Sorting in the Labor Market: Theory and Measurement[*†]

Rafael Lopes de Melo

First Draft: February 15, 2008

This Version: December 9, 2009

Abstract

Are more skilled workers employed by more productive firms? Are complementarities important in production? We provide three contributions to the measurement of sorting. First, we use a standard frictional sorting model to show that the standard empirical method used to measure sorting in the labor market can be biased in favor of not detecting sorting. Second, we isolate the economic mechanism responsible for this bias. Finally, we propose an alternative method to detect sorting that is immune from this bias. According to the model, sorting is prevalent in labor markets, as measured by our alternative method, but the standard method fails to detect it.

[*]This paper was part of my dissertation. I am indebted to Giuseppe Moscarini for the extensive advice and support. I would like to thank Fabian Lange, Fabien Postel-Vinay, Björn Brügemann, Jeremy Lise and Robert Shimer for their valuable comments and help. Very useful feedback was also received from participants at the seminars at Yale, UCL, University of Chicago, University of Chicago School of Business, Northwestern University, Princeton University, University of Pennsylvania, Boston University, University of Western Ontario, University of Montreal, CREI, SED, LACEA, PUC-Rio, and Ipea. I alone am responsible for any shortcomings in the paper.

[†]Department of Economics, University of Chicago (e-mail: lopesdemelo@uchicago.edu).

1 Introduction

Are more skilled workers employed by more (or less) productive firms? Are complementarities important in production? This paper analyzes the assortative matchup between heterogeneous firms and workers in the labor market. An extensive literature in economic theory studies the sorting patterns of heterogeneous agents. Important examples are partners in a marriage, buyers and sellers negotiating a product, students and teachers, players and teams, and workers and firms. A common feature among all these examples is that positive (negative) complementarities in production between agents induce positive (negative) sorting in equilibrium. For example, Becker [6] shows that if the production function is supermodular, the unique equilibrium allocation is efficient and exhibits perfect sorting between the partners in a marriage: the most desirable individuals get together, and vice-versa.[1]

Influential empirical results suggest, however, that little sorting actually takes place between workers and firms in labor markets—which would reflect the absence of complementarities. This empirical methodology adopts the framework of Abowd, Kramarz et al [3, 2, 1] (hereafter, AKM), which consists of estimating a log-linear wage regression with worker and firm fixed effects. The procedure then uses the estimated fixed effects of workers and firms as proxies for the underlying heterogeneities.[2] One standard—and very robust—result in this literature is that the correlation between these fixed effects is zero (or even negative). Taken at face value, this means that there is little sorting in labor markets.

This paper revisits this result by assessing how this empirical methodology performs in measuring sorting in a standard dynamic assignment model. We show that this correlation is consistently biased downwards, and can miss the true degree of sorting by a large extent—i.e. even if we have a large degree of sorting in the model, this empirical measure may still be close to zero. The reason for this bias lies in the economic mechanisms of the model, and does not vanish as sample size goes to infinity.

[1] If $f(x,y)$ is smooth, then supermodularity is equivalent to $f_{xy} > 0$. Also, he uses the core concept of equilibrium.

[2] These fixed effects can capture both unobservable and observable time-fixed traits of workers and firms—e.g. education, gender, race, sector.

We propose an alternative measure of sorting, still based on the same fixed effects methodology, that captures the degree of sorting in the model remarkably well: the correlation between a worker fixed effect and the average fixed effects of his coworkers. Lopes de Melo [13] and Bagger and Lentz [5] take up our proposed measure and compute the value for this correlation respectively in Brazil and Denmark. They find values between 0.3 and 0.4. This suggests that sorting actually plays a big role in labor markets, but the first correlation was just missing it. Moreover, models that explain the AKM 0 correlation by having a no-sorting equilibrium will have a hard time matching this empirical moment.

Our theory builds on model of Shimer and Smith [17], adapted to labor markets. The model has four main ingredients. First, there is heterogeneity of both workers and firms. Second, there are complementarities in production, which induce positive sorting in equilibrium. Third, search frictions add noise to the sorting process, causing the agents to accept suboptimal partners in order to avoid idleness. Finally, firms have limitations in their capacity to post new vacancies. This creates ex-ante rents for vacancies and creates reasons for firms to reject some workers in equilibrium.

The reason why the AKM correlation is biased in this model is due to non-monotonicities in the wage equation caused by the interaction of wage bargaining and limitations in the capacity of the firms to post new vacancies. When jobs are scarce firms demand compensation from the workers to fill a vacancy. This compensation is increasing in firm productivity, which creates non-monotonicities in the wage equation with respect to firm productivity. These non-monotonicities are a natural reflection of the job scarcity and the fact that firms reject workers in equilibrium: the least skilled worker that a given firm hires generates zero surplus, and as a consequence earns his reservation wage.[3] Thus, for this worker it is actually preferable to find a job in a less productive firm.

The alternative sorting moment overcomes this problem for two reasons. First, it is informative about sorting because as the top workers end up in the top firms, they end up with top coworkers as

[3] With positive sorting, this is true for any firm that only hires workers above the minimum skilled worker.

well. Also, this correlation tends to increase as we increase the degree of complementarities in the model. Second, these non-monotonicities do not affect the ordering of wages across workers—i.e, wages are always increasing in worker skill.

Finally, the AKM methodology is a cumbersome methodology that requires large datasets that follow workers and firms over time, and typically involves the estimation of a model with millions of coefficients. We then explore alternatives that could potentially measure sorting well in the model, but would be easier to implement. First, we try a moment that uses only cross-sectional data, based on the index of segregation proposed by Kremer and Maskin [10]. Second, we look at a measure of sorting that is consistent: as the sample size goes to infinity it converges to the true (rank) correlation between worker and firm types. We show that both of these measures consistently overestimate the degree of sorting in short samples. This happens because in the model there is rent-sharing between workers and firms, so even if workers were assigned randomly to firms, there would be still a common element in wages. Thus, despite the fact that AKM method is not correctly specified for this model, it helps correct for this short sample biases brought by the rent sharing.

The rest of the paper is structured as follows. Section 2 describes the related literature, Section 3 the theoretical model, Section 4 describes and evaluates the empirical measures of sorting and Section 5 concludes.

2 Related Literature

This paper relates to several strands of literature. First, the basis of the empirical methodology used throughout the paper comes from a series of papers by Abowd, Kramarz et al [3, 1, 2]. With the advent of datasets that follow workers and firms at the same time, they use the mobility of workers across firms to estimate a wage equation with worker and firm fixed effects. The results from these papers provided a number of insights and puzzles in topics such as industry and size wage differences, and whether the law of one price applies in labor markets.[4]

Second, our model is derived from the theoretical assignment literature. Becker [6] presents a model with two sided heterogeneity, and shows that if there are enough complementarities in production the equilibrium exhibits perfect sorting and is efficient. Shimer and Smith [17], Lu and McAfee [14] and Sattinger [14] introduce search frictions in the Becker model, which adds noise to the equilibrium allocations and requires stronger complementarities for the equilibrium to exhibit positive assortative matching.[5] Eeckhout and Kircher [9] develop a directed search model that is an intermediate case between the frictionless model and the economy with search frictions. That framework requires less complementarities than the search case, but more than the frictionless economy, to induce positive sorting. This is related to the work in Shi [16].

Thirdly, a recent stream of papers applies models with two-sided heterogeneity and bilateral matches to the labor market. Woodcock [19] uses a search model with no complementarities in production and match specific learning, which features negative sorting in equilibrium, as a way to justify the negative correlation in wage fixed effects in the AKM [6] regressions. Lentz [11] and Bagger and Lentz [5] formulate a search model with endogenous search intensity, and structurally estimate it using the Danish matched employer-employee dataset. They need weaker complementarities in production to have positive sorting, as the high-skill worker search with a higher intensity. Lise, Meghir and Robin [12] build on the model of Shimer and Smith [17], introducing on-the-job search, a wage mechanism as in Cahuc, Postel-Vinay and Robin [7] and a dynamic process for firms that causes them to change productivity over time. Eeckhout and Kircher [8] use variations from the Becker model [6] to argue that, from wage data alone, one cannot distinguish a model that features positive sorting from a model of negative sorting. However, they also argue that wages can give information about the strength of sorting, which is consistent with our results. Moreover, they argue that the strength is more important than the sign for questions related to efficiency and inequality.

[4] See Mortensen [15] for a survey.

[5] Shimer and Smith [17] use the random search technology.

[6] His model implies a specification like in AKM in levels, not in logs.

3 The Model

The model shares the basic features of Shimer and Smith [17], plus specific features of labor markets such as wages, unemployment compensation and asymmetries between worker and jobs.[7] It is a continuous time economy with heterogeneous agents, complementarities in production and search frictions. Production in this economy happens in bilateral matches. Workers meet jobs according to a stochastic process, i.e. matching is not instantaneous because of search frictions. They observe each other's types, decide if they want to pair up or not, and, upon matching, produce output and split the proceeds according to a pre-designed rule. Similarly as in their model, matches are dissolved exogenously at a given rate.

One important assumption is that there is a fixed stock of workers of heterogeneous skills and and a fixed stock of jobs of different productivity. Whereas the former assumption is standard, the assumption of a fixed stock of jobs is not. In fact, other authors model vacancy creation by assuming constant returns to scale and free entry at each firm type.[8] We depart from this tradition, and follow strictly Shimer and Smith [17], in order to introduce a mechanism in the model which allows us to better explain the data: when jobs are scarce firms demand compensation from the workers to fill a vacancy, which creates non-monotonicities in the wage equation with respect to firm productivity.

3.1 The Environment

There is a continuum of workers and jobs, each indexed by a type. We exogenously assign a value $h \in \mathbb{R}$ for each worker, which can be interpreted as his/her human capital, and value $p \in \mathbb{R}$ for each job, which can be interpreted as that firm's productivity. These exogenous random variables have atomless distributions $\overline{L} : \mathbb{R} \to [0,1]$ and $\overline{G} : \mathbb{R} \to [0,1]$, respectively, with density functions $\overline{l}(h)$ and $\overline{g}(p)$. These types are assumed to be perfectly observable and the distributions are known to all agents. The mass of workers is normalized to 1, and the mass of jobs per worker is $\overline{N}$.

[7] Shimer and Smith [17] study the problem of generic symmetric agents.

[8] For example, Gautier and Teulings [18] and Lentz [11].

Workers and firms discount time at rate r and have linear preferences.

When a worker of type h and a firm of type p match they produce a flow output of $F(h,p)$ at every instant. The production function is an essential element to our model because it has strong implications for the sorting patterns of workers. Shimer and Smith [17] provide a full characterization of the necessary and sufficient conditions for positive assortative matching in their model. In particular, they show that if the production function, the log of its first partial derivatives and the log of its cross derivative are all supermodular then the economy exhibits positive assortative matching.

Unemployed workers receive a flow value b and vacant jobs receive zero flow value.

We assume a random search technology where worker/jobs meet each other at a finite Poisson rate. Upon meeting, they take a draw from the corresponding idle type distribution—e.g., if an unemployed worker meets an employer he/she takes a draw from the vacancies distribution. Unemployed workers meet vacancies at the rate λ^W. On the firm side, vacancies meet unemployed workers at a rate λ^F. Finally, all matches are exogenously destroyed at rate δ. Moreover, we assume that the economy is in steady state.

The presence of search frictions creates a temporary bilateral monopoly power in a match. We follow several others and adopt the generalized Nash bargaining solution to determine wages. We choose this particular mechanism because it simplifies the solution of the problem, while embedding the economic forces that we believe are important to explain the data.

Before describing how the equilibrium of the model is determined, we introduce the following notation. An employed worker of type h working for a firm of type p has value $W(h,p)$, or, if this workers is unemployed, $U(h)$. A job of type p has value $J(h,p)$ if matched with a worker of type h, or $V(p)$ if vacant. The surplus of a match between worker h and job p is $S(h,p) \equiv [W(h,p)-U(h)]+[J(h,p)-V(p)]$. The wage in any match (h,p) is $w(h,p)$.

Matching sets are characterized by indicators functions. For an unemployed worker of type h who meets a vacant job of type p , the indicator $\alpha(h,p)$ assumes a value of 1 if they match, and 0 otherwise. The steady state distribution of employed matches is $\Gamma(h,p)$, with density $\gamma(h,p)$. The density of unemployed workers for each type is $l(h)$, and the density of vacancies is $g(p)$. Finally, the unemployment rate, u, and the number of vacancies in the economy, v, are also determined endogenously.

3.2 Values and Decisions

An employed worker h at a firm p earns a flow wage of $w(h,p)$. At a rate δ his/her match is destroyed exogenously. This implies the following value equation:

$$rW(h,p) = w(h,p) + \delta\left[U(h) - W(h,p)\right] \tag{1}$$

An unemployed worker of type h enjoys flow unemployment benefits b and receives job offers at rate λ^W:

$$rU(h) = b + \lambda^W \int_{-\infty}^{\infty} \alpha(h,p^*)[W(h,p^*) - U(h)]g(p^*)\,dp^*. \tag{2}$$

A productive job of type p, employing a worker of type h, enjoys flow profits $F(h,p) - w(h,p)$, but can be dissolved by the δ shock:

$$rJ(h,p) = F(h,p) - w(h,p) + \delta\left[V(p) - J(h,p)\right] \tag{3}$$

Finally, a vacancy does not earn any flow values, but can hire unemployed workers:

$$rV(p) = \lambda^F \int_{-\infty}^{\infty} \alpha(h^*,p)[J(h^*,p) - V(p)]l(h^*)\,dh^* \tag{4}$$

As previously described, wages are determined by the Generalized Nash bargaining solution. Standard results then imply that

$$S(h,p) = \frac{[J(h,p) - V(p)]}{(1-\beta)} = \frac{[W(h,p) - U(h)]}{\beta}. \tag{5}$$

Plugging Equations (1) and (3) in this formula and rearranging we solve for the wage

$$w(h,p) = \beta\left[F(h,p) - rV(p)\right] + (1-\beta)\,rU(h). \tag{6}$$

From this expression, we can see that wages are a function of output and the outside options of the worker and the firm.

When idle agents meet, they decide to pair up if doing so increases their values. It is easy to see from Equation (5) that this is equivalent to choosing the option that provides the largest surplus. Thus, when an unemployed worker meets a vacancy, the decision is given by

$$\alpha(h,p) = 1\left[S(h,p) > 0\right] \tag{7}$$

3.3 Steady State Flows

We now describe the equilibrium equations that jointly determine the stationary distributions $\Gamma(h,p)$, $l(h)$ and $g(p)$, and the idleness rates u and v.

The first equilibrium equation that we describe is between the flows in and out of employed matches of type (h,p). If $\alpha(h,p) = 0$ then $\gamma(h,p) = 0$. Otherwise, $\gamma(h,p)$ is determined by equating the inflows to the outflows,

$$u\lambda^{W} l(h)\,g(p) = \delta(1-u)\,\gamma(h,p), \tag{8}$$

Employed workers of type h have to equal the workers of that type minus the unemployed ones. This implies that

$$\bar{l}(h) - ul(h) = (1-u)\int_{-\infty}^{-\infty} \gamma(h,p^{*})\,dp^{*}. \tag{9}$$

The same holds for jobs of type p

$$\overline{N}\overline{g}(p) - vg(p) = \left(\overline{N} - v\right)\int_{-\infty}^{-\infty} \gamma(h^{*},p)\,dh^{*}. \tag{10}$$

Next, the unemployment rate is determined by integrating (8) over the full support of h and p, and its stationarity requires

$$\delta(1-u) = u\lambda^W \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \alpha(h^*,p^*) l(h^*) g(p^*) dh^* dp^*. \tag{11}$$

Finally, the last requirement for equilibrium is that the total number of employed workers has to equal filled jobs.

$$1-u = \overline{N} - v. \tag{12}$$

Equations 8 to 12 determine all endogenous objects in this model. However, for us to reach a steady state we still need to make sure that the number of unemployed workers finding jobs equal to the number of vacancies meeting workers. This is true if λ^F satisfies

$$\lambda^F = \frac{u\lambda^W}{v}.$$

3.4 Equilibrium

Definition 1. *A steady state equilibrium in this economy consists of values for $\gamma(h,p)$, $l(h)$, $g(p)$, u, v, $U(h)$, $V(p)$ and $\alpha(h,p)$ such that Equations (2), (4), (7), (8), (9), (10), (11) and (12) are satisfied.*

Shimer and Smith [17] have a proof of equilibrium existence, but not uniqueness, in the case with a quadratic matching function—which we follow.

3.5 Properties of the Wage Function

It is useful at this point to establish some properties of the wage function. Let the average wages of workers and firms of a given type respectively be $w^W(h) \equiv \frac{\int_{-\infty}^{\infty}\alpha(h,p^*)\gamma(h,p^*)w(h,p^*)dp^*}{\int_{-\infty}^{\infty}\alpha(h,p^*)\gamma(h,p^*)dp^*}$ and $w^F(p) \equiv \frac{\int_{-\infty}^{\infty}\alpha(h^*,p)\gamma(h^*,p)w(h^*,p)dh^*}{\int_{-\infty}^{\infty}\alpha(h^*,p)\gamma(h^*,p)dh^*}$.

Proposition 2. *Let $F(h,p)$ be such that we have an equilibrium with positive assortative matching,* *as in Shimer and Smith [17], and* $F_p(h,p) \longrightarrow_{h \to h_{min}} 0$. *In the Steady State equilibrium of this economy the wage function described by Equation 6 satisfies the following properties:*

1. *Wages are strictly increasing in worker skill:* $w_h(h,p) > 0, \forall \{(h,p) | \alpha(h,p) = 1\}$.

2. *Wages decrease in jobs productivities for some pairs* (h,p)*:* $w_p(h,p) < 0, \exists \left\{(h,p) | \alpha^U(h,p) = 1\right\}$.

3. *Average worker wages are increasing in worker skill:* $w_h^W(h) > 0, \forall h$.

4. *Average job wages are increasing in job productivity:* $w_p^W(p) > 0, \forall p$.

Proof. See Appendix. □

It is worth noting that parts 1-3 of the Proposition hold for any distributions, whereas part 4 only holds for the equilibrium distribution.

Part 2 of Proposition 1 states that wages are non-monotone in firm productivity. These non-monotonicities are a natural reflection of the job scarcity and the fact that firms reject workers in equilibrium: the least skilled worker that a given firm hires generates zero surplus, and as a consequence earns his reservation wage.[9] Thus, for this worker it is actually preferable to find a job in a less productive firm. Moreover, our intuition suggests that this should extend beyond Nash bargaining and be true for any wage mechanism where matching is bilaterally efficient.

3.6 Proposed Algorithm

This model does not have a closed form solution. Thus, we have to solve this model computationally. For our empirical implementation we use the following algorithm to solve the model. First, we discretize the support of the type distributions. Second, we guess an initial value for all endogenous objects. Then we take the following steps:

1. Values and matching sets: given all the guesses we iterate simultaneously on Equations (2) and (4) until finding a fixed point, which yield values for $U(h)$ and $V(p)$. In each step of the iteration, we update the matching set, $\alpha(h,p)$ using (7) .

[9] With positive sorting, this is true for any firm that only hires workers above the minimum skilled worker.

2. Steady state flow equations: given the output from Step 1 we iterate (8) to (12) until we find a fixed point. This gives a value for $\gamma(h,p)$, $l(h)$, $g(p)$, u and v.

3. Repeat Steps 1 and 2 until the endogenous distributions converge.

Because we do not have results showing that these mappings are contractions, we cannot anticipate that this algorithm works. However, in practice, this algorithm does work over a wide range of parameters.

3.7 Illustration of the Equilibrium

As an example, Figures 1 and 2 illustrate the equilibrium of a symmetric economy—same distribution of worker and job types and Nash Bargaining parameter equal to 0.5— and production function $F(h,p) = hp$. [10]

The first panel in Figure 1 illustrates the matching set for the idle agents. The upward sloping bands are a direct consequence of the equilibrium with positive assortative matching: the low types are rejected by the corresponding high types. The bottom left panel contrasts the distribution of all workers in this economy with the distribution of the unemployed. The distribution of the idle types is stochastically dominated by the distribution of all workers. This is a clear implication of the positive assortative matching, where the low types are systematically rejected by the high corresponding types.

The upper right panel shows the outside option of workers (or firms). This is clearly convex, reflecting the nature of the production technology. The bottom right panel illustrates the unemployment rate by worker type. Low type workers have a much higher unemployment rate, since the high productivity firms reject them. The unemployment rate increases for the very high types because the high skill workers prefer not to match with low types in order to assure a good match.

[10] We use the following specification for this graph. The distributions of (log) types are $N(0,0.5)$ for both workers and jobs. We assume that $\overline{N} = 1$, which implies that $v = u$. Also, we normalize the transition rates the following way: $\lambda^W = \lambda u = \lambda^F$. This reduces the transition parameters of the model to λ and δ. We choose $\lambda = 0.75$ and $\delta = 0.025$. Finally, we set the unemployment insurance parameter, b, to 0.

Figure 2 plots log-output and log-wages as functions of the worker and the firm's types. The first thing to notice is that both figures have "valleys". Those represent the regions outside the matching sets , i.e., whenever $\alpha(h,p)=0$.[11] Now, recall that $F(h,p)=hp$. Because of that, log-output grows linearly both in the worker's and the firm's type on the valid regions of the matching set.

The same does not hold true for wages. We can observe from the wage function that wages are unambiguously increasing in h. However, that is not true when increasing p. For workers with low h, increasing p can actually reduce wages. For workers with high h, increasing p always increase wages. This non-monotonicity of wages with respect to p suggest that the firm fixed effects in a wage regression are unlikely to capture the true ordering across firms productivities. However, the same graph suggests that wage data can still be useful to infer sorting. This is because of two reasons. First, in our model the high-skill workers work for the high-productivity firms. A consequence of that is that these workers work with other high skill workers as well.[12] Second, wages are monotonic in h, suggesting that the relationship between worker and coworker wages reflects that of the primitives.

As Proposition 1 shows, the fact that $w(h,p)$ is monotonically increasing in h, but not in p is more general than this example. The reason why the wages can be asymmetric (always increasing in h, but not in p) in a completely symmetric environment is due to the fact that wages are paid by the firms to the workers and not vice versa. Similar asymmetries are present in $F(h,p)-w(h,p)$. Also, Figure 2 gives the impression that these non-monotonicities are only relevant for the bottom part of the skill distribution. This is generally true for this particular setup (as one can see from the proof of Proposition 1 item 2), but this is not robust to extensions to the model such as a higher contact rate for the more productive jobs/workers. This extension can also eliminate the implication that the unemployment rate is increasing for the very top workers.

[11] We input a value slightly smaller than the minimum wage produced by the equilibrium of the model.

[12] So far the model has been formulated in terms of workers and jobs. In Section 4 we show how to introduce the concept of a firm to the model.

4 Measuring Sorting in the Model

In this Section we discuss alternative ways to measure sorting in the model described in Section 3. Our objective is to obtain an empirical measure, which can be calculated using available matched-employer-employee datasets, that correctly captures the extent of sorting in the model. We consider four measures of sorting. Two of those are based on the AKM methodology, while the remaining two not.

4.1 AKM methodology

The most common exercise applied to a matched employer-employee datasets is the linear error decomposition proposed by AKM. Their results motivated much of the subsequent empirical and theoretical literature in this field.

This procedure consists of estimating the following wage equation.

$$w_{it} = x_{it}\beta + \theta_i + \psi_{J(i,t)} + \varepsilon_{it} \tag{13}$$

where: w_{it} denotes the measure of earnings, x_{it} the vector of time-varying covariates, θ the worker's fixed effects and $\psi_{J(i,t)}$ the fixed effect of firm J, which employs worker i at time t. Note that x_{it} only includes time-varying variables, because of the collinearity with the fixed effects.

To estimate this model, a sample with a significant amount of mobility is necessary, because only workers who switch firms allow disentangling fixed effects of workers and firms. Another identifying assumption of this model is that the mobility patterns of workers are exogenous. It is clear that these are strong assumptions, and our model sheds light on them.[13] However, note that this does not affect our analysis since we are not using this econometric model to make inference, but rather to obtain information about the economic model. Moreover, in order to estimate this model using the method of direct inference using this linear regression as an auxiliary model, one would not need this to be correctly specified, but rather just to be informative of certain aspects of

[13] In this model separations are exogenous, but match formation from unemployment is endogenous.

the data.

Under the assumptions described above this statistical model can be estimated by ordinary least squares. However, this can be a challenging mathematical problem by itself, since an usual sample has millions of workers and hundreds of thousands of firms, each of which with a coefficient. To solve that we use the iterative conjugated gradient algorithm described in Abowd et al [2].

4.2 Empirical Measures of Sorting Based on AKM

This fixed effects methodology provides us with potential proxies for worker and firm heterogeneities: the fixed effects. With this measures in hand, then a natural measure of sorting is simply the correlation

Definition 3. *Assume we have a panel dataset that follows workers and firms over time. Let $\hat{\theta}$ and $\widehat{\psi}$ respectively be the estimated worker and firms fixed effects from Equation 13. Our first measure of sorting is then $S1 \equiv Corr\left(\hat{\theta}, \widehat{\psi}\right)$.*

As previously mentioned, a very robust and surprising result in this literature is that $S1$ is close to 0 or sometimes even negative, which is apparently at odds with the equilibrium of our model. However, non-monotonicities in Equation (6), discussed in Section 3, may distort the mapping between the firms' productivities and the wage fixed effects. The same non-monotonicities do not affect the ordering of wages across workers. Moreover, in the model, workers of similar skill tend to work together as they sort in similar firms. This motivates us to use a new moment to measure of sorting: the correlation between θ_i and the average θ among his coworkers. This leads to our second measure of sorting.

Definition 4. *Let $\hat{\theta}$ be the same estimate of the worker fixed effects as described above. We first construct the variable $\hat{\theta}_{i,t}^{other} \equiv \frac{\sum_{k \in \tilde{J}(i,t)} \widehat{\theta}_k}{N_{\tilde{J}(i,t)}}$, where where $\tilde{J}(i,t)$ denotes the set of workers at firm J, who employs i at t, minus worker i and $N_{\tilde{J}(i,t)}$ is the number of coworkers.*[14] *Our second measure*

[14] To avoid issues with singularities we restrict the sample to observations of firms that contain at least two workers at a given time. Alternatively, we tried setting $\theta_{i,t}^{coworkers}$ to zero whenever a firm had one worker, obtaining similar results.

of sorting is then $S2 \equiv Corr\left(\hat{\theta}, \hat{\theta}^{other}\right)$.

We use both the Pearson and the Spearman (rank) definitions of correlations. Most of the previous literature report the Pearson correlation of $S1$. However, there are at least two reasons to believe that the rank correlation is more appropriate. First, if this statistical model is not correctly specified, there is no reason to believe that the estimated fixed effects are a linear mapping of the true worker and firm types. The rank correlation would be more robust in that case since it is invariable to any strictly monotone transformation of the variables. Second, even if the model is correctly specified, in the seminal Becker [6] model the prediction is that the rank correlation of the true types is 1, not the Pearson correlation.

4.3 Evaluating the Empirical Measures of Sorting

In order to measure the performance of these measures we generate simulations from the model and apply these empirical measures to the simulations. The idea is to evaluate how these empirical measures perform in realistic sample sizes for a wide range of parameter values. We take this approach instead of aiming for asymptotic results because there are reasons to believe that these estimators suffer from small sample bias. In particular, Andrews et al. [4] show that the mechanics of the AKM estimator introduces a spurious negative bias on $S1$ in short samples. In Section 4.6 we consider an alternative measure of sorting that converges to the true sorting correlation when the sample size goes to infinity.

Since the model is in continuous time, we approximate the economy with the discrete-time correspondent, taking the periodicity to be a week. We simulate the path of 10000 workers, whose initial labor market state is drawn from the steady state distribution, over 7 years.[15] [16] For tractability we discretize the support of the type distributions using 50 points. We distribute the vector of probabilities of the (exogenous) worker and the job-type distributions proportionally to log-normal

[15] The wage regression only includes the data from workers who are employed, so a typical worker will show in the sample for 5-6 years. This is to mimic the length of a typical matched employer-employee dataset.

[16] This number of workers is smaller than the one found in typical matched employer-employee datasets. However, we found that increasing this number does not affect the results significantly, while it increases the computational time substantially.

distributions: $log(h) \sim N\left(0, \sigma^h\right)$ for the workers and $log(p) \sim N(0, \sigma^p)$ for the firms. We assume a production function that satisfies log-supermodularity: $F(h,p) = hp$. With regard to the the transition parameters, first, we assume that $\overline{N} = 1$, which implies that $v = u$. Also, we normalize the transition rates similarly as in the example of Section 3. Finally, we shut down the unemployment insurance: $b = 0$.[17]

Recall that our model is defined in terms of workers and jobs. However, when moving to the data, we need to introduce the concept of firms in this economy. We assume that firms only own jobs of one type, and are exogenously distributed according to the distribution $\Upsilon(p)$, with density $\upsilon(p)$. The mass of firms per worker is N_F . All jobs of a given type are divided equally among the firms of that type, which implies that each firm owns $\frac{\overline{N}\bar{g}(p)}{N_F \upsilon(p)}$ jobs. In our empirical specification we take $\Upsilon(p)$ to be Uniform over the support of $log(p)$.[18] In practice, we draw 1000 firms proportionally to the distribution $\Upsilon(p)$, in order to have the ratio of workers per firm similar to the data.[19] Thus, when a worker meets a firm, firms of type p have a chance $\frac{g(p)}{\sum_{j=1}^{1000} g\left(p_j\right)}$ of being chosen.[20]

We are left with 6 parameters in the model: the variances of the worker and job heterogeneities, σ^h and σ^p, the transition parameters δ and λ, the Nash bargaining parameter β and the discount rate ρ.[21] We pick five different values for each of these parameters, as shown in Table 1. This implies a total of 15625 combinations.

For each combination of parameters we perform the following steps:

1. We solved for the equilibrium of the model, using the algorithm proposed in Section 3.6. The algorithm worked 100% of the time.

[17] It is well known that structural estimates of search models usually yield values for b smaller than zero (see Mortensen [15]). One interpretation for that is that b includes, in addition to unemployment benefits, non-pecuniary considerations. In this model, increasing b would lead to some workers to exit the market altogether.

[18] This implies that low productivity firms are more frequent. We also tried exponential distributions, obtaining similar results.

[19] This implies an average of about 10 workers per firm. In Abowd et al [1] they report that in their sample there are in average 3.4 workers per firm in France and 18 in the US.

[20] Note that each firm owns $\frac{\overline{N}g(p)}{N_F \upsilon(p)}$ vacancies in steady state, and $\upsilon(p)$ comes from a Uniform distribution.

[21] All time rates are monthly. Also, recall that we assuming a quadratic meeting technology, which implies that $\lambda^W = \lambda^F = u\lambda$.

σ^h	0.25	0.4	0.55	0.7	0.85
σ^p	0.25	0.4	0.55	0.7	0.85
δ	0.034	0.067	0.1	0.13	0.17
λ	0.75	1	1.25	1.5	1.75
β	0.17	0.33	0.5	0.67	0.83
ρ	0.018	0.034	0.05	0.067	0.084

Table 1: Parameter Values for the Simulations.

2. We generated a simulated panel from the model, following the design described above.[22]

3. We applied the AKM fixed effects regression to the simulations from the model.

4. We computed S1 and S2, as well as the true sorting correlations $Corr(h, p)$ and $Corr\left(h, h^{other}\right)$.[23]

Figure 3 shows how our empirical measures of sorting relate to the true measures of sorting they are supposed to capture. We present the scatter plot of our empirical measures along with the respective theoretical counterpart, as well as the joint density of both. We can see that with the rank correlations the observations are closer to the 45^0 line than the Pearson correlations, so we will focus on panel B. The figure shows that $S1$ can be a very unreliable counterpart for $Corr(h, p)$, whereas $S2$ captures $Corr\left(h, h^{other}\right)$ remarkably well. First, comparing $S1$ to it's theoretical analogous, $Corr(h, p)$, we see that it consistently has a downward bias, which can be very severe. For example, we can see cases where $Corr(h, p) \approx 0.5$, but $S1 \approx -0.1$. In fact, for around 20% of the parameter combinations there is a downward bias of at least 0.2. Second, comparing $S2$ to it's theoretical counterpart, $Corr\left(h, h^{other}\right)$, we can see that it does a very good job capturing the true correlation. For all the parameter combinations $S2$ lie within 0.05 of $Corr\left(h, h^{other}\right)$.

4.4 Intuition Behind the Results

Now that we showed that $S2$ is a much more reliable measure of sorting than $S1$, we explain the intuition behind our results. Recall from Section 3 that wages are always increasing in worker skill,

[22] We use the same realization of shocks for all parameter combinations.

[23] $\hat{h}^{other}_{i,t} \equiv \frac{\sum_{k \in \tilde{J}(i,t)} \widehat{h}_k}{N_{\tilde{J}(i,t)}}$

but not necessarily in firm productivity. These non-monotonicities are the reason why the worker-firm correlation is distorted, but the worker-coworker is not. To substantiate this statement, Figure 4 shows how $S1$ and $S2$ perform when we apply the AKM methodology to output, rather than wages, for the same simulations as before. Recall that output in this formulation is $log\left[F\left(h,p\right)\right]=h+p$, so there are no non-monotonicities in that case. We can see that both $S1$ and $S2$ capture the true sorting correlations remarkably well.

Another way to see this is by looking at how the fixed effect estimators capture the unobserved heterogeneity of workers and firms individually. For the same combinations of parameters and simulations described above we compute $Corr\left(h,\theta\right)$ and $Corr\left(p,\psi\right)$.[24] Figure 5 present the CDFs of these correlations. As we can see, the worker fixed effects capture almost perfectly the relative rankings of workers, whereas the firm fixed effect can be a very noisy proxy for the underlying firm productivity.

4.5 How Reliable are These Estimators

In order to measure how reliable are the sorting estimators $S1$ and $S2$, we performed the following exercise. We chose a wide range of parameter combinations and repeated steps 1-4 of Section 4.3.[25] However, instead of generating one simulated panel, we generated 10 simulated panels and computed the relevant statistics for each of them.[26] We proceeded by computing the mean and the standard deviation of $S1$ and $S2$ for each combination of parameters. Figure 6 displays the joint distribution of means and standard deviations. The standard deviations for both $S1$ and $S2$ are relatively small (always below 0.025), which suggests that at these sample sizes these are fairly reliable estimators.

[24] These are rank correlations.

[25] Because of the size of this exercise we pick three values for each parameter, instead of five. The values are σ^h and $\sigma^p \in \{0.35, 0.55, 0.78\}$, $\delta \in \{0.05, 0.1, 0.15\}$, $\lambda \in \{0.87, 1.25, 1.63\}$, $\beta \in \{0.25, 0.5, 0.75\}$ and $\rho \in \{0.001, 0.0045, 0.009\}$.

[26] We keep the same 10 realizations of shocks for all parameter combinations.

4.6 Alternative Measures of Sorting

Thus far we have focused exclusively on the measures $S1$ and $S2$, both of which use the AKM methodology. This is a cumbersome procedure and is mispecified for this model, which invites the question of why using it in the first place. In this Section we consider two alternative measures of sorting, which do not rely on the AKM procedure. The first one, $S3$ uses only cross-section data and is inspired in the index of segregation proposed by Kremer and Maskin [10]. The second measure, which we call $S4$, is a theoretically consistent measure of sorting: the model predicts that as the sample size goes to infinity it converges to the true sorting correlation. We show that both $S3$ and $S4$ suffer from severe short sample bias, overestimating the degree of sorting in the economy.

First, we define

Definition 5. *Given a cross-section of wages let $\bar{w}$ be the average wage in the sample. Also, let Z_j be the set of workers at firm j and z_j the respective cardinality. Then,*

$$S3 \equiv \frac{\sum_{j=1}^{J} \sum_{i \in Z_j} (w_i - \bar{w}) \sum_{k \in Z_j} (w_k - \bar{w}) / z_j}{\sum_{j=1}^{J} \sum_{i \in Z_j} (w_i - \bar{w})^2}.$$

This measure is exactly the index of segregation proposed by Kremer and Maskin [10], applied to wages. A value of $S3$ close to zero indicates that all firms have the same skill-mix of workers—no sorting—, and a correlation of one indicates complete segregation, in which all workers within a firm have the same skill—like in Becker [6]. This can be computed simply by looking at the R^2 of a regression of worker wages on firm dummies. The theoretical counterpart to which $S3$ should be compared is the index of segregation for the worker skill—i.e, replace w with h on $S3$.

Second, we try a measure of sorting that converges to the true sorting correlation as the sample size goes to infinity. The idea is to use the fact that, in the model, average worker wages and average firm wages are respectively increasing in their types.

Definition 6. *Assume we have a panel dataset that follows workers and firms over time. Let T_i $\left(T_j\right)$ be the set of years that worker i (firm j) shows in the data and N_i $\left(N_j\right)$ the respective cardinality.*

Also, let $Z_{j,t}$ be the set of workers at firm j at time t and $z_{j,t}$ the respective cardinality. Finally, let $w_i^W \equiv \frac{\sum_{t\in T_i} w_{it}}{N_i}$ *and* $w_j^F \equiv \frac{\sum_{t\in T_j}\sum_{i\in Z_{j,t}} w_{i,t}}{\sum_{t\in T_j} z_{j,t}}$*. Then,* $S4 \equiv Corr\left(w_i^W, w_j^F\right)$

Recall from Proposition 1 that $w_h^W(h) > 0$ and $w_p^F(p) > 0$. Moreover, since $\gamma(h,p)$ is a stationary distribution it is the case that $lim_{N_i\to\infty}\frac{\sum_{t\in T_i} w_{it}}{N_i} = w^W(h)$ and $lim_{z_{j,t}\to\infty}\frac{\sum_{t\in T_j}\sum_{i\in Z_{j,t}} w_{i,t}}{\sum_{t\in T_j} z_{j,t}} = lim_{N_j\to\infty}\frac{\sum_{t\in T_j}\sum_{i\in Z_{j,t}} w_{i,t}}{\sum_{t\in T_j} z_{j,t}} = w^F(p)$. Thus, the rank correlation $S4$ asymptotically converges to $Corr(h,p)$ as we sample an infinite number of years.

Figure 7 shows how $S3$ and $S4$ perform in capturing their theoretical counterparts. As we can see, both measures consistently overestimate their theoretical counterparts. The reason for why that happens is because the rent sharing in the model introduces a spurious common element to wages—i.e., even if workers were randomly allocated to firms, there would be a common element in wages due to the fact that part of the firm's profits are shared with the workers. Also, the fact that $S4$ does not correct for this spurious element suggests that realistic sample sizes (as in number of years) are far from enough for asymptotic results to apply. Even though the AKM methodology is mispecified for this model, it still corrects fairly well for this spurious element by using the mobility information of workers, which explains why $S2$ outperforms $S3$ and $S4$.

5 Conclusion

We offer three contributions to the measurement of sorting in the labor market.

First, we show that the standard methodology of measuring sorting in the labor market may be biased in favor of no sorting. More specifically, we use an equilibrium model of the labor market with heterogeneous workers and firms that exhibits positive sorting in equilibrium, because of the presence of complementarities in production. We show that for a wide range of parameter values this measure substantially underestimates the true extent of sorting. Our second contribution is to identify the source of this bias in non-monotonicities in the wage equation (when increasing firm's productivity) caused by the interaction of wage bargaining with limited capacity of the firms to post new vacancies. In our model, high productivity firms have better outside options than their low- productivity counterparts, which causes downward pressure on the wages of their workers. This is particularly relevant for low-skilled workers, who can be paid less when working for a more productive firm. All this suggests that the firm wage fixed effects are not good proxies for the firms productivities.[27] Thirdly, we propose a new empirical method to detect sorting that is immune from this bias: the correlation between a worker's wage fixed effect and the average fixed effects of the co-workers at the same firm. We find that this method correctly detects sorting in the simulated data from the model. We also argued this worker-coworker correlation, based on the AKM methodology, is preferable to "simpler" measures of sorting, because it corrects for comovements in wages brought by rent-sharing that are unrelated to sorting.

Finally, one drawback of our worker-coworker measure of sorting is that it cannot detect the "sign of sorting". Recall that the positive complementarities of our model imply that "good" workers end up with other "good" workers as a byproduct of working for "good" firms. Another mechanism could be that these "good" workers are clustering in "bad" firms.[28] In fact, Eeckhout and Kircher [9] use a related model to argue that from wage data alone one cannot distinguish positive from negative sorting. They also argue, however, that for questions related to efficiency and inequality the degree of sorting is more important than the sign. Our work contributes in that direction: the empirical worker-coworker correlation captures very well the true worker-coworker correlation in the model, which tends to increase as we increase the degree of complementarities in the model.

[27] Abowd et al [1] also show an example where a theoretical model with positive sorting can imply a zero/negative correlation in wage fixed effects. However, their example requires wages to be decreasing in worker skill, and a counterfactual negative correlation between employment probability and wages.

[28] A third possibility would be that workers cluster together because of complementarities between workers.

6 Appendix

6.1 Proof of Proposition 1

1. $w_h(h,p) > 0, \forall \{(h,p) \,|\, \alpha(h,p) = 1\}$

   Assume that the primitives of the model are such that the value functions are differentiable.[29] If we differentiate Equation (6) with respect to h we obtain

$$w_h(h,p) = \beta F_h(h,p) + r(1-\beta) U_h(h).$$

   It remains to show that $U_h(h) > 0, \forall h$. Differentiating Equation (2) w.r.t. h, and using Equation (5)

$$rU_h(h) = \beta \lambda^W \int_{-\infty}^{\infty} \alpha(h,p^*) \frac{F_h(h,p^*) - rU_h(h)}{r+\delta} g(p^*)\, dp^*.$$

   In this step, it is not necessary to differentiate the bounds of the integral because the surplus is zero at those points. Rearranging the term with $U_h(h)$ to the LHS it is clear that $U_h(h) > 0, \forall h$.

2. $w_p(h,p) < 0, \exists \{(h,p) \,|\, \alpha(h,p) = 1\}$.

   Without loss of generality, let the support of h and p be $[0,1]$. Let $a_{mi}(h)$ and $a_{ma}(h)$ be respectively the minimum and the maximum job that worker h pair with, and $b_{mi}(p)$ and $b_{ma}(p)$ the same for the firms. We know that both are non-decreasing and $a_{mi}(0) = 0$ and $a_{ma}(1) = 1$. Fix h. Take the (h,p) pairs such that $p = a_{ma}(h) < 1$. In that case we know that $b_{mi}(p) = h$. If we differentiate Equation 6 we obtain

$$w_p(h,p) = \beta F_p(h,p) - r\beta J_p^U(p).$$

[29] See Shimer and Smith [17] for details.

Now, an analogous derivation to the previous one in this Section yields

$$rJ_p^U(p) = \frac{(1-\beta)\lambda^W \int_h^{b_{ma}(p)} F_p(h^*,p)\, l(h^*)\, dh^*}{r+\delta+(1-\beta)\lambda^W \int_h^{b_{ma}(p)} l(h^*)\, dh^*}.$$

Plugging in the previous expression we obtain

$$w_p(h,p) = \beta\left[\frac{F_p(h,p)(r+\delta)+(1-\beta)\lambda^W \int_h^{b_{ma}(p)} [F_p(h,p)-F_p(h^*,p)]\, l(h^*)\, dh^*}{r+\delta+(1-\beta)\lambda^W \int_h^{b_{ma}(p)} l(h^*)\, dh^*}\right].$$

Note that from supermodularity the term$[F_p(h,p)-F_p(h^*,p)]$ is always negative, and increasing in h. Since $F_p(h,p) \longrightarrow_{h\to 0} 0$ it follows that $w_p(h,p) < 0$ for a low enough h.

3. $w_h^W(h) > 0, \forall h$.

   Follows immediately from 1.

4. $w_p^F(p) > 0, \forall p$.

   First, it is useful to note that $\gamma(h,p) = \frac{u\lambda^W l(h) g(p)}{\delta(1-u)}$ .Next, Using Equation 6, the definition of $w^F(p)$ and the derived expression for $V_p(p)$ we obtain

$$\begin{aligned}
w_p^F(p) &= \frac{\int_{-\infty}^{\infty} \alpha(h^*,p)\gamma(h^*,p)\,[\beta F_p(h^*,p) - r\beta V_p(p)]\, dh^*}{\int_{-\infty}^{\infty} \alpha(h^*,p)\gamma(h^*,p) dh^*} \\
&= \frac{\int_{-\infty}^{\infty} \alpha(h^*,p) l(h^*)\, \beta F_p(h^*,p)\, dh^*}{\int_{-\infty}^{\infty} \alpha(h^*,p) l(h^*)\, dh^*} - r\beta V_p(p) \\
&> \frac{\int_{-\infty}^{\infty} \alpha(h^*,p) l(h^*)\, \beta\,[F_p(h^*,p) - F_p(h^*,p)]\, dh^*}{\int_{-\infty}^{\infty} \alpha(h^*,p) l(h^*)\, dh^*} \\
&= 0
\end{aligned}$$

References

[1] J. Abowd, F. Kramarz, P. Lengerman, and S. Perez-Duarte. Are Good Workers Employed by Good Firms? A Test of a simple assortative matching model for France and the United States. *unpublished manuscript*, 2004.

[2] J.M. Abowd, R. Creecy, and F. Kramarz. Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data. *unpublished manuscript*, 2002.

[3] J.M. Abowd, F. Kramarz, and D.N. Margolis. High Wage Workers and High Wage Firms. *Econometrica*, 67(2):251–333, 1999.

[4] M.J. Andrews, L. Gill, T. Schank, and R. Upwarde. High wage workers and low wage firms: Negative assortative matching or statistical artefact?

[5] J. Bagger and R. Lentz. An Equilibrium Model of Wage Dispersion and Sorting. *unpublished manuscript*, 2008.

[6] G.S. Becker. A Theory of Marriage: Part I. *The Journal of Political Economy*, 81(4):813–846, 1973.

[7] P. Cahuc, F. Postel-Vinay, and JM Robin. Wage Bargaining With On-The-Job Search: Theory and Evidence. *Econometrica*, 74(2):323–364, 2006.

[8] J. Eeckhout and P. Kircher. Identifying Sorting, in Theory. *unpublished manuscript*, 2008.

[9] J. Eeckhout and P. Kircher. Sorting and Decentralized Price Competition. *unpublished manuscript*, 2008.

[10] M. Kremer, E. Maskin, L. Center, T. Floor, and E.D.W.B. Room. Wage inequality and segregation by skill. *NBER working paper*, 1996.

[11] Rasmus Lentz. Sorting in a General Equilibrium On-the-Job Search Model. *unpublished manuscript*, 2007.

[12] J. Lise, C. Meghir, and J.M. Robin. Matching, Sorting and Wages. *unpublished manuscript*, 2008.

[13] R. Lopes de Melo. *Estado de uma Nação*, chapter Emparelhamento no Mercado de Trabalho Brasileiro: Evidência e uma Explicação. Ipea, 2009.

[14] X. Lu and R.P. McAfee. Matching and Expectations in a Market with Heterogeneous Agents. *Advances in Applied Microeconomics*, 6:121–156, 1996.

[15] D. Mortensen. *Wage Dispersion: Why are Similar Workers Paid Differently*. MIT Press, 2003.

[16] S. Shi. Frictional assignment. i. efficiency. *Journal of Economic Theory*, 98(2):232–260, 2001.

[17] R. Shimer and L. Smith. Assortative Matching and Search. *Econometrica*, 68(2):343–369, 2000.

[18] C.N. Teulings and P.A. Gautier. The Right Man for the Job. *Review of Economic Studies*, 71(2):553–580, 2004.

[19] S.D. Woodcock. Heterogeneity and Learning in Labor Markets. *unpublished manuscript*, 2005.

Figure 1: Characterization of Equilibrium, Symmetric Case

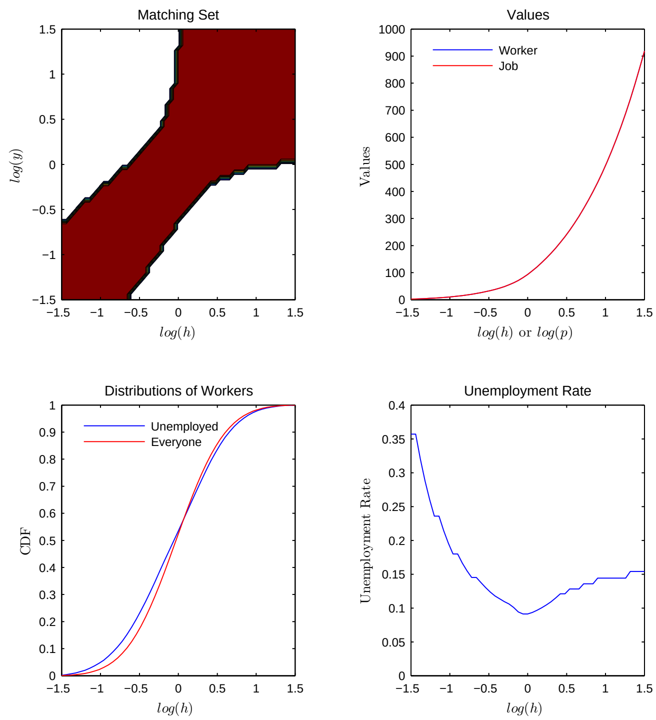


Figure 2: Log-Output and Log-Wages in the Model

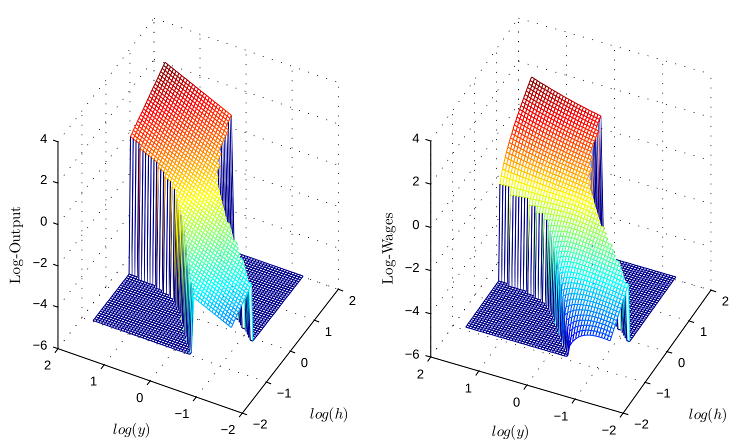
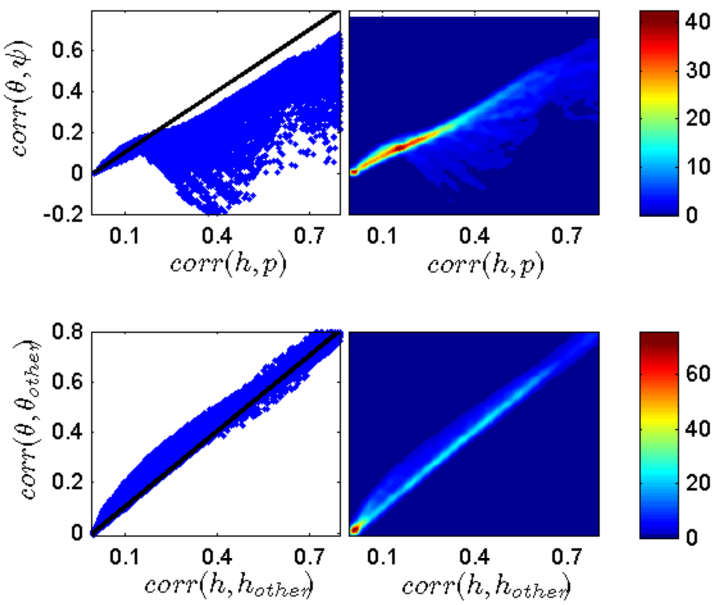


Figure 3: True Model Correlations vs Fixed Effects Correlations: Scatter Plots and Densities

A) Pearson Correlations



B) Rank Correlations

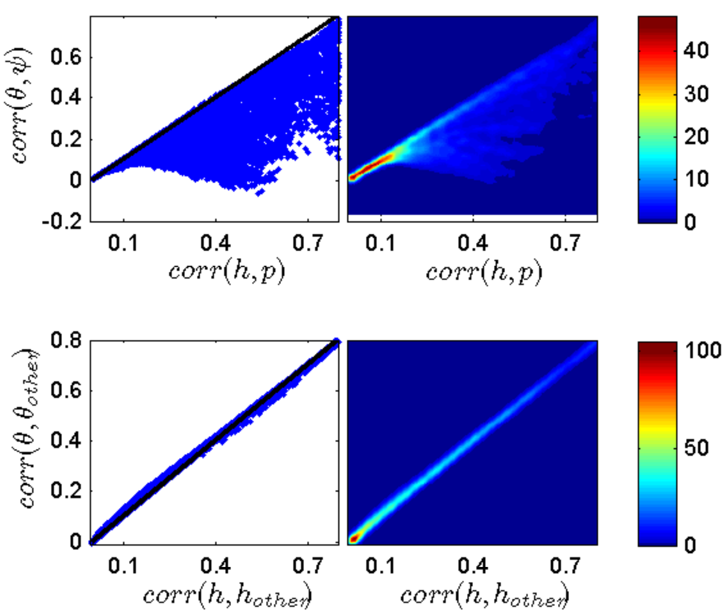


Figure 4: AKM correlations Using Output Instead of Wages: Scatter Plots and Densities Rank Correlations

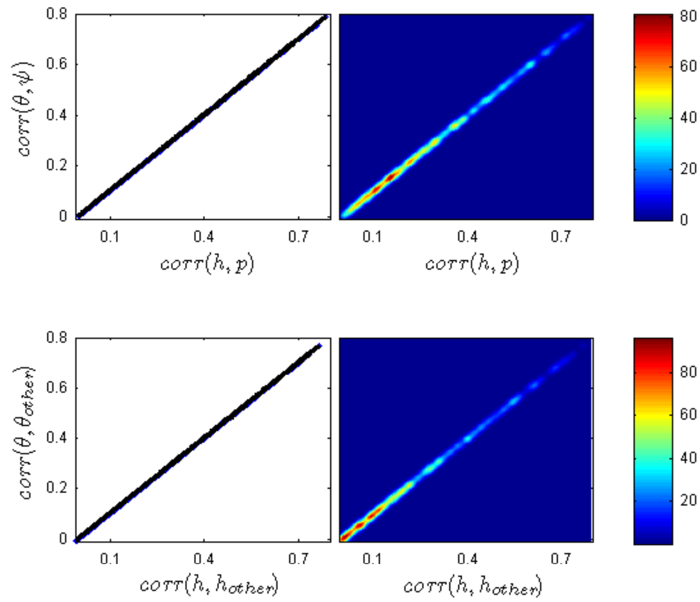


Figure 5: Correlation Between Types and Fixed Effects for Firms and Workers

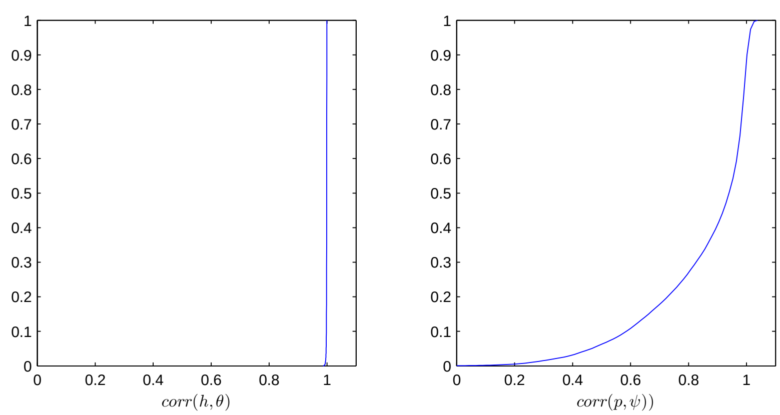


Figure 6: Variability of Estimators

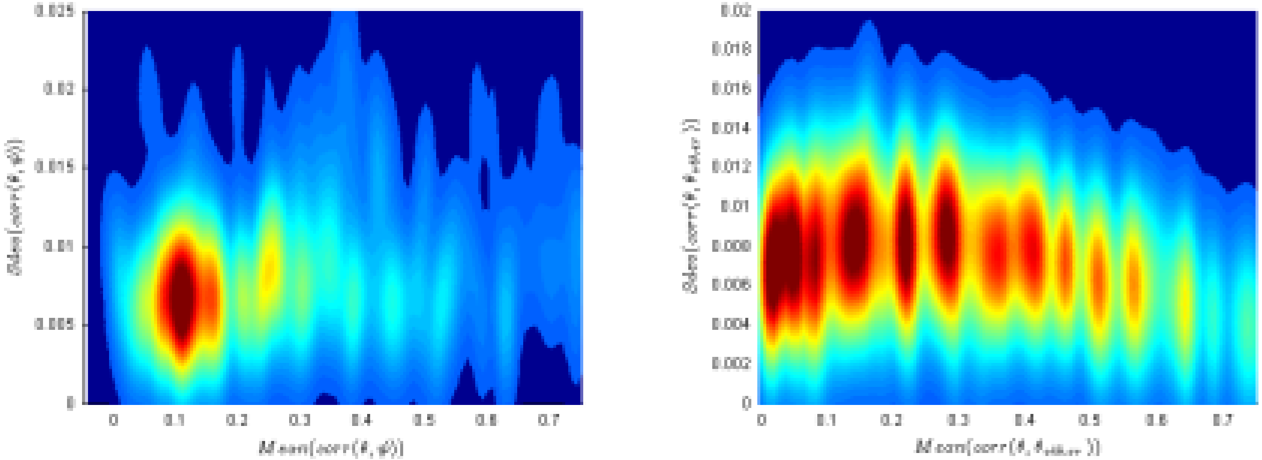


Figure 7: Alternative Measures of Sorting: Scatter Plots and Densities