y^0_{cc}[s,c(s)] = \underline{\xi} \tag{32}$$

A.3 Derivation of $x(\cdot)$

Define:

$$x^0(s,c) \equiv p(c) + \underline{f}(s,c) - r(s)$$

Hence:

$$\begin{aligned} x^0(s,c) &= p(c) - p^0(s,c) = y^0(s,c) - r(s) \\ x^0[s,c(s)] &= x(s) \\ x^0_{ss}[s,c(s)] &= -p^0_{ss}[s,c(s)] = g(s) \\ x^0_{cc}[s,c(s)] &= y^0_{cc}[s,c(s)] = \frac{y_s^2(s)}{g(s)} \end{aligned} \tag{33}$$

Consider the matching sets $m_s[c(s)]$ and $m_c(s)$, which are defined by the condition $x^0(s,c) \geq 0$. In a second order Taylor expansions of $x^0(s,c)$ around its maximum at $c(s)$, these matching sets are symmetric intervals around $c(s)$:[23]

$$\begin{aligned} m_s[c(s)] &\cong \{s \in [s - \Delta_s[c(s)], s + \Delta_s[c(s)]]\} \\ m_c(s) &\cong \{c \in [c(s) - \Delta_c(s), c + \Delta_c(s)]\} \end{aligned} \tag{34}$$

where $\Delta_s[c(s)]$ and $\Delta_c(s)$ are the solutions to:

$$\begin{aligned} x(s) &\cong \frac{1}{2}x^0_{ss}[s,c(s)]\,\Delta_s[c(s)]^2 = \frac{1}{2}g(s)\,\Delta_s[c(s)]^2 \\ x(s) &\cong \frac{1}{2}x^0_{cc}[s,c(s)]\,\Delta_c(s)^2 = \frac{1}{2}\frac{y_s^2(s)}{g(s)}\Delta_c(s)^2 \end{aligned} \tag{35}$$

using equation (33) for the second derivatives. Division of both side of the Bellman equations (10) and (11) by $\underline{\beta}\kappa R(s)$ and by $\left(1-\underline{\beta}\right)\kappa P[c(s)]$, using $e^{y^0(s,c)-r(s)} - 1 = e^{x(s)} - 1 \cong x(s)$ and $\frac{R(s)}{P[c(s)]} \cong F[s,c(s)] \cong 1$, and a Taylor expansion of the integrand, Teulings and Gautier (2003), Proposition 2 for details, yields:

$$\begin{aligned} v[c(s)] &\cong \frac{3}{4\sqrt{2}}\frac{1}{\underline{\beta}\kappa}y_s(s)\,g(s)^{-1/2}\,x(s)^{-3/2} \\ u(s) &\cong \frac{3}{4\sqrt{2}}\frac{\underline{\varphi}}{\left(1-\underline{\beta}\right)\kappa}g(s)^{1/2}\,x(s)^{-3/2} \end{aligned} \tag{36}$$

Similarly, a Taylor expansion of steady state flow condition (12) yields:

$$\frac{1}{\sigma^s}\phi\left(\frac{s-\mu^s}{\sigma^s}\right) \cong 2\sqrt{2}y_s(s)^{-1}\,g(s)^{1/2}\,\kappa u(s)\,v[c(s)]\,x(s)^{1/2} \tag{37}$$

We can ignore the term $u(s)$ on the left hand side, since in a Taylor expansion starting from the Walrasian equilibrium, this term is of higher order relative to total labor supply. Using equation (36) to eliminate $u(s)$ and $v[c(s)]$ from equation (37) yields:

$$x(s)^{5/2} \cong \frac{9}{8\sqrt{2}}\frac{\underline{\varphi}}{\left(1-\underline{\beta}\right)\underline{\beta}\kappa}\sigma^s\phi\left(\frac{s-\mu^s}{\sigma^s}\right)^{-1} g(s)^{1/2} \tag{38}$$

$$\begin{aligned} x(s,\Delta\pi,\pi^c) &\cong \chi\left[\frac{\Delta\sigma-\sigma^c}{\underline{\sigma}}\phi(0)\,\phi\left(\frac{s+\Delta\mu-\mu^c}{\Delta\sigma-\sigma^c}\right)^{-1}\sqrt{\frac{g(s,\Delta\pi,\pi^c)}{\underline{\xi}}}\right]^{2/5} \\ x(s) &\cong \chi\left[\frac{\sigma^s}{\underline{\sigma}}\phi(0)\,\phi\left(\frac{s-\mu^s}{\sigma^s}\right)^{-1}\sqrt{\frac{g(s)}{\underline{\xi}}}\right]^{2/5} \\ \chi &\equiv \underline{\chi}_0\kappa^{-2/5} \end{aligned} \tag{39}$$

where $\chi_0 \equiv \frac{9}{8\sqrt{2}}\frac{\underline{\varphi}}{\underline{\beta}\left(1-\underline{\beta}\right)}\underline{\sigma}\phi(0)^{-1}\sqrt{\underline{\xi}}$

[23] By definition, $c(s)$ maximizes $x^0(s,c)$ for a given s. Implicitly, we assume that s maximizes $x^0(s,c)$ for $c = c(s)$, so that $x^0_s[s,c(s)] = x^0_c[s,c(s)] = 0$. One can proof that this is true in Walrasian equilibrium, but it need no longer to be true in a search equilibrium. However, Teulings and Gautier (2003) proof that this is a higher order phenomenon, see their Proposition XX.

A.4 The derivation of equation (17)

The derivation of equation (17) is based on higher order expansion of equation (13) and (12). We consider Taylor expansions starting from the Walrasian equilibrium in the point $\Delta\pi = 0$, so that $\pi^s = \pi^c$ Hence:

$$g(s) = y_s(s) = \underline{\xi}$$

Consider the integral on the right hand side of equation (13):

$$\begin{aligned}
&\int_{m_s(c)} u(s)\, \underline{F}(s,c)\, ds \qquad (40)\\
\cong\ & \frac{3}{4\sqrt{2}} \frac{\varphi\sqrt{\underline{\xi}}}{\left(1-\underline{\beta}\right)\kappa} \int_{m_s(c)} x(s)^{-3/2}\, \underline{F}(s,c)\, ds\\
\cong\ & \frac{3}{4\sqrt{2}} \frac{\varphi\sqrt{\underline{\xi}}}{\left(1-\underline{\beta}\right)\kappa} \int_{-\Delta_s(c)}^{\Delta_s(c)} H(z,c)\, dz\\
\cong\ & \frac{3}{4\sqrt{2}} \frac{\varphi\sqrt{\underline{\xi}}}{\left(1-\underline{\beta}\right)\kappa} \int_{-\Delta_s(c)}^{\Delta_s(c)} H(0,c) \left(1 + h_z(0,c)\, z + \frac{1}{2} h_{zz}(0,c)\, z^2\right) dz\\
\cong\ & \frac{3}{2} \frac{\varphi}{\left(1-\underline{\beta}\right)\kappa} x\left[s(c)\right]^{-1} \underline{F}\left[s(c), c\right] \left(1 + \frac{1}{3}\underline{\xi}^{-1} x\left[s(c)\right] h_{zz}\left[s(c)\right]\right)
\end{aligned}$$

where $s(c)$ is the inverse of $c(s)$ and where:

$$\begin{aligned}
H(z,c) &\equiv x\left[z + s(c)\right]^{-3/2} \underline{F}\left[z + s(c), c\right]\\
h_z\left[s(c)\right] &\equiv H_z(0,c) / H(0,c)\\
h_{zz}\left[s(c)\right] &\equiv H_{zz}(0,c) / H(0,c)
\end{aligned}$$

The first step in (40) applies equation (36), the second uses the approximation of the domain of integration, see (34), defining $z \equiv s - s(c)$, the third step applies a Taylor expansion to $H(z,c)$ around $z = 0$, while the final step applies equation (35) for $\Delta_s(c)$ and the standard formula for integration of a parabola, using $g(s) = \underline{\xi}$. Note that the term with $h_z\left[s(c)\right]$ drops out by symmetry. Since $\underline{f}_s(s,c) = -\underline{f}_{ss}(s,c) = y_s(s) = \underline{\xi}$, we have:

$$\begin{aligned}
h_{zz}(s) &= -\underline{\xi}\left(1-\underline{\xi}\right) - 3\underline{\xi}\frac{x_s}{x} - \frac{3}{2}\frac{x_{ss}}{x} + \frac{15}{4}\frac{{x_s}^2}{x^2} \Rightarrow\\
1 + \frac{1}{3}\underline{\xi}^{-1} x h_{zz}(s) &\cong \exp\left[-\frac{1}{3}\left(1-\underline{\xi}\right)x - x_s - Q\right]\\
Q &\equiv \underline{\xi}^{-1}\left(\frac{1}{2}x_{ss} - \frac{5}{4}\frac{{x_s}^2}{x}\right)
\end{aligned}$$

where we omit the argument of x for the sake of convenience. Hence, omitting the argument of F and x:

$$\begin{aligned}
&\int_{m_s(c)} u(s)\, \underline{F}(s,c)\, ds\\
\cong\ & \frac{3}{2} \frac{\varphi}{\left(1-\underline{\beta}\right)\kappa} x^{-1} \underline{F} \exp\left[-\frac{1}{3}\left(1-\underline{\xi}\right)x - x_s - Q\right]
\end{aligned}$$

Substitution in equation (13) yields:

$$\begin{aligned}
& \frac{1}{\sigma}\phi\left(\frac{c(s)-\mu^c}{\sigma^c}\right)Y^{\#} & (41)\\
= \;& v[c(s)]\left[\kappa\int_{m_s(c)} u(s)\underline{F}(s,c)\,ds-\underline{\varphi}\right]\\
\cong \;& \frac{9}{8\sqrt{2}}\frac{\underline{\varphi}}{(1-\underline{\beta})\underline{\beta}\kappa}x^{-5/2}\left[\underline{F}\exp\left(-\frac{1}{3}(1-\underline{\xi})x-x_s-Q\right)-\frac{2}{3}(1-\underline{\beta})x\right]\\
\cong \;& \frac{9}{8\sqrt{2}}\frac{\underline{\varphi}}{(1-\underline{\beta})\underline{\beta}\kappa}x^{-5/2}\underline{F}\exp\left(-\frac{2}{3}(1-\underline{\beta})x-\frac{1}{3}(1-\underline{\xi})x-x_s-Q\right)
\end{aligned}$$

We substitute $v[c(s)]$ for equation (36) in the third line and where we apply $e^{ax}-bx=e^{(a-b)x}+O[x^2]$ in the last line.[24] Similarly, consider the integral on the right hand side of equation (12):

$$\begin{aligned}
& \int_{m_c(s)} v(c)\,dc \cong \frac{3}{4\sqrt{2}}\frac{\sqrt{\underline{\xi}}}{\underline{\beta}\kappa}\int_{m_c(s)} x[s(c)]^{-3/2}\,dc & (42)\\
\cong \;& \frac{3}{4\sqrt{2}}\frac{\sqrt{\underline{\xi}}c_s}{\underline{\beta}\kappa}\int_{-\Delta_s}^{\Delta_s} x[z+s]^{-3/2}dz\\
\cong \;& \frac{3}{4\sqrt{2}}\frac{\sqrt{\underline{\xi}}c_s}{\underline{\beta}\kappa}\int_{-\Delta_s}^{\Delta_s}\left[x(s)+x_s(s)z+\frac{1}{2}x_{ss}(s)z^2\right]dz\\
\cong \;& \frac{3}{2}\frac{c_s}{\underline{\beta}\kappa x}\exp(-Q)
\end{aligned}$$

where we apply the transformation $z\equiv s(c)-s$ in the second line; $c_s(s)$ is the Jacobian of this transformation, where we drop its argument. Rewriting equation (12) and substitution of the equations (36) and (42) yields:

$$\begin{aligned}
\frac{1}{\sigma}\phi\left(\frac{s-\mu^c}{\sigma^c}\right) &= u(s)\left[\kappa\int_{m_c(s)} v(c)\,dc+1\right] & (43)\\
&\cong \frac{9}{8\sqrt{2}}\frac{\underline{\varphi}\sqrt{\underline{\xi}}c_s}{\underline{\beta}(1-\underline{\beta})\kappa}x^{-5/2}\left[\exp(-Q)+\frac{2}{3}\underline{\beta}x\right]\\
&\cong \frac{9}{8\sqrt{2}}\frac{\underline{\varphi}\sqrt{\underline{\xi}}c_s}{\underline{\beta}(1-\underline{\beta})\kappa}x^{-5/2}\exp\left(\frac{2}{3}\underline{\beta}x-Q\right)
\end{aligned}$$

since $\pi^s=\pi^c$ and where we apply again $e^{ax}-bx=e^{(a-b)x}+O[x^2]$. Division of equation (41) by equation (43), rearranging terms and taking logs yields:

$$-\frac{1}{2}\left(\frac{s-\mu^c}{\sigma^c}\right)^2+\underline{f}[s,c(s)]-\underline{\omega}x(s)-x_s(s)\cong-\frac{1}{2}\left(\frac{c(s)-\mu^c}{\sigma^c}\right)^2+y^{\#}+\ln c_s(s) \qquad (44)$$

where $\underline{\omega}\equiv 1-\frac{1}{3}\underline{\xi}$.[25]

[24] Furthermore, we use $\underline{F}=1+O[x]$ for $\Delta\pi=0$, so that $e^{ax}-b\underline{F}x=e^{(a-b)x}+O[x]$.

[25] For the sake of transparency, we simplify the one but least term on the left hand side, $\underline{\omega}x(s)$ by setting $\underline{\omega}\equiv 1-\frac{1}{3}\underline{\xi}\cong 1$. The term $-\frac{1}{3}\underline{\xi}$ captures a small offsetting effect: task type $c(s)$ is done by both better and lower

A.5 The GE effect of search on $y(s)$

Equation (44) holds identically for all s. Hence, its derivatives with respect to s have to apply. We evaluate these derivatives for $\Delta\pi = 0$ at $s = \mu^c$:

$$\begin{aligned} -\underline{\xi}\left(c_s - 1\right) - x_{ss} &\cong -\sigma^{c^{-2}}\left(c - \mu^c\right) c_s \\ -\underline{\xi}\left(c_s - 1\right)^2 - \sigma^{c^{-2}} - \underline{\omega} x_{ss} &\cong -\sigma^{c^{-2}} c_s^2 \end{aligned} \tag{45}$$

leaving out the argument of c_s and c and using $x_s = x_{sss} = c_{ss} = 0$. In the Walrasian equilibrium with $\Delta\pi = 0$, $c(s) = s$. Hence:

$$\begin{aligned} c - \mu &= O[x] \\ c_s - 1 &= O[x] \Rightarrow \\ c_s^2 - 1 &= 2\left(c_s - 1\right) + O\left[x^2\right] \\ \left(c_s - 1\right)^2 &= O\left[x^2\right] \end{aligned}$$

Substitution of these relations and dropping terms of order $O\left[x^2\right]$, and substitution of $x_{ss} = \frac{2}{5}\sigma^{c^{-2}}\chi$ in equation (45), and solving for y_s and y_{ss} yields:

$$\begin{aligned} y_s &\cong \underline{\xi} + \frac{1}{5}\left(\underline{\omega}\underline{\xi}^2\sigma^{c^2} + 2\underline{\xi}\right)\chi \\ y_{ss} &\cong \frac{1}{5}\underline{\omega}\underline{\xi}\chi \end{aligned} \tag{46}$$

B Tables and Figures

Table 1: Cross regional WLS estimates of key variables on labor market density

mean $(\hat{\chi})$	0.635
stdev $(\hat{\chi})$	0.210
$\hat{\mu}^s_{\hat{\chi}}$	$-0.058(3.17)$
$\hat{\mu}^c_{\hat{\chi}}$	$-0.078(4.25)$
$\hat{\sigma}^s_{\hat{\chi}}$	$-0.018(2.37)$
$\hat{\sigma}^c_{\hat{\chi}}$	$-0.011(1.89)$
$w_{\hat{\chi}}$	$-0.390(11.10)$

Note: t-values in brackets. The wage regression includes a full set of controls and t-values are adjusted for heteroskedasticity. The other regressions are weighted by the inverse of the standard error of the lhs variable.

skilled workers than type s. Due to the log linear structure, the extra productivity of the better skilled worker carries a greater weight than the lower productivity of the less skilled worker, yielding a net output gain. For realistic values of $\underline{\xi}$, this effect can be safely ignored.

Table 2: Derivatives of y, y_s, y_{ss} at $\Delta\pi = s = 0$ and $\pi^c = \underline{\pi}$

	$\Delta\mu$	$\Delta\sigma$
y	$\underline{\Sigma}^2$	$\frac{1}{2}\underline{\Sigma}\left(1+3\underline{\Sigma}^2\right)$
y_s	$-\underline{\xi}$	$-\underline{\xi}\underline{\Sigma}$
y_{ss}	0	$-\underline{\xi}^2\underline{\Sigma}^{-1}$
y_{μ^c}	$\underline{\xi}$	$\underline{\xi}\underline{\Sigma}$
y_{σ^c}	$2\underline{\xi}\underline{\Sigma}$	$\frac{1}{2}\underline{\xi}(1+9\underline{\Sigma}^2)$

Table 3: Derivatives of $x_\chi(s)$ at $\Delta\pi = s = 0$ and $\pi^c = \underline{\pi}$

	level	μ^s	σ^s	$\Delta\mu$	$\Delta\sigma$	s	ss
x_χ	1	0	$\frac{2}{5}\underline{\xi}\underline{\Sigma}^{-1}$	$\frac{1}{5}\left(2\underline{\xi}-1\right)$	$-\frac{1}{5}\underline{\Sigma}\left(1+\underline{\xi}\underline{\Sigma}^{-2}\right)$	0	$\frac{2}{5}\underline{\xi}^2\underline{\Sigma}^{-2}$

Table 4: Direct and general equilibrium effects of search frictions in a single region at $\Delta\pi = s = 0$ and $\pi^c = \underline{\pi}$

Function	value
y_χ	$-\frac{2}{5}\underline{\Sigma}^2 - \frac{1}{10}\underline{\xi}^{-1}\underline{\Sigma}^2\left(1+3\underline{\Sigma}^2\right)$
$y_{s\chi}$	$\frac{1}{5}\left(2\underline{\xi}+\underline{\Sigma}^2\right)$
$y_{ss\chi}$	$\frac{1}{5}\underline{\xi}$
x_χ	1
$x_{s\chi}$	0
$x_{ss\chi}$	$\frac{2}{5}\underline{\xi}^2\underline{\Sigma}^{-2}$
$y_{\mu^c\chi}$	$-\frac{2}{5}\left(\underline{\xi}+\underline{\Sigma}^2\right)$
$y_{\sigma^c\chi}$	$-\frac{1}{5}\underline{\Sigma}\left(1+6\underline{\Sigma}^2+4\underline{\xi}\right)$

Table 5: The equilibrium with labour mobility

$\Delta\mu_\chi$	$\Delta\sigma_\chi$	l_χ	w_χ
$\frac{4}{5}$	$\frac{1}{5}\left(\underline{\xi}^{-1}\underline{\Sigma}-2\underline{\Sigma}^{-1}\right)$	$-\frac{6}{5}-\underline{\Sigma}^2$	$-\frac{6}{5}-\underline{\Sigma}^2+\frac{2}{3}\underline{\beta}$

Table 6: Benchmark parameter values

$\underline{\xi}$	$\underline{\Sigma}$	$\underline{\beta}$	$\mathrm{E}[\chi]$	$\chi_{\hat{\chi}}$	$\mathrm{cv}[\chi]$	$\mathrm{cv}[\kappa]$	$\mathrm{cv}[\sigma^c]$	$\lvert p_{\sigma^c\sigma^c}\rvert$
0.25	0.50	0.50	0.30	6.7	0.11	0.28	0.023	3

Note:cv stands for coeffcient of variation.

Table 7: Cross regional WLS estimates of key variables on labor market density

Sample	unrestricted	age $<$ 30 yrs	unrestricted
# regions	82 (min obs=800)	62 (min obs=400)	62 (same as age<30)
mean ($\hat{\chi}$)	0.635	0.637	0.637
stdev ($\hat{\chi}$)	0.210	0.219	0.219
mean $\hat{\sigma}^c$	0.353	0.293	0.353
mean $\Delta\hat{\sigma}$	0.020	0.019	0.019
$\hat{\mu}^s_{\hat{\chi}}$	-0.058 (3.17)	-0.089 (5.23)	-0.043 (2.23)
$\hat{\mu}^c_{\hat{\chi}}$	-0.078 (4.25)	-0.119 (6.79)	-0.072 (3.50)
$\hat{\sigma}^s_{\hat{\chi}}$	-0.018 (2.37)	-0.025 (2.95)	-0.019 (2.38)
$\hat{\sigma}^c_{\hat{\chi}}$	-0.011 (1.89)	-0.037 (4.08)	-0.011 (1.65)
$\Delta\hat{\mu}_{\hat{\chi}}$	0.020 (1.84)	0.030 (2.67)	0.029 (2.74)
$\Delta\hat{\sigma}_{\hat{\chi}}$	-0.006 (0.93)	-0.013 (1.85)	-0.008 (1.21)
$\hat{w}_{\hat{\chi}}$	-0.370 (8.90)	-0.350 (7.08)	0.372 (7.89)

Note: T-values within brackets. The regressions are weighted by the inverse of the standard error of the lhs variable , $\hat{\mu}^s$ and $\hat{\mu}^s$ are 0.

Table 8: Cross regional WLS estimates of key variables on labor market density continued

Sample	cost of living areas	age $<$ 30 yrs, (C)MSA
# regions	67 (min obs=500)	50 (min obs = 250)
mean ($\hat{\chi}$)	0.617	0.612
stdev ($\hat{\chi}$)	0.187	0.198
mean $\hat{\sigma}^c$	0.354	0.294
mean $\Delta\hat{\sigma}$	0.020	0.018
$\hat{\mu}^s_{\hat{\chi}}$	-0.056 (2.28)	-0.072 (2.96)
$\hat{\mu}^c_{\hat{\chi}}$	-0.088 (3.48)	-0.102 (4.43)
$\hat{\sigma}^s_{\hat{\chi}}$	-0.018 (1.97)	-0.033 (3.12)
$\hat{\sigma}^c_{\hat{\chi}}$	-0.018 (2.52)	-0.034 (3.18)
$\Delta\hat{\mu}_{\hat{\chi}}$	0.032 (2.12)	0.031 (2.02)
$\Delta\hat{\sigma}_{\hat{\chi}}$	-0.002 (0.02)	-0.002 (0.15)
$\hat{l}_{\hat{\chi}}$	-0.266 (3.69)	-0.300 (3.59)
$\hat{w}_{\hat{\chi}}$	-0.288 (7.09)	-0.310 (5.91)

Note: T-values within brackets.The regressions are weighted by the inverse of the standard error of the lhs variable, $\hat{\mu}^s$ and $\hat{\mu}^s$ are 0 by construction.

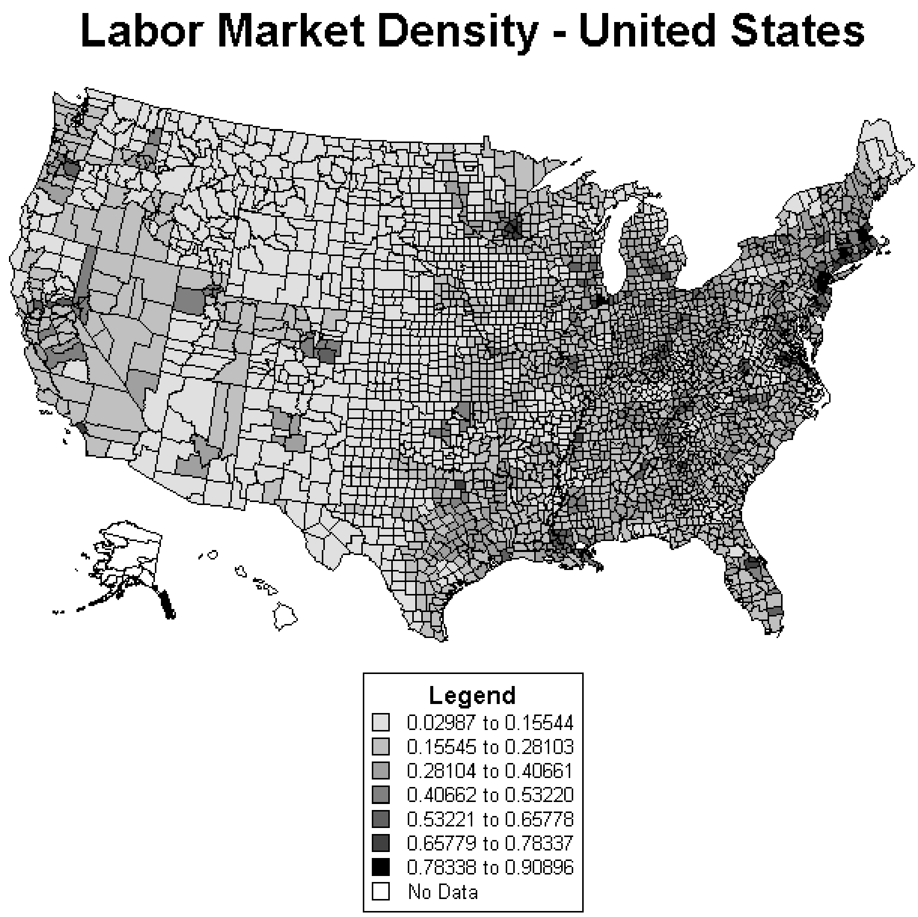
Labor Market Density - United States
Legend
0.02987 to 0.15544
0.15545 to 0.28103
0.28104 to 0.40661
0.40662 to 0.53220
0.53221 to 0.65778
0.65779 to 0.78337
0.78338 to 0.90896
No Data

Figure 1:

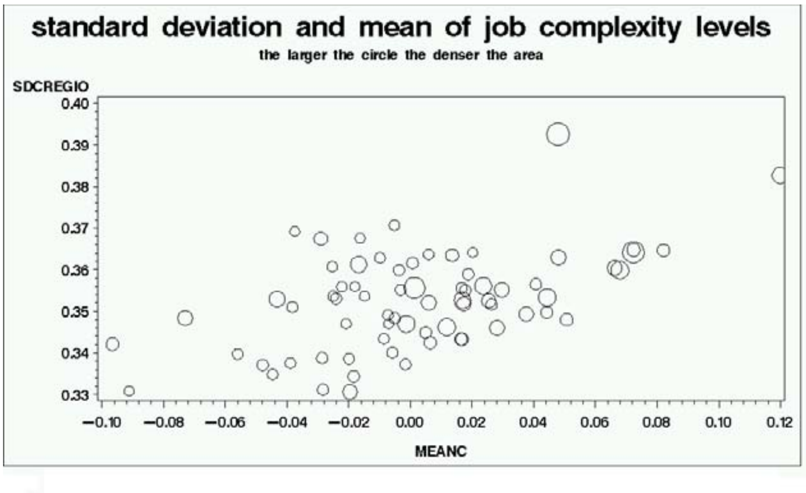

Figure 2:

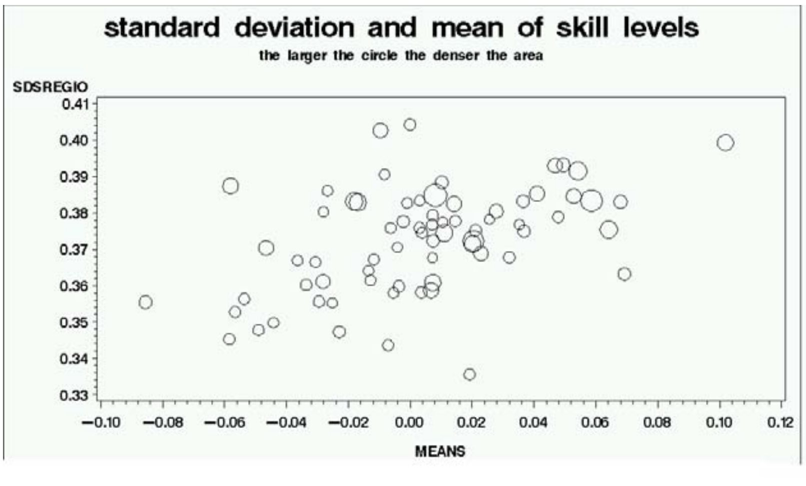

Figure 3:

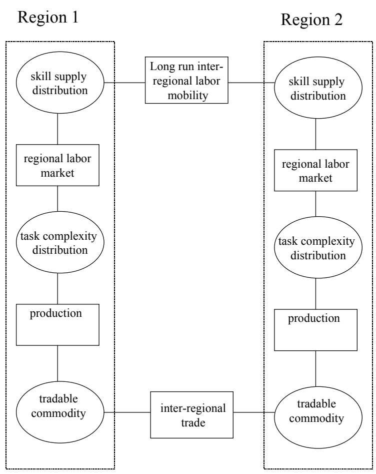

Figure 4:

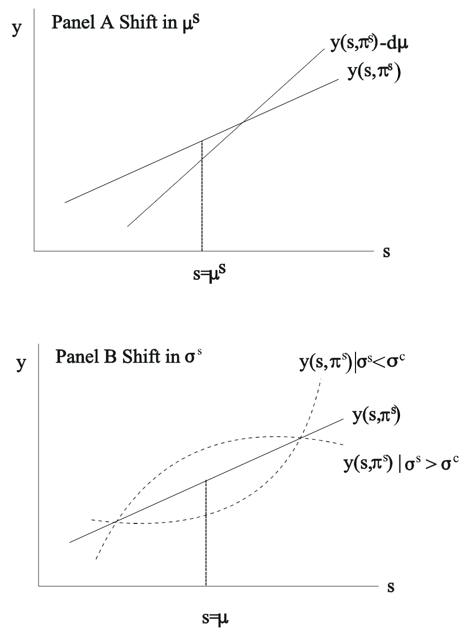

Figure 5: Shifts in μ^s and σ^s

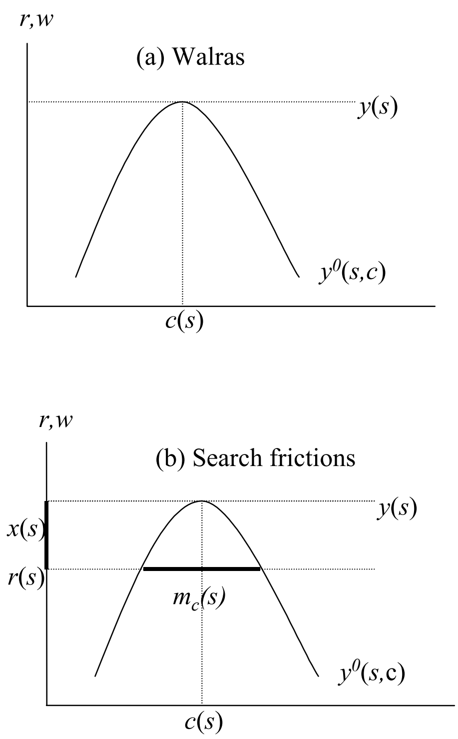

Figure 6: Walras versus search frictions

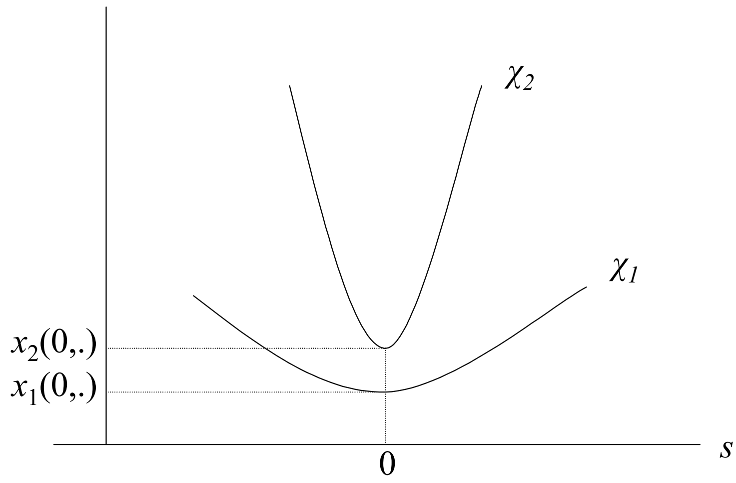

Figure 7: Search frictions in dense and non dense areas

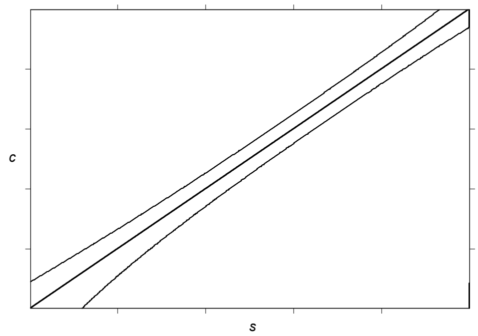

Figure 8: Matching sets: the aggregate outcome

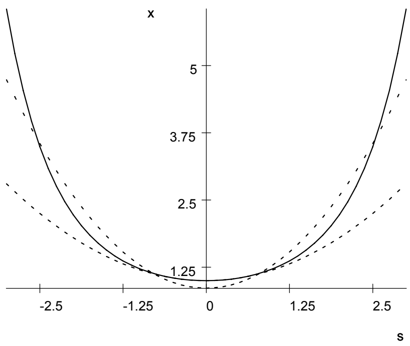

Figure 9: The function $x(s, \cdot)$: true value, 2^e order Taylor expansion, means square approximation

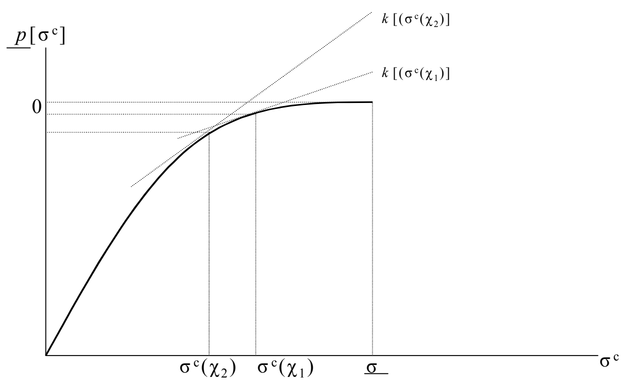

Figure 10: Equilibrium price formation of the tradable