CPB Netherlands Bureau for Economic Policy Analysis

CPB Discussion Paper | 206

Sorting and the output loss due to search frictions

Pieter A. Gautier
Coen N. Teulings

Sorting and the output loss due to search frictions*

Pieter A. Gautier† and Coen N. Teulings ‡

January, 2012

Abstract

We analyze a general search model with on-the-job search and sorting of heterogeneous workers into heterogeneous jobs. This model yields a simple relationship between (i) the unemployment rate, (ii) the value of non-market time, and (iii) the max-mean wage differential. The latter measure of wage dispersion is more robust than measures based on the reservation wage, due to the long left tail of the wage distribution. We estimate this wage differential using data on match quality and allow for measurement error. The estimated wage dispersion and mismatch for the US is consistent with an unemployment rate of 5%. Finally, we find that without search frictions, output would be 6.6% higher.

JEL codes: E24, J62, J63, J64

*We thank seminar participants at MIT, the 2009 Sandjberg conference on search models of the labor market, the Sciences Po conference on sorting, University of Munich, SED (2011) and the 2011 Tinbergen conference in Amsterdam for useful comments and discussions. Finally, we thank Bart Hobijn for sharing his labor-market flow data with us and Xiaoming Cai for excellent research assistance.

†VU University Amsterdam, Tinbergen Institute, CEPR, Ces-Ifo, IZA, email: pgautier@feweb.vu.nl

‡Netherlands Bureau of Economic Policy Analysis, University of Amsterdam, Tinbergen Institute, CEPR,.CeS-Ifo, IZA, email: C.N.Teulings@cpb.nl

1 Introduction

Relative to a competitive economy, an economy with search frictions generates less output because (i) there are idle resources like unemployed workers, (ii) resources are spent on recruitment activities, and (iii) the assignment of workers to jobs is sub-optimal. The interaction between the importance of a precize match and frictions is a key determinant for the allocation of workers to jobs. If all unemployed workers and jobs were alike, it would be hard to imagine why it would take workers months to find a suitable job. But also if workers and jobs are heterogeneous but search frictions are absent, workers will instantaneously be employed at their optimal job types and it does not matter how much output falls if a worker is matched with a sub optimal job type. This paper derives and estimates the output loss due to search frictions allowing for this interaction between frictions and the substitutability of worker types.

Specifically, we analyze a class of search models with on-the-job search (OJS) and heterogeneous workers and jobs. As a starting point we take the framework of Gautier, Teulings, and Van Vuuren (2010) where the productivity of a match is decreasing in the distance between worker and job type and where employed workers continue moving towards the most productive jobs. We use a production technology that can be interpreted as a second order Taylor approximation of a more general production function. Within this framework, various wage mechanisms can be analyzed like wage posting with full commitment as in Burdett and Mortensen (1998) and wage mechanisms without commitment as in Coles (2001) and Shimer (2006). The key difference between wage setting with and without commitment is that in the first case, firms pay both hiring and no quit premia to hire/ keep workers whereas in the latter case, the only reason for firms to pay workers above their reservation wage is to prevent them from quitting. The equilibrium is characterized by a relationship between just three statistics: (i) the unemployment rate, (ii) the value of non-market time, and (iii) a summary statistic for wage dispersion between identical workers, the *max-mean* wage differential. We show that this statistic is more informative and more robust than alternatives like the ratio of the mean wage to the reservation wage (i.e. the *mean-min* ratio of Hornstein, Krusell and Violante, 2010). This relation hardly depends on any of the model's parameters, except for the relative efficiency of on- versus off-the-job search, ψ, and the value of non-market time. For the calculation of the total output loss due to search frictions for a given unemployment rate, even the effect of ψ is a higher order phenomenon.

The combination of two-sided heterogeneity and search frictions relates our model to the literature on hedonic pricing in the spirit of Rosen (1974), Sattinger (1975) and Teulings (1995,2005). The last three of those models are hierarchical, in the sense that better skilled workers have a comparative advantage in more complex jobs. In a Walrasian equilibrium, there is perfect sorting. With search frictions, this perfect correlation between worker and job types breaks down, since workers cannot afford to wait for ever till the optimal match comes along. Shimer and Smith (2000) and Teulings and Gautier (2004) are early examples of assignment models with search frictions. Hierarchical models are difficult to solve because matching sets in the corners of the type space do not have interior boundaries. We therefore first transform the hierarchical model into a circular model in the spirit of Marimon and Zilibotti (1999) and Gautier, Teulings, and Van Vuuren (2008). The idea is that output is decreasing in the distance between worker and job types along the circumference of the circle. This makes it possible to derive a closed form solution of the equilibrium. When turning to the empirical analysis of data on individual wages, and on worker and job characteristics, we have to reintroduce the hierarchical aspect of the model.

Ultimately, the usefulness of our model depends on how well it can simultaneously describe the observed wage dispersion for a given skill type due to mismatch, the unemployment rate, and the ratio of job-to-job versus unemployment-to-job worker flows. We show that the equilibrium unemployment rate in our model that is consistent with the observed amount of wage dispersion is 5% which seems reasonable. Given that our model performs well, it can be used to calculate the total output loss due to search frictions which we estimate to be about 6.5%. The majority of the output loss is due to recruitment activities and mismatch.

Hornstein et.al. (2010) also derive a simple relationship between the unemployment rate and wage dispersion that holds for a large class of search models. They argue that search models without OJS cannot explain the coexistence of low unemployment rates and substantial wage dispersion because the first suggests low frictions and the latter suggests high frictions. Gautier and Teulings (2006) made the related point that without

OJS, estimates of output losses due to search frictions based on the unemployment rate are substantially lower than estimates based on wage dispersion. Allowing for OJS and unobserved heterogeneity can resolve this issue. OJS lowers the reservation wage which increases wage dispersion for a given unemployment rate. As in Hornstein et al. (2010) this requires a sufficiently large contact rate for employed workers.

Similarly, Eeckhout and Kircher (2010) construct a measure based on the distance between the lowest and highest wage. One disadvantage of relating wage dispersion to the lowest observed wage (see also the *mean-min* ratio of Hornstein et.al.,2010) is that the wage distribution for a given skill type has a long left tail. This long tail is consistent with OJS for the reasons spelled out in Burdett and Mortensen (1998): (i) OJS reduces the lowest acceptable job type because less option value of continued search has to be given up when accepting a job, and (ii) less workers quit from good matches and more workers accept good matches. Empirically, it matters a lot whether one takes the 1^{th} or the 2^{nd} percentile of the wage distribution as a proxy for the reservation wage. Therefore, the difference between the highest wage at the optimal assignment and the mean wage (the max-mean differential) is a more robust measure for wage dispersion.

In the equilibrium where firms are unable to commit, quasi-rents per worker are higher than in the social optimum due to a business-stealing externality. Under free entry, these quasi-rents are spent on (excess) vacancy creation, see Gautier, Teulings and van Vuuren (2010). We estimate the output loss due to this business-stealing externality to be 1% (if no firm commits relative to the case where all firms commit and if frictions are such that the unemployment rate under commitment is 5%).

The empirical estimate of the max-mean wage differential for identical workers in heterogeneous jobs is an extension of the framework of Gautier and Teulings (2006) with OJS. Our estimate is based on the intercept of a simple quadratic wage regression with appropriately normalized measures for worker and job characteristics. This type of inference is highly sensitive to measurement error in both characteristics because an observed sub-optimal matched worker can either reflect true mismatch, or simply imply measurement error. Our estimation procedure accounts for this problem. Gautier and Teulings (2006) use second order terms in worker and job characteristics to capture the concavity of wages around the optimal assignment that is implied by search models with sorting.

However, there is a crucial difference between a model with and without OJS. In a model without OJS, wages are a linear transformation of the match surplus. Since the match surplus is a differentiable function of the match quality indicator, so is the wage function. This simple relation no longer applies with OJS, see Shimer (2006). The wage function turns out to be non-differentiable at the optimal assignment. At that point, firms are prepared to pay the highest premiums to raise hiring and to reduce quitting. In our empirical application, we take this into account. Allowing for OJS is important since Fallick and Fleischman (2004) and Nagypal (2005) show that job-to-job flows are substantial.

Lise, Meghir and Robin (2008) and Lopes de Melo (2008) also look at sorting in models with OJS. Their focus is on interpreting the correlations between worker and firm fixed effects and how this relates to complementarities between worker and job types in the production technology. Finally, Eeckhout and Kircher (2009) consider a simple model based on Atakan (2006) where workers and jobs are randomly matched and have the option to dissolve and at some cost move to a competitive sector with perfect sorting. They derive a similar hump shaped relation where productivity is highest at the optimal assignment and decreases in the distance from this optimal assignment. This framework is however less suitable to bring to the data.

The structure of the rest of this paper is as follows. Section 2 presents the basic framework. Section 3 discusses the basic steps in our argument. In this section, we also reinstall the hierarchical features of the model and derive the wage function that comes with it. We also show how we can normalize worker and job skills such that we can relate the constant in a simple wage regression to the max-mean wage differential. Finally, section 4 concludes.

2 The model

2.1 Why a circular model?

Shimer and Smith (2000) and Teulings and Gautier (2004) analyze an assignment model with heterogeneous workers and jobs with search frictions. Though the idea is appealing, the analysis of this type of models is complicated. Figure 1 provides an intuition for why this is the case. The figure shows the space of potential matches between skill types,

s, and worker types, c. The Walrasian equilibrium assignment is depicted by the main diagonal. Comparative advantage of skilled workers in complex jobs implies that it is upward sloping. Perfect sorting implies that it is a one-to-one correspondence. Search frictions make that the equilibrium assignment is a set rather than a single point where the mismatch indicator x measures the distance of worker type s to her optimal assignment. The optimal match is some interior point of the matching set. Teulings and Gautier (2004) use second order Taylor expansions to characterize the equilibrium. Figure 2 shows why this approach works. The figure depicts Rosen's hedonic equilibrium in the wage-skill space. The upper panel is the Walrasian case. The upward sloping line is the market wage for a worker with skill type s, the curves represent the value of output in various c-type jobs when occupied by workers of skill type s. The point of tangency is the optimal assignment, the only assignment that is relevant in the Walrasian case. The lower panel is the case with search frictions. The upward sloping line is now the reservation wage of the worker. All values of s enveloped by this reservation wage curve and the value-of-output curve are now part of the matching set of that job type. The option value of search for a firm is proportional to the surface enveloped by the wage curve and the value-of-output curve. A second order Taylor expansion provides a good approximation for this surface.[1] The same approach can be applied to the option value of search for a job seeker. Figure 1 shows why this approach works fine for the middle part of the distribution, but does not work well for the tails of the distribution: in the tails there is no interior solution for either the upper or the lower bound of the matching set. This "corner problem" complicates the characterization of the equilibrium in this type of models tremendously. Teulings and Gautier (2004) show that their Taylor expansion provide a fairly accurate characterization of the equilibrium for not too large search frictions, that is, for not too high unemployment rates.

Gautier, Teulings, and Van Vuuren (2005) use an alternative approach. They show that by taking out the hierarchical aspect of the model, the south-west and the north-

[1] Theoretically, the surface should be weighted with the density of job seekers for each worker type s. When this density function is differentiable, the variations in the density have only third or higher order effects on the option value, see Teulings and Gautier (2004), Proposition 2. The intuition is that the density weighted matching set has more mass at the right and less mass at the left and both approximately offset each other. Hence, using a uniform vacancy distribution (as we do in the paper) is not restrictive.

east corner of the matching space can be "glued" together. Then, a circular model in the spirit of Marimon and Zilibotti (1999) can be used, see the lower part of Figure 1. The distance x to the optimal assignment is now measured along the circumference of the circle. The advantage of using a circular model is that it has closed form analytical solution. All conclusions regarding search frictions from the analysis of the hierarchical model in Teulings and Gautier (2004) based on Taylor expansions carry over to the circular model. Hence, one can use the closed form analytical solution as a starting point for doing comparative statics in the hierarchical model. The log wage function can be approximated by using a linear decomposition in a Walrasian term that is derived from the hierarchical model, and a search frictions term that is derived from the circular model. Taking out the hierarchical aspect of the model implies that the (reservation) wage function becomes horizontal, that is, we take out the first order of skill differentials. What is crucial is that the differential between the reservation wage curve and the value-of-output curve remains unaffected. Hence, it is differentiable in its maximum, that is, in the optimal assignment. We impose this feature in the theoretical model

The Taylor expansions in Teulings and Gautier (2004) work fine in a model without OJS. But they cannot be extended easily to a model with OJS for technical reasons, which are unrelated to the corner problem. The circular model in Gautier, Teulings, and Van Vuuren (2005) works well with and without OJS. We conjecture that the circular model provides a good approximation of the equilibrium in the hierarchical model in the case with OJS as it did in the case without. Hence, we apply the same decomposition of the log wage function as we did in the model without OJS, namely in a Walrasian term derived from the hierarchical model and a term for search frictions derived from the circular model. First, we provide a closed form analysis of a search equilibrium with OJS for the circular model. Next, we reintroduce the hierarchical aspect for the empirical analysis of wage differentials by adding a hierarchical term based on the characterization of the Walrasian equilibrium in Section 3. This procedure allows us to use data on worker and job characteristics to derive an empirical estimate of workers' skill s and jobs' complexity c, and hence the mismatch indicator x. Those data on s and c are hierarchical by nature.

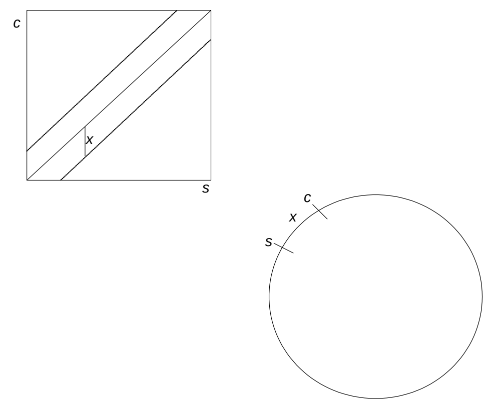

Figure 1: The hierarchical versus the circular model

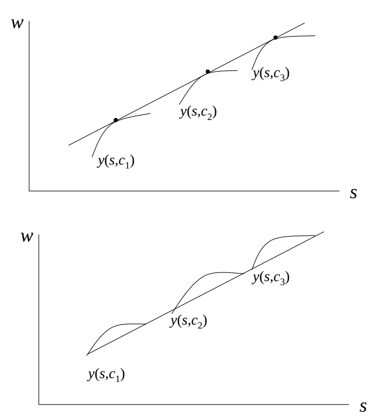

Figure 2: Identifying the strength of sorting: Walras vs frictions

2.2 Assumptions

Production

There is a continuum of worker types s and job types c; s and c are locations on a circle. Workers can only produce output when matched to a job. The productivity of a match of worker type s to job type c depends on the "distance" $|x|$ between s and c along the circumference of the circle, where x is defined as $x \equiv s - c$. $Y(x)$ has an interior maximum at $x = 0$ and is symmetric around this maximum (which is normalized to one) Finally, $Y(x)$ is twice differentiable and strictly concave. We consider the simplest functional form that meets these criteria:

$$Y(x) = 1 - \frac{1}{2}\gamma x^2. \tag{1}$$

We will call x is the *mismatch* indicator. The parameter γ determines the substitutability of worker types. $Y(x)$ can be interpreted as a second order Taylor approximation around the optimal assignment of a more general production technology. Since, the first derivative of a continuous production function equals zero in the optimal assignment, $Y'(0) = 0$, the first order term drops out. We are interested in equilibria where unemployed job seekers do not accept all job offers, which imposes a minimum constraint on γ.[2]

Labor supply and the value of non market time

Labor supply per s-type is uniformly distributed over the circumference of the circle. Total labor supply in period t equals $L(t)$. Unemployed workers receive the value of non market time B. Employed workers supply a fixed amount of labor (normalized to unity) and their payoff is equal to the wage they receive. Workers live for ever. They maximize the discounted value of their expected lifetime payoffs.

Labor demand

There is free entry of vacancies for all c-types. The cost of maintaining a vacancy is equal to K per period. After a vacancy is filled, the firm's only cost is the worker's wage. The supply of vacancies is determined by a zero profit condition. Vacancies are uniformly distributed over the circumference of the circle; $v(c) = v$ is the measure of vacancies per

[2] A sufficient condition for this is that $Y(x) < 0$ for at least some x. Let C be the circumference of the circle, then $0 \leq x \leq \frac{1}{2}C$. Hence $\gamma > 8C$.

unit of c.

Job search technology

We use a reduced form specification of the job search technology:

$$\lambda = \lambda\left(u, v\right), \lambda_v > 0,$$

and assume that the rate at which unemployed workers meet jobs is λ and the rate at which employed workers meet jobs is $\psi\lambda$. The parameter $\psi, 0 \leq \psi \leq 1$, measures the relative efficiency of on- relative to off-the-job search; $\psi = 0$ is the case without OJS; for $\psi = 1$, on- and off-the-job search are equally efficient. When a worker quits her old job, this job disappears.

Job destruction

Matches between workers and jobs are destroyed at an exogenous rate $\delta > 0$.

Golden-growth path

We study the economy while it is on a golden-growth path, where the discount rate $\rho > 0$ is equal to the growth rate of the labor force. We normalize the labor force at $t = 0$ to one. Hence, the size of the labor force is $L(t) = \exp(\rho t)$. The assumption of a golden-growth path buys us a lot in terms of transparency and tractability. The implications of the golden growth assumptions are equivalent to those that follow from the assumption that the discount rate ρ converges to zero, an assumption that is often applied in the wage posting literature, see for example Burdett and Mortensen (1998) because discounting reduces future output while population growth increases it. New workers enter the labor force as unemployed.

Labor supply per worker type and the productivity in the optimal assignment $Y\left(0\right)$ are normalized to one. Hence, in the absence of search frictions, the output of this economy equals one.

Wage setting

Wages, denoted by $W\left(x\right)$, are set unilaterally by the firm, conditional on the mismatch indicator x in the current job. We analyze wage setting under two different assumptions. Under the first assumption, firms can commit to a future wage payment contingent on x.

Then, firms pay both no-quit and hiring premiums, that is, they account for the positive effect of a higher wage offer on reduced quitting and increased hiring. Under the second assumption, firms are unable to commit to future wage payments. In this case, hiring premiums are non-credible because immediately after the worker has accepted the job, the firm has no incentive to continue paying a hiring premium, since the worker cannot return to her previous job. Workers anticipate this, and will therefore not respond to this premium in the first place, and hence, firms will not offer it. No-quit premiums are credible even without commitment because it is in the firm's interest to pay them as soon as the worker has accepted the job.[3]

2.3 The asset values of (un)employment and vacancies

The golden-growth assumption is particularly useful for the derivation of the asset values of employment, unemployment, and vacancies.

Asset value of unemployment

The asset value of unemployment, denoted by V^U is a weighted average of the worker's payoff while unemployed, B, and the expected wage when employed, $\mathrm{E}_x W$, the weights being the unemployment and the employment rate, respectively:

$$\rho V^U = uB + (1-u)\,\mathrm{E}_x W. \tag{2}$$

The derivation of this and the subsequent Bellman equations can be found in the technical appendix A.1 to this paper. Why does this relation take such a simple form? The reason is that the growth rate of the workforce is equal to the worker's discount rate. Therefore, the expected payoff of a worker with one year of experience is equal to the average payoff of the cohort of workers that entered the labor force one year ago. Likewise, the expected payoff of a worker with two years of experience is equal to the average payoff of the cohort of workers that entered the labor market two years ago, etc. The asset value of unemployment is equal to the weighted sum of expected payoffs for each level of experience, future payoffs being discounted at a rate ρ per year. This weighted sum is exactly equal to the sum of payoffs for the current workforce. The fact that older cohorts are smaller than

[3] See also Bontemps van den Berg and Robin (2000) for wage setting with and Coles (2001) and Shimer (2006) for wage setting without commitment.

younger cohorts due to the growth of the labor force at a rate ρ exactly offsets the effect of discounting future payoffs for the calculation of the asset value of unemployment. The term $(1-u)\mathrm{E}_x W$ can be interpreted as the option value of finding a job. Alternatively, when $\rho \to 0$, workers spend a fraction u of their life as unemployed and the rest of the time they are employed.

Asset value of employment in the marginal job

Let $V^E(x)$ be the asset value of holding a job with mismatch indicator x. At $\bar{x}$, an unemployed job seeker is indifferent between accepting the job or waiting for a better offer: $V^E(\bar{x}) = V^U$. Again, the asset value for this job is a weighted average of $W(\bar{x})$ and $\mathrm{E}_x W$,

$$\rho V^E(\bar{x}) = \frac{uW(\bar{x}) + \psi(1-u)\,\mathrm{E}_x W}{u + \psi(1-u)}, \tag{3}$$

The factor $u + \psi(1-u)$ is the effective supply of job seekers, namely u unemployed and $(1-u)$ employed job seekers. The latter are discounted by a factor ψ due to their lower search effectiveness. Hence, the weights in equation (3) are the shares of unemployed and employed respectively in the effective supply. The intuition for this equation is that the acceptance sets are the same for an unemployed and a employed job seeker at his marginal job, both accept any job: $0 \le |x| < \bar{x}$. However, the option value of search for an employed worker is only a fraction ψ of the option value of an unemployed due to their lower contact rate. The value of non market time B does not show up in the equation since $V^E(\bar{x}) = V^U$ which implies that we can substitute $V^E(\bar{x})$ for V^U, thereby eliminating B from the equation. For $0 < \psi < 1$, unemployed job seekers give up part of the option value of search by accepting a job and they need to be compensated for this, implying that $W(\bar{x}) > B$. For $\psi = 1$, on- and off-the-job search are equally efficient.

Asset value of vacancies

Adding up the zero profit condition for all vacancies implies that the total cost of maintaining vacancies must be equal to the aggregate rents that firms make in currently filled jobs.

$$vK = (1-u)\left(\mathrm{E}_x Y - \mathrm{E}_x W\right). \tag{4}$$

The left hand side of this equation is the total cost of vacancies at $t = 0$. The right hand side is employment, $1-u$, times the quasi-rents per worker, which is equal to expected output $\mathrm{E}_x Y$ minus expected wages $\mathrm{E}_x W$.

The reservation value of the mismatch indicators

The definition of $\bar{x}$ as the mismatch indicator of a job which is just acceptable to an unemployed job seeker implies:

$$W(\bar{x}) = Y(\bar{x}) = 1 - \frac{1}{2}\gamma\bar{x}^2, \tag{5}$$

Substitution of equation (2) and (4) in the condition $V^E(\bar{x}) = V^U$ yields,

$$W(\bar{x}) = [u + \psi(1-u)]\,B + (1-\psi)(1-u)\,\mathrm{E}_x W. \tag{6}$$

When on and off the job search are equally efficient, $\psi = 1$, equation (6) simplifies to:

$$W(\bar{x}) = B = 1 - \frac{1}{2}\gamma\bar{x}^2. \tag{7}$$

where the last step follows from (5). Hence, the relation between γ and $\bar{x}$ does not depend on expected wages in this case, and consequently neither on whether or not firms can commit on paying hiring premiums.

The output loss due to search frictions

The definition of the output loss due to search frictions is given by:

$$X = (1-u)(1-\mathrm{E}_x Y) + u(1-B) + vK = 1 - (1-u)\,\mathrm{E}_x Y - uB + vK. \tag{8}$$

The output loss is equal to employment, $1-u$, times the difference between productivity in the optimal assignment, $Y(0) = 1$, and the expected productivity in the actual assignment, $\mathrm{E}_x Y$, plus unemployment, u, times the difference between the productivity in the optimal assignment and the value of non market time, $1-B$, plus the cost of keeping vacancies open, vK. Substitution of equation (2) and (4) in (8) yields:

$$X = 1 - \rho V^U = u(1-B) + (1-u)(1-\mathrm{E}_x W). \tag{9}$$

The last step follows from the fact that by the zero profit condition, the cost of maintaining vacancies is equal to the surplus of expected productivity over expected wages, see equation (4). This is useful because both v and K are hard to measure. The first equality tells us that the output loss is equal to the output in the optimal assignment ($Y(0) = 1$)

minus the asset value of unemployment. The second equality in (9) tells us that the output loss is equal to the sum of the output loss for unemployed and for employed workers. The former is equal to the lost output in the optimal assignment minus the value of non market time, while the latter is equal to the foregone wage income. Under free entry, the difference between wages and productivity is spent on vacancies.

2.4 Equilibrium flow conditions

Under both assumptions for wage setting, commitment and no-commitment, wages are a decreasing function of x for $x \geq 0$, $W_x(x) > 0$, implying that workers accept any job-offer with a lower mismatch indicator $|x|$ than their current job. Hence, we can analyze job-to-job flows independent of the exact form of the wage policy $W(x)$. Let $G(x), x \geq 0$, be the fraction of workers employed in jobs at smaller distance from their optimal job than $|x|$. This implies that $G(0) = 0$ and $G(\bar{x}) = 1$, since $\bar{x}$ is the largest acceptable value of $|x|$. The golden growth assumption requires that the number of workers employed in jobs with a mismatch indicator lower than x grows at a rate ρ:

$$2\lambda x \{u + \psi(1-u)[1 - G(x)]\} - \delta(1-u)G(x) = \rho(1-u)G(x). \tag{10}$$

The first term on the left-hand side is the number of people that find a job with mismatch indicator lower than x, either from unemployment (the first term in square brackets), or by mobility from jobs with a larger mismatch indicator (the second term in brackets). The number of better jobs is given by $2x$, since the worker can accept jobs both to the left and to the right of her favorite job type $x = 0$. The second term in brackets is weighted by the factor ψ, reflecting the efficiency of on- relative to off-the-job search. The final term on the left-hand side is the outflow of workers due to job-destruction. The right-hand side reflects that at the balanced growth path, employment grows at a rate ρ at all levels including the class of workers with a mismatch indicator smaller than x, $G(x)$. Mobility within this class is irrelevant because the disappearance of the old match and the emergence of the new one cancel. Evaluating (10) at $\bar{x}$ and solving for u yields,

$$u = \frac{1}{1 + \kappa\bar{x}}, \tag{11}$$

where: $\kappa \equiv \frac{2\lambda}{\rho+\delta}$. By a proper choice of the unit of measurement of the mismatch indicator x, the following normalization can be applied without loss of generality:

Normalization: $\kappa = 1$

This normalization requires an adjustment in the parameter γ to account for the redefinition of the unit of measurement of x.[4] Apart from this adjustment, all other parameters and variables are unaffected. We apply this normalization in what follows for the sake of notational convenience. Substitution of (11) into condition (10) gives:

$$G(x) = 1 - \frac{\bar{x} - x}{(1 + \psi x)\,\bar{x}},\ g(x) = \frac{1 + \psi \bar{x}}{\bar{x}\,(1 + \psi x)^2}. \tag{12}$$

2.5 Wage formation

The wage formation processes are the same as in Gautier, Teulings, and Van Vuuren (2010). Since the model is symmetric around $x = 0$, we can focus on the analysis of $W(x)$ for $x \geq 0$.

Wage setting with commitment

When firms can commit on future wage payments, the optimal wage policy of the firm maximizes the expected value of a vacancy,

$$W(x) = \arg\max_{W} \left(\left[u + \psi (1 - u)\, \widehat{G}(W) \right] \frac{Y(x) - W}{\rho + \delta + \psi \lambda F(W)} \right), \tag{13}$$

where $\widehat{G}[W(x)] \equiv 1 - G(x)$ is the distribution of wages among employed workers and where $F[W(x)] \equiv 1 - \frac{x}{\bar{x}}$ is the wage offer distribution, using the fact that the distribution of x is uniform by assumption. The effect of $\widehat{G}(W)$ on the optimal wage offer is the hiring premium, the effect of $F(W)$ is the no-quit premium. The first order condition of this problem reads:

$$W_x(x) = -2 \frac{\psi}{1 + \psi x} [Y(x) - W(x)]. \tag{14}$$

This differential equation can be solved analytically for $W(x)$, using equation (5) as an initial condition, see Appendix A.4.

[4] Let $x_{\kappa=1}$ be the normalized version of the mismatch indicator x, $x_{k=1} \equiv \kappa x$ and let $\overline{x}_{\kappa=1} \equiv \kappa \overline{x}$. Define $\gamma_{\kappa=1} \equiv \kappa^{-2}\gamma$. Then:

$$Y(x) = 1 - \frac{1}{2}\gamma x^2 = 1 - \frac{1}{2}\gamma_{\kappa=1} x_{\kappa=1}^2,$$
$$u = 1/(1 + \overline{x}_{\kappa=1}).$$

We suppress the subscript $\kappa = 1$ in what follows.

Wage setting without commitment

When firms cannot commit on future wage payments, hiring premiums are non-credible and. Hence, the term $\widehat{G}(W)$ in equation (13) is replaced by $1 - G(x)$ reflecting that the first term in brackets does not depend on the wage and that the wage maximizes the value of a filled job rather than the value of a vacancy. Then, the first order condition reads,

$$W_x(x) = -\frac{\psi}{1+\psi x}\left[Y(x) - W(x)\right]. \tag{15}$$

The only difference with equation (14) is a factor two, reflecting the fact that firms pay hiring and no quit premia in the case of commitment, while they pay only a no quit premium in the case without commitment. Again, an analytical solution for $W(x)$ is available, see Appendix A.4.

2.6 Characterization of the equilibrium

The shape of the wage function

Figure 3 depicts $Y(x)$ and $W(x)$ both for the case with and without commitment, setting the value of non market time at $B = 0.6$ (which we do in all subsequent plots). We motivate the choice for B in section 3.2. Contrary to $Y(x)$, $W(x)$ is non-differentiable at $x = 0$. This is due to the hiring and no-quit premiums that firms pay. Since the density of employment is highest for low values of $|x|$, the elasticity of labor supply is high for these types of job. A slight variation in wages has large effects both on the probability that workers accept an outside job offer and on the number of workers who are prepared to accept the wage offer (the latter being relevant in the case with commitment only). Hence, firms will bid up wages aggressively for these types of jobs.

Figure 3 shows that the wage in the optimal assignment is higher when firms can commit than when they cannot, since the ability to commit increases competition between firms for workers. Figure 3 also reveals that for $x = 0$ the slope of the wage function is smaller (in absolute value) for the case with than without commitment. This is remarkable, since the only difference between the expressions for the slope is a factor 2 in the differential equations for wages for the case with commitment, compare equation (14) and (15). However, the slope is proportional to the flow profits per worker, i.e. the difference between the productivity and wages, $Y(x) - W(x)$ which is more than twice as large

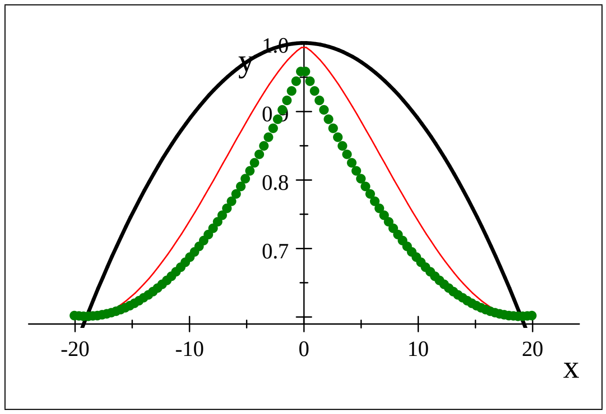

Figure 3: Productivity $Y(x)$ (thick) and wages $W(x)$ with (thin) and without (dotted) commitment

in the case without commitment. The latter effect dominates yielding a steeper wage function in that case.

The distribution of x among employed workers

Figure 4 depicts the density and distribution function of the mismatch indicator $|x|$ conditional on employment, for the case $u = 5\%$ and $\psi = 0.5$ (we use these values in all subsequent plots, unless stated otherwise). For a given unemployment rate, $G(x)$ is identical with or without commitment since workers climb the job ladder equally fast in both cases (because wages are in both cases strictly decreasing in x). The main message from Figure 3 is that the distribution of $|x|$ has a large probability mass close to zero (the optimal assignment) and a long right tail of bad matches. The median value of x is equal to $(1-u)/(1+u) < 1$, far smaller than $\overline{x} = (1-u)/u = 24.0$ (the mismatch indicator in the worst match). The reason for this pattern is that workers who are matched badly quit their jobs fast. The reverse holds for good matches, so their density is high. The skewness of the distribution of $|x|$ has a number of counter intuitive implications for the wage distribution that are spelled out in greater detail below.

The impact of ψ on wage differentials

What wage differential and output loss is consistent with a particular value of ψ and γ (and the corresponding unemployment rate)? The max-mean wage differential is hardly affected by ψ while the mean-min ratio is highly sensitive to it. The reason why ψ does not

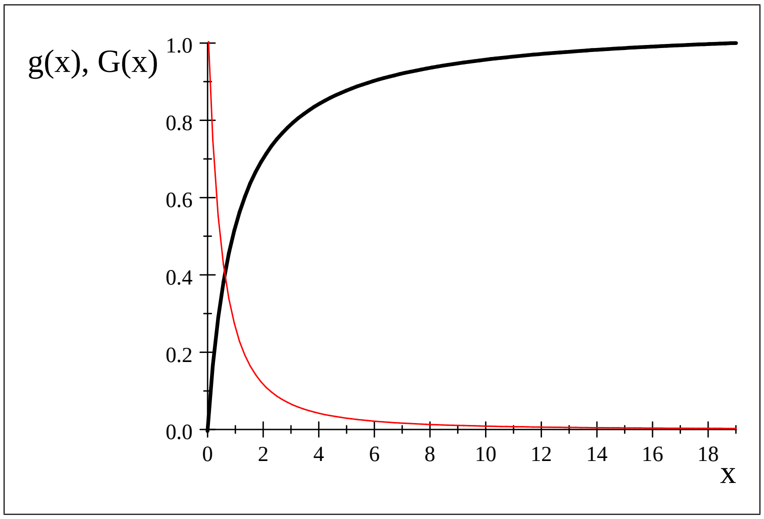

Figure 4: The distribution (thick black) and density function (thin red) of x conditional on employment

matter for the max-mean wage differential is that lowering the value of ψ while keeping u constant has two offsetting effects on wage differentials near the optimal assignment. The density of x at the optimal assignment is equal to $g(0) = \psi + u/(1-u)$, see equation (12). Hence, a lower value of ψ implies that there are fewer workers close to the optimal assignment since search by employed workers is less efficient and since employed job seekers are a particularly relevant source of labour supply for an $x = 0$. This reduces the mean wage. However, unemployed job seekers also become more choosy if ψ goes down since they give up a share $1-\psi$ of the option value of search by accepting a job. Therefore, the lowest wage $W(\overline{x})$ increases and consequently the mean wage increases a bit as well; the total effect is that wage differentials become smaller. For the mean-max differential, the negative and positive effects on the mean wage almost cancel while the maximum wage is unaffected since there are always some workers around their optimal job.

The max-mean differential on the one hand and the mean-min and max-min wage differentials on the other hand also tell opposite stories about the effect of commitment on wage differentials. Commitment makes firms compete more fiercely for workers, driving up the maximum wage. Since the minimum wage is the same for the case with and without commitment, this implies that the mean-min ratio is larger under commitment. However, since the slope of the wage function close to the optimal assignment is smaller under commitment, see Figure 3, the max-mean differential can be smaller with commitment.

Wrapping up

Both versions of the model -commitment and no commitment- yield two analytical relations between the expected wage and the productivity loss due to search frictions on the one hand and the parameters B and ψ and the unemployment rate u on the other hand:

$$W(0) - \mathrm{E}_x W = \widetilde{W}(u; B, \psi), \qquad (16)$$
$$Y(0) - \mathrm{E}_x Y = \widetilde{Y}(u; B, \psi) = \frac{1}{2}\gamma \mathrm{Var}[x],$$

where $W(0) - \mathrm{E}_x W$ is the expected wage loss compared to the wage in the optimal assignment, $W(0)$, and mutatis mutandis the same for the expected productivity loss $Y(0) - \mathrm{E}_x Y$. These relations are derived in Appendix A.4. Hence, keeping B and ψ constant, there is an increasing relation between wage differentials $W(0) - \mathrm{E}_x W$ and the productivity loss $Y(0) - \mathrm{E}_x Y$:

$$\frac{d[W(0) - \mathrm{E}_x W]}{d[Y(0) - \mathrm{E}_x Y]} = \frac{\widetilde{W}_u(u; B, \psi)}{\widetilde{Y}_u(u; B, \psi)} > 0. \qquad (17)$$

A higher unemployment rate goes hand in hand with both more wage dispersion and a larger productivity loss since all are manifestations of search frictions. In the next section, we set out a method to obtain empirical estimates for $W(0) - \mathrm{E}_x W$ and $Y(0) - \mathrm{E}_x Y$. Conditional on the values for B and ψ, these estimates imply a value for the unemployment rate u by equation (17). When this unemployment rate is in line with the natural rate of unemployment, this fact can be interpreted as empirical evidence in favour of the model. This test is the topic of the next section.

3 Empirics

3.1 The hierarchical model and the value of γ

Our strategy for the estimation of $W(0) - \mathrm{E}_x W$ or $Y(0) - \mathrm{E}_x Y$ is based on an empirical proxy of the mismatch indicator x. This mismatch indicator is derived from information on workers' skills s and the complexity of jobs c. The mismatch indicator is defined as the difference between both, $x = s - c$. Since the data are hierarchical by nature we have to adjust the model accordingly. A key step in our derivation is the interpretation of the parameter γ. This parameter measures the second derivative of the production function with respect to workers' skill level;. In other words, it measures the curvature of Rosen's "kissing" offer and utility curves, see Figure 2. We can therefore establish a relation between our model of search frictions and the elasticity of substitution between low and high skilled workers. It turns out that there exists a direct correspondence between γ and the elasticity of substitution between low and high skilled workers, see Teulings (2005) and Teulings and Van Rens (2008). The hierarchical counterpart of our production function is $\widehat{Y}(s,c)$ which denotes the productivity of an s-type worker in a c-type job, where the hat above the function distinguishes it from the single argument function $Y(x)$ of the circular model. Instead of equation (1), $\widehat{Y}(s,c)$ satisfies the following relation:

$$\ln \widehat{Y}(s,c) = s - \frac{1}{2}\gamma (s-c)^2, \tag{18}$$

This production functions has two features, reflected by the two terms on the right hand side. First, better skilled workers haven an absolute advantage over less skilled workers (the first term s). Hence, better skilled workers receive higher wages when employed in their optimal assignment.[5] Second, better skilled workers have a comparative advantage in more complex jobs (the second term $-\frac{1}{2}\gamma(s-c)^2$). Hence, in a Walrasian equilibrium, better skilled workers will be employed in more complex jobs.[6] A second order Taylor expansion of the second term of (18) in the optimal assignment $x = s - c = 0$ is equal to

[5]Absolute advantage of better skilled workers applies only locally for this specification, since for high values of s the second order term dominates. Teulings and Gautier (2004) use the production function:

$$\ln \hat{Y}(s,c) = -\frac{1}{\gamma} e^{-\gamma(s-c)} + h(c),$$

where $h(c)$ is some arbitrary function. This function features global absolute advantage: $\hat{Y}_s(s,c) > 0$ for any s and c. Equation (18) can be interpreted as a second order Taylor expansion of this relation in the market equilibrium where $s = c$ with $h(c) = c + \gamma^{-1}$:

$$\ln \hat{Y}(s,c) = s - \frac{1}{2}\gamma (s-c)^2 + O\left[(s-c)^3\right].$$

[6]Formally, comparative advantage implies:

$$\widehat{Y}_{sc}(s,c)\,\widehat{Y}(s,c) > \widehat{Y}_s(s,c)\,\widehat{Y}_c(s,c),$$

equation (1): $Y''(0)/Y(0) = \widehat{Y}_{cc}(c,c)/\widehat{Y}(c,c) = \gamma$.

Let $c(s)$ be the assignment of worker type s in a Walrasian economy without search frictions. In this economy, everybody is assigned to her optimal assignment, that is, the c-type that yields the highest output $Y(s,c)$. Hence, $c(s) = s$ and the second term of equation (18) vanishes. At first sight the linearity of equation (18) in s seems to be a serious limitation to its generality. However, Gautier and Teulings (2006) show that it is not. Since we have not yet defined the units of measurement of s yet, the linearity of the first term is just a matter of a proper scaling of the skill index. By a similar argument, the fact that equation (18) is constructed such that the optimal assignment is characterized by the simple identity $c(s) = s$ instead of a more general function, is not a restriction to the model, but just a matter of proper scaling of the complexity index c. Hence, the restrictive nature of equation (18) is not in the first but in the second term, namely that the coefficient of the second term, γ, does not vary with s. Implicitly, equation (18) defines the units of measurement of s and c (and consequently of the mismatch indicator $x = s - c$).[7] The Mincerian return to the skill index $d\ln Y/ds$ is equal to one in the optimal assignment, where $s = c$. Later on, we use this normalization when constructing our measures for s and c.

In our discussion of the model, we keep constant the distribution of worker-skills and job-complexities at such a level that the Mincerian rate of return to the skill index is equal to unity and that the optimal assignment of worker type is $c(s) = s$. In reality, the distribution of s and c might shift over time or differ between regions. For example, when the average skill level of the workforce increases over time while the distribution of product demand over c-types remains constant, the equilibrium between demand and supply of each c-type requires the optimal assignment and the Mincerian return to human capital to adjust: workers of a particular s type will be assigned to less complex jobs and the Mincerian rate of return to skill will go down. The formal derivation of this argument is relegated to Appendix A.5. This mechanism relates our model to the literature on the

the assignment of worker types s to job types c, $c(s)$ is strictly increasing in a Walrasian economy without search frictions, see Teulings (2005) for a more formal treatment.

[7]The fact that for the empirical analysis we choose a convenient unit of measurement of the x (such that $dw/ds = 1$) implies that the normalization $\kappa = 1$ no longer applies. However, we don't need this normalization in the subsequent argument.

change of relative wages due to shifts in the distribution of human capital, see Katz and Murphy (1992). They estimate the elasticity of complementarity between high and low skilled workers $\eta_{\text{low-high}}$ to be 1.4: a 1% increase in the ratio between high and low skilled workers yields a 1.4% fall in the relative wages of high skilled workers. In Appendix A.5 we show that there is a one-to-one correspondence between this elasticity $\eta_{\text{low-high}}$ and the parameter γ:

$$\gamma = \frac{1}{\text{Var}\left[\ln W\right]\eta_{\text{low-high}}} \cong \frac{1}{0.36 \times 1.4} \cong 2.$$

3.2 The value of B and ψ

Hall and Milgrom derive a value for B based on UI benefits (of 0.25) and an estimated a Frisch elasticity of labor supply of 1. This implies a value of $B = 0.71$. Since micro studies (see the discussion in Hall, 2009) typically find a somewhat smaller labor supply elasticity, we set B at 0.6.[8] We will however also do simulations with $B = 0.71$.

The value of ψ is identified from the relation between the ratio of the employment-to-employment, f_{ee}, and the unemployment-to-employment-hazard rate, f_{ue}. In Appendix A.3 we derive that

$$\frac{f_{ee}}{f_{ue}} = \frac{u}{1-u}\left[\frac{u+\psi\left(1-u\right)}{\psi\left(1-u\right)}\ln\left(\psi\frac{1-u}{u}+1\right)-1\right]. \tag{19}$$

Unfortunately, the evidence on f_{ee} and f_{ue} for the US yields values of ψ that range from above 1, see Nagypal (2008), to values close to 0, see Hornstein et.al. (2010). What explains this wide range? [9] Hornstein et al. (2011) use a similar relation as we derived but then onf_{ee}/f_{eu}. They use Shimer's (2005) estimate of the employment-to-unemployment

[8]Hagedorn and Manovskii (2008) and Hall (2009) want to explain the cyclical behavior of unemployment so they use larger values for B. For these studies, the value of non-market time of the *marginal worker* is relevant whereas here we are interested in the value of non-market time for the *average worker* so a lower value of B is justified.

[9]The value of f_{ee} is 2.7% according to Fallick and Fleischman (2004), 2.9% according to Nagypal and after a correction for missing records in the CPS, Moscarini and Vella (2008) get a value of 3.3%. Nagypal's values come from the SIPP and the CPS. She argues that those estimates are downwardly biased because when workers find a new job when employed but there is a lag in starting this new job, it is not uncommon to briefly register to be unemployed. In the data this implies an employment-unemployment transition followed by an unemployment-employment transition. She argues that this bias is larger than the time aggregation bias in the unemployment outflow rate (some workers loose and find a job between the interview dates).

hazard rate f_{eu}, which equals 3% per month. This high value of f_{eu} is problematic. According to the BLS statistics, median tenure is 53 months. In the absence of duration dependence and ignoring the flow out of the labor forces, this implies that the total hazard out of the current job, $f_{ee}+f_{eu}$, is 1.3%.[10] The transition rate f_{eu} (the equivalent of δ in our model) is assumed to be constant in our model, while the unconditional transition rate, f_{ee}, exhibits negative duration dependence due to heterogeneity in the match quality x: high quality matches survive. Negative duration dependence implies that the hazard rate for low-tenure workers is above 1.3% and the rate for high tenure workers is below 1.3%. Since f_{eu} is constant, it must be smaller than 1.3%, much lower than the value reported by Shimer. This implies that the assumption of the absence of duration dependence of f_{eu} is rejected by the data. Apparently, a small group of weakly attached workers frequently flow between un- and employment. In order to capture this feature of reality, other mechanisms have to be introduced, like learning, see Moscarini (2005), or random growth, see Buhai and Teulings (2006). This falls outside the scope of this paper. Hence, our model is ill suited to explain this particular feature of reality. Reasonable estimates of f_{ee}/f_{ue} vary from 0.076 to 0.132.[11] A value for u of about 5% and ψ of 0.5 yields $f_{ee}/f_{ue} = 0.085$ while a value of ψ of 0.75 yields $f_{ee}/f_{ue} = 0.100$. In our main analysis we set $\psi = 0.5$ but we also run simulations for $\psi = 0.75$. Those values are not sensitive to small changes in u.

3.3 Inference on wages and mismatch

This section describes how we can use data on wages and worker and job characteristics to construct empirical measures of s and c, and hence of the mismatch indicator $x = s - c$, and how we can use these data for inference on the magnitude of wage differentials $W(0) - \mathrm{E}_x W$. The challenging aspect of this exercise is that any empirical measure of s and c will be contaminated by measurement error. Hence, when we use these measures to construct a measure of mismatch $x = s - c$, the observed mismatch can be either

[10] Ignoring flows out of the labor force, the total hazard outof employment can be solved from $1 - \exp[-53(f_{ee} + f_{eu})] = 0.5$.

[11] Monthly transition rates from 1967-2010 similar to Shimer (2007) imply that $f_{ue} = 38\%$ for the mean worker, and hence $f_{ee}/f_{ue} = \frac{2.9}{38} = 0.076$. Tenure data imply $f_{ue} = 22\%$ for the median worker, and hence $f_{ee}/f_{ue} = \frac{2.9}{22} = 0.132$. We thank Bart Hobijn for sharing his data.

the result of measurement error in s or c, or it can be a reflection of real mismatch, or any combination of these two. Even in a perfect Walrasian equilibrium, we would still observe "mismatch" due to measurement error though in fact everybody is assigned to her optimal assignment. The observed variation in $s-c$ therefore exceeds its true variation. Our procedure should therefore explicitly allow for measurement error. This subsection sets out the main line of the argument, using Taylor expansions of the wage function. In the next subsection we provide numerical simulations of the exact relation between our empirical statistics and the model and check the precision of these approximations.

The wage function that goes with the production function (18) of the hierarchical model is:

$$\ln \widehat{W}(s,c) = s + \ln W(s-c),$$

where again we use the hat to distinguish this function from the single argument function of the circular model. The first term s measures the Walrasian, hierarchical part of the model: better skilled workers earn higher wages. The second part measures the wage loss due to search frictions as described by the function $W(x)$ derived in Section 2. Since the function $W(x)$ is symmetric and non-differentiable around $x=0$, a first order Taylor expansion of this function reads,

$$\hat{w}(s,c) \cong \omega_0 + \omega_1 s - \omega_2 |s-c|, \tag{20}$$

where w denotes the log wage in deviation from its mean, $\omega_0 > 0, \omega_1 > 0$ and $\omega_2 > 0$ and s is orthogonal on $s-c$. The parameter ω_2 is interpreted as an estimate of $d\ln W(x)/dx_{|x=0} = W_x(0)/W(0)$. We conveniently define skill indexes s and c in deviation from their mean. Hence, if there were no search frictions, so that $s=c$ for all jobs, then the intercept ω_0 would be equal to zero. In the presence of search frictions, the intercept measures the expected wage loss due to frictions:

$$\omega_0 = \omega_2 \mathrm{E}[|s-c|] \cong \ln W(0) - \mathrm{E}_x \ln W \gtrsim \ln W(0) - \ln \mathrm{E}_x W \gtrsim W(0) - \mathrm{E}_x W.$$

The first equality follows from taking expectations at the left and right hand side of (20) and noting that $\mathrm{E}[\hat{w}(s,c)] = \mathrm{E}[s] = 0$. The second equality follows from evaluating (20) at $x \equiv s-c=0$. The next inequality is due to Jensen's inequality, $\ln\mathrm{E}_x W > \mathrm{E}_x \ln W$. For the final step, note that for small search frictions, and hence small wage differentials,

$W(x) \lesssim W(0) \lesssim 1$, so that the approximation $\ln W \lesssim W - 1$ applies. The idea is to use the intercept ω_0 as a proxy for the max-mean difference, $W(0) - \mathrm{E}_x W$. The numerical simulations that we will present in section 3.5 provide a test of the precision of this approximation. We find that almost no precision is lost by using (20). Furthermore, since $Y(0) - \mathrm{E}_x Y = \frac{1}{2}\gamma \mathrm{Var}[x]$ by equation (16), the variance of x provides an estimate of the productivity loss due to search frictions. These estimates would then allow us to evaluate equation (17). The only remaining problem is how to deal with measurement error.

Suppose we estimate equation (20) by simply replacing the true mismatch indicator $x = s - c$ by the signal $\widehat{x} = \widehat{s} - \widehat{c}$. A bar on top of a variable denotes its observed part. Let ε_x be the measurement error in $\widehat{x}$, where ε_x is assumed to have zero mean and to be uncorrelated with x: $\widehat{x} = x + \varepsilon_x$, and hence: $\mathrm{Var}[\widehat{x}] = \mathrm{Var}[x] + \mathrm{Var}[\varepsilon_x]$. Even if $-\omega_2 |s - c|$ were a perfect approximation of $\ln W(s - c)$, the estimated value of ω_2 will be a biased estimator of $\omega_2 \mathrm{E}[|s - c|]$ due to the strong convexity of the wage function at $x = 0$, since the expected value of $|x|$ conditional on its observed part $\widehat{x}$ is always greater or equal to $|\widehat{x}|$:

$$\mathrm{E}\left[|x| \, |\widehat{x}\right] \geq |\widehat{x}| \, .$$

The closer $\widehat{x}$ is to zero, the stronger this inequality. This is documented in Figure 5, where we present three functions, $|\widehat{x}|$, $\mathrm{E}[|x| \, |\widehat{x}]$, and the least squares estimation of $\mathrm{E}[|x| \, |\widehat{x}] = \beta_0 + \beta_2 \widehat{x}^2 + \varepsilon$, for the case that both the true value x and measurement error ε_x are normally distributed with equal variance equal to unity. The least square approximation of $\mathrm{E}[|x| \, |\widehat{x}]$ is fairly precise for $|\widehat{x}| < 2.5$. Since the variance of $\widehat{x}$ is normalized to unity, this covers the complete relevant range. The following proposition relates the least squares coefficients to the underlying structural model, see Appendix 7 for the proof:

Proposition 1 *Suppose that the true model reads:*

$$w = \omega_0 - \omega_2 |x| \, ,$$

where both w and x are normalized to have a zero mean. Suppose that we observe only $\widehat{x} = x + \varepsilon_x$ where both x and ε_x are distributed normally with $Var[x] \, / \, Var[\widehat{x}] \equiv R$. Suppose that we estimate the following model by OLS:

$$w = \overline{\omega}_0 - \overline{\omega}_2 \widehat{x}^2 + \varepsilon,$$

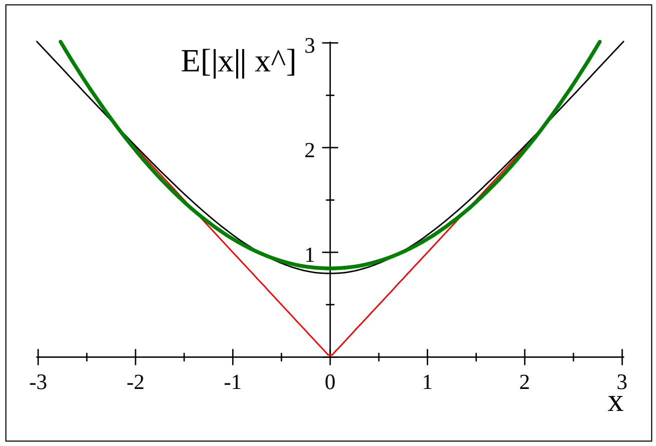

Figure 5: Smoothing of an absolute value function by random mixing for $\sigma_x^2 = \sigma_{\hat{x}}^2 = 1$: $|x|$ (thin), $\mathrm{E}[|x|\,|\hat{x}]$ (thin), least squares estimate (thick)

where ε is a zero mean error term. Then:

$$p \lim \overline{\omega}_0 = \frac{1}{2} R\omega_2 E\left[|x|\right].$$

Hence, when there is no measurement error in the observed signal $\widehat{x}$, so that the signal-noise ratio, $R = 1$, $\overline{\omega}_0$ is equal to half the expected wage loss due to frictions. This underestimation by a factor two is due to the imperfect approximation of the absolute value by the quadratic specification. When on top of this imperfection in the functional form, there is also measurement error in the signal $\widehat{x}$, the underestimation becomes even more severe. The estimate of $\overline{\omega}_0$ is proportional to the signal-to-noise ratio R.

This justifies the idea of approximating the underlying model (20) by a regression model of w with a quadratic term $(\widehat{s} - \widehat{c})^2$:

$$w = \overline{\omega}_0 + \overline{\omega}_1 \widehat{s} - \overline{\omega}_2 \left(\widehat{s} - \widehat{c}\right)^2 + \varepsilon. \tag{21}$$

Under the assumption of joint normality of $\widehat{x}$ and ε_x, Proposition 1 implies the following relation between $W(0) - \mathrm{E}_x W$ and the estimated parameters $\overline{\omega}_0$ and $\overline{\omega}_2$ (see Appendix A.6),

$$\overline{\omega}_0 = \frac{1}{2} R\omega_2 \mathrm{E}\left[|x|\right] \cong \frac{1}{2} R\left[W(0) - \mathrm{E}_x W\right],$$

The higher the variance of the measurement error ϵ_x the smaller the signal-to-noise ratio R, the smaller the estimated value of $\overline{\omega}_0$ and the more $W(0) - \mathrm{E}_x W$ is underestimated by

$\overline{\omega}_0$. Furthermore, the productivity loss due to search frictions satisfies

$$Y(0) - \mathrm{E}_x Y = \frac{1}{2}\gamma R \mathrm{Var}[\widehat{x}].$$

When $\mathrm{Var}[\widehat{x}]$ is used as a proxy for $\mathrm{Var}[x]$, the productivity loss is overestimated by a factor R^{-1}, exactly the reverse of the underestimation of the expected wage loss compared to the wage in the optimal assignment. We can eliminate the signal noise ratio from these two expressions to obtain a relation between the expected wage loss and the expected productivity loss:

$$\gamma\overline{\omega}_0 \mathrm{Var}[\widehat{x}] \cong [W(0) - \mathrm{E}_x W][Y(0) - E_x Y] \tag{22}$$

We have estimates for all variables on the left hand side of this approximate equality. Apart from the Taylor expansion in equation (20), this expression relies on the approximate normality of the distribution of x.

Figures 6 and 7 depict the equality version of equation (22) and equation (17) from section 2. Equation (22) has a negative slope, equation (17) has a positive slope, so there is at most one intersection. The locus of equation (22) is based on $\gamma = 2$ and the estimated values of $\overline{\omega}_0 = 0.0241$ and $\mathrm{Var}[\widehat{x}] = 0.0306$, see the discussion in the next section. Given the values for $B = 0.60$ and $\psi = 0.5$, there is a one-to-one correspondence between every point on the curve associated with equation (17) and the unemployment rate u. Hence, this intersection implies a value for the unemployment rate. Note that there are two versions of equation (17), one for the case where firms can commit to pay hiring premiums and one where they cannot. The intersection for the case with commitment corresponds to u is 5% (for the case without commitment we find similar values).

Table 1 presents the theoretical relation between (true) wage dispersion, mismatch and unemployment for increasing frictions as predicted by the model. So for $B = 0.6$ and given an unemployment rate of 4%, the table shows that the difference between the maximum and mean wage is 3.47%.and the average worker produces 1.32% less than he would at his optimal job type.

3.4 Measuring wages and mismatch

For the construction of empirical measures for the observed part of workers' skill level $\widehat{s}$ and the level of job complexity $\widehat{c}$ we use a methodology spelled out in Gautier and

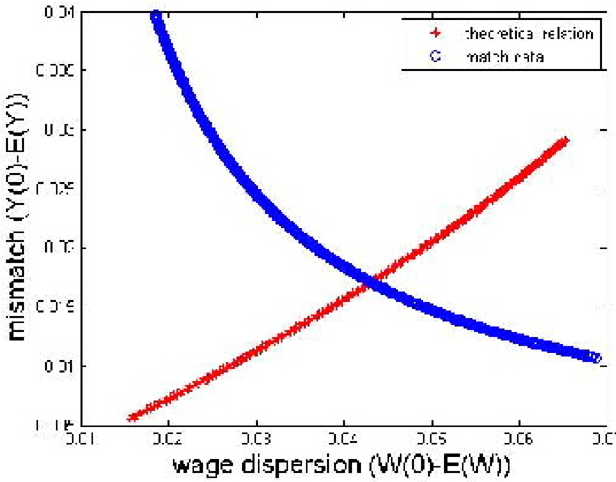

Figure 6: Identification

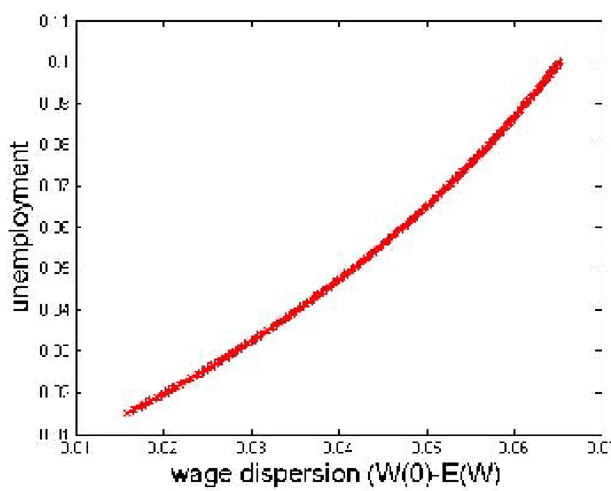

Figure 7: Unemployment and wage dispersion

$u(\%)$	B	$[W(0) - \mathrm{E}_x W]\,(\%)$	$[Y(0) - E_x Y]\,(\%)$
4.00	0.4	5.24	2.01
	0.6	3.47	1.32
5.00	0.4	6.18	2.43
	0.6	4.09	1.61
6.00	0.4	7.03	2.84
	0.6	4.66	1.88

Table 1: Wage dispersion, Mismatch and Unemployment for $\psi = 0.5$

Teulings (2006). We apply this methodology to data for the United States taken from the March supplements of the CPS 1989-1992, see Gautier and Teulings (2006), for details:

$$w = \vec{x}'\vec{\beta} + \varepsilon_s + \varepsilon_w, \tag{23}$$

$$w = \vec{z}'\vec{\alpha} + \varepsilon_c + \varepsilon_w,$$

where $\vec{x}$ and $\vec{z}$ are vectors of observed worker and job characteristics respectively,[12] where ε_w captures measurement error in log wages, and where ε_s and ε_c capture (i) unobserved characteristics of workers and jobs respectively and (ii) the effect non-optimal assignment on wages. It is convenient to normalize our data on $\vec{x}$ and $\vec{z}$ such that they have zero mean. Since we have defined w to have a zero mean too, it does not make sense to include a constant in this regression. The estimated parameter vector can then be used to construct indices for the observed worker and job characteristics:

$$\widehat{s} = \vec{x}'\vec{\beta},$$

$$\widehat{c} = \vec{z}'\vec{\alpha}.$$

Again, both indices have zero mean by construction.[13] So, the skill measure is the predicted wage conditional on worker characteristics and the job complexity level is the predicted wage conditional on job characteristics. The skill measure satisfies our normalization of the Mincerian rate-of-return-to-skill index ($dw/ds = 1$), see equation (18).

[12] We apply the following personal characteristics: gender, total years of schooling, a third-order polynomial in experience, highest completed education, being married, having a full- or part-time contract as well as various cross terms of experience, education, and being married. As job characteristics, 520 occupation and 242 industry dummies are applied.

[13] We also normalize $\overline{s}$ and $\overline{c}$ such that in a regression: $w = \beta_1 \overline{s} + \beta_2 \overline{s}^{\ 2} + \varepsilon_w$, $\beta_2 = 0$ and the same for $\overline{c}$, see Gautier and Teulings (2006) for details.

Next, we use these indices and estimate

$$w = \overline{\omega}_0 + \omega_s \widehat{s} + \omega_c \widehat{c} - \omega_{ss} \widehat{s}^2 + 2\omega_{sc} \widehat{s}\widehat{c} - \omega_{cc} \widehat{c}^2 + \varepsilon, \tag{24}$$
$$w = \overline{\omega}_0 + \omega_s \widehat{s} + \omega_c \widehat{c} - \omega_2 \left(\widehat{s} - \widehat{c}\right)^2 + \varepsilon.$$

The second regression imposes two restrictions $\omega_{ss} = \omega_{cc} = \omega_{sc}$. At first sight, this equation seems inadequate to capture model (21). The model includes first order terms for both worker and job characteristics, $\widehat{s}$ and $\widehat{c}$, where model (21) has only a first order term for the worker's skill s. However, since worker and job characteristics are correlated and since worker characteristics are only partially observed, observed job characteristics will serve as a proxy for unobserved worker characteristics and vice versa, so that we can expect both ω_s and ω_c to be positive. The problem of establishing the "structural" value of ω_s and ω_c has fascinated labor economists ever since the publication of Krueger and Summers' (1998) seminal paper on efficiency wages, see e.g. Bound and Katz (1988) or Abowd, Kramarz, and Margolis (1998). The issue is whether ω_c truly measures the effect of job characteristics, or whether it is merely a proxy for unobserved worker characteristics, see Gautier and Teulings (2006) and Eeckhout and Kircher (2011) for a more elaborate discussion. For now, we adhere to equation (20), which does not allow for a 'true' first order effect of job characteristics. Estimating (24) yields

$$\begin{aligned} w &= \underset{(8.86)}{0.0125} + \underset{(182.4)}{0.61}\widehat{s} + \underset{(207.7)}{0.66}\widehat{c} - \underset{(21.2)}{0.17}\widehat{s}^2 - \underset{(21.6)}{0.17}\widehat{c}^2 + \underset{(36.6)}{0.43}\widehat{s}\widehat{c}, \\ w &= \underset{(14.66)}{0.0241} + \underset{(182.2)}{0.61}\widehat{s} + \underset{(207.5)}{0.66}\widehat{c} - \underset{(35.11)}{0.2007}\left(\widehat{s} - \widehat{c}\right)^2, \\ \mathrm{Var}\left[\widehat{x}\right] &= 0.0306 \end{aligned} \tag{25}$$

Hence, $\overline{\omega}_0 = 0.0241$. Note that the sign restrictions are satisfied for all three second order terms. The F-test[14] rejects the restrictions $\omega_{ss} = \omega_{cc} = \omega_{sc}$ due to the large number of observations that we use but the difference in R^2 between the restricted and the non-restricted model is only 0.0003.

What legitimizes us to interpret the first order term ω_c as capturing unobserved heterogeneity in workers' skills and why can we not interpret the second order terms in the same manner? Why would second order terms really capture the concavity of the wage

[14] $\frac{\left(R_u^2 - R_{restr}^2\right)/2}{(1-R_u^2)/(222179-6)} = \frac{(0.4479-0.4476)/2}{(1-0.4479)/(222179-6)} = 60.4.$

function in the mismatch indicator x and not unobserved components in either s or c? Gautier and Teulings (2006) argue that while the argument of unobserved heterogeneity applies for ω_c, it is much less likely to apply to the second order terms. They provide three arguments. First, when observed and unobserved worker and job characteristics are distributed jointly normal, it is impossible for second order terms to be a proxy for the unobserved component of a first order term, because the correlation of a second order term in $\widehat{s}$ and/or $\widehat{c}$ with the unobserved skill index is a third moment, and third moments of a normal distribution are equal to zero. A simple empirical test for this assertion is that including the second order terms should not affect the first order terms, which turns out to be the case. Second, the interpretation of these coefficients as capturing the concavity of the wage function implies sign restrictions, $\omega_{sc} > 0, \omega_{ss} < 0$, and $\omega_{cc} < 0$, which are met for all three coefficients for all six countries for which the model is estimated in Gautier and Teulings (2006). There is no reason why these restrictions should hold when the second order terms are to be rationalized from unobserved worker heterogeneity. A final argument relies on the grouping worker characteristics in $\widehat{s}$ and job characteristics in $\widehat{c}$.[15] If the significance of the second order terms is indeed driven by the concavity of the wage function in the mismatch indicator, then their sign would depend on the vectors $\vec{x}$ and $\vec{z}$ capturing worker and job characteristics respectively. If both vectors were composed out of mixtures of job and worker characteristics, e.g. experience and occupation dummies in $\vec{x}$ and education and industry dummies in $\vec{z}$, then the concavity result should not come out. Table 2 shows that for both alternative combinations, putting education and occupation in $\vec{x}$ or putting education and industry in $\vec{x}$ (and the remaining variables in $\vec{z}$), the concavity result becomes either much weaker or even breaks down. Hence, the concavity result only survives when worker and job characteristics are separated in two groups and it is not a statistical artifact.

3.5 Numerical simulations

The derivation of equation (22) in section 3.3 relies upon Taylor approximations and the approximation of the distribution of x by a normal. In this section, we simulate or model

[15] We thank Jean Marc Robin for the idea of this test.

$\widehat{s}$ includes:	$\widehat{s}$	$\widehat{c}$	$\widehat{s}^2$	$\widehat{c}$	$\widehat{sc}$
educ, occ	0.517	0.650	-0.009	-0.037	0.094
(t)	(132.6)	(175.24)	(-0.86)	(-4.10)	6.17)
experience, ind	0.324	0.805	-0.050	0.010	0.053
(t)	(79.9)	(233.2)	(-4.44)	(1.10)	(3.46)

Table 2: Test of concave relation between wages and s and c

with the exact expressions and the true distribution of x. In Appendix A.7 we describe how we calibrate our model, conditional on the values of $B = 0.6, \gamma = 2$ and $\psi = 0.5$ as discussed in the previous sections. Basically, we add noise to the model and calibrate the unemployment rate u and the signal-noise ratio R to match the estimated mismatch $\mathrm{Var}[\widehat{x}] = 0.0306$ and intercept $\overline{\omega}_0 = 0.0241$ (the match is exact).

This procedure yields an unemployment rate of 5.0% for the case with commitment and 4.7% for the case without commitment for our base line parameters. We view those as reasonable estimates of the natural rate of unemployment. Tables 3 and 4 show the implied values of unemployment, the signal-noise ratio (R), worker flows, wage dispersion measures (both the max-mean difference and Hornstein et al.'s (2011)mean-min ratio) and the expected productivity loss due to mismatch, $\mathrm{E}[Y(0) - Y(x)] = \mathrm{E}\left[\frac{1}{2}\gamma x^2\right] = \frac{1}{2}\gamma\sigma_x^2$ of the above procedure for different values of B and ψ.

For the calibrations with $\psi = 0.75$, the reservation wage drops and therefore the mean-min ratio increases but since the mean wage goes up (and the highest wage remains the same), the max-mean ratio goes down. Note that this is not a comparative statics exercise but rather a calibration on observed statistics, since $\overline{\omega}_0$ is an outcome, not a primitive of the model. The Hall-Milgrom calibration of $B = 0.71$ with $\psi = 0.5$ generates an unemployment rate of 10% which seems too high. For $\psi = 0.75$, u is about 7.5% which is still acceptable. The model also generates substantial wage dispersion even when unemployment is low. Only when B is high and ψ is low, the mean-min ratio is small (see the discussion in Hornstein et al. 2011) but the max mean ratio is still large.

The expected productivity loss due to mismatch ($\frac{1}{2}\gamma\sigma_x^2$) is 1.6% for our baseline case with commitment and 1.7% without commitment. It is stable within a reasonable range of parameter values (of ψ and B), varying between 0.7% and 2.4%.

$\psi = 0.50$	B	$u(\%)$		$\frac{f_{ee}}{f_{ue}}(\%)$		$R(\%)$		$W(0) - E_x W$ x100		$\frac{\mathrm{E}_x W}{W(\overline{x})}$		$\frac{1}{2}\gamma\sigma_x^2(\%)$	
commitment		yes	no	yes	no	yes	no	yes	no	yes	no	yes	no
	0.40	2.1	1.8	5.0	4.5	36.3	32.7	3.1	2.8	1.43	1.47	1.1	1.0
	0.60	5.0	4.7	8.4	8.1	52.9	55.2	4.1	4.3	1.24	1.27	1.6	1.7
	0.71	9.5	10.1	11.8	12.1	68.2	79.6	4.7	5.4	1.15	1.12	2.1	2.4

Table 3: Calibration results for $\psi = 0.5$

$\psi = 0.75$	B	$u(\%)$		$\frac{f_{ee}}{f_{ue}}(\%)$		$R(\%)$		$W(0) - E_x W$ x100		$\frac{\mathrm{E}_x W}{W(\overline{x})}$		$\frac{1}{2}\gamma\sigma_x^2(\%)$	
commitment		yes	no	yes	no	yes	no	yes	no	yes	no	yes	no
	0.40	1.3	1.1	4.2	3.6	23.6	19.7	2.1	1.7	1.80	1.83	0.7	0.6
	0.60	3.4	3.1	7.9	7.5	37.7	35.8	3.2	3.0	1.40	1.42	1.2	1.1
	0.71	7.2	7.8	12.5	13.1	52.7	58.8	4.1	4.5	1.24	1.25	1.6	1.8

Table 4: Calibration results for $\psi = 0.75$

3.6 Unemployment and composition of the output loss

Table 5 shows for our baseline parameters that if firms can commit to wages, the output loss due to search friction is 6.6% while if firms cannot commit, it is 6.7%. So interestingly, the estimated output loss is not very sensitive to whether firms can or cannot commit to wages. If they can commit, a larger share of the loss is due to unemployment and mismatch. In the no-commitment case because of a business-stealing externality (see Gautier et.al. (2010), a larger share is due to vacancy creation cost. The idea is that without commitment, when opening a vacancy, individual firms do not internalize the future output loss of the firm they will poach a worker from. Although the transitions of workers to better matches are always efficient, the expected productivity gains are too small to justify the entry cost of the marginal firm from a social point of view.[16]

Table 5 is not suitable to estimate the business-stealing effect because again it reflects given calibration exercises. In order to estimate the business stealing effect, we keep all primitives constant and compare X with and without commitment. Note that we also no longer can keep κ constant and have to assume an explicit matching function. We take the quadratic of Gautier et al. (2009). Starting from the commitment case with

[16] Elliot (2010) discusses other wage mechanisms in a network framework that also internalize the business-stealing externality.

ψ	0.50						0.75					
B	0.4		0.6		0.71		0.4		0.6		0.71	
Commitment	yes	no	yes	no	yes	no	yes	no	yes	no	yes	no
$u(1-B)$	1.27	1.06	1.99	1.88	2.76	2.94	0.79	0.64	1.36	1.24	2.09	2.26
vK	2.18	1.96	3.08	3.22	3.77	4.38	1.42	1.19	2.23	2.12	2.99	3.32
$(1-u)\left[Y(0)-E_xY\right]$	1.09	0.98	1.54	1.61	1.89	2.19	0.71	0.60	1.12	1.06	1.50	1.66
X	4.53	4.00	6.60	6.71	8.42	9.50	2.93	2.43	4.71	4.43	6.57	7.24

Table 5: Decomposition of output loss due to frictions

frictions such that $u = 5.0\%$ and $X = 6.6$ (which can be shown to be almost first best) and switching to the no commitment case will lead to excessive vacancy creation such that u drops to 2.9% and $X = 7.5\%$. If firms cannot commit, the welfare loss due to the business-stealing externality would therefore be almost 1%. Table 5 also shows that the output loss due to mismatch is about the same in magnitude as the output loss due to unemployment which together are responsible for a bit more than half of the total output loss.

4 Conclusion

Because of frictions only a subset of the contacts between vacancies and workers results in a match and this creates (i) unemployment, (ii) wage dispersion amongst identical workers and (iii) mismatch. Our contribution is that we measure all those manifestations of search frictions (allowing for measurement error) and that we present a model that can jointly explain the observed values. Specifically, we ask our model which unemployment rate corresponds to observed wage dispersion and mismatch. The model predicts an unemployment rate of around 5%, which is evidence for its validity and therefore we feel confident to use it to estimate the output loss due to search frictions.

Search frictions directly generate output losses due to the fact that resources are allocated sub-optimally and indirectly because decentralized wage mechanisms potentially come with distortions. Allowing for two-sided heterogeneity is extremely important because it is the interaction between the search frictions, the type distributions and the production technology that determines how important these frictions are. If workers are identical and firms are identical then all contacts result in a match. Under two-sided heterogeneity, the production technology matters because it specifies how much output is lost when a given job type is occupied by a sub-optimal worker type. Search frictions generate a lot of output loss if a precize match is very important while if worker types are almost perfect substitutes, the output loss will be modest. By combining information on wage dispersion and the substitutability of worker types we can learn about the actual amount of frictions and the importance of a precize match. We then use our model to quantify and decompose this total output loss. Traditionally, most of the macro labor literature focussed on unemployment but our results imply that mismatch and job creation cost are at least as important. We find this total loss is between 5% and 10% (for the value of non-market time being between 0.4 and 0.71, the relative efficiency of on-the-job search between 0.50 and 0.75 and depending on whether firms can or cannot ex ante commit to wages as a function of match quality). In Gautier and Teulings (2006) we did not allow for on-the-job search and therefore substantially overestimated the output loss due to frictions.

Other contributions of our paper are that we show that the max-mean wage differential is a more robust measure for wage dispersion than measures based on the reservation wage because workers move towards the best jobs so the density around the highest wage is a lot higher than around the lowest wage. We also discuss a simple and tractable method for estimating the size of wage differentials allowing for measurement error. Finally, we derive a relation between Katz and Murphy's (1992) elasticity of substitution between high and low-skilled workers and the second derivative of the production function of our model.

References

[1] Abowd, J.M., F. Kramarz, and D.N. Margolis (1998), High wage workers and high wage firms, *Econometrica*, 67, 251-333.

[2] Atakan, A. (2006), Assortative Matching with Explicit Search Costs, *Econometrica* 74, 667–680.

[3] Becker, G. S. (1973), A Theory of Marriage: Part I, *Journal of Political Economy* 81(4), 813-46.

[4] Bontemps, C. J.M. Robin and G.J. Van den Berg (2000), Equilibrium search with continuous productivity dispersion: Theory and nonparametric estimation, *International Economic Review*, 41, 305-58.

[5] Bound, J. and A. Krueger (1991), The Extent of Measurement Error in Longitudinal Earnings Data: Do two wrongs make a right?, *Journal of Labor Economics*, 9, 1-24.

[6] Buhai, S. and C.N. Teulings (2006), *Tenure profiles and efficient separation in a stochastic productivity model*, CESinfo Working Paper Series, No.1688.

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

[8] Coles M. (2001), Equilibrium Wage Dispersion, Firm Size, and Growth, *Review of Economic Dynamics*, 4-1,159-187.

[9] Dickens, W. and L.F. Katz (1987). Inter-Industry wage differences and theories of wage determination, *NBER* working paper No. 2271.

[10] Eeckhout, J. and P. Kircher (2011). Identifying sorting, in theory, *Review of Economic Studies.* Forthcoming.

[11] Eeckhout, J. and P. Kircher (2010). Sorting and decentralized price competition, *Econometrica 78-2, 539-574.*

[12] Elliott, M. (2010). Inefficiencies in networked markets, mimeo Stanford.

[13] Elsby, M., B. Hobijn and A. Sahin (2008),. Unemployment Dynamics in the OECD, NBER Working Paper 14617.

[14] Fallick, B. and C. Fleischman (2004), The importance of employer-to-employer flows in the U.S. labor market, working paper, Federal Reserve Bank Board of Governors, Washington DC.

[15] Gautier, P.A and C.N. Teulings (2006), How large are search frictions? *Journal of the European Economic Association*, 4-6, 1193-1225.

[16] Gautier, P.A, C.N. Teulings and A.P. Van Vuuren (2005), Labor market search with two sided heterogeneity: hierarchical versus circular models, *in:* H. Bunzel, B.J. Christensen, G.R. Neumann, and J.M. Robin, eds., *Structural Models of Wage and Employment Dynamics, conference volume in honor of Dale Mortensen*, Elsevier, Amsterdam.

[17] Gautier, P.A, C.N. Teulings and A.P. Van Vuuren (2008), On the job search and sorting, Tinbergen Institute discussion paper 070/3.

[18] Hall, R.E. (2009), Reconciling Cyclical Movements in the Marginal Value of Time and the Marginal Product of Labor, *Journal of Political Economy*, 117, 281-322.

[19] Hall, R.E. and Paul R. Milgrom (2008), The Limited Influence of Unemployment on the Wage Bargain, *American Economic Review*, 98(4), 1653-1674.

[20] Hagedorn J. and I. Manovskii (2008), The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited, *American Economic Review*, 98, 1692-1706.

[21] Hornstein, A., P. Krusell and G.L. Violante (2010), Frictional wage dispersion in search models: a quantitative assessment, mimeo, Federal Bank of Richmond.

[22] 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.

[23] Lise, J., C. Meghir, and J.M. Robin (2008), "Matching, Sorting and Wages ", UCL mimeo, 2008.

[24] Lopes de Melo (2008), "Sorting in the Labor Market: Theory and Measurement".

[25] Marimon R. and F. Zilibotti (1999), Unemployment vs. mismatch of talents: reconsidering unemployment benefits, *Economic Journal*, 109, 266-291.

[26] Moscarini, G. (2005), Job matching and the wage distribution, *Econometrica*, 73(2), 481-516.

[27] Moscarini, G., and F. Vella (2008), Occupational Mobility and the Business Cycle, NBER WP 13819.

[28] Mortensen, D.T. (2000), Equilibrium unemployment with wage posting: "Burdett-Mortensen meets Pissarides", in B.J. Christensen, P. Jensen, N. Kiefer and D.T. Mortensen, eds., *Panel data and structural labor market models*, Amsterdam, Elsevier science.

[29] Mortensen, D.T. and C. Pissarides (1999), New developments in models of search in the labor market, *in:* O.C. Ashenfelter and D. Card, *Handbook of Labor Economics* 3b, North-Holland, Amsterdam.

[30] Nagypal E. (2005), On the extent of job-to-job transitions, working paper, Northwestern University, Evanston.

[31] Pissarides, C.A. (1994), Search unemployment with on-the-job search, *Review of Economic Studies*, 61, 457-475.

[32] Pissarides, C.A. (2000), *Equilibrium unemployment theory*, 2nd edition, MIT Press, Cambridge.

[33] Postel-Vinay, F. and J.M. Robin (2002), Equilibrium wage dispersion with worker and employer heterogeneity, *Econometrica*, 70, 2295–2350.

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

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

[36] Shimer, R. (2005), The Cyclical Behavior of Equilibrium Unemployment and Vacancies,"*American Economic Review*, 95, 25-49.

[37] Shimer, R. (2006), On-the-job Search and Strategic Bargaining, *in:* H. Bunzel, B.J. Christensen, G.R. Neumann, and J.M. Robin, eds., *Structural Models of Wage and Employment Dynamics, conference volume in honor of Dale Mortensen*, Elsevier, Amsterdam.

[38] Shimer, Robert (2007), Reassessing the Ins and Outs of Unemployment, NBER Working Paper 13421.

[39] Shimer, R. and L. Smith (2000), Assortative matching and search", *Econometrica*, 68, 343-369.

[40] Teulings, C.N. (1995), The wage distribution in a model of the assignment of skills to jobs, *Journal of Political Economy*, 103, 280-315.

[41] Teulings, C.N. (2005), Comparative advantage, relative wages, and the accumulation of human capital, *Journal of Political Economy*, 113, 425-461.

[42] Teulings, C.N. and P.A. Gautier (2004), The right man for the job, *Review of Economic Studies*, 71, 553-580.

[43] Teulings, C.N. and T. Van Rens (2008), Education, growth, and income inequality, *Review of Economics and Statistics*, 90, 89-104.

A Appendix

A.1 Derivation of the asset values

The Bellman equation for the asset value of employment reads

$$\rho V^E(x) = W(x) + 2\psi\lambda \int_0^x \left[V^E(z) - V^E(x)\right] dz - \delta \left[V^E(x) - V^U\right]. \tag{26}$$

Totally differentiating (26) yields

$$V_x^E(x) = \frac{W_x(x)}{\rho + \delta + 2\psi\lambda x}. \tag{27}$$

The solution to this differential equation is

$$V^E(x) = \int_0^x \frac{W_x(z)}{\rho + \delta + 2\psi\lambda z} dz + C_0.$$

Integrating by parts yields

$$V^E(x) = \frac{W(x)}{\rho + \delta + 2\psi\lambda x} - \frac{W(0)}{\rho + \delta} + 2\lambda\psi \int_0^x \frac{W(z)}{(\rho + \delta + 2\psi\lambda z)^2} dz + C_0. \tag{28}$$

Evaluating (28) at $x=0$ gives an initial condition that can be used to solve for C_0

$$C_0 = V^E(0) = \frac{W(0)}{\rho+\delta} + \frac{\delta}{\rho+\delta} V^U.$$

Substitution of this equation into (28) yields the desired expression. Let $\mathrm{E}_x W \equiv \int_0^{\bar{x}} g(x) W(x)\, dx$ be the expected wage of a filled job. Evaluate (3) at $\bar{x}$ and use the definition of g in (12) to get

$$\rho V^E(\bar{x}) = \rho V^U = \frac{W(z) + \psi \bar{x} \mathrm{E}_x W}{1+\psi \bar{x}} = \frac{u W(\bar{x}) + \psi (1-u) \mathrm{E}_x W}{u + \psi (1-u)}. \tag{29}$$

Next, note that the right-hand side of (26) and (29) are equal which can be used to get an expression for $\int_0^z \left[V^E(x) - V^U\right] dx$. Substitution of this expression into (2) gives ρV^U as a function of $W(\bar{x})$, which can be eliminated by solving (29) for $W(\bar{x})$. This gives

$$\rho V^U = \frac{B + \bar{x} \mathrm{E}_x W}{1+\bar{x}} = uB + (1-u) \mathrm{E}_x W, \tag{30}$$

where the final step uses (11).

The free entry condition implies that the option value of a vacancy of type c must be equal to K. Hence, by defining $\mathrm{E}_x Y \equiv \int_0^{\overline{x}} g(x) Y(x)\, dx$, we obtain:

$$\begin{aligned} K &= 2\lambda \int_0^{\overline{x}} \{u + \psi (1-u) [1 - G(x)]\} \frac{Y(x) - W(x)}{\rho + \delta + 2\psi\lambda x} dx \\ &= (1-u)(\mathrm{E}_x Y - \mathrm{E}_x W). \end{aligned}$$

The first term in the integrand is the effective labor supply, $u + \psi (1-u) [1 - G(x)]$ for a vacancy of type x. It is equal to the number of unemployed, u plus the number of workers employed in jobs with a mismatch indicator that exceeds x, $(1-u)[1-G(x)]$. The second factor is the discounted value of a filled vacancy. Just as in the wage equation, we discount current revenue $Y(x) - W(x)$ by the discount rate ρ plus the separation rate δ plus the quit rate $2\psi\lambda x$. The second line follows from substituting (11) and (12) in.

A.2 Variance of x

$$\begin{aligned} \mathrm{Var}[x] &= \int_0^{\overline{x}} x^2 g(x)\, dx = \frac{1+\psi\overline{x}}{\overline{x}} \int_0^{\overline{x}} \frac{x^2}{(1+\psi x)^2} dx = \frac{1+\psi\overline{x}}{\psi^3 \overline{x}} \int_0^{\psi\overline{x}} \frac{y^2}{(1+y)^2} dy \\ &= \frac{\psi\overline{x}(2+\psi\overline{x}) - 2(1+\psi\overline{x}) \log(1+\psi\overline{x})}{\psi^3\overline{x}}. \end{aligned}$$

Note that by the definition of $g(x)$, the support of $G(x)$ is $[0, \overline{x}]$ and not $[-\overline{x}, \overline{x}]$

A.3 $\frac{f_{ee}}{f_{ue}}$ ratio

$$
\begin{aligned}
uf_{ue} &= 2\lambda\left(1-u\right), \\
(1-u)f_{ee} &= 2\int_0^{\overline{x}} \psi\lambda\left(1-u\right)g\left(x\right)xdx = 2\psi\lambda\left(1-u\right)\frac{\left(1+\psi\overline{x}\right)}{\psi^2\overline{x}}\int_0^{\psi\overline{x}}\frac{q}{\left(1+q\right)^2}dq \\
&= 2\lambda\left(1-u\right)\left(\frac{1+\psi\overline{x}}{\psi\overline{x}}\ln\left(\psi\overline{x}+1\right)-1\right), \\
\frac{f_{ee}}{f_{ue}} &= \frac{u}{1-u}\left(\frac{1+\psi\overline{x}}{\psi\overline{x}}\ln\left(\psi\overline{x}+1\right)-1\right) = \frac{1+\psi\overline{x}}{\psi\overline{x}^2}\ln\left(\psi\overline{x}+1\right)-\frac{1}{\overline{x}}.
\end{aligned}
$$

A.4 Wages and expected wages

$$
W(x) = 1-\gamma\left[\left(\frac{1+\psi x}{\psi}\right)^2\log\left(\frac{1+\psi\overline{x}}{1+\psi x}\right)-\frac{\overline{x}}{\psi}\frac{\left(1+\psi x\right)^2}{1+\psi\overline{x}}+\frac{x}{\psi}+\frac{3}{2}x^2\right], \tag{31}
$$

$$
\begin{aligned}
W_x\left(x\right) &= -2\gamma\frac{\psi}{1+\psi x}\left[\left(\frac{1+\psi x}{\psi}\right)^2\ln\left(\frac{1+\psi x}{1+\psi\overline{x}}\right)-\frac{\overline{x}}{\psi}\frac{\left(1+\psi x\right)^2}{1+\psi\overline{x}}+\frac{x}{\psi}+x^2\right], \\
\mathrm{E}_x W &= \int_0^{\overline{x}} g\left(x\right)W\left(x\right)dx \\
&= 1-\int_0^{\overline{x}}\frac{1+\psi\overline{x}}{\overline{x}\left(1+\psi x\right)^2}\gamma\left[-\left(\frac{1+\psi x}{\psi}\right)^2\ln\left(\frac{1+\psi x}{1+\psi\overline{x}}\right)-\frac{\overline{x}}{\psi}\frac{\left(1+\psi x\right)^2}{1+\psi\overline{x}}+\frac{x}{\psi}+\frac{3}{2}x^2\right]dx \\
&= 1-\gamma\frac{1+\psi\overline{x}}{\psi^3\overline{x}}\int_0^{\psi\overline{x}}\frac{1}{\left(1+q\right)^2}\left[-\left(1+q\right)^2\ln\left(\frac{1+q}{1+\psi\overline{x}}\right)-\psi\overline{x}\frac{\left(1+q\right)^2}{1+\psi\overline{x}}+q+\frac{3}{2}q^2\right]dq \\
&= 1-\gamma\frac{3}{\psi^3\overline{x}}\left[\psi\overline{x}+\frac{1}{2}\psi^2\overline{x}^2-\left(1+\psi\overline{x}\right)\ln\left(1+\psi\overline{x}\right)\right],
\end{aligned}
$$

$$
\gamma = \frac{2\left(-1+B\right)\psi^3(\psi(-1+u)-u)\bar{x}}{\begin{array}{c}\psi\bar{x}(6(-1+u)+3\psi(-1+u)(-2+\bar{x})+\psi^2\bar{x}(3-3u+\bar{x}) \\ +6(-1+\psi)(-1+u)(1+\psi\bar{x})\log(1+\psi\bar{x})\end{array}}. \tag{32}
$$

where we first apply change of variables, $q=\psi x$, $dx=\frac{1}{\psi}dq$ and then repeatedly use partial integration. The last equation follows from (6) and (7) (the expression here is written as the Mathematica output)

no commitment

$$
\begin{aligned}
W(x) &= 1-\gamma\left[\frac{1+\psi x}{\psi^2}\ln\left(\frac{1+\psi x}{1+\psi\overline{x}}\right)-\frac{x-\overline{x}}{\psi}-\frac{1}{2}x\left(x-2\overline{x}\right)\right], \qquad (33)\\
\mathrm{E}_x W &= \int_0^{\overline{x}} g\left(x\right)W\left(x\right)dx\\
&= 1-\int_0^{\overline{x}}\frac{1+\psi\overline{x}}{\overline{x}\left(1+\psi x\right)^2}\gamma\left[\frac{1+\psi x}{\psi^2}\ln\left(\frac{1+\psi x}{1+\psi\overline{x}}\right)-\frac{x-\overline{x}}{\psi}-\frac{1}{2}x\left(x-2\overline{x}\right)\right]dx\\
&= 1-\gamma\frac{1+\psi\overline{x}}{\psi^3\overline{x}}\int_0^{\psi\overline{x}}\frac{1}{\left(1+q\right)^2}\left[\left(1+q\right)\ln\left(\frac{1+q}{1+\psi\overline{x}}\right)-q+\psi\overline{x}-\frac{1}{2}q\left(q-2\psi\overline{x}\right)\right]dq\\
&= 1-\gamma\frac{1+\psi\overline{x}}{\psi^3\overline{x}}\left[-\frac{1}{2}\ln^2\left(1+\psi\overline{x}\right)+\psi\overline{x}\ln\left(1+\psi\overline{x}\right)-\frac{1}{2}\frac{\psi^2\overline{x}^2}{1+\psi\overline{x}}\right],\\
\gamma &= \frac{2\left(-1+B\right)\left(\psi\left(-1+u\right)-u\right)\bar{x}}{\frac{(-1+\psi)(-1+u)\bar{x}^2}{\psi}+\overline{x}^3-\frac{(-1+\psi)(-1+u)(1+\bar{x})(2\psi\bar{x}-\ln(1+\psi\bar{x}))\ln(1+\psi\bar{x})}{\psi^3}}. \qquad (34)
\end{aligned}
$$

where the last equation follows again from (6)and (7) and is here expressed as in Mathematica output. In all relations presented above, $\overline{x}$ can be eliminated using

$$\overline{x}=\frac{1-u}{u},$$

compare equation (11), while γ can be eliminated using its expressions.

A.4.1 Wage differentials and the output loss due to search for $\psi = 1$

commitment

$$
\begin{aligned}
W\left(0\right)-\mathrm{E}_x W &= 2\left(1-B\right)\left(\frac{u}{1-u}\right)^2\left(\frac{5}{2}-u+\frac{3}{2}u^{-1}+\frac{4-u}{1-u}\ln u\right),\\
\mathrm{E}_x W-W\left(\overline{x}\right) &= 2\left(1-B\right)\left(\frac{u}{1-u}\right)^2\left(\frac{1}{2}\left(\frac{1-u}{u}\right)^2-\frac{3}{2}\frac{1+u}{u}-3\frac{1}{1-u}\ln u\right),\\
X &= 6\left(1-B\right)\left(\frac{u}{1-u}\right)^2\left(1-u+\frac{1}{2}u^{-1}\left(1-u\right)^2+\ln u\right).
\end{aligned}
$$

no commitment

$$
\begin{aligned}
W\left(0\right)-\mathrm{E}_x W &= 2\left(1-B\right)\left(\frac{u}{1-u}\right)^2\left(-\frac{1}{2}\frac{1}{1-u}\left(\ln u\right)^2-\frac{1+u}{u}\ln u-\frac{3}{2}\frac{1-u}{u}\right),\\
\mathrm{E}_x W-W\left(\overline{x}\right) &= 2\left(1-B\right)\left(\frac{u}{1-u}\right)^2\left(\frac{1}{2}\left(\frac{1-u}{u}\right)^2+\frac{1}{2}\frac{1}{1-u}\left(\ln u\right)^2+\frac{1}{u}\ln u+\frac{1}{2}\frac{1-u}{u}\right),\\
X &= -2\left(1-B\right)\left(\frac{u}{1-u}\right)^2\left(\frac{1}{2}\left(\ln u\right)^2+\frac{1-u}{u}\ln u+\frac{1}{2}\frac{1-u}{u}\right).
\end{aligned}
$$

In all these equations, we use equation (7) to eliminate γ.

A.5 The derivation of γ

We follow the derivation in Teulings and Van Rens (2008). Due to the assumption of comparative advantage of skilled workers in complex jobs, the optimal assignment is an increasing function, $c'(s) > 0$. Let Y^* be aggregate output per worker. We assume that this output is produced by a Leontieff technology, requiring the input of all c-type jobs in fixed proportions. Let $h(c)$ be the density of the input of a c-type job required to produce one unit of aggregate output. Equilibrium on the commodity market for job type $c(s)$ requires the equality of supply and demand for each s-type:

$$Y^* h[c(s)] = g(s) Y[s, c(s)] / c'(s), \tag{35}$$

where $g(s)$ is the skill density function. The left hand side is the demand for the output of job type $c(s)$; it is equal to aggregate output Y^* times the density of job type $c(s)$ required per unit of aggregate output, $h[c(s)]$. The right hand side is the supply of output of job type $c(s)$; it is equal to the density of worker type s, times its productivity in job type $c(s)$, $Y[s, c(s)]$ times the Jacobian $ds/dc = 1/c'(s)$. We assume s and c to be distributed normally with mean μ_s and μ_c respectively and identical standard deviations $\sigma_s = \sigma_c = \sigma$. Teulings and Gautier (2004) show that locally (around the optimal assignment) the distribution of s and c can be approximated by a uniform. Without loss of generality, we normalize $\mu_c = \sigma^2$; the only thing that matters in this model turns out to be the difference between μ_s and μ_c. Taking logs in equation (35) and using the density function for $h(\cdot)$ and $g(\cdot)$ yields,

$$\ln Y^* - \frac{1}{2}\left(\frac{c(s) - \sigma^2}{\sigma}\right)^2 = -\frac{1}{2}\left(\frac{s - \mu_s}{\sigma}\right)^2 + s - \frac{1}{2}\gamma [s - c(s)]^2 - \ln c'(s). \tag{36}$$

This equation should hold identically for all s.

Let $\ln W(s)$ be the log wage for worker type s in equilibrium. The zero profit condition implies $W(s) = Y[s, c(s)]$. Firms offering jobs of type c choose their preferred worker type s as to maximize profits, or equivalently, to minimize the log of the cost of production per unit of output, $\ln W(s) - \ln Y(s, c)$. Hence, the equilibrium wage function $W(s)$ satisfies the first order condition

$$\frac{d \ln W(s)}{ds} = \frac{d \ln Y(s, c)}{ds}|_{c=c(s)} = 1 - \gamma [s - c(s)], \tag{37}$$

where we use equation (18) in the second equality. The system of equations (36) and (37) is solved by the following expressions for $c(s)$ and $W(s)$:

$$c(s) = s - \mu_s,$$
$$\frac{d \ln W(s)}{ds} = 1 - \gamma\mu_s.$$

The wage function $\ln W(s)$ is linear in s. $d \ln W(s)/ds = 1 - \gamma\mu_s$ is the return to the human capital index s. This return depends on the supply of human capital, that is, on the mean of the skill distribution. The equilibrium assignment of section 3.4, $c(s) = s$, implies that $\mu_s = 0$. In that case $d \ln W(s)/ds = 1$, as is implied by equation (18). A percentage point upward shift in the mean of the skill distribution, μ_s, reduces the return to human capital by γ % point. Katz and Murphy split labour into two skill groups, low and high, and consider the effect on relative wages of a shift in labour supply from the one to the other. Let s^* be the cut off level. All worker types with $s > s^*$ are classified as high skilled; all other workers as low skilled. Hence, $\Phi[\sigma^{-1}(s^* - \mu_s)]$ is the share of low skilled workers and $\Phi[-\sigma^{-1}(s^* - \mu_s)]$ is the share of high skilled workers, where $\Phi(\cdot)$ denotes the standard normal distribution function. Katz and Murphy estimate the elasticity η, which is the ratio of the change in relative supply of high and low skilled workers to the change in relative wages.

$$\eta \equiv -\frac{d\left(\ln \Phi[\sigma^{-1}(s^* - \mu_s)] - \ln \Phi[-\sigma^{-1}(s^* - \mu_s)]\right)/d\mu_s}{d\left(\mathrm{E}[\ln W(s)|s > s^*] - \mathrm{E}[\ln W(s)|s < s^*]\right)/d\mu_s} = \frac{1}{\sigma^2\gamma}. \tag{38}$$

Let $z \equiv \frac{s-\mu_s}{\sigma}$, $z^* \equiv \frac{s^*-\mu_s}{\sigma}$. The effect of an increase in μ_s on the mean log wage of low and high skilled workers respectively reads:

$$\begin{aligned}
\frac{d}{d\mu_s}\mathrm{E}[\ln W(s)|s < s^*] &= \frac{d}{d\mu_s}\mathrm{E}[\ln W(0) + (1-\gamma\mu_s)s|s < s^*] \\
&= \frac{d}{d\mu_s}(1-\gamma\mu_s)s\sigma\mathrm{E}[z|z < z^*] - \frac{d}{d\mu_s}(1-\gamma\mu_s)s\mu_s \\
&= \gamma\sigma\frac{\phi(z^*)}{\Phi(z^*)} - \frac{1}{2}\gamma\mu_s, \\
\frac{d}{d\mu_s}\mathrm{E}[\ln W(s)|s > s^*] &= -\gamma\sigma\frac{\phi(z^*)}{1-\Phi(z^*)} - \frac{1}{2}\gamma\mu_s,
\end{aligned}$$

Hence:

$$\frac{d}{d\mu_s}\left(\mathrm{E}[\ln W(s)|s > s^*] - \mathrm{E}[\ln W(s)|s < s^*]\right) = -\gamma\sigma\frac{\phi(z^*)}{\Phi(z^*)[1-\Phi(z^*)]}$$

The effect on the number of low and high skilled workers reads:

$$\frac{d \ln \Phi\left[\sigma^{-1}\left(s^*-\mu_s\right)\right]}{d \mu_s} = \frac{\phi\left(z^*\right)}{\sigma \Phi\left(z^*\right)},$$
$$\frac{d \ln\left[1-\Phi\left[\sigma^{-1}\left(s^*-\mu_s\right)\right]\right]}{d \mu_s} = \frac{\phi\left(z^*\right)}{\sigma\left[1-\Phi\left(z^*\right)\right]}.$$

Hence:

$$\frac{d}{d \mu_s}\left(\ln \Phi\left[\sigma^{-1}\left(s^*-\mu_s\right)\right]-\ln\left[1-\Phi\left[\sigma^{-1}\left(s^*-\mu_s\right)\right]\right]\right) = \frac{1}{\sigma} \frac{\phi\left(z^*\right)}{\Phi\left(z^*\right)\left[1-\Phi\left(z^*\right)\right]},$$
$$-\frac{d\left(\ln \Phi\left[\sigma^{-1}\left(s^*-\mu_s\right)\right]-\ln\left[1-\Phi\left[\sigma^{-1}\left(s^*-\mu_s\right)\right]\right]\right) / d \mu_s}{d\left(\mathrm{E}\left[\ln W(s) \mid s>s^*\right]-\mathrm{E}\left[\ln W(s) \mid s<s^*\right]\right) / d \mu_s} = \frac{1}{\sigma^2 \gamma}.$$

A.6 The proof of Proposition 1

Assume: $r \sim N\left(0, \sigma_r^2\right)$ and $q \sim N\left(0, \sigma_q^2\right)$ with $\mathrm{Cov}(r, q)=0$. Define: $x \equiv r+q$ and $\sigma_x^2 \equiv \sigma_r^2+\sigma_q^2$.[17] We have:

$$\mathrm{E}[|x|]=\sqrt{2 \pi}^{-1} 2 \sigma_x.$$

Consider the OLS regression:

$$|x|=\beta_0+\beta_2 r^2+v$$

where v is a zero mean error term.

Lemma 2 *The expected value of the OLS estimators for β_0 and β_2 reads:*

$$\beta_0=\sqrt{2 \pi}^{-1}\left(\sigma_x+\frac{\sigma_q^2}{\sigma_x}\right),$$
$$\beta_2=\sqrt{2 \pi}^{-1} \sigma_x^{-1}.$$

Consider the model:

$$w=\omega_0-\omega_2|x|$$
$$w=\bar{\omega}_0-\bar{\omega}_2 r^2+v,$$

where $\mathrm{E}[w]=0$ and v is a zero mean error term. Hence, we have:

$$\omega_0=\sqrt{2 \pi}^{-1} 2 \sigma_x \omega_2$$
$$\bar{\omega}_0=\omega_0-\beta_0 \omega_2=\sqrt{2 \pi}^{-1} \frac{\sigma_r^2}{\sigma_x} \omega_2=\frac{1}{2} R \omega_2 \mathrm{E}[|x|],$$

[17] In the main text, we use the model of a regressor contaminated with measurement error, $\widehat{x}=x+\varepsilon_x$ with $\mathrm{Cov}[x, \varepsilon_x]=0$. Here, we use the model of partial observation of x where r is the regressor, $x=r+q$ with $\mathrm{Cov}[r, q]=0$. Both models yield the same expression for $\bar{\omega}_0$ for the case: $\mathrm{Var}[x]$ / $/\mathrm{Var}[\widehat{x}]=\mathrm{Var}[r]/\mathrm{Var}[x]=R$. The proof of the lemma is simpler for the case of partial observation of x.

where $R \equiv \sigma_r^2/\sigma_x^2$, which is equivalent to the definition of R in section 3.3.

Proof of the lemma:

The sum of least squares reads:

$$\Sigma = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{\sigma_r\sigma_q}\left(|r+q| - \beta_0 - \beta_2 r^2\right)\phi\left(\frac{q}{\sigma_q}\right)dq\phi\left(\frac{r}{\sigma_r}\right)dr.$$

Define: $\sigma \equiv \sigma_x^{-1}\sigma_r\sigma_q$. The first order condition for β_0 implies:

$$\begin{aligned}
0 &= \int_{-\infty}^{\infty}\frac{1}{\sigma_r\sigma_q}\left[\begin{array}{c} -\int_{-\infty}^{-r}(r+q)\phi\left(\frac{q}{\sigma_q}\right)dq + \int_{-r}^{\infty}(r+q)\phi\left(\frac{q}{\sigma_q}\right)dq \\ -\int_{-\infty}^{\infty}\left(\beta_0+\beta_2 r^2\right)\phi\left(\frac{q}{\sigma_q}\right)dq \end{array}\right]\phi\left(\frac{r}{\sigma_r}\right)dr \\
&= \int_{-\infty}^{\infty}\frac{1}{\sigma_r}\left[r\left[2\Phi\left(\frac{r}{\sigma_q}\right)-1\right] + 2\sigma_q\phi\left(\frac{r}{\sigma_q}\right) - \left(\beta_0+\beta_2 r^2\right)\right]\phi\left(\frac{r}{\sigma_r}\right)dr \\
&= \int_{-\infty}^{\infty}\frac{2}{\sigma_r}r\Phi\left(\frac{r}{\sigma_q}\right)\phi\left(\frac{r}{\sigma_r}\right)dr + \int_{-\infty}^{\infty}\sqrt{\frac{2}{\pi}}\frac{\sigma_q}{\sigma_r}\phi\left(\frac{r}{\sigma}\right)dr - \left(\beta_0+\beta_2\sigma_r^2\right) \\
&= \int_{-\infty}^{\infty}2\frac{\sigma_r}{\sigma_q}\phi\left(\frac{r}{\sigma_q}\right)\phi\left(\frac{r}{\sigma_r}\right)dr + \sqrt{\frac{2}{\pi}}\frac{\sigma_q}{\sigma_r}\sigma - \left(\beta_0+\beta_2\sigma_r^2\right) = \sqrt{\frac{2}{\pi}}\sigma_x - \left(\beta_0+\beta_2\sigma_r^2\right),
\end{aligned}$$

where in the third line we apply integration by parts. For β_2 we obtain:

$$\begin{aligned}
0 &= \int_{-\infty}^{\infty}\frac{r^2}{\sigma_r\sigma_q}\left[\begin{array}{c} -\int_{-\infty}^{-r}(r+q)\phi\left(\frac{q}{\sigma_q}\right)dq + \int_{-r}^{\infty}(r+q)\phi\left(\frac{q}{\sigma_q}\right)dq \\ -\int_{-\infty}^{\infty}\left(\beta_0+\beta_2 r^2\right)\phi\left(\frac{q}{\sigma_q}\right)dq \end{array}\right]\phi\left(\frac{r}{\sigma_r}\right)dr \\
&= \int_{-\infty}^{\infty}\frac{2}{\sigma_r}r^3\Phi\left(\frac{r}{\sigma_q}\right)\phi\left(\frac{r}{\sigma_r}\right)dr + \int_{-\infty}^{\infty}\sqrt{\frac{2}{\pi}}\frac{\sigma_q}{\sigma_r}r^2\phi\left(\frac{r}{\sigma}\right)dr - \left(\beta_0\sigma_r^2+3\beta_2\sigma_r^4\right) \\
&= \int_{-\infty}^{\infty}\frac{2}{\sigma_r}\left(r^3-2\sigma_r^2 r\right)\Phi\left(\frac{r}{\sigma_q}\right)\phi\left(\frac{r}{\sigma_r}\right)dr + \int_{-\infty}^{\infty}4\sigma_r r\Phi\left(\frac{r}{\sigma_q}\right)\phi\left(\frac{r}{\sigma_r}\right)dr \\
&+ \sqrt{\frac{2}{\pi}}\sigma^3 - \left(\beta_0\sigma_r^2+3\beta_2\sigma_r^4\right) \\
&= \int_{-\infty}^{\infty}\sqrt{\frac{2}{\pi}}\frac{\sigma_r}{\sigma_q}r^2\phi\left(\frac{r}{\sigma}\right)dr + \int_{-\infty}^{\infty}2\sqrt{\frac{2}{\pi}}\frac{\sigma_r^3}{\sigma_q}\phi\left(\frac{r}{\sigma}\right)dr + \sqrt{\frac{2}{\pi}}\frac{\sigma_q\sigma^3}{\sigma_r} - \left(\beta_0\sigma_r^2+3\beta_2\sigma_r^4\right) \\
&= \sqrt{\frac{2}{\pi}}\frac{\sigma_r^2\sigma_q^2}{\sigma_x} + 2\sqrt{\frac{2}{\pi}}\frac{\sigma_r^4}{\sigma_x} - \left(\beta_0\sigma_r^2+3\beta_2\sigma_r^4\right),
\end{aligned}$$

where in the third line we repeatedly apply integration by parts. Solving these first order conditions for β_0 and β_2 proves the lemma.■

A.7 Calibrating the model

First, we calibrate the unemployment rate u and the signal-noise ratio R to exactly match the estimated mismatch $\mathrm{Var}[\widehat{x}] = 0.0306$ and intercept $\overline{\omega}_0 = 0.0241$. In step 1, we choose a starting value for u, which implies a value for $\overline{x}$ and γ conditional on the normalization $\kappa = 1$, see equation (11) for $\overline{x}$ and equation (32) or (34) in Appendix A.4 for γ (depending on whether we assume commitment or not). In step 2 we generate 100,000 draws from the distribution $G(x)$, see equation (12) and calculate the corresponding value of $w = \ln W(x)$, using equation (31) or (33), taking deviations from the mean of w (as we did with or data) so that $\mathrm{E}[w] = 0$. We then renormalize the values of x according to the normalization $dw/ds = 1$ and the associated value $\gamma_{dw/ds=1} = 2$.[18] In step 3, we add normally distributed measurement error to these values of x and choose the variance of the measurement error such that we get the estimated mismatch, $\mathrm{Var}[\widehat{x}] = 0.0306$. Then, we "estimate" by OLS:

$$w = \overline{\omega}_0 - \omega_2 \widehat{x}^2 + \varepsilon.$$

If the "estimated" value of $\overline{\omega}_0$ is below 0.0241 (the value in 25), the model generates too little wage dispersion which implies that we should increase u, and the other way around if $\overline{\omega}_0$ is above 0.0241. Then, we return to step 1. This iteration process is continued till u converges.[19] In both our baseline calibration and our robustness checks we were able to match $\overline{\omega}_0$ and $\mathrm{Var}[\widehat{x}]$ exactly.

[18]

$$x_{dw/ds=1} = x_{\kappa=1}\sqrt{\gamma_{\kappa=1}/\gamma_{dw/ds=1}},$$

see footnote 4.

[19] This usually takes about 20 steps and lasts a few seconds.

Publisher:

CPB Netherlands Bureau for Economic Policy Analysis
P.O. Box 80510 | 2508 GM The Hague
T (070) 3383 380

March 2012 | ISBN 978-90-5833-548-7