TI 2002-061/3
Tinbergen Institute Discussion Paper

Search and the City

C.N. Teulings
P.A. Gautier

Faculty of Economics, Erasmus University Rotterdam, and Tinbergen Institute

Tinbergen Institute Amsterdam
Keizersgracht 482
1017 EG Amsterdam
The Netherlands
Tel.: +31.(0)20.5513500
Fax: +31.(0)20.5513555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31.(0)10.4088900
Fax: +31.(0)10.4089031

Search and the City*

C.N. Teulings and P.A. Gautier†

June 14, 2002

Abstract

Can increasing returns to scale in search explain regional differentiation between cities and rural areas? To answer this question, we develop a model of an economy that consists of several regions. Within each region, jobs and workers are heterogeneous by respectively skill and job complexity type. Because of the search frictions, firms and workers in each region must trade-off a better expected match quality against a longer period of non-production. Labor mobility between regions induces the equalization of reservation wages for each skill type and interregional trade of end products yields regional specialization in production. The model predicts that high density areas make use of their scale advantage by producing end products with a high dispersion of skill requirements. Empirical evidence for the United States corroborates the implications of the model.

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

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

1 Introduction

The question why people live in cities, where housing prices are typically higher than in rural areas, has received a lot of attention in the literature. All explanations require some form of increasing returns to scale (IRS). Well known examples are: Marshall (1890), Arrow (1962), Romer (1988), Lucas (1988, 1999), Glaeser et al. (1992). Those models imply that people have incentives to cluster together because the diffusion of knowledge is higher in densely populated areas. IRS can also be caused by urbanization spillovers as in Jacobs (1969) who argues that the most important knowledge transfers are between industry rather than within industry spillovers.[1] Finally, Krugman (1991), Ciccone and Hall (1996) and Fujita et al. (1999) emphasize transportation costs in combination with "love for variety demand" to explain why people prefer to locate close to each other.

We develop a model of clustering based on IRS in job search. Under IRS in search, the available set of jobs is larger in cities than in the country side. The high number of potential trading partners increases the amount of contacts per unit of time. A higher contact rate (all else equal) shortens unemployment durations and improves the expected match quality. This works as a centripetal force.[2] However, the stock of real estate in a city is scarce and this puts a limit on the number of people that can move into the city. The real estate owners capture the rents of the advantages that cities offer. Hence, the cost of living is higher in densely populated (large scale) areas and it therefore works as a centrifugal force. Our model generates sharp predictions about the composition of workers and jobs in cities relative to the country side. Because cities have a comparative advantage in search intensive activities, we expect industries that require a large mix of job and worker inputs to move to cities. This is consistent with the empirical observation that cities are places where there are concentrations of both very high and very low skilled workers. In some respects, our model is similar to the knowledge spillover model. In both approaches, productivity is increasing in the number of interactions between agents. The aim of this paper is to explicitly model these interactions and to provide empirical evidence supporting the relevance of the mechanisms described. As a rough estimate, we find that search frictions can explain about one third of the wage differentials between high and low density areas.

The model that we propose is an extension of the Shimer and Smith (2000) and Teulings

[1] The empirical evidence for knowledge spillovers in cities is limited. Henderson et al.(1995) finds some evidence. Glaeser et al. (1992) find more evidence for between industry spillovers than for within industry spillovers. Finally, Acemoglu and Angrist (2000) use variation in child labor laws and compulsory attendance laws to investigate whether there exists a causal relation between average schooling and state wage levels. They find the external returns to schooling to be statistical insignificant.

[2] In our model we focus on the labor market but search frictions on the marriage or the goods market are also lower in cities, so similar arguments can be made for those markets.

and Gautier (2000) matching models with a continuum of worker and job types. These models extend the hedonic pricing and assignment models of Rosen (1974), Sattinger (1975, 1995), and Teulings (1995,2001) with search frictions characterized by IRS. Job seekers and firms trade the cost of prolonged search against the quality of the match. In a frictionless Walrasian market, they end up in the optimal match but in a market with frictions, they are prepared to also accept suboptimal matches in order to save on search cost. The higher the cost of search, the more willing job seekers are to accept suboptimal matches. We study the implications of regional differentiation in labor market density where regions are connected through both labor mobility and commodity trade. Interregional labor mobility equalizes real reservation wages across regions. Trade forces dense areas to specialize in the production of goods which require a large variety of task inputs (workers and job types), and which are therefore search intensive.

Our analysis extends previous work, on the effects of shifts in the skill distribution on equilibrium wage rates in assignment models, see Teulings (2001), and on the cost of search in assignment models when we take value added in the optimal assignment as given, see Teulings and Gautier (2000). The cost of search are defined as the difference between log value added in the optimal assignment (which equals the reservation wage in a frictionless world) and the log reservation wage. Here we combine both papers to show how search frictions not only yield a wedge between this optimal value added and the reservation wage, but also affect the optimal value added itself, by their general equilibrium effect on task prices. There are essentially two effects, *skill compression* and *skill spoiling*. 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 skill distribution relative to the actual skill distribution. The skill spoiling effect is due to the fact that 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. Since we assume that 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. Hence, search frictions yield a drop in the mean of the effective skill distribution. When we allow for interregional labor mobility, these general equilibrium effects of search frictions induce shifts in the regional skill distribution. This yields the prediction that the difference between the means of the skill and task distribution is larger in non-dense areas to offset the skill spoiling effect. Our empirical evidence provides support for this surprising implication. As in Teulings and Gautier

(2000), a full analytical characterization of the equilibrium is not available, so that we have to rely on Taylor expansions of the equilibrium conditions. We provide numerical simulations to show that our expansions offer a reasonable description of the equilibrium.

An obvious objection against our model is that most empirical studies based on aggregate data do not reject CRS in matching, see the overview of Petrongola and Pissarides (2000). In general, returns to scale are difficult to measure because in equilibrium these returns tend to be exploited till the point where they no longer exist. The same applies here. Aggregate matching functions give only limited information on the returns to scale in the contact process for at least two reasons. In Teulings and Gautier (2000) we argue that part of the returns to scale in the contact process are absorbed by a greater selectivity on the side of both job seekers and firms. Here, we show that in equilibrium, dense areas are more heterogeneous because they produce "search intensive" goods. This reduces the probability that a given contact results in a profitable match. We are able to test for the presence of this mechanism by explicitly modeling the endogenous location of job and worker types. On the job search puts further limits on the relevance of this type of evidence for establishing the nature of the matching function. When a large part of search is done while working, unemployment is a bad proxy for the stock of job seekers, compare Shimer's (1999) analysis of the courses of low unemployment in regions with a high inflow of youngsters. Nevertheless, most of the empirical research on returns to scale in matching has focussed solely on the relation between aggregate matches and aggregate stocks of unemployment and vacancies.

In a world with worker and job heterogeneity, there are predictions for two alternative sets of variables. First, since search frictions force workers and firms to accept suboptimal matches, the output loss relative to the optimal match must be shared by both players in some way. As long as workers take some part of this loss, for example by Nash bargaining over wages, we must be able to track the size of this loss in wage data by using measures for match quality. This route is travelled in a companion paper, Gautier and Teulings (2002), where we find strong support for the presence of search frictions. Second, when search frictions differ between regions, they induce interregional labor mobility and specialization in production, yielding predictions for the distributions of skill supply and task demand in regions of various scale. This paper applies the latter type of evidence.

At a more general level, the message of the paper is that for many purposes it is important to explicitly model the heterogeneity of the population. We show that the second moments of the skill and job complexity level distributions have important effects on the structure of the economy. Moreover, by taking heterogeneity into account we show that it is actually possible to bring the matching models of the Diamond (1982), Mortensen (1982) and Pissarides (1990) type to the data and investigate their structural implications.

Related arguments have been put forward by others. Kim Sunwoong (1989) has argued that as the size of the market increases, firms offer a larger variety of job requirements, workers specialize more and the average match quality improves. Glaeser (1999) argued that for labor market pooling to work, workers must be able to change employers without changing residencies. This is typically an advantage that cities have above the country side. Dumais et al. (1997) find that industry location is far more driven by labor mix than by any other explanatory variable and that firms locate near one another so that workers can move from one firm to another in the event of a firm specific downturn. Glaeser and Maré (2001) show 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 real benefit. Rotemberg and Saloner (1991) argue for a different type of spillover derived from the relation between market power and labor market density: multiple firms protect workers against ex post appropriation of investments in human capital.

The plan of the paper is the following. Section 2 gives the basic structure of the model. Section 3 solves the assignment problem for a Walrasian world. In Section 4 we discuss the direct effect of search frictions, while Section 5 discusses their general equilibrium effect. Section 6 presents the effects of interregional trade and labor mobility. Finally, in Section 7 we show that the testable implications of the model are supported by empirical cross regional evidence for the US.

2 The model

2.1 Basic assumptions

The economy considered in this paper is made up of a large number of regions, of which two are shown in Figure 1. Each region produces a single composite tradable commodity. The regions are related to each other by interregional trade in those commodities and by labor mobility. Transportation of commodities and labor mobility are costless, so that real wages for worker types and commodity prices are equalized between regions. Each region is small relative to the economy as a whole, so that it takes economy wide wages and commodity prices as given.

Workers supply a fixed amount of labor and receive no unemployment benefits or utility of leisure. They are risk neutral. The number of workers in a region is determined by the available stock of real estate. This is the quasi-fixed factor in a region. Hence, real estate owners capture any region specific rents via the rental price of their property. We refer to a region with a lot of real estate as a large region. Both tradable commodities and the cost of real estate enter into the workers utility function. This utility function is homothetic, so that all workers 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.

Workers differ by their level of skill, s. They can only produce output when employed by a firm. Each firm performs one specific type of task. The tasks are indexed by their level of complexity c. We consider an infinitum of worker and task types: both s and c can take any real value. The output of a specific task cannot be consumed directly but is used for the production of the region specific tradable commodity. There exists a unique set of task prices that clears task markets within each region. 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.

The production technology of tasks is equal across regions and is such that, within each region, each worker type s has a unique task type where it produces the highest value added. We label the optimal task type $c(s,\cdot)$, where "$\cdot$" denotes other arguments that will be discussed below. Hence, each worker type s would be assigned to this "optimal" task $c(s,\cdot)$ in a hypothetical Walrasian equilibrium. However, regional labor markets are not Walrasian. Due to search frictions, workers do not wait forever till they find their favorite task, $c(s,\cdot)$. Instead, they also accept tasks that yield a lower value added than task $c(s,\cdot)$. The contact technology within a region between vacant tasks and unemployed workers exhibits increasing returns to scale (IRS). In large regions, more contacts take place per unit of time, so job search is more efficient.

The skill distribution within a region is approximately normal:

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

where μ^s is the mean and σ^s is the log standard deviation of the skill distribution. 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: $\pi^s \equiv [\mu^s, \sigma^s]$. At the fundamental level, there is no rational for the normality of the skill distribution, since the distribution in each region is determined endogenously by labor mobility. The normality assumption can be thought of as a second order Taylor expansion of a log density function from a more general family of distributions, with π^s being the parameters of this Taylor expansion.

The region specific tradable commodities are produced from tasks by a Leontieff technology; production does not require other inputs. So, 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. The regional distribution of type c tasks specifies how much of each task is required for the production of one unit of the region specific tradable commodity. This distribution is normal with mean μ^c and standard deviation σ^c. These parameters are therefore counterparts of μ^s and σ^s, but now for the demand side of the regional labor markets. Analogous to $\pi^s \equiv [\mu^s, \sigma^s]$, we define: $\pi^c \equiv [\mu^c, \sigma^c]$. A tradable commodity is therefore fully characterized by the parameters π^c of the complexity distribution of the tasks that are required for its production. Like π^s, π^c is determined endogenously. Each region specializes in that tradable which makes best use of its comparative advantage.

The following example clarifies this. Consider Detroit, a region that produces cars. The production of cars requires the input of many c-type tasks which are used in fixed proportions (wheels are no substitute for the window or the motor). These tasks are on average probably close to the nation wide mean level of complexity. However, producing cars requires both very simple tasks (assemblage) and very complex tasks (design). Hence, 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 . In our data, Boston and Washington DC are examples of regions having both a high μ^c and σ^c, while Houston and Seattle have a high μ^c but a low σ^c. Los Angeles and Miami have a low μ^c and a high σ^c, while in the rural areas of Arkansas both μ^c and σ^c have a low value.

Figures 2 and 3 plot the empirical relation between π^s and π^c and labor market density for the US. [3]. The larger the circle, the denser the area. Those Figures suggest that dense areas produce on average more complex and more dispersed goods and have a more skilled and heterogeneous labor force.

Due to IRS in search, the market equilibrium in an infinitely large region would be equal to the hypothetical Walrasian equilibrium. We treat this hypothetical region as the benchmark. It is useful to define: $\Delta\pi \equiv \pi^s - \pi^c$. Then:

Assumption: for the economy wide benchmark: $\pi^s = \pi^c \equiv \underline{\pi}$ *or:* $\Delta\pi = 0$, *where we normalize:* $\underline{\mu} = 0$

The equality $\underline{\mu}^s = \underline{\mu}^c = 0$ is just a matter of a proper normalization and not restrictive. The equality $\underline{\sigma}^s = \underline{\sigma}^c$ is however an important assumption that greatly simplifies the analysis, as will become clear below. The general equilibrium effects of search frictions will induce shifts in π^s

[3] We postpone the discussion of the measurement of those variables to section 7.

and π^c, due to interregional labor mobility on the one hand and due to regional specialization and interregional trade on the other hand. A nice feature of our model is that labor mobility affects $\Delta\pi$, while specialization affects only π^c, and not $\Delta\pi$. Any shift in π^c due to specialization translates one for one into a shift in π^s, leaving $\Delta\pi$ unchanged. Hence, it is most useful to cast the analysis in $\Delta\pi$ and π^c instead of π^s and π^c, for this allows us to analyze the effects of labor mobility and trade separately.

The discussion proceeds as follows. In Section 3, we discuss the equilibrium assignment in a Walrasian equilibrium, following Teulings (2001). 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 4, we discuss the implications of the introduction of search frictions, while keeping value added in the optimal assignment constant, following Teulings and Gautier (2000). Sections 5 analyzes the general equilibrium effects of search frictions. Search frictions change the effective skill distribution, causing a shift in value added in the optimal assignment of skill type s. This provides the missing link for a complete characterization of the equilibrium assignment problem in a single region with search frictions. This model will then be the starting point for the analysis of interregional labor mobility and trade in Section 6.

3 Assignment of workers to jobs under Walras

3.1 Technology

Let $\underline{p}(\pi^c)$ be the nation wide equilibrium price of a tradable commodity of type π^c; $\underline{p}(\pi^c)$ is a hedonic price index in the sense of Rosen (1974). Since regions are small compared to the economy as a whole, they take $\underline{p}(\pi^c)$ as given. There is positive demand for all tradable commodities. We normalize $\underline{p}(\cdot)$ and its derivatives in the benchmark region to zero without loss of generality:

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

The normalization $\underline{p}(\underline{\pi}) = 0$ is equivalent to using the benchmark commodity as the numeraire, the normalization $\underline{p}_{\pi}(\underline{\pi}) = 0$ is equivalent to a definition of the relative units of measurement of commodities of various types. In Section 5.4, it will be shown that this normalization is the only choice that is consistent with $\pi^c = \underline{\pi}$ being the optimal choice in the benchmark region. For the higher order derivatives we apply a second order polynomial:

Assumption: $\underline{p}(\pi^c) = -\frac{1}{2}(\pi^c - \underline{\pi})' \, \underline{\Psi} \, (\pi^c - \underline{\pi})$

where $\underline{\Psi}$ is a 2×2 symmetric and (semi-) positive definite matrix. Obviously, this specification is consistent with the previous normalization. $\underline{p}(\pi^c)$ can be interpreted as a utility function (transformed to exhibit a constant marginal utility of money), measuring the value of a unit of consumption with characteristics π^c. Then, its matrix of second derivatives, $\underline{\Psi}$, measures the elasticities of substitution between these characteristics.

Let $y^0(s, c, \Delta\pi, \pi^c)$ be log value added of an s-type worker employed at a c-type job. By definition it satisfies:

$$y^0(s, c, \Delta\pi, \pi^c) \equiv p(c, \Delta\pi, \pi^c) + \underline{f}(s, c) \tag{1}$$

where $p(c, \Delta\pi, \pi^c)$ denotes log task prices as function of the skill and complexity distribution and where $\underline{f}(s, c)$ is 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:[4]

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

where $\underline{\xi}$ is a nation wide technology parameter. We return to its interpretation below. 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.

The production of the tradable commodity of a region from tasks is fully characterized by the parameters π^c of the task complexity distribution, reflecting the Leontieff coefficients of this production process.

3.1.1 The Walrasian equilibrium

In equilibrium, the commodity price equals the sum of task prices weighted by their Leontieff input shares:

$$\underline{p}(\pi^c) = \ln\left[\int_{-\infty}^{\infty} \frac{1}{\sigma^c}\phi\left(\frac{c - \mu^c}{\sigma^c}\right)\exp\left[p(c, \cdot)\right] dc\right] \tag{2}$$

where $\phi(\cdot)$ is the standard normal density function. Because of the zero profit condition and since there are no other factors of production than labor, wages are equal to value added. Each worker type is therefore assigned to the task-type where it produces the highest value added.

[4] Teulings and Gautier (2001) 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 $\underline{\xi}$).

Let $c(s,\Delta\pi,\pi^c)$ be the value of c that maximizes $y^0(s,c,\cdot)$ for that s. Hence, assuming $p(c,\cdot)$ to be differentiable, $c(s,\cdot)$ satisfies:

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

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

$$y(s,\Delta\pi,\pi^c) \equiv y^0[s,c(s,\cdot),\cdot]$$

The situation is depicted in y,c-space for a particular s-type in panel A of Figure 5. The curve $y^0(s,c,\cdot)$ denotes log value added of this particular s-type for various c-types. By definition, its maximum is $y(s,\cdot)$, which is attained at $c = c(s,\cdot)$. The difference $y(s,\cdot) - y^0(s,c,\cdot)$ 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 are 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. By the envelope theorem, the first order effect of deviations $c \neq c(s,\cdot)$ vanish. Hence, the second derivative $y_{cc}^0(s,c,\cdot)$ is an appropriate statistic for the sensitivity of output to suboptimal assignment. It will therefore be discussed more extensively later on.

The optimal assignment must clear the market for tasks, in logs:

$$-\ln\sigma^s - \frac{1}{2}\left(\frac{s-\mu^s}{\sigma^s}\right)^2 + \underline{f}[s,c(s,\cdot)] = -\ln\sigma^c - \frac{1}{2}\left(\frac{c(s,\cdot)-\mu^c}{\sigma^c}\right)^2 + y^{\#} + \ln c_s(s,\cdot) \tag{4}$$

where $y^{\#}$ is log output of the composite tradable commodity. The left hand side is the log supply of labor of type s (the normal density function) plus its log productivity in task type $c(s,\cdot)$. The first two terms on the right hand side are the log demand for task type $c(s,\cdot)$ (the log of the Leontieff coefficient plus log output) and the last term is the log Jacobian $\frac{dc}{ds} = c_s(s,\cdot)$.

In general, the differential equations (3) and (4), which determine $p(c,\cdot)$ and $c(s,\cdot)$, do not have an analytical solution. However, the solution is simple in the special case where $\Delta\sigma = 0$:

$$\begin{aligned}
c(s,\cdot) &= s - \Delta\mu \\
p(c,\cdot) &= \underline{p}(\pi^c) + \Pi + \underline{\xi}e^{-\Delta\mu}c \\
y(s,\cdot) &= \underline{p}(\pi^c) + \Pi + \underline{\xi}e^{-\Delta\mu}s + \underline{\xi}\left(1 - e^{-\Delta\mu} - \Delta\mu\right) \\
\Pi &\equiv \underline{\xi}\left(1 - e^{-\Delta\mu}\right)(\Delta\mu + \mu^c) + \frac{1}{2}\underline{\xi}^2\left(1 - e^{-2\Delta\mu}\right)\underline{\sigma}^2
\end{aligned} \tag{5}$$

The expression for Π follows from substitution of the $p(c,\cdot)$ expression in (2). Relative wages 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 2, Panel A. The return to skill $y_s(s,\cdot)$ is inversely related to μ^s. 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 is at $s = \mu^s + \frac{1}{2}\sigma^{s^2}$.

This mechanism provides an interpretation for the parameter $\underline{\xi}$. Consider an increase of the skill level s of all workers by h. At constant relative wages, this raises their wages by $\underline{\xi}h \times 100\%$, see equation (5). However, since s goes up by h for all workers, μ^s will also increase by h. This shift in μ^s changes relative wages. It reduces the return to skill $y_s(s,\cdot)$ by $h \times 100\%$. Hence, $1/\underline{\xi}$ is the *compression elasticity*: 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. Obviously, this change in relative prices is due to imperfect substitution between skill types in the production of the composite commodity. Alternatively, $1/\underline{\xi}$ can therefore be interpreted as a measure of the substitutability between skill types, as in Teulings (2001).

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 the median value of $s = \mu^s$ in the nation-wide benchmark region, $y(0,0,\underline{\pi})$. An approximate solution for the general case is presented in Appendix A.1. In Table 1 we present this solution in the form of a set of derivatives (where $\underline{\psi}_{xx}$ denotes the relevant element of the matrix $\underline{\Psi}$). All derivatives are checked easily from equation (5), except those with respect to $\Delta\sigma$ for which we refer to Appendix A.1. Note that $y_{\pi^c} = 0$: relative wages depend only on $\Delta\pi$, and not on π^c. This offers a first intuition for why trade only affects π^c, and not $\Delta\pi$. The general equilibrium effects of search frictions, which shift task prices $p(c,\cdot)$, are fully offset by changes in $\Delta\pi$. Hence, a change in π^c due to regional specialization leaves task prices unaffected if and only if it leaves $\Delta\pi$ unaffected.

Figure 5, panel B provides an intuition for the derivatives for the case $\Delta\sigma > 0$. An increase in σ^s raises 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 extremes of the wage distribution and rise for intermediate skill levels: $y_{ss\Delta\sigma}(s,\cdot) < 0$. Furthermore: $y_{s\Delta\sigma}(s,\cdot) < 0$. 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 scale. The opposite applies for $\Delta\sigma < 0$.

Equivalent to the definition of log value added $y^0(s,c,\cdot)$ of worker type s in job type c in equation (1), we can define the log cost per unit of output of employing an s-type worker in a c-type job:

$$p^0(s,c,\cdot) \equiv y(s,\cdot) - \underline{f}(s,c) \tag{6}$$

Just like $y^0_{cc}(s,c,\cdot)$ is an appropriate measure for the cost of suboptimal assignment, $c \neq c(s,\cdot)$, from the point of view of the worker, so is $p^0_{ss}(s,c,\cdot)$ an appropriate statistic for the cost of suboptimal assignment, $s \neq s(c,\cdot)$, $s(c,\cdot)$ being the optimal assignment that maximizes $p^0_{ss}(s,c,\cdot)$. Since the optimal assignments $s(c,\cdot)$ and $c(s,\cdot)$ are unique and upward sloping in the Walrasian equilibrium, the one is the inverse of the other: $s[c(s,\cdot),\cdot] = s$.[5] For future reference, we refer to this concept as the *mismatch coefficient*, denoted by $g(s,\Delta\pi,\pi^c)$.[6] The mismatch coefficient will be used as a summary statistic for the cost of suboptimal assignment in our analysis of search frictions. In the Appendix A.3, we proof the following relations:

$$p^0_{ss}[s,c(s,\cdot),\cdot] \equiv g(s,\cdot) = y_{ss}(s,\cdot) + y_s(s,\cdot) \tag{7}$$

$$y^0_{cc}[s,c(s,\cdot),\cdot] = -y_s(s,\cdot)^2\, g(s,\cdot)^{-1} \tag{8}$$

$y^0_{cc}[s,c(s),\cdot]$ is the mirror image of $p^0_{ss}[s,c(s,\cdot),\cdot]$, in that it measures the relative loss in wages for a worker of being employed at a suboptimal firm type, $c \neq c(s)$.We conclude this section with some observations regarding the benchmark equilibrium ($\Delta\pi = 0, \pi^c = \underline{\pi}$), since it is the starting point of our Taylor expansions and since the benchmark determines real reservation wages in all regions, $\underline{r}(s)$. By equation (5) and (7) we have:

$$\begin{aligned}
\underline{r}(s) &= y(s,0,\underline{\pi}) = \underline{\xi}s \\
\underline{f}[s,c(s,0,\underline{\pi}] &= 0 \\
p^0_{ss}[s,c(s,0,\underline{\pi}),0,\underline{\pi}] &\equiv g(s,0,\underline{\pi}) = \underline{\xi} \\
y^0_{cc}[s,c(s,0,\underline{\pi}),0,\underline{\pi}] &= \underline{\xi}
\end{aligned}$$

For $\Delta\sigma = 0$ (as in the benchmark equilibrium), $y(s,\cdot)$ is linear, so that the mismatch coefficient is independent of s. This feature simplifies our analysis considerably. The mismatch coefficient is directly related to the compression elasticity, $^1/_{\underline{\xi}}$. In general, the more difficult it is to substitute between types of workers, the greater the cost of suboptimal assignment, and the more harmful

[5] This feature applies only in the Walrasian equilibrium, see Teulings and Gautier (2001).
[6] This concept is closely related to the complexity dispersion parameter γ, see Teulings (2001).

are search frictions. Finally, the normality of the skill distribution and the linearity $\underline{r}(s)$ imply that the wage distribution is log normal, which is a reasonable description of empirical wage distributions. Its variance is:

$$\text{Var}[w] = \underline{\sigma}^2 \underline{\xi}^2 \equiv \underline{\Sigma}^2 \tag{9}$$

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

4 Search Frictions

4.1 Search technology

When there are search frictions, s-type workers meet only a limited amount of c-type firms per unit of time. After a contact takes place, both have to decide whether to match or continue searching. When they decide to match, the surplus is shared by Nash bargaining as in Pissarides (1990), 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.[7] There is free entry of firms, so that the asset value of a vacancy is zero in equilibrium. Central to our analysis is the notion that there is IRS in the contact technology:

$$M = UV$$

where M is the number of contacts, U is the number of job seekers, and V is the number of vacancies. In rates, the matching function reads:

$$m = uv\Lambda \tag{10}$$

where Λ is the size of the labor force, m is the number of contacts per period of time per unit of the work force so that $M = m\Lambda$, u is the unemployment rate ($U = u\Lambda$) and v is the number of vacancies per unit of the workforce ($V = v\Lambda$).

The contact rates $\lambda_{i \to j}$ for worker (job) type i to run into job (worker) type j that go with equation (10) are:

$$\begin{aligned} \lambda_{s \to c} &\equiv v(c)\Lambda \\ \lambda_{c \to s} &\equiv u(s)\Lambda \end{aligned} \tag{11}$$

[7] 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.

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. Hence: $\int v(c)\,dc = v$ and $\int u(s)\,ds = u$. 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.

4.2 Search equilibrium

Since there are no unemployment benefits or utility of leisure, the instantaneous pay off of unemployment is zero. The current pay off of a firm with a vacancy is 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 $R(s,\cdot)$ be the reservation wage of worker type s (upper cases are the exponents of the corresponding lower cases) and $\underline{\kappa}$ the flow cost of maintaining a vacancy of type c in terms of foregone output of this task type. Then, the Bellman equations for the reservation wage of the worker and for the value of a vacancy read:

$$R(s,\cdot) = \underline{\beta}\lambda \int_{m_c(s,\cdot)} v(c)\left[Y^0(s,c,\cdot) - R(s,\cdot)\right]\,dc \tag{12}$$

$$P(c,\cdot)\,\underline{\kappa} = (1-\underline{\beta})\lambda \int_{m_s(c,\cdot)} u(s)\left[Y^0(s,c,\cdot) - R(s,\cdot)\right]\,ds \tag{13}$$

where $\lambda \equiv \frac{\Lambda}{\underline{\delta}+\underline{\rho}}$; $\underline{\delta}+\underline{\rho}$ is the unemployment inflow rate into job seeking, $\underline{\delta}$ for destroyed matches and $\underline{\rho}$ for the growth of the workforce, so $\lambda v(c)$ measures the contact rate of to a job of type c divided by the inflow rate into job seeking. The left hand side of (13) is the monetary flow cost of maintaining a vacancy of type c, that is $\underline{\kappa}$ units of task type c times the task price $P(c,\cdot)$. The sets $m_s(c,\cdot)$ (or: $m_c(s,\cdot)$) are the subsets of c (or: s) with whom s (or: c) is willing to match. These subsets are determined by the condition that the match surplus is positive:

$$Y^0(s,c,\cdot) - R(s,\cdot) > 0 \Leftrightarrow y^0(s,c,\cdot) > r(s,\cdot) \Leftrightarrow s \in m_s(c,\cdot) \Leftrightarrow c \in m_s(s,\cdot)$$

The steady state flow equilibrium condition reads:

$$\frac{1}{\sigma^s}\phi\left(\frac{s-\mu^s}{\sigma^s}\right) - u(s) = \lambda u(s) \int_{m_c(s,\cdot)} v(c)\,dc \tag{14}$$

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:

$$\lambda v(c) \int_{m_s(c,\cdot)} u(s)\,\underline{F}(s,c)\,ds - \underline{\kappa} v(c) = \frac{1}{\sigma^c}\phi\left(\frac{c-\mu^c}{\sigma^c}\right)\exp\left(y^{\#}\right) \tag{15}$$

which is the equivalent of equation (4) for search equilibria. The left hand side is the output of the inflow of filled vacancies. The right hand side is task demand for type c, the Leontieff coefficient times output of the composite commodity.

4.3 The cost of search

The notation introduced previously for the Walrasian case can be extended to the analysis of search frictions. We express the (relative) cost of search $x(s,\cdot)$ as the difference between the maximum of log value added in the optimal task $y(s,\cdot)$ and the log reservation wage $r(s,\cdot)$:

$$x(s,\cdot) \equiv y(s,\cdot) - r(s,\cdot) \tag{16}$$

We will be specific on the other arguments of $x(s,\cdot)$ below. $x(s,\cdot)$ is the relative difference between the maximum value added and the reservation wage of a type s worker. The situation is depicted in the lower panel of Figure 5. Panel A represents the Walrasian case, where the log reservation wage $r(s,\cdot)$ is equal to $y(s,\cdot)$; hence, $x(s,\cdot) = 0$. All workers of type s will be assigned to task type $c(s,\cdot)$. Panel B represents the case with search frictions. The log reservation wage $r(s,\cdot)$ is less than the maximum log value added $y(s,\cdot)$, the difference being $x(s,\cdot)$. An s-type job seeker accepts all c-type tasks for which $y^0(s,c,\cdot) > r(s,\cdot)$. In the Walrasian equilibrium, the worker is assigned to the unique job type $c(s,\cdot)$ that maximizes $y^0(s,c,\cdot)$ but under search frictions, she accepts all c-types satisfying the constraint $y^0(s,c,\cdot) > r(s,\cdot)$. Since wages are set by Nash bargaining, a fraction β of the difference $y^0(s,c,\cdot) - r(s,\cdot)$ translates into wages. Hence, wages are a decreasing function of $[c - c(s)]^2$, that is, the deviation of the actual value of c in a particular match for worker type s from its optimal value $c(s)$. Gautier and Teulings (2002) find strong empirical support for this implication. Figure 7 depicts both situations in s,c-space: the Walrasian equilibrium is represented by the diagonal, the search equilibrium by a band around it.

An approximation of $x(s,\cdot)$ can be derived from the system (12)-(14), see Teulings and Gautier (2001) and Appendix A.2 for the derivation:

$$x(s,\cdot) = \left[\underline{\theta}\frac{\sigma^s}{\lambda}\phi\left(\frac{s-\mu^s}{\sigma^s}\right)^{-1} \underline{F}\left[s,c(s,\cdot)\right]^{-1}\sqrt{g(s,\cdot)}\right]^{2/5} + O\left(x^2\right) \tag{17}$$

where $\underline{\theta} \equiv \frac{9}{8\sqrt{2}}\frac{\kappa}{\underline{\beta}(1-\underline{\beta})}$; $\underline{\theta}$ is of little interest for the purpose of this paper, so we do not discuss it here. The other three effects in the equation play an important role. First, $\frac{\lambda}{\sigma^s}\phi\left(\frac{s-\mu^s}{\sigma^s}\right)$ measures the effect of labor supply of type s. This effect is equal to the size of the total labor supply, λ, times the density function for type s. This effect is due to IRS in search: the larger the supply at a particular interval of the skill distribution, the lower the cost of search at that interval. Second, labor productivity $\underline{F}[s,c(s)]$ comes in through the cost of vacancies: the lower the productivity, the heavier the burden of the cost of a vacancy $\underline{\kappa}$ relative to the potential output of that job. The final factor in equation (17) is the mismatch coefficient $g(s,\cdot)$, measuring the curvature of $y^0(s,c,\cdot)$, see Figure 5, panel B. The larger this coefficient, the larger the cost of out of equilibrium assignment, and the larger therefore $x(s,\cdot)$.

Since the first effect, $\frac{\lambda}{\sigma^s}\phi\left(\frac{s-\mu^s}{\sigma^s}\right)$, is central to our analysis, we discuss it in somewhat greater detail. The size of the labor force comes in with an elasticity of 2/5. This elasticity follows from the order of the first non-vanishing terms in the Taylor expansions for the equations (12)-(14), see Teulings and Gautier (2001): $x(s,\cdot)$ enters both Bellman equations (12) and (13) with order 3/2 (adding up to 6/2), and it enters the flow equilibrium condition (14) 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 contacts, λ, or by the density of the skill distribution for that s-type, $\frac{1}{\sigma^s}\phi\left(\frac{s-\mu^s}{\sigma^s}\right)$, is irrelevant for the effect on $x(s,\cdot)$. This trade off between scale and density will be crucial in our analysis of interregional trade as non-dense regions specialize in the production of homogeneous commodities which reduces σ^s-and (partially) offsets the initial negative effect on $x(s,\cdot)$ of having a low λ.

Further insights into the nature of scale effects are obtained by rewriting equation (17). Define:

$$\chi(\lambda) \equiv \underline{\lambda}_0^{2/5}\lambda^{-2/5} \tag{18}$$

where $\underline{\lambda}_0 \equiv \underline{\theta}\underline{\sigma}\phi(0)^{-1}\sqrt{\underline{\xi}}$. Then, equation (17) can be written as:

$$x(s,\chi,\pi^s,\Delta\pi,\pi^c) = \chi\left[\frac{\sigma^s}{\underline{\sigma}}\frac{\phi(0)}{\phi([s-\mu^s]/\sigma^s)}\underline{F}[s,c(s,\Delta\pi,\pi^c)]^{-1}\sqrt{\frac{g(s,\Delta\pi,\pi^c)}{\underline{\xi}}}\right]^{2/5} + O\left(\chi^2\right) \tag{19}$$

where we drop the argument of $\chi(\lambda)$. The reason for having separate arguments for $\Delta\pi$ and π^s will become clear in Section 5. Note that

$$x(0,\chi,\underline{\pi},0,\underline{\pi}) = \chi + O\left(\chi^2\right)$$

χ is therefore the baseline cost of search. In the hypothetical Walrasian region, $\lambda \to \infty$ and the baseline cost of search are zero: $\chi(\infty) = 0$. So, instead of characterizing a region by its size of the labor force, we can also characterize it by its baseline search frictions, using the definition of $\chi(\lambda)$ to switch from the one to the other. This is what we do in what follows: we drop λ, and focus on χ, since this leads to simpler expressions.

The density $\frac{1}{\sigma^s}\phi\left(\frac{s-\mu^s}{\sigma^s}\right)$ reaches a maximum at the median skill type $s=\mu^s$, and hence $x(s,\cdot)$ reaches a minimum at that point. The cost of search is higher in the tails of the distribution. This would also be true under a CRS contact technology because extreme s-types meet a limited amount of extreme $c-$types. The factor between square brackets in equation (19) reflects the excess cost of search for $s \neq \mu^s$ as well as the effect of deviations of $\underline{F}[s, c(s,\cdot)]$ and $g(s,\cdot)$ from their benchmark value. The multiplicative structure of equation (19) implies that all derivatives of $x(s,\cdot)$ to other arguments than χ are of order $O(\chi)$. This has two important implications. First, the effect of χ on $\pi^s, \Delta\pi$, and π^c will be shown to be of $O(\chi)$. Then, the indirect effect of χ via $\pi^s, \Delta\pi$, and π^c on $x(s,\cdot)$ is of order $O(\chi^2)$. Hence, we only have to account for the direct effect of χ in a first order approximation. Second, the larger χ, the more $x(s,\cdot)$ differs between the median and the tails of the distribution. The situation is depicted in Figure 8, where the locus of $x(s,\cdot)$ is drawn for two different values of χ. We can think of New York as having χ_1 and Kansas as having $\chi_2 > \chi_1$. For the median worker type $s=0$, the log cost of search are somewhat higher in Kansas. If we move towards the tails where the market becomes thinner and thinner, differences in search frictions between New York and Kansas get larger and larger. This drives small regions with a high value of χ to specialize in tradable commodities that use predominantly skill types around the median, resulting in a low σ^c. Since we will apply Taylor expansions to $x(s,\cdot)$, we summarize its derivatives in Table 2 for future reference.

The final column in Table 2 presents the "adjusted" second derivative of $x(s,\cdot)$, $x_{\widehat{ss}}$, which is proportional to $\phi(s/\underline{\sigma})^{-2/5}$. Our approach in this paper will be to approximate this function by a parabola, starting from its minimum at $s=0$. Since all even derivatives of this function are positive, this second order expansion underestimates the value of $x(s,\cdot)$ in the tails. The "adjusted" second derivative accounts for this fact, by setting the parameters of the parabola such that the expected squared deviations between the true value of $x(s,\cdot)$ and its expansion is minimized, see the Appendix for details. This adjustment matters, as the "adjusted" second derivative is more than twice as large as its unadjusted counterpart: $x_{\widehat{ss}} = 2.15 x_{ss}$. Figure 9 provides a graphical comparison of the true value (the continuous curve), the Taylor expansion (the small dotted curve) and the "adjusted" expansion (the large dotted curve) for $\sigma_s = 1$. In all future calculations, we use $x_{\widehat{ss}}$ instead of x_{ss}. This does not change our conclusions qualitatively but since all results depend crucially on the differences in the magnitude of search frictions between the median and the tails, it does have a substantial quantitative effect.

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. In general, workers are not assigned to the job type $c(s)$ that maximizes their log value added $y^0(s,c,\cdot)$ 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 $x(s,\cdot)$, while unemployment and the cost of vacancies account for the other two thirds, see Teulings and Gautier (2001). The distribution of the latter two thirds is proportional to the bargaining power of workers and firms.

$$\begin{aligned}\text{unemployment rate} &\cong \frac{2}{3}\underline{\beta}x(s,\cdot) \\ \text{vacancy rate} &\cong \frac{2}{3}\left(1-\underline{\beta}\right)x(s,\cdot) \\ \text{sub-optimal assignment} &\cong \frac{1}{3}x(s,\cdot)\end{aligned} \tag{20}$$

5 The general equilibrium effects of search

Equation (16) defines log reservation wages in search equilibrium as log value added $y(s,\cdot)$ minus the cost of search $x(s,\cdot)$. We have derived a relation for $x(s,\cdot)$ in the previous section. This section deals with $y(s,\cdot)$. Section 2.2 provides a relation for $y(s,\cdot)$ in the Walrasian case. However, this relation is influenced by general equilibrium effects of search frictions on task prices, $p(c,\cdot)$, which in turn affect $y(s,\cdot)$. For this purpose, we employ the following definition:

Definition: *the effective skill distribution with parameters* $\pi^s+\vec{\pi}^s(\chi)$ *is that skill distribution that in the absence of search frictions,* $\chi=0$*, yields the same set of tasks prices* $p(c,\cdot)$ *as the actual distribution with parameters* π^s *yields in the presence of search frictions* χ

The general equilibrium effects of search frictions impose a wedge between the actual and effective skill distribution, captured by a shift in its parameters, $\vec{\pi}^s(\chi)$. This section aims at deriving an expression for $\vec{\pi}^s(\chi)$ for the benchmark economy $\Delta\pi=0,\pi^c=\underline{\pi}$.

The simplest way to get a feeling for this wedge is the *skill compression effect.* The cost of search $x(s,\cdot)$ measures the fraction of labor supply that is lost due to search frictions, either by unemployment, or the cost of maintaining vacancies, or by the productivity loss due to suboptimal assignment. $x(s,\cdot)$ is larger in the tails of the distribution than around the median, $x_{ss}(s,\cdot)>0$. The effective skill distribution is therefore more compressed than the actual skill distribution since effective supply is reduced by search frictions more heavily in the tails than around the median, $\vec{\sigma}^s(\chi)<0$. The *skill compression effect* leads to regression to the mean. In a perfect Walrasian world, the equilibrium assignment is characterized by a one-to-one correspondence between skills and task complexities, $s=s(c,\cdot)$. Search frictions introduce noise in that relation, so that on average resources are directed more towards the production of the median complexity type.

Besides *skill compression*, there is also *skill spoiling*. In a search equilibrium, a particular c-type task is not only undertaken by the optimal $s(c,\cdot)$-type worker but by all the s-types in the matching set $m_s(c)$. Absolute advantage, $\underline{f}_s(c,s) > 0$, implies that higher s-types are more productive and therefore replace more than one worker of type $s(c,\cdot)$. The reverse holds for lower s-types. Consider a c-type for which $s(c,\cdot)$ is below the mode of the skill distribution. Since the density function is increasing below the mode, the likelihood of meeting an $s > s(c)$ type is larger than an $s < s(c)$ type. The reverse holds for a c-type job for which $s(c,\cdot)$ is above the mode of the skill distribution. Since higher s-types have an absolute advantage on all jobs, search frictions raise output for low c-types relative to high c-types. To offset this shift in output, workers must be on average assigned to more complex tasks. This is equivalent to a fall in the mean of the effective skill distribution: $\vec{\mu}^s(\chi) < 0$.

A full analytical characterization of the effective skill distribution is not available. Our strategy is to apply a second order Taylor expansion to the log density function of the effective skill distribution to establish its parameters. The derivation is in Appendix A.4. We provide an intuition for the main steps. Consider equation (15) that gives the market equilibrium for task type c. The integral on the right hand side is approximated by a Taylor expansion, using a symmetric integration interval around its midpoint $s(c,\cdot)$. Taking logs and using the benchmark assumption $\Delta\pi = 0, \pi^c = \underline{\pi}$ yields:[8]

$$-\frac{1}{2}\left(\frac{s(c,\cdot)}{\underline{\sigma}}\right)^2 + \underline{f}\left[s(c,\cdot), c\right] - \underline{\omega}x - x_s \cong -\frac{1}{2}\left(\frac{c}{\underline{\sigma}}\right)^2 + \text{constant} \tag{21}$$

where $\underline{\omega} \equiv 1 - \frac{1}{3}\underline{\xi}$. The final two terms on the left hand side capture the effect of search frictions. They are of order $O(x)$. Without these terms, the equation is identical to the market clearing condition (4) since in the Walrasian case, $s(c,\cdot)$ is the inverse of $c(s,\cdot)$. The term $-\underline{\omega}x$ captures the *skill compression effect*: a fraction $x(s,\cdot)$ of the actual skill distribution is "lost" due to search frictions.[9] The term $-x_s$ reflects the *skill spoiling effect*.

Equation (21) holds identically for all s. Hence, its first two derivatives with respect to s evaluated at $s = 0$ must apply. This yields two conditions.[10] The *skill compression effect* drops out in the first derivative (since $x_s = 0$), but survives in the second (since $x_{\hat{s}\hat{s}} > 0$): *skill compression* works by the larger impact of frictions on skill levels in the tails than at the median.

[8] See equation (15) in the appendix. We use that $s(c)$ is approximately the inverse of $c(s)$, see Teulings and Gautier (2001), Proposition 4.

[9] The term $-\frac{1}{3}\underline{\xi}$ captures a small offsetting effect. With search frictions, task type c is done by both better and lower skilled workers than type $s(c,\cdot)$. 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. Hence, a wider matching set, as is observed in the tails, yields a higher output gain.

[10] Equation (21) itself at the point $s = 0$ is not informative since it includes a constant containing the level of aggregate output $y^{\#}$. However, this constant drops out after differentiation.

The *skill spoiling effect* survives in the first derivative of equation (21) (again, since $x_{\hat{s}\hat{s}} > 0$), but vanishes in the second (since $x_{sss} = 0$): *skill spoiling* works by its proportional impact on all skill levels. We obtain the following expressions for skill spoiling and skill compression:

$$\begin{aligned} \vec{\mu}^s(\chi) &= -\frac{2}{3}\sqrt{\frac{5}{3}}\chi + O\left(\chi^2\right) \\ \vec{\sigma}^s(\chi) &= -\frac{1}{3}\sqrt{\frac{5}{3}}\underline{\xi}^{-1}\underline{\Sigma\omega}\chi + O\left(\chi^2\right) \end{aligned} \tag{22}$$

5.1 Benchmark values for the exogenous variables

In order to get an idea of the numerical implications of the model we set the parameters at realistic values from an empirical point of view. Evidence from Teulings and Vierra (1999) suggests that $\underline{\xi} \simeq 0.25$. For the US, the variance of log wages, $\underline{\Sigma}$, 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. When we set the natural rate of unemployment at about 5 %, $\chi \cong 0.15$ according to equation (20). Table 3 summarizes those benchmark values of the model's parameters.

5.2 Combining parts: equilibrium in a single region

We are now in a position that we can characterize a search equilibrium in the benchmark economy with $\Delta\pi = 0$, and $\pi^c = \underline{\pi}$, completely. Reservation wages in this economy satisfy:

$$r\left(s, \chi, \underline{\pi}\right) = y\left(s, \Delta\vec{\pi}, \underline{\pi}\right) - x\left(s, \chi, \underline{\pi}, \Delta\vec{\pi}, \underline{\pi}\right) \tag{23}$$

where $r\left(s, \chi, \pi^c\right)$ is the log reservation wage of worker type s as a function of the baseline cost of search χ and the characteristics π^c of the tradable produced in that region and $\Delta\vec{\pi} \equiv \pi^s + .\vec{\pi}^s - \pi^c$. The argument $\Delta\vec{\pi}$ in the functions $y\left(s, \cdot\right)$ and $x\left(s, \cdot\right)$ takes the value it has for the *effective* skill distribution, in order to account for the general equilibrium effects of search frictions on task prices $p\left(c, \cdot\right)$. Since we consider the benchmark case where $\pi^s = \pi^c$, we have $\Delta\vec{\pi} = \vec{\pi}^s$ Equation (23) makes clear why we need to have three arguments, $\pi^s, \Delta\vec{\pi}$, and π^c in the function $x\left(s, \cdot\right)$, where normally two arguments would suffice since $\pi^s = \pi^c + \Delta\pi$. The first argument π^s refers to the parameters of the actual skill distribution, at it captures the effect of the actual density function on the cost of search, while the second argument $\Delta\vec{\pi}$ refers to the parameters of the effective skill distribution relative to π^c, as it captures the effect of task prices on productivity and on the mismatch coefficient, see equation (19).

Table 4 summarizes the effects of search frictions on the maximum log value added $y(s,\cdot)$ and the cost of search $x(s,\cdot)$, starting from the Walrasian benchmark, $\chi = 0, \Delta\pi = 0, \pi^c = \underline{\pi}$. Note that $y_{\Delta\pi}$ is a 2×1 vector, and $y'_{\Delta\pi}$ denotes its transpose. The effect on reservation wages $r(s,\cdot)$ can be obtained by simply subtracting the one from the other since $r(s,\cdot) = y(s,\cdot) - x(s,\cdot)$. Since a full analytical characterization is unavailable, the effects are presented by their impact on the level and the first two derivatives of $y(s,\cdot)$ and $x(s,\cdot)$, all evaluated for the median skill type, $s = 0$. The effects on higher order derivatives are ignored. Note that these higher order derivatives are zero in the Walrasian benchmark (since $y(s,\cdot) = \underline{\xi}s$), so that the approximation is reasonably precise for small search frictions, χ close to zero. Furthermore, we restrict the attention to effects of order $O(\chi)$, which is again reasonable for small frictions.

The second column presents the level and derivatives of $y(s,\cdot)$ in the Walrasian benchmark, and the direct effect of the baseline cost of search. The third column lists the general formulas for the calculation of the indirect effects via the shifts in $\vec{\pi}^s$ on $y(s,\cdot)$, while the fourth and fifth column specify separately the indirect effects of the skill compression effect $\vec{\sigma}^s$ and the skill spoiling effect $\vec{\mu}^s$, using Table 1 of derivatives of $y(s,\cdot)$. The indirect effects of $\vec{\pi}^s$ on $x(s,\cdot)$ drop out since $x_{\Delta\pi} = x_{s\Delta\pi} = x_{ss\Delta\pi} = O(\chi)$ and $\vec{\pi}^s = O(\chi)$, so that those indirect effects are $O(\chi^2)$. Both skill compression and skill spoiling lead to an increase in the return to skill y_s. This rise in the return to skill causes y for the median worker to go down. Substitution effects sum to zero, so that the mean value added $\mathrm{E}[Y(s,\cdot)]$ must be constant. Since there is more value above than below the median of s (due to the skewness of the log normal distribution), 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. Relative wages in both tales of the distribution are pushed up. Since y_{ss} goes up, y must go down to keep $\mathrm{E}[Y(s,\cdot)]$ constant. For reasonable values of $\underline{\Sigma}^2$, the effect of skill compression and skill spoiling at y is about equal. However, skill spoiling has a much smaller effect on y_s than skill compression.

6 Interregional labor mobility and trade

6.1 Interregional labor mobility

Since the long run cost of labor mobility are assumed to be zero, workers will continue to migrate to other regions till real reservation wages are equal across regions for each s-type. 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. Let $l(\chi, \pi^c)$ denote this log cost of living index for a region with baseline search cost χ that produces a tradable with characteristics π^c. We consider a region that produces the benchmark commodity, $\pi^c = \underline{\pi}$. By the choice of numeraire: $\underline{p}(\underline{\pi}) = 0$. Migration continues till the cost of living index and the skill distribution are such that real reservation wages for all χ and s are equal to reservation wages in the Walrasian benchmark region:

$$r(s, \chi, \underline{\pi}) - l(\chi, \underline{\pi}) = y(s, \Delta\vec{\pi}, \underline{\pi}) - x(s, \chi, \pi^s, \Delta\vec{\pi}, \underline{\pi}) - l(\chi, \underline{\pi}) = \underline{r}(s) \tag{24}$$

where $\Delta\vec{\pi}(\chi) \equiv \pi^s(\chi, \underline{\pi}) - \underline{\pi} + \vec{\pi}^s(\chi)$ measures labor mobility and $\vec{\pi}^s(\chi)$ measures the difference between the actual and the effective skill distribution) and where $\pi^s(\chi, \pi^c)$ are the parameters of the skill distribution as a function of χ and π^c. Equation (24) is a straightforward extension of equation (**??**). It is verified easily that equation (24) applies in the Walrasian benchmark.[11] Again, the parameter $\Delta\vec{\pi}$ refers to the *effective*, and not the *actual* skill distribution. Apart from the general equilibrium effects $\vec{\pi}^s(\chi)$, it contains also the effect of labor mobility, $\pi^s(\chi, \underline{\pi}) - \underline{\pi}$. As in the previous section, we can ignore the effect of $\Delta\vec{\pi}(\chi)$ on x, since it is of higher order. The same holds for the effect of $\pi^s(\chi, \underline{\pi})$ on x.

Our approach in the analysis of labor mobility is to apply a first order Taylor expansion in $\Delta\vec{\pi}$ to the market clearing condition and its first two derivatives with respect to s, starting from the Walrasian benchmark, $\chi = 0, \Delta\pi = 0, \pi^c = \underline{\pi}$, and for the median worker type, $s = 0$. This yields three conditions, determining the cost of living and the parameters of the actual skill distribution for a region with cost of search χ:

$$\begin{aligned} r(0, \chi, \underline{\pi}) &= l(\chi, \underline{\pi}) \\ r_s(0, \chi, \underline{\pi}) &= \underline{\xi} \\ r_{ss}(0, \chi, \underline{\pi}) &= 0 \end{aligned} \tag{25}$$

where we use equation (5). Since these equations must apply identically for all χ, their first derivatives with respect to χ must also apply. Using Table 1 we obtain:

$$\begin{aligned} l_\chi &= y'_{\Delta\pi}\Delta\vec{\pi}_\chi - 1 \\ 0 &= y'_{s\Delta\pi}\Delta\vec{\pi}_\chi \\ 0 &= y'_{ss\Delta\sigma}\Delta\vec{\pi}_\chi - x_{\widehat{ss}}/\chi \end{aligned} \tag{26}$$

where the subscript denotes the partial derivative with respect to χ evaluated at $\chi = 0$. The final two equations can be solved recursively for the parameters of the effective skill distribution $\Delta\vec{\pi}_\chi$:

$$\Delta\vec{\pi}_\chi = \frac{x_{\widehat{ss}}}{\chi y_{ss\Delta\sigma}} \begin{bmatrix} -\frac{y_{s\Delta\sigma}}{y_{s\Delta\mu}} \\ 1 \end{bmatrix} = \frac{2}{3}\sqrt{\frac{5}{3}} \begin{bmatrix} 1 \\ -\underline{\Sigma}^{-1} \end{bmatrix} \tag{27}$$

[11] There, $\chi = 0$, and hence $\vec{\pi}^s(\chi) = 0$, so that $\pi^s(\chi) - \underline{\pi} = \Delta\vec{\pi} = 0$. Then, $y(s, 0, \underline{\pi}) = \underline{r}(s)$ solves the equation, implying that $l(0, \underline{\pi}) = 0$.

where we use Tables 1 and 2. Then, the parameters of the actual skill distribution $\Delta\pi_\chi$ can be derived by the substraction of the skill spoiling and compression effect:

$$\Delta\pi_\chi = \Delta\vec{\pi}_\chi - \vec{\pi}^s_\chi = \frac{1}{3}\sqrt{\frac{5}{3}}\left[\begin{array}{c} 4 \\ \frac{1}{3}\underline{\xi}^{-1}\underline{\Sigma}\underline{\omega} - 2\underline{\Sigma}^{-1} \end{array}\right] \tag{28}$$

For a given complexity distribution π^c, the mean of the skill distribution is an increasing function of the cost of search: $\mu^s_\chi > 0$. This is due to the *skill spoiling effect* $\vec{\mu}^s_\chi$: search frictions cause a loss in the effective use of skill, which raises the return to skill y_s. The *skill compression 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 χ on the standard deviation of the skill distribution depends on two effects in the tails which are of opposite sign. The direct effect of larger search frictions on reservation wages (the second term between square brackets) is negative, leading to an emigration of skill types in the tail (a reduction of σ^s). The indirect general equilibrium effect (the final term) works in the opposite direction. The reduction in effective supply in the tails of the distribution pushes up the wages for these skill types ($y_{ss\Delta\pi}\vec{\pi}^s_\chi > 0$). The direct effect dominates when worker types are good substitutes, that is for high values of $\underline{\xi}$:

$$2\underline{\xi} > \underline{\Sigma}^2\underline{\omega} \tag{29}$$

This condition is satisfied for our benchmark parameter values $\underline{\xi} \cong 0.25$ and $\underline{\Sigma} \cong 0.50$. Then, smaller regions with larger χ export worker types from the tails of the skill distribution: $\sigma^s_\chi < 0$.

Substitution of the expressions $\Delta\vec{\pi}_\chi$ and π^s_χ from (27) and (28) in (26) yields:

$$l_\chi = -1 - \frac{1}{3}\sqrt{\frac{5}{3}}\left(1 + \underline{\Sigma}^2\right) \tag{30}$$

The first term on the right hand side measures the direct effect of the baseline cost of search. The second term measures the general equilibrium effects. A simple way to interpret these effects is to realize that χ is the minimal value of $x(s,\cdot)$ attained for the median skill type. Since the no arbitrage condition (24) dictates that all log reservation wages have to be reduced by the same amount, the higher value of $x(s,\cdot)$ in the tails has to be offset by substitution effects that increase $y(s,\cdot)$ in the tails. Since the value weighted sum of substitution effects equals zero, wages have to go down in the median by an amount approximately equal to $\mathrm{E}\left[e^{x(s,\cdot)}\right] - e^\chi$. The greater $\underline{\Sigma}$ the greater this effect, for the usual reason that there is more value above than below the median and that substitution effects have to sum to zero.

We can apply equation (30) to calculate the effect of χ on the expected log wages for the median worker type. 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, $\mathrm{E}\big[y - y^0\big] = \frac{1}{3}x$, the average log productivity surplus above the log reservation wage must satisfy: $\mathrm{E}\big[y^0 - r\big] = \frac{2}{3}x$. Workers get a share $\underline{\beta}$ of this. The average log nominal wage of the median worker, w, therefore satisfies: $w = l + \frac{2}{3}\underline{\beta}x$. Hence:

$$w_\chi = l_\chi + \frac{2}{3}\underline{\beta} = -1 - \frac{1}{3}\sqrt{\frac{5}{3}}\left(1 + \underline{\Sigma}^2\right) + \frac{2}{3}\underline{\beta} \tag{31}$$

Log nominal wages decrease by more than the baseline cost of search χ. The cost of living decreases even more than wages because wages must compensate for the longer unemployment spells in high χ regions.

In the appendix we show some simulation results to check how accurate our Taylor expansions are. They show that our model captures about 89% of the cost of search.

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 of a region with search frictions χ. The results for $\Delta\vec{\pi}(\chi)$ and $\Delta\pi(\chi)$ derived in equation (27) and (28) carry over unchanged to the case with trade.[12] Hence, the parameters of the actual skill distribution satisfy:

$$\pi^s(\chi, \pi^c) = \pi^c + \Delta\pi(\chi) \tag{32}$$

This result has an important implication. Any kind of specialization of a region in the production of a specific type of tradable commodity, π^c, will translate one-for-one in a change in the skill distribution by labor mobility. Regional specialization and trade therefore leave task prices $p(c, \cdot)$ unaffected, since the implied changes in task demand are exactly offset by labor mobility. Hence, the effect of labor mobility and of trade on the skill distribution are additively separable. The value of the optimal specialization of a region with baseline search frictions χ, denoted $\pi^c(\chi)$, is derived from rewriting the first equation of the system (25):

$$l(\chi, \pi^c) = y\left[0, \Delta\vec{\pi}(\chi), \pi^c\right] - x\left[0, \chi, \pi^c + \Delta\pi(\chi), \Delta\vec{\pi}(\chi), \pi^c\right] \tag{33}$$

Equation (33) specifies the cost of living in a region as a function of its baseline search frictions χ and the characteristics of its tradable commodity π^c. Competition forces firms in a region to produce that tradable commodity that maximizes the log reservation wage for the median skill type $s = 0$, $y(0, \cdot) - x(0, \cdot)$. Since labor mobility equalizes reservation wages for all skill types

[12] $\Delta\vec{\pi}_\chi$ and $\Delta\pi_\chi$ do not depend on π^c because $y_{\pi^c} = x_{\pi^c} = 0$ in the benchmark equilibrium, see the tables of derivatives 1 and 2. We used this fact to simplify notation by not including an argument π^c in $\Delta\vec{\pi}(\chi)$ and $\Delta\pi(\chi)$. In a higher order approximation, this simplification would no longer apply.

s, this will simultaneously maximize the reservation wages of all worker types. The rents from living in high density area go to the real estate owners, who push up the cost of living till real reservation wages are equalized between regions, $l(\cdot) = y(0,\cdot) - x(0,) = \underline{r}(0) = 0$. Similar to the hedonic price index model in Rosen (1974), an equilibrium value $\pi^c(\chi)$ solves the first order condition:

$$
\begin{aligned}
l_{\pi^c}[\chi, \pi^c(\chi)] &= 0 \\
l_{\pi^c\pi^c}[\chi, \pi^c(\chi)] &\leq 0
\end{aligned}
\tag{34}
$$

where the inequality is the second order condition of the problem. A second order Taylor expansion of $y(0,\cdot)$ and a first order expansion of $x(0,\cdot)$ with respect to $\Delta\vec{\pi}$ and π^c yields:[13]

$$
\begin{aligned}
y(0, \Delta\vec{\pi}, \pi^c) &= y'_{\Delta\pi}\Delta\vec{\pi} + \frac{1}{2}\Delta\pi' y_{\Delta\pi\Delta\pi}\Delta\vec{\pi} \\
&\quad + (\pi^c - \underline{\pi})' y_{\pi^c\Delta\pi}\Delta\vec{\pi} - \frac{1}{2}(\pi^c - \underline{\pi})' \underline{\Psi}(\pi^c - \underline{\pi}) + O(\chi^3) \\
x(0, \chi, \pi^c + \Delta\pi, \Delta\vec{\pi}, \pi^c) &= \chi + x_{\sigma^s}(\sigma^c + \Delta\pi) + x'_{\Delta\pi}\Delta\vec{\pi} + O(\chi^3)
\end{aligned}
$$

Differentiation with respect to π^c and realizing that $\Delta\vec{\pi}$ does not depend on π^c up to an order $O(\chi^2)$ yields:

$$
\begin{aligned}
y_{\pi^c}(0, \Delta\vec{\pi}, \pi^c) &= 2y_{\pi^c\Delta\pi}\Delta\vec{\pi} - \underline{\Psi}(\pi^c - \underline{\pi}) + O(\chi^2) \\
\frac{dx(0, \chi, \pi^c + \Delta\pi, \Delta\vec{\pi}, \pi^c)}{d\sigma^c} &= x_{\sigma^s} + O(\chi^2)
\end{aligned}
$$

The second equation reveals an important point. The direct effect of σ^c on x vanishes since $x_{\sigma^c} = 0$, but there is an indirect effect via σ^s. This implication follows directly from the fact that a region's specialization in the production of tradable commodities, π^c, translates one-for-one into the skill distribution, see equation (32) and leaves task price unaffected. Substitution of these equations in the first order condition (34), rearranging terms, applying equation (27) and Tables 1 and 2 yields:

$$
\begin{aligned}
\underline{\Psi}(\pi^c - \underline{\pi}) &= 2\chi y_{\pi^c\Delta\pi}\Delta\vec{\pi}_\chi - x_{\sigma^s} + O(\chi^2) \\
&= \underline{\xi}\chi\left\{\frac{4}{3}\sqrt{\frac{5}{3}}\begin{bmatrix} 1 & \underline{\Sigma} \\ 2\underline{\Sigma} & \frac{1}{2}(1+9\underline{\Sigma}^2)\end{bmatrix}\begin{bmatrix} 1 \\ -\underline{\Sigma}^{-1}\end{bmatrix} - \frac{2}{5}\begin{bmatrix} 0 \\ \underline{\Sigma}^{-1}\end{bmatrix}\right\} \\
&= -\frac{\underline{\xi}\chi}{\underline{\Sigma}}\begin{bmatrix} 0 \\ \frac{2}{3}\sqrt{\frac{5}{3}}\left(1+5\underline{\Sigma}^2\right) + \frac{2}{5}\end{bmatrix} + O(\chi^2)
\end{aligned}
\tag{35}
$$

[13] Again, we use we use $y_{\pi^c} = x_{\pi^c} = x_{\mu^s} = 0$, see Tables 1 and 2. Furthermore, $y(0,0,\underline{\pi}) = 0$ and $x(0,\chi,\underline{\pi},0,\underline{\pi}) = \chi$. Note that $y_{\Delta\pi\Delta\pi}$ and $y_{\pi^c\Delta\pi}$ are matrices.

If the matrix $\underline{\Psi}$ were positive definite, equation (35) would define for each value of χ a set $\pi^c(\chi)$ of dimension zero. The equilibrium set of all $\pi^c(\chi)$ 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.[14] Hence, $\underline{p}(\pi^c)$ can only be a market equilibrium if $\underline{\Psi}$ is positive semi-definite. Competition will force it to be that way. The only matrix $\underline{\Psi}$ that is consistent with this requirement reads:

$$\underline{\Psi} = \begin{bmatrix} 0 & 0 \\ 0 & \underline{\psi}_{\sigma\sigma} \end{bmatrix} \tag{36}$$

Combining (35) and (36) yields:

$$\sigma_\chi^c = -\left[\frac{2}{3}\sqrt{\frac{5}{3}}\left(1+5\underline{\Sigma}^2\right)+\frac{2}{5}\right]\frac{\xi}{\underline{\Sigma}\underline{\psi}_{\sigma\sigma}} \tag{37}$$

The model does not yield any prediction regarding μ^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 benchmark region. This choice implies that the mismatch coefficient $g(s,\cdot)$ 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. However, the model does yield predictions regarding σ^c, which depends negatively on χ, and hence positively on the scale of the labor market, λ. A reduction in σ^c by specialization translates one-for-one in a reduction in σ^s. 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 a general level of supply λ and the density function of the skill distribution, $\frac{1}{\sigma^s}\phi(0)$. 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. 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 extensive job search. High density regions do the opposite, they produce tradables with heterogeneous inputs, in order to explore their comparative advantage in search intensive production.

The smaller the elasticity $\underline{\psi}_{\sigma\sigma}$, the larger the degree of regional specialization in σ^c, that is, the higher is σ_χ^c. Figure 4 provides a graphical illustration. The parabola is the function $\underline{p}(\sigma^c)$.

[14] 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 $l_{\pi^c\pi^c}(\cdot)$ negative semi-definite, since these π^c do not exist in equilibrium and hence the value of $\underline{p}(\pi^c)$ is irrelevant. Hence, this case is embedded in the case where the constraint of negative semi-definiteness is imposed.

By construction, it reaches a maximum for the benchmark value, $\sigma^c = \underline{\sigma}$. Define the function

$$k(\sigma^c, \chi) \equiv y(0, \cdot) - x(0, \cdot) - \underline{p}(\sigma^c)$$

which is the cost of producing a tradable of type σ^c in a region with frictions χ. The right hand side of equation (35) is the partial derivative of this function, $k_{\sigma^c}(\sigma^c, \chi)$. As this derivative itself does not depend on σ^c, $k(\sigma^c, \chi)$ is linear in σ^c. Equation (35) reveals that $k_{\sigma^c}(\sigma^c, \chi) < 0$ and $k_{\sigma^c \chi}(\sigma^c, \chi) < 0$. Figure 4 shows this function for two values of χ ($\chi_2 > \chi_1$). The optimal choice of σ^c sets marginal revenue $\underline{p}_{\sigma^c}(\sigma^c)$ equal to marginal cost $k_{\sigma^c}(\sigma^c, \chi)$. Cost of living adjusts such that it makes this value of σ^c a point of tangency between both curves; $k_{\sigma^c \chi}(\sigma^c, \chi) < 0$ implies that regions with large χ choose low values of σ^c. The fact that this cross-derivative is non-zero is therefore crucial for our conclusion. If it were zero (as it is for μ^c), then σ^c would be independent of χ, and so would be the pattern of specialization. However, this cross derivative is non-zero for two reasons: $x_{\sigma^s} \neq 0$ and $2y_{\pi^c \Delta\pi} \Delta\vec{\pi} \neq 0$. The first mechanism relates directly to the cost of search, which are larger in the tails of the distribution, see Figure 8 in Section 2.4. 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. The second mechanism is due to the general equilibrium effects of this differential impact of search frictions. Skill compression raises task prices in the tails of the distribution and therefore the production cost of dispersed commodities.

6.3 Do the selection effects offset IRS?

The results in this paper strongly depend 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 (2001) 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. Hence, direct measurement of the returns by regressing the flow of realized matches on the stocks of job seekers and vacancies underestimates the returns to scale by one third. 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 alternative mechanisms which dampen IRS in contacts. Metropoles specialize in the production of the most heterogeneous commodities in terms of task inputs. That is their comparative advantage. A greater heterogeneity in task demand puts a burden on the search process, by reducing the share of match offers that is acceptable.

This mechanism can even more than offset the initial advantage in contact technology of large metropoles.

In order to analyze the relevance of this issue in somewhat greater detail, we extend the order of approximation of $x(0,\cdot)$, to terms of $O\left(\chi^2\right)$:

$$
\begin{aligned}
& x\left[0,\chi,\pi^s,\Delta\vec{\pi},\pi^c\right] \\
= \;& \chi + x'_{\Delta\pi}\Delta\vec{\pi} + x_{\sigma^s}\left(\sigma^c+\Delta\sigma\right) + O(\chi^3) \\
= \;& \chi - \frac{2}{15}\sqrt{\frac{5}{3}}\chi^2 \begin{bmatrix} 1-2\underline{\xi} \\ \underline{\Sigma}+\underline{\xi}\underline{\Sigma}^{-1} \end{bmatrix}' \begin{bmatrix} 1 \\ -\underline{\Sigma}^{-1} \end{bmatrix} \\
& + \frac{2}{5}\underline{\xi}\underline{\Sigma}^{-1}\chi\left[\sigma^c - \frac{2}{3}\sqrt{\frac{5}{3}}\underline{\Sigma}^{-1}\chi + \frac{1}{3}\sqrt{\frac{5}{3}}\underline{\xi}^{-1}\underline{\Sigma\omega}\chi\right] + O(\chi^3) \\
= \;& \chi + \frac{2}{15}\sqrt{\frac{5}{3}}\left[1+\frac{5}{3}\underline{\xi}-\frac{\underline{\xi}}{\underline{\Sigma}^2}\right]\chi^2 - \frac{4}{15}\left[\sqrt{\frac{5}{3}}\left(1+5\underline{\Sigma}^2\right)+\frac{3}{5}\right]\frac{\underline{\xi}^2}{\underline{\Sigma}^2\underline{\psi}_{\sigma\sigma}}\chi^2 + O(\chi^3)
\end{aligned}
$$

where we use $x_{\pi^c} = x_{\mu^s} = 0$ and $\sigma^s = \sigma^c + \Delta\sigma$ in the first equality, we replace $\Delta\pi$ and $\Delta\vec{\pi}$ by (27) and (28) and apply Tables 1 and 2 in the second equality, and where we use (37) to eliminate σ^c. The first term, χ, captures the direct effect of search frictions. The second term captures the indirect effects, via the dispersion of the skill distribution and via the curvature of $y^0\left[s,c\left(s\right),\cdot\right]$, as measured by the mismatch parameter $g\left(s,\cdot\right)$. The skill compression effect $\vec{\sigma}^s$ unequivocally raises the mismatch parameter and therefore search frictions. However, the skill compression effect is largely offset by labor mobility, which raises σ^s. Finally, labor mobility in response to the skill spoiling effect raises productivity in regions with high search frictions due to the inflow of high skilled workers. This reduces $x\left(s,\cdot\right)$, since the cost of vacancies absorbs a smaller share of output. The third term captures the effect of regional specialization. This shifts σ^c, which translates one for one into a shift in σ^s.

Whether or not an increase in the scale of the labor market, reduces $x\left(s,\cdot\right)$ depends therefore on the sign of:

$$
\begin{aligned}
\frac{dx\left(0,\cdot\right)}{d\chi} &= 1 + \frac{4}{15}\sqrt{\frac{5}{3}}\left[1+\frac{5}{3}\underline{\xi}-\frac{\underline{\xi}}{\underline{\Sigma}^2}\right]\chi - \frac{8}{15}\left[\sqrt{\frac{5}{3}}\left(1+5\underline{\Sigma}^2\right)+\frac{3}{5}\right]\frac{\underline{\xi}^2}{\underline{\Sigma}^2\underline{\psi}_{\sigma\sigma}}\chi + O(\chi^2) \quad (38) \\
&= 1 + 0.02 - 0.062\underline{\psi}_{\sigma\sigma}^{-1} + O(\chi^2)
\end{aligned}
$$

where we use the benchmark parameters and $\chi = E\left[\chi\right] = 0.15$ in the second line. Hence, the elasticity of specialization should therefore be of the order of magnitude $\underline{\psi}_{\sigma\sigma}^{-1} \cong 16$ for returns to scale to be fully absorbed by regional specialization. We will test this issue empirically in the next section.

7 Empirical evidence

7.1 Introduction

The model discussed in the previous section 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. In this section we set out to test these implications. Since large scale regions like metropoles are more efficient in matching workers and jobs, wages are higher in these regions, as we observe empirically. One of our goals is to get an idea about the fraction of those regional wage differences that can be explained by differences in search frictions. We will use empirical measures for s, c and the baseline cost of search, χ. We discuss these measures in the next section.

First it is useful to summarize the implications of the model, both qualitatively and quantitatively. For some of the parameters of our theoretical model, like the standard deviation of the wage distribution, $\underline{\Sigma}$, there exists plenty of information. For others, like the benchmark value of the mismatch parameter $\underline{\xi}$ we have to rely on some limited empirical evidence. Only, for the value of the elasticity $\underline{\psi}_{\sigma\sigma}$, which determines the degree of regional specialization in the production of the tradable commodities, we have no information at all. Below, we use the parameter values of Table 3, except for $\underline{\psi}_{\sigma\sigma}$, where we use our estimation results to get an impression about its value. The testable implications are:

1. from (28): $\Delta\mu_\chi > 0$, due to the skill spoiling effect. Quantitatively $\Delta\mu_\chi = \frac{4}{3}\sqrt{\frac{5}{3}} = 1.72.$
2. also from (28): $\Delta\sigma_\chi \gtrless 0$, since the direct and the indirect effect of the skill compression effect are of opposite sign; using Table 3 we obtain $\Delta\sigma_\chi = \frac{1}{3}\sqrt{\frac{5}{3}}\left(\frac{1}{3}\underline{\xi}^{-1}\underline{\Sigma\omega} - 2\underline{\Sigma}^{-1}\right) = -0.93$
3. $\sigma^c_\chi < 0$ due to regional specialization and commodity trade: using the benchmark parameter values from Table 3 and equation (37) yields $\sigma^c_\chi = -\left[\frac{2}{3}\sqrt{\frac{5}{3}}\left(1+5\underline{\Sigma}^2\right) + \frac{2}{5}\right]\frac{\underline{\xi}}{\underline{\Sigma}}\underline{\psi}^{-1}_{\sigma\sigma} = -1.17\underline{\psi}^{-1}_{\sigma\sigma}$
4. from equation (31): $w_\chi < 0$, since large areas are more efficient in matching workers to jobs; from Table 3 $w_\chi = -1.20$
5. from the equations (30) and (30): $l_\chi < w_\chi < 0$, we get $l_\chi < 0$, because areas with low search frictions attract many workers which pushes up the rents, $l_\chi < w_\chi$ because wages

must offset the longer duration of unemployment in small scale regions, and are therefore less depressed in small areas than the cost of living; from Table 3 we get $l_\chi = -1.53 < w_\chi$.

7.2 Operationalization of variables

Estimation requires that the three crucial variables in our analysis are operationalized, the skill index s, the complexity index c, and the cost of search χ. For the skill and complexity, we make use of the positive association between wages on the one hand and both indices on the other hand. We take the Walrasian equilibrium as a point of reference for the subsequent discussion. As discussed in Section 3, log wages w are a continuous strictly increasing function of s, due to absolute advantage, $w = w(s), w'(s) > 0$. Furthermore, the equilibrium assignment of skill type s to job type c is a strictly increasing, continuous function, $s = s(c), s'(c) > 0$. A combination of these results yields $\frac{dw[s(c)]}{dc} = w'[s(c)]\,s'(c) > 0$. Summing up, log wages are an increasing function of both the skill level s and the level of job complexity c, while the skill level s is an increasing function of the complexity level c (and vice versa). These results provide the framework for the operationalization of the concepts skill and complexity. The worker skill index s is assumed to be a function of the usual variables showing up in an earnings equation: 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, and an error term capturing unobserved worker characteristics. Since s is positively related to log wages, we can use the regression coefficients of the log earnings equation as the weights for the aggregation of these variables in a single skill index.[15] Similarly, job complexity is assumed to be proxied by 520 occupation and 242 industry dummies, and an error term for unobserved characteristics.[16] Again, since $\frac{dw[s(c)]}{dc} > 0$, the regression coefficients of economy wide log earnings equations are used as the weights for the aggregation of these dummies in a single complexity index. Call those empirical measures respectively $\widehat{s}$ and $\widehat{c}$. Note that information on wages does not enter directly into the calculated values of $\widehat{s}$ and $\widehat{c}$. Information on wages is only used to obtain the coefficients for the aggregation of either personal or job characteristics. Estimates for π^s and π^c, denoted $\hat{\pi}^s$ and $\hat{\pi}^c$, are obtained by the calculation of the mean and standard deviation of $\widehat{s}$ and $\widehat{c}$ for each region (we discuss the classification of regions below). By construction, $dw/d\widehat{s} = dw/d\widehat{c} = 1$. In terms of the model, this implies that the empirical skill measure is equal to $\widehat{s} = y_s s = \underline{\xi} s$ and hence $\hat{\pi}^s = \underline{\xi}\pi^s$, and

[15] The empirical results in the next section do not change qualitatively when we use the skill index of Portela (2001) which is independent of the wage.

[16] 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.

the same for $\widehat{c}$, since $c'(s,\cdot) = 1$ for the Walrasian benchmark. We use the value for $\underline{\xi}$ from Table 3 to translate between π^s and π^c on the one hand and $\hat{\pi}^s$ and $\hat{\pi}^c$ on the other hand.

The index of labor market density developed in Gautier and Teulings (2000) is applied as a proxy for χ. We label this empirical measure, $\hat{\chi}$. The 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 the job is relatively high. Alternatively, in a high density metropolitan labor market, many workers are potential candidates to fill the job and hence the probability for each individual worker to actually occupy the job is low. Gautier and Teulings (2000) show how to estimate this type of measure for the United States, using a regional classification of both home and work location into 1138 areas, as is available in the 5% PUMS of the 1990 Census, to construct an index for 1138 public use micro data area's. Next, we aggregate up our measure to 82 regions with more than 800 observations (the larger (C)M(S)As and states for the remaining non-(C)M(S)A areas) and merge it to the CPS March supplements (89-92). The mean value of our index is 0.64, the standard deviation is 0.21, the densest areas are Washington D.C. ($\hat{\chi} = 0.18$) and Boston, the area which is the least dense is rural Montana ($\hat{\chi} = 0.95$). Our claim that $\hat{\chi}$ is a good proxy for χ implies that both variables are highly correlated. It does not imply that they have the same metric. We assume that there exists a linear relation between them, $\hat{\chi} = \Theta\chi + \Theta_0$, so that $d\hat{\chi}/d\chi = \Theta$. The parameter Θ_0 will be irrelevant for our analysis. For Θ, we have no way to establish its value a priori, so there is no alternative than deriving it empirically.

We use the Dumond et al.(1999) cost of living index 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.

7.3 Estimation results

Table 6 and 7 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.[17] The implications 1 till 4 are supported by the data. All coefficients have

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

the predicted sign and are significant at the 95% level (except $\Delta\widehat{\mu}_{\widehat{\chi}}$, which is only significant at the 90% level). We were particularly surprised that we find support for implication 1, $\Delta\widehat{\mu}_{\widehat{\chi}} < 0$ 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 complexity of labor demand, the mean skill level is higher in low density regions. For the dispersion of the complexity distribution, this is predicted by the model, see implication 3. We also find a strong effect for the mean of the complexity distribution $\widehat{\mu}^c_{\widehat{\chi}} < 0$. This result provides evidence neither against nor in favor of our model. The model does not yield 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 flavour of our model, is that the mismatch parameter $g(s,\cdot)$ 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 on 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 7 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 8 and 9. Again, the estimates are consistent with implications 1 till 4 and part of implication 5. We find that $l_\chi < 0$, but not $l_\chi < w_\chi$.[18] 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 Θ can be derived both from implication 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 $\Delta\mu_\chi$. We have, (see the discussion in Section 3.1 and 3.2):

$$0.030 = \Delta\hat{\mu}_{\hat{\chi}} = \frac{\xi}{\Theta}\Delta\mu_\chi \cong \frac{\xi}{\Theta}1.72 \Rightarrow \Theta \cong 15$$

This number can be used to calculate the coefficient of variation of χ, using the benchmark value of $\mathrm{E}(\chi) = 0.16$ from Table 3

$$\frac{\mathrm{stdev}\,(\chi)}{\mathrm{E}\,(\chi)} = \frac{\mathrm{stdev}\,(\hat{\chi})}{\Theta \mathrm{E}\,(\chi)} \cong 0.1$$

From the definition of $\chi\,(\lambda)$, see equation (18) we know that $\ln\chi = \frac{2}{5}\ln\lambda_0 - \frac{2}{5}\ln\lambda$, where λ is the scale or density of the market. Hence, the coefficient of variation of λ is about $\frac{5}{2} \times 0.1 = 25\%$. Our estimation results therefore suggest that a four standard deviation interval of the log of the scale of the market has a width of $4 \times 0.25 = 1$, or λ varies with a factor e. Search is e times as effective in the top 5% metropolitan areas than in the 5% least densely populated rural regions. Similarly, we can calculate the coefficient of variation for σ^c, which is $\hat{\sigma}^c_\chi \cdot \mathrm{stdev}(\hat{\chi})\,/\hat{\sigma}^c = 0.034(0.198/0.293) = 2.3\%$. Hence, regional specialization in tradable commodities offsets about 2.3/25=9.2% of the initial differences in the scale of the labor market by making dense areas more heterogeneous.

Implication 3 yields an estimate for $\underline{\psi}^{-1}_{\sigma\sigma}$

$$0.037 = -\hat{\sigma}^c_{\hat{\chi}} = -\frac{\xi}{\Theta}\sigma^c_\chi \cong \frac{\xi}{\Theta}1.17\underline{\psi}^{-1}_{\sigma\sigma} \Rightarrow \underline{\psi}^{-1}_{\sigma\sigma} \cong 2$$

This number is a factor eight lower than the number $\underline{\psi}^{-1}_{\sigma\sigma} \cong 16$ quoted in Section 6 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 (38). Substituting $\underline{\psi}^{-1}_{\sigma\sigma} \cong 2$ in (38) yields:

$$\frac{dx\,(0,\cdot)}{d\chi} = 1.02 - 0.06 * 2 = 0.90$$

[18] 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.

Hence, the comparative advantage of large areas in search is absorbed for 10% 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 7 with the theoretically predicted values that attribute all cost of living and wage differences to differences in search frictions. This gives:

$$\begin{aligned} 0.27 &= -l_{\hat{\chi}} \stackrel{?}{=} -\frac{l_{\chi}}{\Theta} = \frac{1.53}{15} \cong 0.10 \\ 0.29 &= -w_{\hat{\chi}} \stackrel{?}{=} -\frac{w_{\chi}}{\Theta} = \frac{1.20}{15} \cong 0.08 \end{aligned}$$

Hence, search frictions account for $28-37\%$ of the cross regional earnings differentials.

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 one eighth 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 (2000) An empirical measure for labor market density, Discussion paper TI 2000-036/3, Tinbergen Institute.

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. (1999 Externalities and Cities, *Mimeo,* University of Chicago.

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.

ROTEMBERG J. AND G. SALONER (1991) Competition and human capital accumulation: A theory of Inter-regional specialization and trade, NBER working paper nr. 3228.

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. (2001) Comparative advantage, relative wages, and the accumulation of human capital, Mimeo,Tinbergen Institute.

TEULINGS C.N. AND P.A. GAUTIER (2000) The right man for the job, Discussion paper TI 2000-038/3, Tinbergen Institute.

TEULINGS C.N. AND J. VIEIRA (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 (5) for the case $\Delta\sigma \neq 0$ reads, see Teulings (2001):

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

A.2 Linear approximation of the relation between $s(c,\cdot)$ and $\Delta\pi$

Analogous to equation (3), the first order condition for the cost minimization reads:

$$p_s^0\left[s(c,\cdot),c,\underline{\xi},\cdot\right] = y_s\left[s(c,\cdot),\underline{\xi},\cdot\right] - \underline{f}_s\left[s(c,\cdot),c\right] = 0 \tag{39}$$

Writing (39) as $y_s[s(c,\cdot),\cdot] = \underline{f}_s[s(c,\cdot),c]$ and taking logs yields:

$$\ln y_s[s(c,\cdot),\cdot] = ln\underline{\xi} + c - s(c,\cdot)$$

where we use $\underline{f}_s[s(c,\cdot),c] \equiv \underline{\xi}e^{c-s(c,\cdot)}$. Since this equation applies identically for all c, its first derivative with respect to c also applies. This yields:

$$s'(c,\cdot)\frac{y_{ss}[s(c,\cdot),\cdot]}{y_s[s(c,\cdot),\cdot]} = 1 - s'(c,\cdot) \tag{40}$$

Define c_0 such that $s(c_0,\cdot) \equiv 0$. Taylor expansions in $\Delta\pi$ around the point $\Delta\pi = 0, \pi = \underline{\pi}, c = c_0$ of the left hand side of both equations yields:

$$\begin{aligned} \frac{y_{s\Delta\pi}(0,0,\underline{\pi})'\Delta\pi}{y_s(0,0,\underline{\pi})} &= -\Delta\mu - \underline{\Sigma}\Delta\sigma = c_0 + O\left(\Delta\sigma^2\right) \\ s'(c_0,\cdot)\frac{y_{ss\Delta\pi}(0,0,\underline{\pi})'\Delta\pi}{y_s(0,0,\underline{\pi})} &= -s'(c_0,\cdot)\frac{\underline{\xi}}{\underline{\Sigma}}\Delta\sigma = 1 - s'(c_0,\cdot) + O\left(\Delta\sigma^2\right) \end{aligned} \tag{41}$$

where we use Table 1 of derivatives. In particular, we use $y_s(0,0,\pi) = \underline{\xi}$ and $y_{ss}(0,0,\pi) = 0$.

A.3 Derivation of (8)

In the subsequent derivation we omit all arguments of functions other than s and c. Partially differentiating equation (6) with respect to s twice and using $\underline{f}_{ss}[s,c(s)] = -\underline{f}_s[s,c(s)] = -y_s(s,\cdot)$, yields:

$$p_{ss}^0[s,c(s,\cdot),\cdot] \equiv g(s,\cdot) = y_{ss}(s,\cdot) + y_s(s,\cdot) \tag{42}$$

The relation between $p^0_{ss}\left[s, c\left(s, 0, \underline{\pi}\right), 0, \underline{\pi}\right]$ and $y^0_{cc}\left[s, c\left(s, 0, \underline{\pi}\right), 0, \underline{\pi}\right]$ can be obtained from the conditions (3) and (39). Differentiation yields:

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

Multiplying both conditions and using $c_s\left(s\right)s_c\left(c\right)=1$ (they are inverse functions) and $\underline{f}_{cc}\left(s,c\right)=-\underline{f}_{sc}\left(s,c\right)=\underline{f}_{ss}\left(s,c\right)$ yields:

$$p_{cc}\left[c\left(s\right),\cdot\right]=\frac{\underline{f}_{ss}\left[s,c\left(s\right)\right]y_{ss}\left(s,\cdot\right)}{\underline{f}_{ss}\left[s,c\left(s\right)\right]-y_{ss}\left(s,\cdot\right)}$$

Partially differentiating equation (1) with respect to c twice, substitution of previous result, and using $\underline{f}_{ss}\left[s,c\left(s\right)\right]=\underline{f}_{s}\left[s,c\left(s\right)\right]=y_s\left(s,\cdot\right)$ yields equation (8).

A.4 Derivation of $x(\cdot)$ from Teulings and Gautier (2001)

In the subsequent derivation we omit all arguments of functions other than s and c. Teulings and Gautier (2001) apply skill and complexity indices $\bar{s}$ and $\bar{c}$ respectively:

$$\begin{aligned}\bar{s} &\equiv -e^{-s}\\ \bar{c} &\equiv e^{c}\end{aligned}$$

Point of departure is their equation (23):

$$x\left(\bar{s}\right)=\left(\underline{\theta}^*\frac{1}{\lambda}\bar{l}(\bar{s})^{-1}\frac{\underline{\kappa}P\left[\bar{c}\left(\bar{s}\right)\right]}{R(\bar{s})}\sqrt{y_{\bar{s}\bar{s}}(\bar{s})}\right)^{2/5}+O\left(x^2\right)$$

where $\underline{\theta}^*$ is a constant, $\bar{l}(\bar{s})$ is the density function of $\bar{s}$, $y_{\bar{s}\bar{s}}\left(\bar{s}\right)$ is the second derivative of log value added in the optimal assignment,[19] and where $\bar{c}\left(\bar{s}\right)$ is defined analogous to $c\left(s\right)$ in this paper. The term $B^*\left(\bar{s}\right)$ in their equation drops out, since here the value of leisure is zero. Their term $K^*\left(\bar{s}\right)=\frac{\underline{\kappa}}{R(\bar{s})}$ is replaced here by $\frac{\underline{\kappa}P[\bar{c}(\bar{s})]}{R(\bar{s})}$, since the cost of vacancies are in terms of task type c here, instead of in terms of the composite commodity. Applying a transformation of variables yields:

$$\begin{aligned}y_{\bar{s}\bar{s}}(\bar{s}) &= y_{ss}\left(s\right)\left(\frac{ds}{d\bar{s}}\right)^2+y_s\left(s\right)\frac{d^2s}{\left(d\bar{s}\right)^2}=\left[y_{ss}\left(s\right)+y_s\left(s\right)\right]e^{-2s}\\ \bar{l}\left(\bar{s}\right) &= l\left(s\right)\frac{ds}{d\bar{s}}=l\left(s\right)e^{s}\end{aligned}$$

where $l\left(s\right)$ is the density function of s. In the Walrasian benchmark, equation (5) applies. Hence, $\frac{R(\bar{s})}{P[\bar{c}(\bar{s})]}=\underline{F}\left(s,c\right)+O\left(x\right)$. Defining: $\underline{\theta}\equiv\underline{\theta}^*\underline{\kappa}$, using: $g\left(s\right)\equiv y_{ss}\left(s\right)+y_s\left(s\right)$, and substitution of the previous relations yields equation (17).

[19] Their equation (23xx) has $\bar{c}_{\bar{s}}\left(\bar{s}\right)$ instead of $y_{\bar{s}\bar{s}}\left(\bar{s}\right)$. However, these are shown to be equal, see their equation (XX).

A.5 The derivation of $x_{\widehat{ss}}$

For simplicity, we consider the case $\sigma^s = \chi = 1$. The proof can easily be generalized. Then: $x(s) = \left(\frac{\phi(s)}{\phi(0)}\right)^{-2/5}$. Define its approximation: $\hat{x}(s) = X_0 + \frac{1}{2}X_1 x_{ss} s^2 = X_0 + \frac{1}{5}X_1 s^2$, where we use the value for x_{ss} from Table 2 in the second step; $X_0 = X_1 = 1$ yields the standard second order Taylor expansion in the point $s = 0$. For the "adjusted" approximation, the coefficients X_0 and X_1 are set such that they minimize the integral of squared deviations, weighted by the density of s:

$$\begin{aligned}
\Omega &= \int_{-\infty}^{\infty} \phi(s)\left[x(s) - \widehat{x}(s)\right]^2 ds \\
&= \int_{-\infty}^{\infty} \phi(s)^{1/5}\phi(0)^{4/5} - 2\phi(s)^{3/5}\phi(0)^{2/5}\widehat{x}(s) + \phi(s)\widehat{x}(s)^2 ds \\
&= \int_{-\infty}^{\infty} \phi\left(\frac{s}{\sqrt{5}}\right) - 2\phi\left(\frac{s}{\sqrt{5/3}}\right)\left[X_0 + \frac{1}{5}X_1 s^2\right] + \phi(s)\left[X_0 + \frac{1}{5}X_1 s^2\right]^2 ds \\
&= \sqrt{5} - 2\sqrt{\frac{5}{3}}\left[X_0 + \frac{1}{3}X_1 s^2\right] + X_0^2 + \frac{2}{5}X_0 X_1 + \frac{3}{25}X_1^2
\end{aligned}$$

The first order conditions for X_0 and X_1 read:

$$\begin{aligned}
-2\sqrt{\frac{5}{3}} + 2X_0 + \frac{2}{5}X_1 &= 0 \\
-\frac{2}{3}\sqrt{\frac{5}{3}} + \frac{2}{5}X_0 + \frac{6}{25}X_1 &= 0
\end{aligned}$$

Hence:

$$\begin{aligned}
X_0 &= \frac{2}{3}\sqrt{\frac{5}{3}} \cong 0.86 \\
X_1 &= \frac{5}{3}\sqrt{\frac{5}{3}} \cong 2.15
\end{aligned}$$

A.6 The derivation of $\vec{\pi}^s$

In most of the subsequent derivation we omit all arguments of functions other than s and c. Consider Taylor expansions in x around the Walrasian equilibrium for the point $\Delta\pi = \pi^s + \vec{\pi}^s - \pi^c = 0$, $\pi^c = \underline{\pi}$, and hence $\mu^s = \vec{\mu}^s = 0$. In that point, $g(s) = \underline{\xi}$, $c(s) = s$ and $\underline{F}[s(c), c] = \underline{F}[s, c(s)] = 1$. The domains of interaction $m_s(c)$ and $m_c(s)$ are given by the conditions $p^0(s,c) \leq p(c)$ and $y^0(s,c) \geq r(s)$ respectively. Define:

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

Hence: $x^*[s, c(s)] = x(s)$, $x^*_{ss} = -p^0_{ss}$ and $x^0_{cc} = y^0_{cc}$. We use second order Taylor expansions of $x^*(s,c)$ around its maximum in s and c, so that $m_s(c)$ and $m_c(s)$ are symmetric intervals around $s(c)$ and $c(s)$ respectively:

$$\begin{aligned}
m_s(c) &\cong \{s \in [s(c) - \Delta_s(c), s + \Delta_s(c)]\} \\
m_c(s) &\cong \{c \in [c(s) - \Delta_c(s), c + \Delta_c(s)]\}
\end{aligned} \qquad (44)$$

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

$$\begin{aligned} x[s(c)] &= \frac{1}{2}x^*_{ss}\Delta_s(c)^2 = \frac{1}{2}\underline{\xi}\Delta_s(c)^2 \\ x(s) &= -\frac{1}{2}x^*_{cc}\Delta_c(s)^2 = \frac{1}{2}\underline{\xi}\Delta_c(s)^2 \end{aligned} \tag{45}$$

using equation (7) and (8) for the second derivatives. Division of both side of the Bellman equations (12) by $\underline{\beta}\lambda R(s)$, using $e^x - 1 = x + O(x^2)$ and $\frac{R(\overline{s})}{P[\overline{c}(\overline{s})]} = \underline{F}(s,c) + O(x) = 1 + O(x)$, and a Taylor expansion of the integrand yields:

$$\begin{aligned} v(c) &\cong \frac{3}{4\sqrt{2}}\frac{\sqrt{\underline{\xi}}}{\underline{\beta}\lambda}x[s(c)]^{-3/2} \\ u(s) &\cong \frac{3}{4\sqrt{2}}\frac{\underline{\kappa}\sqrt{\underline{\xi}}}{(1-\underline{\beta})\lambda}\frac{P[c(s)]}{R(s)}x(s)^{-3/2} \\ &\cong \frac{3}{4\sqrt{2}}\frac{\underline{\kappa}\sqrt{\underline{\xi}}}{(1-\underline{\beta})\lambda}x(s)^{-3/2} \end{aligned} \tag{46}$$

see Teulings and Gautier (2001), Proposition 2. Consider the integral on the right hand side of equation (15).

$$\begin{aligned} &\int_{m_s(c)} u(s)\,\underline{F}(s,c)\,ds \\ \cong\ &\frac{3}{4\sqrt{2}}\frac{\underline{\kappa}\sqrt{\underline{\xi}}}{(1-\underline{\beta})\lambda}\int_{m_s(c)} x(s)^{-3/2}\,\underline{F}(s,c)\,ds \\ \cong\ &\frac{3}{4\sqrt{2}}\frac{\underline{\kappa}\sqrt{\underline{\xi}}}{(1-\underline{\beta})\lambda}\int_{-\Delta_s(c)}^{\Delta_s(c)} H(z,c)\,dz \\ \cong\ &\frac{3}{4\sqrt{2}}\frac{\underline{\kappa}\sqrt{\underline{\xi}}}{(1-\underline{\beta})\lambda}\int_{-\Delta_s(c)}^{\Delta_s(c)} \left(H(0,c) + H_z(0,c)\,z + \frac{1}{2}H_{zz}(0,c)\,z^2\right)dz \\ \cong\ &\frac{3}{2}\frac{\underline{\kappa}}{(1-\underline{\beta})\lambda x[s(c)]}\left(1 + \frac{1}{3}\underline{\xi}^{-1}x[s(c)]\,a[s(c)]\right) \end{aligned} \tag{47}$$

where $H(z,c) \equiv x[z+s(c)]^{-3/2}\,\underline{F}[z+s(c),c]$ (hence, $H(0,c) = x[s(c)]^{-3/2}$) and $a[s(c)] \equiv H_{zz}(0,c)/H(0,c)$. The first step in (47) applies equation (46), the second uses the approximation of the domain of integration, see (44), the third applies a Taylor expansion to $H(z,c)$ around $z=0$, while the final step applies equation (45) for $\Delta_s(c)$ and the standard formula for integration of a parabola. Since $\underline{f}_s(s,c) = -\underline{f}_{ss}(s,c)$, we have:

$$\begin{aligned} a(s) &= -\underline{f}_s\left(1-\underline{f}_s\right) - 3\underline{f}_s\frac{x_s}{x} - \frac{3}{2}\frac{x_{ss}}{x} + \frac{15}{4}\frac{{x_s}^2}{x^2} \\ 1 + \frac{1}{3}\underline{\xi}^{-1}x(s)\,a(s) &\cong \exp\left[-\underline{\xi}^{-1}\left(\frac{1}{3}\underline{f}_s\left(1-\underline{f}_s\right)x + \underline{f}_s x_s + \frac{1}{2}x_{ss} - \frac{5}{4}\frac{{x_s}^2}{x}\right)\right] \end{aligned}$$

where we omit the argument(s) of x and $\underline{f}_s$ for the sake of convenience. Hence:

$$
\begin{aligned}
&\lambda \int_{m_s(c)} u(s)\,\underline{F}(s,c)\,ds - \underline{\kappa} \qquad (48)\\
\cong\;& \underline{\kappa}\left(\frac{3}{2}\frac{1}{\left(1-\underline{\beta}\right)x}\exp\left[-\underline{\xi}^{-1}\left(\frac{1}{3}\underline{f}_s\left(1-\underline{f}_s\right)x+\underline{f}_s x_s+\frac{1}{2}x_{ss}-\frac{5}{4}\frac{{x_s}^2}{x}\right)\right]-1\right)\\
\cong\;& \frac{3}{2}\frac{\underline{\kappa}}{\left(1-\underline{\beta}\right)x}\exp\left[-\frac{2}{3}\left(1-\underline{\beta}\right)x-\frac{1}{3}\left(1-\underline{\xi}\right)x-x_s-\underline{\xi}^{-1}\left(\frac{1}{2}x_{ss}-\frac{5}{4}\frac{{x_s}^2}{x}\right)\right]
\end{aligned}
$$

where we use $e^{ax}-bx=e^{(a-b)x}+O(x)$ and $\underline{f}_s=\underline{\xi}$ in the last line. Similarly, consider the integral on the right hand side of equation (14):

$$
\begin{aligned}
&\int_{m_c(s)} v(c)\,dc \cong \frac{3}{4\sqrt{2}}\frac{\sqrt{\underline{\xi}}}{\underline{\beta}\lambda}\int_{m_c(s)} x\left[s(c)\right]^{-3/2}dc \qquad (49)\\
\cong\;& \frac{3}{4\sqrt{2}}\frac{\sqrt{\underline{\xi}}}{\underline{\beta}\lambda}\int_{-\Delta_s}^{\Delta_s} x[z+s]^{-3/2}dz \simeq \frac{3}{4\sqrt{2}}\frac{\sqrt{\underline{\xi}}}{\underline{\beta}\lambda}\int_{-\Delta_s}^{\Delta_s}\left[x(s)+x_s(s)z+\frac{1}{2}x_{ss}(s)z^2\right]dz\\
\cong\;& \frac{3}{2}\frac{1}{\underline{\beta}\lambda x}\exp\left[-\underline{\xi}^{-1}\left(\frac{1}{2}x_{ss}-\frac{5}{4}\frac{{x_s}^2}{x}\right)\right]
\end{aligned}
$$

Rewriting equation (14) and substitution of the equations (46) and (49) yields:

$$
\begin{aligned}
l(s) &= u(s)\left[\lambda\int_{m_c(s)} v(c)\,dc+1\right] \qquad (50)\\
&\cong \frac{9}{8\sqrt{2}}\frac{\underline{\kappa}\sqrt{\underline{\xi}}}{\underline{\beta}\left(1-\underline{\beta}\right)\lambda}x^{-5/2}\exp\left[\frac{2}{3}\underline{\beta}x-\underline{\xi}^{-1}\left(\frac{1}{2}x_{ss}-\frac{5}{4}\frac{{x_s}^2}{x}\right)\right]
\end{aligned}
$$

where we apply again $e^{ax}-bx=e^{(a-b)x}+O(x)$. Using this relation we can rewrite equation (46) for $v(c)$ as:

$$
v(c) \cong \frac{2}{3}\frac{\left(1-\underline{\beta}\right)x}{\underline{\kappa}}\exp\left[-\frac{2}{3}\underline{\beta}x+\underline{\xi}^{-1}\left(\frac{1}{2}x_{ss}-\frac{5}{4}\frac{{x_s}^2}{x}\right)\right]l\left[s(c)\right] \qquad (51)
$$

Substitution of equation (48) and (51) in (15) and taking logs yields:

$$
-\frac{1}{2}\left(\frac{c}{\underline{\sigma}}\right)^2+\text{constant} \cong \underline{f}\left[s(c),c\right]-\frac{1}{2}\left(\frac{s(c)}{\underline{\sigma}}\right)^2-\underline{\omega}x-x_s
$$

where $\underline{\omega}\equiv 1-\frac{1}{3}\underline{\xi}$. This equation holds identically for all c. Hence, its derivatives with respect to c have to apply. Define c_0 such that $s(c_0)\equiv\underline{\mu}_s=0$. Then:[20]

$$
\begin{aligned}
-\frac{c_0}{\underline{\sigma}^2} &\cong \underline{f}_s\left[s'(c_0)-1\right]-x_{\hat{s}\hat{s}}s'(c_0)\\
-\underline{\sigma}^{-2} &\cong \underline{f}_s\left[s'(c_0)-1\right]^2-\underline{\sigma}^{-2}s'(c_0)^2-\underline{\omega}x_{\hat{s}\hat{s}}s'(c_0)^2
\end{aligned}
$$

[20] Implicitly, by substituting $\underline{f}_s\left[s(c),c\right]$ for $\underline{\xi}$ in the terms in the second line of equation (**??**), we ignore its derivative with respect to c. However, since these terms include a factor of order $O(x)$ and since taking the derivative introduces a factor $s'(c_0)-1$, which is also of order $O(x)$, these terms are of order $O\left(x^2\right)$.

since $\underline{f}_c = -\underline{f}_s$, $\underline{f}_{ss} = \underline{f}_s$, $x_s = x_{sss} = 0$, and $s"(c) = g'(s) = 0$ (due to $\Delta\sigma = 0$). In the benchmark equilibrium, $s(c) = c$. The terms c_0 and $s'(c_0) - 1$ are therefore of order $O(x)$, and hence: $s'(c_0)^2 = 1 + 2[s'(c_0) - 1] + O(x^2)$ and: $e^{c_0} = 1 + O(x)$. Rearranging terms, dropping all terms of order $O(x^2)$, using Table 2, and applying $x(0) = \chi$ yields:

$$\begin{aligned} -c_0 &\cong \underline{\xi}^{-1}\underline{\Sigma}^2 \left[s'(c_0) - 1\right] - \frac{2}{3}\sqrt{\frac{5}{3}}\chi \\ s'(c_0) - 1 &\cong -\frac{1}{3}\sqrt{\frac{5}{3}}\underline{\omega}\chi \end{aligned} \tag{53}$$

Since the analysis starts in the point $\Delta\pi = 0$, we have $\Delta\vec{\pi} = \vec{\pi}^s$. Substitution in equation (41) yields:

$$\begin{aligned} 0 &= c_0 + \vec{\mu}^s + \underline{\Sigma}\vec{\sigma}^s \\ s'(c_0) &= 1 + \frac{\underline{\xi}}{\underline{\Sigma}}\vec{\sigma}^s s'(c_0) \end{aligned} \tag{54}$$

Substitution of (54) in (53), rearranging terms, and using $s'(c_0) - 1 = O(x)$ yields equation (22).

B Numerical simulations

Section 4.3 provides a characterization of the search equilibrium with labor mobility by means of Taylor expansions. However, one can never be sure how accurate these expansions are in practice, in particular when a multitude of expansions are combined in a complicated manner. Hence, we offer some numerical simulations of the model. Our calculations are based on a 600 x 600 grid, where we let s and c vary from $-3\underline{\sigma}$ to $3\underline{\sigma}$, using the values of π^s as derived from the Taylor expansions in the previous sections.[21] Table 5 provides simulation results for the derivatives of r and x, next to their "theoretical" values for $\chi = 0.038$ and 0.116 (generating an average unemployment rate of respectively 1.23 and 3.74 %). In the latter case, we underestimate $x[s, \chi, \pi^s(\chi, \pi^c), g(s, \Delta\pi, \pi^c)]$ by 11% and underestimate the drop in $r(s, \Delta\pi, \pi^c)$ due to search frictions by 30% whereas when $\chi = 0.038$, those figures are respectively: 11% and 23%.

[21] We use the benchmark values of the exogenous variables for our simulations. Teulings and Gautier (2001) have presented similar simulations. However, their ambition is more limited in the sense that their Taylor expansion expression for x takes r in the search equilibrium as given, while here we use the Taylor expansion for skill compression and skill spoiling effects to charaterize r, so that we calculate r and x simultaneously.

C Tables and Figures

Table 1: Derivatives of the function $y\left(s, \Delta\pi, \pi^c\right)$

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

Table 2: Derivatives of the function $x\left(s, \chi, \pi^s, \Delta\pi, \pi^c\right)$

	level	μ^s	σ^s	$\Delta\mu$	$\Delta\sigma$	s	ss	$\widehat{ss}$
x	χ	0	$\frac{2}{5}\underline{\sigma}^{-1}\chi$	$\frac{1}{5}\left(2\underline{\xi}-1\right)\chi$	$-\frac{1}{5}\underline{\Sigma}\left(1+\underline{\xi}\underline{\Sigma}^{-2}\right)\chi$	0	$\frac{2}{5}\underline{\sigma}^{-2}\chi$	$\frac{2}{3}\sqrt{\frac{5}{3}}\underline{\sigma}^{-2}\chi$

Table 3: Benchmark parameter values

$\underline{\xi}$	$\underline{\Sigma}$	$\underline{\beta}$	E[χ]	$\underline{\omega}$
0.25	0.50	0.50	0.15	0.92

Table 4: Direct and indirect effects of search frictions in a single region up to an order $O\left(\chi^2\right)$

Function	Walras/ direct effect	General equilibrium effect	Skill spoiling (via $\vec{\mu}^s$)	Skill compression (via $\vec{\sigma}^s$)
y	0	$y'_{\Delta\pi}\vec{\pi}^s$	$-\frac{2}{3}\sqrt{\frac{5}{3}}\underline{\Sigma}^2\chi$	$-\frac{1}{6}\sqrt{\frac{5}{3}}\underline{\xi}^{-1}\underline{\omega\Sigma}^2\left(1+3\underline{\Sigma}^2\right)\chi$
y_s	$\underline{\xi}$	$y'_{s\Delta\pi}\vec{\pi}^s$	$\frac{2}{3}\sqrt{\frac{5}{3}}\underline{\xi}\chi$	$\frac{1}{3}\sqrt{\frac{5}{3}}\underline{\omega\Sigma}^2\chi$
y_{ss}	0	$y'_{ss\Delta\sigma}\vec{\sigma}^s$	0	$\frac{1}{3}\sqrt{\frac{5}{3}}\underline{\omega}\underline{\xi}\chi$
x	χ	0	0	0
x_s	0	0	0	0
$x_{\widehat{ss}}$	$\frac{2}{3}\sqrt{\frac{5}{3}}\underline{\xi}^2\underline{\Sigma}^{-2}\chi$	0	0	0

Table 5: Simulation results in %

	$\lambda = 12500$ ($U = 1.23$)				$\lambda = 3125$ ($U = 3.74$)			
derivative	0	1	2	4	0	1	2	4
$x(\cdot)$								
expansion	3.81	0.00	0.38	0.11	6.64	0.00	0.66	0.20
numerical	4.22	0.04	0.61	-0.32	7.46	0.12	0.96	0.19
$r(\cdot)$								
expansion	-5.86	25.00	0.00	0.00	-10.21	25.00	0.00	0.00
numerical	-7.62	25.58	0.50	0.06	-14.50	26.54	1.02	-0.01
χ	3.38				11.6			
$\vec{\mu}^s$	-3.3				-9.9			
$\vec{\sigma}^s$	-2.5				-7.5			

Table 6: 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
$\hat{\mu}^s_{\hat{\chi}}$	-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.

Table 7: 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
$\hat{\mu}^s_{\hat{\chi}}$	-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.

Table 8: Means of the estimated vaiables

Sample	unrestricted	age < 30 yrs	unrestricted
# regions	82 (min obs=800)	62 (min obs=400)	62 (same as age<30)
$\hat{\sigma}^c$	0.353	0.293	0.353
$\Delta\hat{\sigma}$	0.020	0.0193	0.0191

Note: $\hat{\mu}^s$ and $\hat{\mu}^s$ are by construction 0.

Table 9: Means continued

Sample	cost of living areas	age < 30 yrs, (C)MSA
# regions	(min obs=500)	60 (min obs = 250)
$\hat{\sigma}^c$	0.354	0.294
$\Delta\hat{\sigma}$	0.020	0.018

Note: $\hat{\mu}^s$ and $\hat{\mu}^s$ are 0.

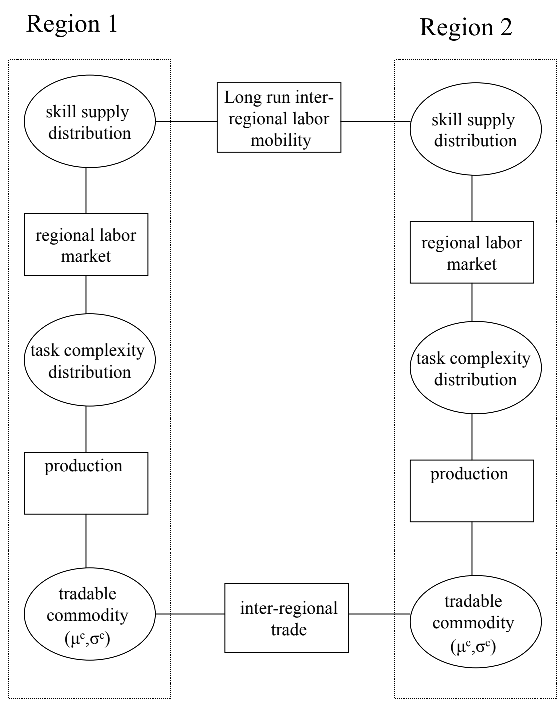

Figure 1: Two region example

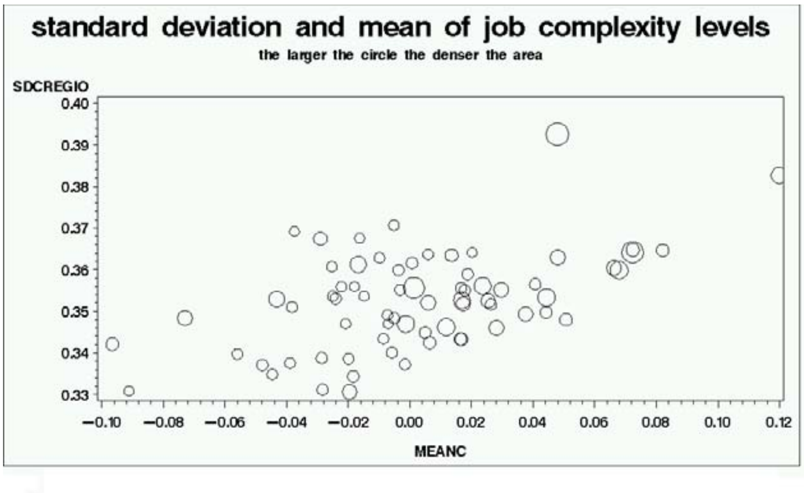

Figure 2:

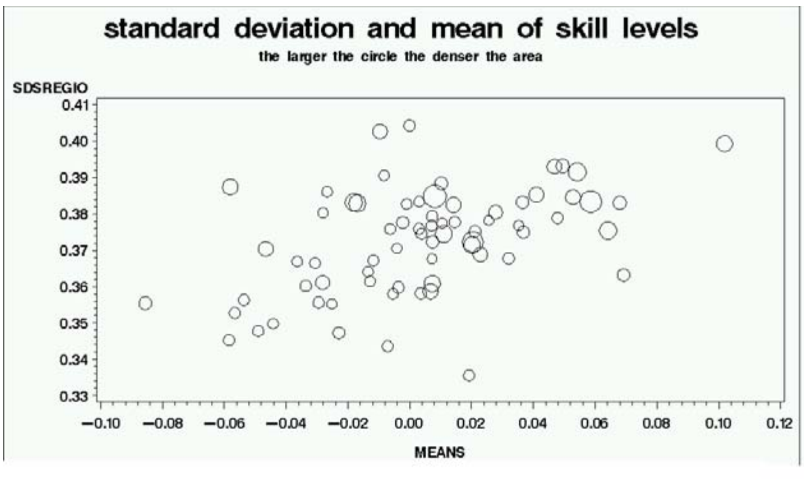

Figure 3:

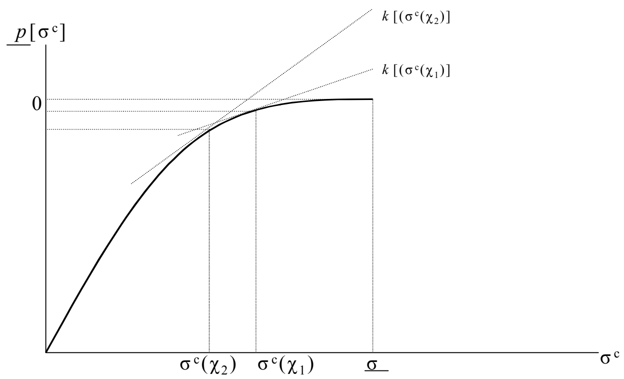

Figure 4: Equilibrium price formation of the tradable

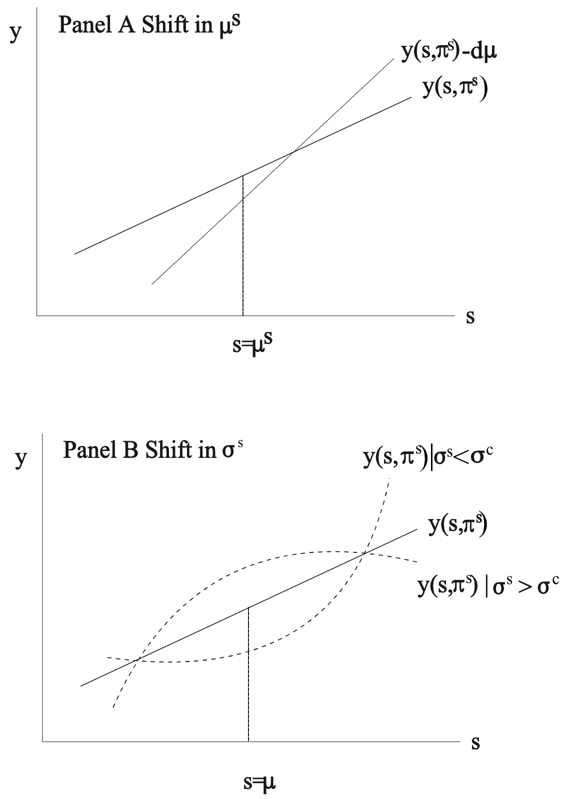

Figure 5: Shifts in μ^s and σ^s

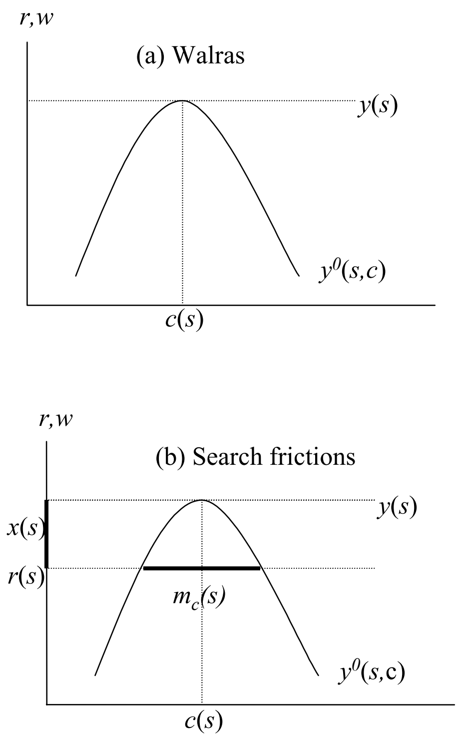

Figure 6: Walras versus search frictions

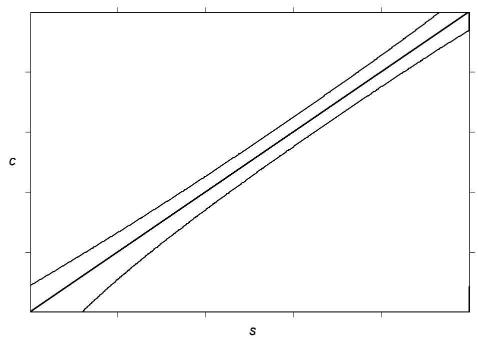

Figure 7: Matching sets: the aggregate outcome

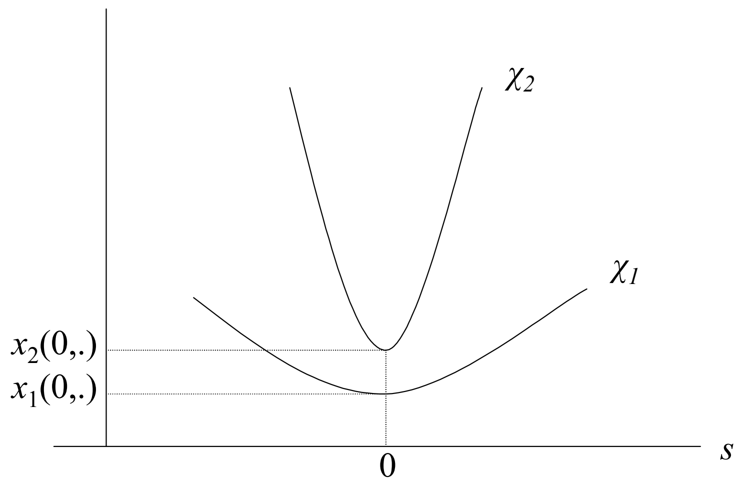

Figure 8: Search frictions in dense and non dense areas

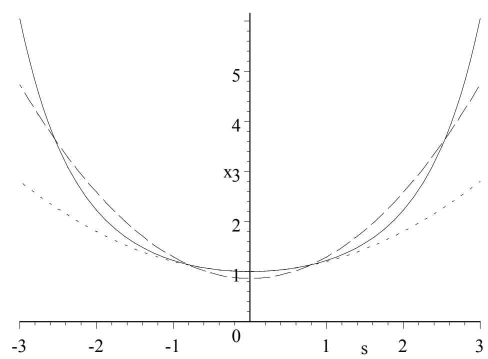

Figure 9: The function $x(s,\cdot)$: true value, 2^e order Taylor expansion, means square approximation