Firm wage differentials and labor market sorting

Reconciling theory and evidence

Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence

Rafael Lopes de Melo

University of Edinburgh

Why do firms pay different wages? Empirical evidence suggests the presence of substantial differences in firm pay controlling for worker skill. Moreover, these differences are uncorrelated with skills, indicating the absence of sorting. I show that the face value interpretation is inconsistent with evidence on coworker segregation. I interpret the evidence by applying a sorting model and show that the correlation is biased. I identify nonmonotonicities in wages as the reason for this bias and show that a measure of worker-coworker sorting is more accurate. By calibrating the model to US data, I confirm that the model matches many job market characteristics.

I. Introduction

It is well documented that a substantial share of wage dispersion takes place between firms. For example, Davis and Haltiwanger (1991) performed a variance decomposition of US manufacturing wages and found that around

000

60 percent of dispersion occurs between rather than within firms. Similarly, the documented large firm-size wage differentials (e.g., Oi and Idson 1999) and firm industry differentials (e.g., Krueger and Summers 1988) have been the focus of a number of studies. These differentials could be attributed to differences in worker skill levels across firms as well as discrepancies in wage policies between firms. The latter could arise because of noncompetitive features of a labor market. Abowd, Kramarz, and Margolis (1999), Abowd, Creecy, and Kramarz (2002), and Abowd et al. (2004) (hereafter AKM) attempted to quantify the relative importance of worker versus firm components in determining wages. The authors focused on estimating a wage regression that includes both worker and firm fixed effects, which was achieved through the use of longitudinal matched employer-employee data sets. Their methodology was subsequently applied to the data sets of several different countries with consistent results. Authors of extant works based on this approach reported substantial firm fixed-effects differentials, which account for a sizable share of firm-size differentials and industry wage differentials. Another robust result arising from this methodology is that the correlation between the two sets of fixed effects is close to zero, or sometimes even negative. Card, Heining, and Kline (2013) recently quantified how changes in the dispersion of worker and firm fixed effects, and their correlation (referred to as sorting), explain changes in inequality over time.

This work aims to provide structural interpretation of the facts obtained by applying the AKM methodology, using explicit models of labor market dynamics. I examine models that include both worker and firm heterogeneity, as well as a job-matching process that yields mobility and assortative matching. My first aim is to make explicit the connection between the AKM model and the widely used equilibrium job search model, which I label "the piece-rate model." The model features a dynamic matching process between heterogeneous workers and firms in a frictional environment and has been shown to successfully explain several characteristics pertaining to labor market transitions and wage dynamics. According to its postulates, the frictions in the economy cause more productive firms to pay higher wages to identical workers, resulting in a self-selection process of workers into firms. However, this process is independent of worker skill level, which implies the absence of sorting in equilibrium. As a result, this model provides a "face value" interpretation for the facts. More specifically, the worker fixed effects capture differences in skill levels across individual employees, whereas the firm fixed effects capture wage differentials paid by more productive companies as a result of frictions. Under this interpretation, the high dispersion of firm fixed effects found in the data is indicative of a sizable degree of frictional wage dispersion. Moreover, the fact that the correlation between fixed effects is close to zero is supportive of an equilibrium with no sorting between worker and firm types.

This paper's first contribution stems from demonstrating that the piece-rate model is inconsistent with new evidence, which I sourced from a Brazilian matched employer-employee data set. Applying the AKM methodology, I obtain results largely similar to those reported in the extant literature with respect to the facts already described. In addition, I also provide evidence of a high correlation between the fixed effects of workers and their coworkers, which suggests a strong degree of worker segregation among firms.[1] This violates the model's implication that the distribution of workers is the same across firms. Furthermore, I use data on education, occupation, sectors, and locations to demonstrate that this correlation is not driven by composition of observables, a possibility that has been allowed for in empirical applications of the model.

After presenting the evidence of coworker sorting, I shift the focus toward a variant of the job search model that actually features assortative matching in equilibrium. A clear challenge stems from the need for the model to account for the zero correlation between fixed effects while permitting sorting in equilibrium. Thus, I build on the frictional matching model with assortative matching proposed by Shimer and Smith (2000). The main feature of the model is the addition of on-the-job search (both voluntary and involuntary), which is a pervasive feature of labor markets and is shown to substantially affect the extent of frictional wage dispersion in the model (e.g., Hornstein, Krusell, and Violante 2011).

The proposed model incorporates an equilibrium wage function that does not conform to the AKM specification. In AKM, firms give an extra bonus to all their employees (the fixed effect), whereas this model yields wages that are nonmonotone with respect to firm productivity. In the sorting economy, each worker's skills and preferences can be matched with an ideal firm type, which is reflected in wages. If workers gain a position in a more productive firm than their ideal one (because of frictions), they will earn less because of the need to compensate the more productive firm for giving up on finding a more skilled worker. These nonmonotonicities have two implications with respect to how the model maps into the AKM methodology. First, since they distort the mapping between firm types and firm fixed effects, the correlation between fixed effects no longer captures correctly the extent of sorting in the economy. More specifically, even if the model allows for a large degree of sorting, this empirical measure may still be negligible. A similar finding was reported independently by Bagger and Lentz (2014) and Lise, Meghir, and Robin (2016). However, to the best of my knowledge, this work is the first to relate this finding

---

[1] Bagger and Lentz (2014) computed this measure on Danish data, and Tzuo Hann Law computed it on German data, both obtaining results similar to mine. I also relate the findings to those of Iranzo, Schivardi, and Tosetti (2008), who report a measure of segregation of worker skill.

to the nonmonotonicities in wages.[2] Second, the same nonmonotonicity that helps explain the correlation between fixed effects may cause problems when attempting to elucidate the observed variance of firm fixed effects. More specifically, since each worker has a different ideal firm type, a given firm will increase the wage of some workers (relative to another firm) and decrease the wage of others, which may lead to a small average "firm effect." To my knowledge, I am the first to make this observation.

I calibrate the model by choosing a set of parameters in order to match a number of features of the US economy. I show that the model is able to explain a number of facts about labor market transitions, the distribution of wages, and wage dynamics. The model also matches well the correlations between fixed effects in AKM, generating a sizable correlation between workers and coworkers, while generating a correlation between workers and firms that is negligible. However, the model generates less variance in firm effects than can be discerned from the data analysis. This is consistent with the tension induced by the nonmonotonicity described earlier. Nonetheless, I perform a second and much less parsimonious calibration, verifying that the results are robust to the more flexible parameterization. This leads to the conclusion that neither the piece-rate model nor this version of the sorting model is fully consistent with the reduced-form evidence, suggesting that additional features should be incorporated. I discuss some possibilities for the modification of the model in the conclusion. However, I note a route pursued in a companion paper (Lopes de Melo 2015) as a particularly fruitful one, as I extend the Shimer and Smith (2000) model to include a second dimension of firm heterogeneity: compensating differentials.

The remainder of the paper is structured as follows. Section II summarizes the related literature, while the basics of the piece-rate model are presented in Section III, where I also summarize the empirical evidence and describe my own work. Section IV is designated for the detailed description of the theoretical model. In Section V, I present the model calibration, aiming to match its output to a number of labor market facts. Finally, Section VI concludes the paper, along with a discussion of future research.

II.   Related Literature

The work described here relates to several strands of literature. First, the essence of the empirical methodology used throughout the paper is derived from the extensive work of Abowd et al. (1999, 2002, 2004), described earlier in this paper.

---

[2] Eeckhout and Kircher (2011) show that this mechanism does not rely on frictions and applies to the frictionless matching model as well.

Second, the model is grounded in the theoretical assignment literature. Becker (1973) presents a model with two-sided heterogeneity, demonstrating that, in the presence of complementarities in production, the equilibrium exhibits perfect sorting. Shimer and Smith (2000) introduced search frictions (via random search) in the Becker model. This approach adds noise to the equilibrium allocations and thus requires stronger complementarities for the equilibrium to exhibit positive assortative matching. Eeckhout and Kircher (2010) developed a directed search model that can be viewed as an intermediate case between the frictionless model and the economy with search frictions. That framework requires fewer complementarities than the search case, but more than the frictionless economy, to induce positive sorting. This approach is related to the work reported by Shi (2001).

Third, authors of a recent stream of studies applied models with firm and worker heterogeneity and matching to study labor markets. In some of these works, matching models with worker and firm heterogeneity are proposed while failing to capture assortative matching in equilibrium. Notable examples of studies based on this framework are those conducted by Postel-Vinay and Robin (2002), Christensen et al. (2005), and Barlevy (2008), among many others.

Finally, another recent stream of research is based on models that allow for incorporating more elaborate patterns of matching between workers and firms to understand labor markets. I believe that this work should be classified in this category. In one such study, Bagger and Lentz (2014) formulate a search model with endogenous search intensity and structurally estimate it using the Danish matched employer-employee data set. This strategy requires weaker complementarities in production to achieve positive sorting, characterized by a highly skilled worker searching with a higher intensity. In their work, Lise et al. (2016) build on the model developed by Shimer and Smith (2000), introducing on-the-job search, a wage mechanism akin to that introduced by Cahuc, Postel-Vinay, and Robin (2006), and a dynamic process that causes firms to change productivity over time. These authors estimate the model using National Longitudinal Study of Youth data and refer to its output to explain certain features of wage dynamics and labor market transitions. In related work, Eeckhout and Kircher (2011) used variations of Becker (1973) to argue that, by analyzing the wage data alone, it is not possible to distinguish a model that features positive sorting from that based on negative sorting. However, they also argue that wages can provide information about the strength of sorting, which is related to the underlying mechanisms of the worker-coworker measure described in this paper. Finally, Hagedorn, Law, and Manovskii (2012) provide results pertaining to the nonparametric identification of the model developed by Shimer and Smith (2000) using matched employer-employee data.

III. Wage Dispersion and Sorting at the Firm Level: The Old View

In this section, I present facts related to wage dispersion and sorting at the firm level, as well as the "face value" interpretation of these facts. First, I introduce a matching model with frictions that features no sorting in equilibrium, in line with models that are widely used in the related literature. The model yields a number of testable predictions, and its structure is directly related to the wage decomposition of AKM, which I make explicit. I proceed with summarizing the related empirical evidence, as well as presenting my own results obtained using a Brazilian matched employer-employee data set. I argue that the face value interpretation is incongruent with the facts, which motivates the analyses and discussions in the remainder of the paper.

A. A Matching Model with Heterogeneity and No Sorting

This model is based on the islands model developed by Lucas and Prescott (1974), whereby islands are reinterpreted as firms. The model is set in continuous time and is populated by a unit mass of workers, each with a type x drawn from distribution $L(x)$. At each point in time, the worker either is "in transit" or occupies an island of type y. Workers in transit are unemployed and earn a flow benefit xb. These workers sample islands at rate λ^0 by taking draws from a distribution $G(y)$. Within islands, there are competitive labor markets, where workers immediately find work and earn their marginal product $w(x, y) = F(x, y) = xy$. Employed workers have the opportunity to move to island y' at rate λ^1, where the new island is drawn from the same distribution $G(y')$. These workers also face a risk of exogenous displacement at rate δ. The value function of unemployed, $U(x)$, and employed workers, $W(x, y)$, can be expressed as follows:

$$rU(x) = xb + \lambda^0 \int [W(x, y') - U(x)] dG(y'),$$

$$(r + \delta)W(x, y) = xy + \delta U(x) + \lambda^1 \int [W(x, y') - W(x, y)] dG(y').$$

The problem the workers face stems from the need to decide when to accept job opportunities in this economy. Employed workers move from one island to another if the new island is more productive than the current one, $y' > y$. Unemployed workers accept jobs above a reservation productivity level $R(x)$, which is such that $W(x, R(x)) = U(x)$. It can be easily demonstrated that the solution to this problem has the following properties:

1. $U(x) = x\overline{U}$ and $W(x, y) = x\overline{W}(y)$.
2. $R(x) = \overline{R}$ and is such that $\overline{W}(R) = \overline{U}$.

Since every worker adopts the same reservation strategy, it follows that $x \perp y$ in equilibrium. More specifically, workers self-select into better jobs, y; however, that matching process does not depend on the type of the worker, resulting in no assortative matching between firm and worker types.

This framework encompasses a class of models of bargaining and competition with firm heterogeneity, which I label the "piece-rate" model.[3] In those models, y does not represent the marginal product per unit of skill, but rather the piece rate, that is, the amount of flow output that the worker appropriates in negotiations per unit of skill. Mortensen (2003) discusses two wage-setting mechanisms that generate a piece-rate offer distribution, one in which firms make wage offers under full commitment (i.e., as described by Burdett and Mortensen [1998]) and another in which firms and workers negotiate a wage in every period. The details are different, but both sets of assumptions have the same implication: piece rates are a strictly increasing function of firm productivity. That assertion implies that those models have the same implications for sorting and the same log-additive wage equation analyzed in Section III.B.

B.   AKM Methodology

The wage equation implied by this model is closely related to the AKM empirical methodology, which I make explicit. I assume that each island in the economy represents a firm and that log wages are measured with error: $\log(w) = \log(xy) + \varepsilon$, where ε denotes classical measurement error. The log wage of a worker i at firm J at time t can be decomposed into two sets of fixed effects, pertinent to the worker and to the firm, respectively, as well as an error term:

$$
\begin{aligned}
\log[w_{iJ(i,t)}] &= \log(x_i) + \log[y_{J(i,t)}] + \varepsilon_{it} \\
&= \theta_i + \psi_{J(i,t)} + \varepsilon_{it},
\end{aligned}
\tag{1}
$$

where $\theta_i = \log(x_i)$ and $\psi_{J(i,t)} = \log[y_{J(i,t)}]$.

In order to estimate equation (1), it is necessary to have access to a matched employer-employee data set, which follows workers and firms over time. This equation has been estimated using data sets from several countries, including that provided by Abowd et al. (1999), who applied the method to French data. The econometric model is based on a num-

---

[3] This model is very similar to the ones presented in the works of Christensen et al. (2005), Barlevy (2008), and others. Postel-Vinay and Robin (2002), Cahuc et al. (2006), and Bagger et al. (2014) use variations of this model, allowing wage contracts to be renegotiated while on the job, akin to the model presented in Sec. IV, while still featuring the no-sorting implication.

ber of identifying assumptions. First, it is assumed that the assignment of workers to firms is uncorrelated with the error term ε_{it}. This condition is satisfied in the baseline model if ε_{it} represents classical measurement error. Second, identification in this model is assured only within groups of workers and firms that are connected. According to Abowd et al. (2002), a group of persons and firms is connected when it comprises all the workers who have ever worked for any of the firms in the group and all the firms at which any of the workers have been employed at some point in time. In typical matched employer-employee data sets, the largest group consists of over 95 percent of the observations. Thus, restricting attention to the largest group of the sample is a common solution. Under these assumptions, this statistical model can be estimated by ordinary least squares, which is a challenging mathematical problem, since usual samples comprise data pertaining to millions of workers and hundreds of thousands of firms. Abowd et al. (2002) propose the use of an iterative conjugated gradient algorithm, which is used by authors of most extant studies, and which I adopt for my own calculations. Finally, it is worth noting that typical implementations of this equation include an extra term $x_{it}\beta$, which captures the effects of time-varying observables (e.g., the returns to experience and time effects).

The piece-rate model allows predictions of assignment patterns and provides a structural interpretation for the firm fixed effects. Recall that, in the model, worker and firm types are independent in equilibrium. One implication of this relation is that worker and firm fixed effects should be uncorrelated in the cross section: $\text{Corr}(\theta_i, \psi_{J(i,t)}) = 0$. Another testable implication of the model is that the distribution of worker types is the same across firms. In particular, the model predicts that worker types should be uncorrelated to the average type of their coworkers: $\text{Corr}(\theta_i, \tilde{\theta}_{J(i,t)}) = 0$, where $\tilde{\theta}_{J(i,t)}$ is the mean value of θ among the coworkers of worker i.[4] Finally, it is worth noting that the model provides a structural interpretation for the fixed effects. More specifically, θ reflects differences in pay due to worker skill level, while ψ reflects frictional wage dispersion due to workers finding better or worse jobs. Consequently, I also highlight the relative variance between worker and firm fixed effects, $\text{Var}(\psi)/\text{Var}(\theta + \psi)$.

C. Evidence across Data Sets

This methodology has been applied to data sets from many different countries with consistent results. Here, I summarize the results of studies from five countries (the United States, France, Germany, Italy, and Den-

---

[4] In samples comprising large firms, this moment is identical to the index of segregation proposed by Kremer and Maskin (1996), using θ as the measure of skill: $\text{Var}(\theta|j)/\text{Var}(\theta) = \text{Corr}(\theta_i, \bar{\theta}_{J(i,t)})$, where $\bar{\theta}_J$ denotes the average value of θ in firm J.

mark), alongside the results pertaining to the Brazilian matched employer-employee data set. I use the labor market census RAIS (Relacao Anual de Informacoes Sociais), an administrative data set collected annually by the Brazilian labor ministry, which includes all firms in the Brazilian formal sector and provides information for all their workers. I restrict the focus to the country's richest state, Sao Paulo, as well as to workers who have at least a high school diploma. In addition, the analyses pertain only to workers who have at least a moderate degree of formal labor market attachment and are included in the sample in at least 5 of the 11 years. These restrictions are intended to minimize the problem of the lack of coverage of the informal sector of the economy. I provide further details pertinent to the data, as well as a description of the sampling selection criteria, in Section A of the appendix.

Table 1 summarizes the results from six countries, focusing on the moments described in Section III.A. It should first be noted that the AKM regression has a very high explanatory power, accounting for approximately 90 percent of wage variation in most studies. It is well known that typical observable terms in Mincer regressions (experience, gender, etc.) have a low explanatory power; hence, most of the explained variation arises from the fixed-effects terms θ and ψ. Second, across all studies the firm fixed effects also account for a sizable share of wage dispersion, whereby $\text{Var}(\psi)/\text{Var}(\theta + \psi)$ ranges from 0.19 to 0.32. Third, across all data sets the value of $\text{Corr}(\theta, \psi)$ either is very close to zero (as in the case of the United States, Italy, or Brazil) or takes a small negative value (as in France or Germany). If we interpret this evidence using the piece-rate model, most of the wage dispersion is driven by worker heterogeneity, while frictions still account for a significant fraction of wage dispersion, as reflected by the value of $\text{Var}(\psi)/\text{Var}(\theta + \psi)$ at around 20–30 percent. In addition, the evidence that $\text{Corr}(\theta, \psi)$ is close to zero in most data sets seems supportive of the model implication of no sorting in labor markets.[5] However, this last conclusion is incongruent with the last piece of evidence, as $\text{Corr}(\theta, \tilde{\theta})$ ranges from .17 to .52 across data sets. This correlation suggests a substantial degree of clustering of workers across firms based on their skill level, which contradicts the implication that the distribution of worker skills is the same across firms.

D.   Sorting across Submarkets

In Section III.C, I concluded that the piece-rate model is inconsistent with the high degree of clustering of workers and coworkers reflected in the values of $\text{Corr}(\theta, \tilde{\theta})$ across a number of data sets. In order for the piece-rate

---

[5] Postel-Vinay and Robin (2002) and Christensen et al. (2005) use this moment to justify absence of a sorting equilibrium.

TABLE 1
SUMMARY OF EVIDENCE BASED ON AKM REGRESSIONS

	United States		France	Germany	Italy	Denmark	Brazil
	(1)	(2)					
$Var(x\beta)$	.03	.14	.02	…	.01	…	.02
$Var(\theta)$	.29	.23	.21	.05	.05	.07	.40
$Var(\psi)$	.08	.053	.08	.013	.01	.014	.18
$Var(\psi)/Var(\theta + \psi)$	.22	.19	.32	.22	.23	.16	.31
$Corr(\theta, \psi)$	−.01	−.03	−.28	−.10	.04	…	.04
$Corr(\theta, \theta)$ [a]	…	…	…	.51	.17	.40	.52
R^2	.89	.9	.84	.79	…	…	.93
Sample statistics:							
Years	1990–99	1984–93	1976–87	1993–97	1981–97	1985–2003	1995–2005
Observations	37.7 million	4.3 million	5.3 million	4.8 million	…	600,000	16.0 million
Workers	5.2 million	293,000	1.2 million	1.8 million	1.7 million	66,000	2.0 million
Firms	476,000	80,000	500,000	1,821	421,000	25,000	137,000
% 1st group [b]	…	99.1%	88.3%	94.9%	99.5%	…	98.6%

SOURCE.—United States in col. 1 is sourced from Woodcock (2015), which covers two nonidentified states and includes all workers who were employed in 1997. United States in col. 2 and France are derived from Abowd et al. (2002). The US data cover 1/10 of workers in the state of Washington, whereas the French data pertain to 1/25 of all workers. Germany is obtained from Andrews et al. (2008) and is based on data from around 2,000 establishments in West Germany. Italy is sourced from Iranzo et al. (2008), which covers 1,200 plants with at least 50 workers. Denmark is sourced from Bagger et al. (2014) and includes all workers with 15–20 years of education. Brazil refers to my own calculations.

 [a] The coworker correlation for Germany was provided by Tzuo Hann Law, using a sample different from the one employed by Andrews et al. (2008). For Italy, Iranzo et al. (2008) compute the index of segregation proposed by Kremer and Maskin (1996), using θ as their measure of skill. When firms are large (as in their sample), that measure is very similar to our worker-coworker measure. However, these authors use Pearson correlations instead of rank correlations. The coworker correlation for Denmark was provided by Bagger and Lentz (2014).

 [b] This denotes the fraction of the sample in the largest connected group.

model to explain those facts, the clustering can be assumed to be driven mechanically by composition. The argument supporting this strategy is as follows. Assume that the data are generated by the model presented in Section III.A but separately across "groups" of workers and firms. A firm under this interpretation consists of the set of workers who belong to the respective "group" in a given establishment. Using an example of lawyers and law firms and call centers and their workers as two separate submarkets, it can be assumed that, on average, the lawyers have a higher skill level relative to call center employees. Even though within groups there would be no firm-worker sorting, the law firms are in greater need of highly skilled workers, which would result in clusters of such workers in these firms. Most empirical implementation of piece-rate models allow for this type of sorting patterns by assuming existence of submarkets divided by education levels or occupation categories.[6] It is worth noting that, in a more general setting with assortative matching, the correlations within groups do not need to be smaller than the unconditional correlation. This can be shown on an example of managers and maintenance staff as the two groups, assuming, once again, that most managers are more skilled than maintenance workers. Assume that firm 1 hires both better managers and maintenance staff than firm 2, but it has a higher ratio of staff to managers. It is easy to construct cases in which the average skill level in the two firms is identical, even though segregation within groups exists.

I show evidence on this hypothesis using the Brazilian sample. To do so, I first divide the data into groups consisting of combinations of education categories and occupations, sectors, or location data. I compute the value of the AKM moments for each of the groups before calculating the average across groups, weighting by the size of each group: $E[\text{Corr}(\theta, \psi G)]$, $E[\text{Corr}(\theta, \tilde{\theta}|G)]$, and $E[\text{Var}(\psi|G)/\text{Var}(\theta + \psi|G)]$. I restrict the sample to firms with at least two workers of a given group and groups with at least 1,000 observations and multiple firms. The results are displayed in table 2. Since the resulting sample is different for each case, I also include the unconditional moments. In all cases, the worker-coworker measure is well above zero, and in the case of education/occupations, it is even greater than the unconditional one, which suggests that the results are not driven by composition. It is also worth pointing out that the relative dispersion in firm fixed effects, $\text{Var}(\psi)/\text{Var}(\theta + \psi)$, is higher within groups than overall, irrespective of the choice of groups. This is consistent with the results obtained by Bagger et al. (2014), who report values of $\text{Var}(\psi)/\text{Var}(\theta)$ between 1.13 and 2 using Danish data and three educational groups. This congruence in findings suggests that some forces that are not captured

---

[6] For example, Postel-Vinay and Robin (2002) allow for seven occupation classes using French data, and Christensen et al. (2005) allow for six occupation classes using Danish data. In addition, see the discussion preceding Postel-Vinay and Robin's n. 19.

TABLE 2
CORRELATION WITHIN GROUPS

GROUP	NUMBER OF GROUPS	OBSERVATIONS	UNCONDITIONAL			WITHIN GROUPS		
			(A)	(B)	(C)	(A)	(B)	(C)
Education	2	15.8 million	.05	.52	.33	.07	.48	.45
Education/occupations[a]	351	9.5 million	.03	.55	.32	.00	.57	.41
Education/sectors	788	15.6 million	.05	.52	.33	.02	.40	.36
Education/location	510	15.5 million	.05	.52	.33	.06	.44	.43

NOTE.—In the sample, two levels of education (high school diploma and college degree) are included. Occupation data are given at a three-digit level, while the sectors are presented using five digits. Location data are at the municipality level. $(A) = \mathrm{Corr}(\theta, \psi)$, $(B) = \mathrm{Corr}(\theta, \tilde{\theta})$, and $(C) = \mathrm{Var}(\psi)/\mathrm{Var}(\theta + \psi)$.

[a] Occupation classification changed after 2002, so I use these data only for the period 1995–2002.

by the model, such as compensating differentials between sectors or differences in local prices/amenities, are not the main component behind the variability in ψ.

IV. A Sorting Model with Frictions

Having concluded that the sorting patterns in the data are inconsistent with the piece-rate model, I introduce a model that incorporates sorting in equilibrium. In this section, I outline the main features of the model, leaving the details to the appendix. The model shares the basic characteristics of that proposed by Shimer and Smith (2000) in that it describes an economy comprising heterogeneous workers and firms, as well as search frictions. The main innovation that I introduce into the model stems from the addition of on-the-job search, which occurs in both a voluntary and involuntary manner. This is important because a large fraction of employment separations involve job-to-job transitions (e.g., Fallick and Fleischman 2004) and because it substantially affects the way frictions shape wage inequality (e.g., Hornstein et al. 2011).

A. The Environment

This is a continuous-time economy, with a unit mass of workers and a measure J of jobs. I assign a value $x \in [\underline{x}, \bar{x}]$ to each worker and $y \in [\underline{y}, \bar{y}]$ to each job, where the worker index is supposed to capture his or her level of human capital, whereas the job index reflects entrepreneurial talent, differences in the stock of capital, and so forth. It is assumed that these types are observable and the distributions are known to all agents. Workers and jobs are distributed according to density $l(x)$ and $g(y)$, respectively. In this economy, each firm represents a collection of jobs of a certain type without cross-complementarities. In line with the work of Shimer and Smith

(2000), the equilibrium characterization is initially described in terms of jobs. However, when the model is applied to the data at the firm level, the notion of firms will be applied. I further assume existence of a measure E of firms, with distribution $v(y)$. For each value of y, all jobs are evenly distributed across the firms of such a type such that

$$E \int_{\underline{y}}^{y} n(y') v(y') \, dy' = J \int_{\underline{y}}^{y} g(y') \, dy',$$

where $n(y)$ is the mass of jobs per firm of type y.

We can study a stationary environment in which workers and jobs are matched via a random search technology. Workers search for jobs while both employed and unemployed, whereas firms offering jobs search for suitable candidates only while these positions are vacant. In addition, unemployed workers encounter vacancies at a rate λ^{U}, while employed workers come across vacancies at rate λ^{E}. Matches are dissolved either exogenously at rate δ or when workers face "reallocation shocks" at rate λ^{R}, which forces the worker to leave his or her current employer and immediately encounter a new firm. As discussed by Jolivet, Postel-Vinay, and Robin (2006) and Lopes de Melo (2007), these shocks help explain the high fraction of job-to-job transitions that involve a wage cut. On the firm side, vacancies are compared with the characteristics of unemployed workers at a rate λ^{F}, employed workers at rate λ^{FE}, and workers subjected to reallocation shocks at rate λ^{FR}. Even though these rates are exogenous for our purposes, they need to satisfy steady-state conditions, which I describe in the appendix.

When a worker of type x is matched with a job of type y, this results in flow output $F(x, y)$, which is assumed to be increasing in both arguments. Unemployed workers receive a flow value $b(x)$ and vacant jobs are assigned zero flow value. Workers and jobs discount time at rate r and have linear preferences over income and flow profits, respectively.

The presence of search frictions yields match rents, which requires one to specify a rule to divide the match surplus between the two parties. I follow the approach developed by Cahuc et al. (2006) and others and assume that wages are determined via a sequential auctions bargaining game. In this game, the surplus of the match is apportioned by applying a generalized Nash bargaining rule, where the worker receives a share β, and the outside options available to workers depend on their labor market history. I describe more details of the game in the next section as well as in the appendix.

B.   Description of the Equilibrium

The equilibrium of this economy consists of value functions and distributions such that workers and firms match optimally, and the distributions satisfy steady-state flow conditions. It is useful to introduce the fol-

lowing notation. An employed worker of type x working for a firm of type y with a contracted wage of w has value $W(x, y, w)$. Alternatively, if this worker is unemployed, his or her value is $U(x)$. A job of type y that pays w has value $J(x, y, w)$ if matched with a worker of type x or $V(y)$ if vacant. The surplus of a match between worker x and job y is given by $S(x, y) \equiv [W(x, y, w) - U(x)] + [J(x, y, w) - V(y)]$. Note that surplus of the match does not depend on wages because workers and firms have linear preferences, in line with our wage-setting mechanism. The steady-state density of employed matches is $e(x, y)$. The density of unemployed workers for each type is $u(x)$, and the density of vacancies is $v(y)$. Finally, the unemployment rate, u, and the number of vacancies in the economy, v, are also determined endogenously.

In this frictional economy, workers and jobs are paired at a Poisson rate, whereby they can match or not. Moreover, upon matching, the surplus of the match is split between the two via a Nash bargaining protocol, whereby a fraction β of the surplus is given to workers and the remainder is received by firms. One important element of the bargaining protocol is the presence of outside options for both workers and firms partaking in negotiations. For firms, this option is always the value of keeping a vacancy, $V(y)$. For workers, it will depend on their labor market status. When coming from unemployment or after facing a reallocation shock, the outside option is the value of search while unemployed, $U(x)$. Then, as employed workers find new matches, they use them as leverage in negotiations, gaining the full surplus of that match as bargaining power. This can be explained using the example of an individual currently working for firm y with second-best option y'. In this case, his or her outside option is $U(x) + S(x, y')$. When this worker is offered a job by a new employer y'', three possibilities emerge. First, if $S(x, y'') < S(x, y')$, there is no change. Second, if $S(x, y') < S(x, y'') < S(x, y)$, the worker remains with the current employer but uses the new firm to leverage his or her negotiation position, and the new outside option becomes $U(x) + S(x, y'')$. Finally, if $S(x, y'') > S(x, y)$, then the worker leaves the current employer, whereby his or her alternative option becomes $U(x) + S(x, y)$. Note that workers and firms match only if they generate positive surplus $S(x, y) > 0$, and when workers are offered employment in other firms, they choose to work at the firm that generates the highest surplus.

Our assumptions imply that the expected value of unemployed workers is equal to

$$rU(x) = b(x) + \beta\lambda^U \int_{\underline{y}}^{\bar{y}} [S(x, y')]^+ v(y') \, dy', \qquad (2)$$

where $[A]^+ = AI[A > 0]$. In the appendix, I show the analogous equation for $V(y)$, which is similar to this one but accounts for the fact that vacan-

cies are also considered by employed workers. I also show in the appendix that the surplus of the match can be computed as

$$(r + \delta + \lambda^R)S(x, y) = F(x, y) - rU(x) - rV(y)$$

$$+ \beta\lambda^R \int_{\underline{y}}^{\bar{y}} [S(x, y')]^+ v(y') \, dy' \tag{3}$$

$$+ \beta\lambda^E \int_{\underline{y}}^{\bar{y}} [S(x, y') - S(x, y)]^+ v(y') \, dy'.$$

The first three terms on the right-hand side of the equation are standard and are presented in the work of Shimer and Smith (2000). In addition, although matches subject to reallocation shocks are eliminated, the worker can appropriate a share β of the surplus of a new match if he or she finds a suitable one. Finally, workers who find better jobs are able to extract a share β of the excess surplus of the new job versus the old one.

In addition to introducing the asset value equations, the equilibrium of this model includes the endogenous distributions $e(x, y)$, $v(y)$, and $u(x)$. In a steady state, those are fixed by equating the inbound and outbound flows of matches of particular types, as well as market clearing conditions. These are described in the appendix.

One very useful property of this framework is that $S(x, y)$, $U(x)$, and $V(y)$ can be computed jointly with the endogenous distributions without the need to compute wages. This facilitates the computation of the equilibrium, since it is not necessary to keep track of the second-best job as a state variable. With knowledge of the value functions and equilibrium distributions, the implied wage function can be retrieved by using the surplus sharing conditions imposed by Nash bargaining. The wage of worker x, working for firm y, with second-best option q satisfies the condition

$$W(x, y, w(x, y, q)) - [U(x) + S(x, q)] = \beta[S(x, y) - S(x, q)], \tag{4}$$

where $W(x, y, w(x, y, y'))$ is computed via an asset value equation described in the appendix. Having the wage function allows the flow profits of a job to be computed as

$$\pi(x, y, q) = F(x, y) - w(x, y, q).$$

C.   *Sorting, Wages, and Intuition*

Next, I illustrate some properties of equilibrium. As emphasized by Becker (1973) and Shimer and Smith (2000), complementarities in production are the main driver behind assortative matching: positive (negative) complementarities induce positive (negative) assortative matching, PAM (NAM).

I consider both types of complementarity and illustrate the case of an economy without on-the-job search, $\lambda^E = \lambda^R = 0$, which helps in developing intuition. In this case, there are no poaching offers, and the wage function depends on the worker and firm types as follows:

$$w^{\mathrm{NOJ}}(x, y) = \beta[F(x, y) - rV(y)] + (1 - \beta)U(x). \qquad (5)$$

It is easy to see that $U'(x) > 0$ and $V'(y) > 0$, which implies that wages are increasing with worker skill level, whereas they may be nonmonotone with respect to firm productivity. We can rearrange equation (5) and obtain $w^{\mathrm{NOJ}}(x, y) = U(x) + \beta S(x, y)$. This implies that wages are at their minimum at the edges of the matching set (where $S(x, y) = 0$), which ensures nonmonotone wages for all workers with interior matching sets. These nonmonotonicities are a natural reflection of job scarcity and optimal assignment. Indeed, in the matching economy, each worker has an ideal job (the one that generates the highest surplus), and workers choose them on the basis of wages. Thus, if the ideal match of a worker with mid-level skills is a mid-level productivity firm, then that is the firm that will pay him or her the most. Highly productive firms would pay that worker less because hiring him or her entails giving up on the opportunity to hire a more suitable worker, which is taken into account in negotiations. It is worth noting that similar nonmonotonicities also emerge in the profit function, $\pi(x, y, q)$, this time with respect to x.

I show two examples of symmetric economies in which $\beta = 0.5$, $J = 1$, $b(h) = 0$, and $l = g$, one with supermodular technology and another with submodular technology.[7] Figure 1 depicts plots of log output and log wages as a function of job productivity, where each line represents a different percentile of worker skill. For each worker type, the support of the plot includes only the firms that belong to his or her matching set, which provides a visual representation of these sets. We can see that highly skilled workers match with highly productive firms in the case of PAM and firms characterized by low productivity in the case of NAM, while the converse is true for low-skilled workers. It is also evident that in both NAM and PAM cases wages are increasing in x, whereas they are nonmonotone with respect to y, as described above.

This nonmonotonicity of wages with respect to y suggests that the firm fixed effects in a wage regression, ψ, are unlikely to capture the underlying level of firm productivity well. This has two implications with respect to mapping the primitives of the model to the empirical fixed-effects methodology. First, the correlation $\mathrm{Corr}(\theta, \psi)$ is distorted and thus no longer

---

[7] I use a constant elasticity of substitution (CES) production function $F(x, y) = (0.5x^\eta + 0.5y^\eta)^{1/\eta}$, where $\eta \to 0$ (Cobb-Douglas) on the supermodular example and $\eta = 3$ on the submodular example. I also assume that the distributions are lognormal with variance 1, $\rho = 0.005$, $\delta = 0.02$, and $\lambda^U = 0.3$.
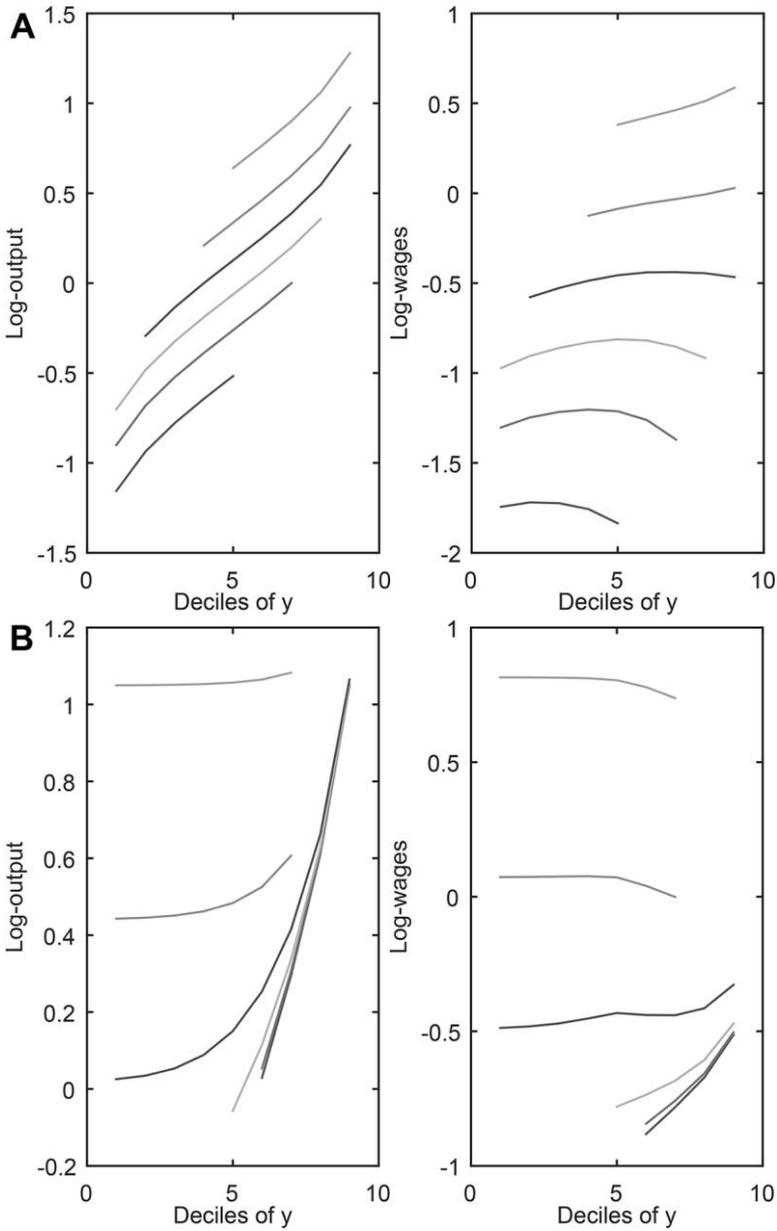
Fig. 1.—Equilibrium output and wages with PAM (panel A) and NAM (panel B), without on-the-job search. Each line refers to a different percentile of worker skill, and for each worker the line includes only the firms that belong to his or her matching set.

accurately reflects the degree of sorting between the model primitives x and y. Second, the same nonmonotonicity tends to suppress the amount of dispersion in ψ generated by the model. As noted previously, workers each have their ideal firm type. Thus, when the difference in wages for a worker in two firms is compared, $w^{\text{NOJ}}(x, y) - w^{\text{NOJ}}(x, y')$, the outcome may be positive for some workers and negative for others, introducing ambiguity into the average "firm effect" on wages. This second point is well illustrated by an example described by Eeckhout and Kircher (2011), using a simplified version of the frictional matching model with uniform distribution of types. In that example, when applying the AKM methods to the equilibrium wage function of the model, the implied firm fixed effects do not vary with their type regardless of the degree of frictions, implying that $\text{Var}(\psi) = 0$.[8] Recall that $\text{Var}(\psi) \gg 0$ on the data, which may be a challenge for the model to match. In general, the variability of ψ will depend on the joint distribution of worker skill and productivity in equilibrium. Consequently, when parameterizing the model, in our approach, we can allow for very flexible shapes for the distributions of worker skill and job productivity.

The same graph suggests that wage data can still be useful for inferring the strength of sorting for two reasons. First, in this model, the highly skilled workers work for the high-productivity firms in the case of PAM (or the low-productivity in the case of NAM) and thus have highly skilled coworkers. Second, wages are monotonic in x, suggesting that the relationship between worker and coworker wages reflects that of the primitives. Thus, the moment $\text{Corr}(\theta, \tilde{\theta})$ is a promising way to measure the intensity of sorting in the economy. One limitation of the measure is that it cannot distinguish the sign of sorting (PAM or NAM), just its intensity.

In sum, to ensure that this model is consistent with the empirical evidence presented in Section III, the economy should exhibit a substantial degree of sorting, as reflected by $\text{Corr}(\theta, \tilde{\theta})$, whereas $\text{Corr}(\theta, \psi)$ would be biased downward because of nonmonotonicity in wages. However, this nonmonotonicity cannot be too strong, such that the model still yields $\text{Var}(\psi) \gg 0$. Whether the model can meet those targets simultaneously is a quantitative matter and is investigated in Section V.

On-the-job search and sequential auctions make this relationship more complex because the wage function includes extra terms that reflect the potential gains to the worker from on-the-job search that are included in the negotiations and depress workers' wages. This mechanism can invalidate the monotonicity in x described above. To explain this argument, we

[8] See their eq. (18) and the discussion in Sec. IV.A. However, the authors do not discuss the implications for the variance of those fixed effects, only the implications for sorting.

can consider the example of the wage function of workers hired from un-employment, $q = \varnothing$, in an economy in which workers have low bargaining power, $\beta = 0$, and $b(x) = b$. Equations (2), (7), and (4) imply that

$$w^{\beta=0}(x, y, \varnothing) = b - \lambda^E \int_{\underline{y}}^{\bar{y}} [S(x, y') > 0]$$

$$\times \min\{S(x, y'), S(x, y)\} v(y') \, dy'.$$

Now, let us consider any firm y^0 with an interior matching set and the least-skilled employee of that firm, $\underline{x}(y^0)$. Note that by construction $S(\underline{x}(y^0), y^0) = 0$, which implies that this worker will have the highest wage among the workers just hired from unemployment to take positions in that firm. Higher values of β, an increasing $b(x)$, and the arrival of competing offers, $q \neq \varnothing$, all counteract this force and restore monotonicity in worker skill level.

V.   Calibration

In this section, I select a parameterization of the model and set the parameters in order to match a number of labor market facts emphasized by the empirical job search literature (e.g., Jolivet et al. 2006). I perform two calibrations. First, I select a "parsimonious" parameterization and do not include the moments described in Section III.C as part of the target moments. I show that, while the model succeeds in matching the targeted moments, it fails to explain some of the facts highlighted in Section III.C. In particular, the model generates less variation in firm fixed effects ψ than observed in the data. For this reason, in the method presented in Section E of the appendix, I adopt a more flexible parameterization and include the aforementioned facts as target moments. This approach improves the fit of the model but still fails to reproduce some aspects of the data.

A.   Calibration

In this section, I describe model calibration, performed by selecting parameters that guarantee that the model matches a number of features of the US data. In order to do so, I need to make certain parametric assumptions. First, I assume that the production function takes a CES form $F(x, y) = [\varphi x^\eta + (1 - \varphi) y^\eta]^{1/\eta}$, which allows for a wide degree of complementarities. If $\eta < 1$, the production function is supermodular, whereas if $\eta > 1$, it is submodular. Second, I assume that the flow value of leisure for the worker takes the following form: $b(x) = bF(x, \phi(x))$, where $b \in [0, 1]$, and $\phi(x)$ denotes the optimal job assigned to worker x in the frictionless

economy.[9] I posit that this is a parsimonious way to allow more productive workers to be more productive in nonmarket activities as well. I assume that worker and firm types follow a lognormal distribution with zero mean and, respectively, σ^W and σ^F standard deviations. It is worth noting that the mean of these distributions cannot be distinguished from φ, which is why they are set to zero.[10]

Following the discussion in Section IV.C, I restrict $\eta < 1$ (positive complementarities), as the empirical measure of sorting helps identify only the strength of sorting, not its sign. I set the discount rate to match a 5 percent annual discount rate. I also set $J = 1$ to ensure that, if frictions are removed from the economy, every worker and job would find a match. Finally, I also set $\varphi = 0.5$. Thus, nine parameters remain, which are selected in order to match a series of facts about labor markets emphasized in the job search literature. Although the model performs in continuous time, in practice, I approximate the economy with the discrete-time correspondent, setting the periodicity to 1 week. I simulate the path of 10,000 workers who are matched to 1,000 firms over a 7-year period. These numbers are chosen to match the length and the ratio of workers per firm of the US matched employer-employee data set, as shown by Woodcock (2015). The number of workers is smaller than that found in typical matched employer-employee data sets. However, I found that increasing this number has little effect on the results, while it increases the computational time substantially. I present results using a much larger sample (300,000 workers) for the calibrated parameter values. I compute the moments on a yearly basis before calculating the average across years. The conditional moments are computed (both data and model), restricting attention to workers who remain employed throughout the year and experience at least one job transition.

In order to initialize the sample, I begin with a sample of unemployed workers drawn from distribution $l(x)$ discretized to a 100-point grid and simulate their paths for a 10-year period, in order to start the economy in a steady state. From that point, I collect 7 years of data, which is the basis of the calculations. On the firm side, I discretize the firm type grid to 500 points and allocate the firms to those grid points according to distribution $v(y)$, which I assume is such that firms have an equal number of total jobs (va-

---

[9] The function $\phi(x)$ is such that $\int_x^1 l(x')\,dx' = J \int_{\phi(x)}^1 g(y')\,dy'$.
[10] This holds because we can always rewrite

$$F(x,y) = [\varphi(ax)^\eta + (1-\varphi)(by)^\eta]^{1/\eta} \propto \left[\frac{\varphi a^\eta}{\varphi a^\eta + (1-\varphi)b^\eta} x^\eta + \frac{(1-\varphi)b^\eta}{\varphi a^\eta + (1-\varphi)b^\eta} y^\eta \right]^{1/\eta}$$

$$= [\varphi' x^\eta + (1-\varphi')y^\eta]^{1/\eta}.$$

cant and filled).[11] Whenever the worker finds a firm, he or she takes a draw
from the (endogenous) distribution of vacancies and is allocated ran-
domly to one of the firms at the corresponding type. Note that the choice
of $v(y)$ does not affect the wage function, matching decisions, the distribu-
tion of vacancies, or the joint densities of worker and job types. As a result,
I have found that this distribution has little impact on the moments under
consideration.

I perform two sets of calibrations, one for workers with a high school
education and another one for workers with at least a college degree.
It is well documented that there are substantial differences in unemploy-
ment inflow rates across workers with different levels of education (e.g.,
Elsby, Hobijn, and Şahin 2010), and this model is not designed to explain
those endogenously. I select a number of data moments as targets for the
calibration. Next, I discuss identification informally, thus providing some
level of intuition as to why that moment should be directly affected by
that parameter.

First, I have the transition rates λ^U, δ, and λ^E, which have a clear associ-
ation with the observed unemployment inflow rates, unemployment out-
flow rates, and job-to-job transition rates. These can be computed in the
model as follows:

$$\text{UErate} = \lambda^U \int_{\underline{y}}^{\bar{y}} \int_{\underline{x}}^{\bar{x}} [S(x', y') > 0]\, u(x')\, v(y')\, dx'\, dy',$$

$$\text{EUrate} = \delta + \lambda^R \int_{\underline{y}}^{\bar{y}} \int_{\underline{y}}^{\bar{y}} \int_{\underline{x}}^{\bar{x}} [S(x', y'') \leq 0]\, e(x', y')\, v(y'')\, dx'\, dy'\, dy'',$$

$$\text{EErate} = \int_{\underline{y}}^{\bar{y}} \int_{\underline{y}}^{\bar{y}} \int_{\underline{x}}^{\bar{x}} \{\lambda^R [S(x', y'') > 0] + \lambda^E [S(x', y'') > S(x', y')]\}$$
$$\times\, e(x', y')\, v(y'')\, dx'\, dy'\, dy''.$$

I use Current Population Survey microdata for the period 1996–2003 and
compute the UErate corrected for time aggregation for both educational
groups following the methodology proposed by Shimer (2012). I focus
on this period to match the time span used to compute the EErate target.
Next, I use unemployment rates by education provided by the Bureau of
Labor Statistics (4.3 percent for high school graduates and 2.2 percent
for college graduates) and adopt the standard steady-state approximation
$u \approx \text{EUrate}/(\text{EUrate} + \text{EUrate})$ to compute the EUrate of each group.
Finally, I make use of the results reported by Nagypál (2008), who pro-

---

[11] There are more grid points on the support of firm types to prevent ties on the worker
job decisions, including job-to-job transitions. I use interpolation methods to obtain values
on the expanded grid points.

vides EErates by education level for the same period.[12] The values are shown in table 3.

Next, I use a series of moments related to the distribution of wages and wage dynamics to calibrate the parameters σ^W, σ^F, λ^R, and β. I employ data from the SIPP 1996 panel, which tracks workers at a weekly frequency and is well suited for measuring job transitions. First, increasing σ^W increases the overall variability of log wages $\mathrm{Var}(w)$, which I use as a target moment. In this environment, characterized by frictions and firm heterogeneity, the same worker can be paid differently in different firms, and this is affected by the extent of firm heterogeneity, σ^F. Consequently, I use as a target moment the variance of wage changes for workers who experience a job-to-job transition, $\mathrm{Var}(\Delta w | \mathrm{JJT})$. As documented by Postel-Vinay and Robin (2002), Jolivet et al. (2006), and Lopes de Melo (2007), a substantial fraction of job-to-job transitions involve a wage cut, which is hard to match in a model in which workers are permitted to switch jobs only for monetary gains. Thus, I set the rate of involuntary reallocation shocks, λ^R, to match the fraction of wage cuts that involve job-to-job transitions, $\mathrm{Pr}(\Delta w < 0 | \mathrm{JJT})$. In addition, recall that in this model, there is a certain extent of backloading in wages as workers start jobs at low wages and negotiate increases as they consider roles in other firms. The extent of backloading is affected by the parameter β, whereby higher β corresponds to the more immediate wage gains. In fact, as shown by Postel-Vinay and Robin (2002), if β is very low, when switching jobs, some workers may accept a wage cut in exchange for future wage growth. Thus, I set β to match the mean wage gain in job-to-job transitions, $E(\Delta w | \mathrm{JJT})$.

At this point, only b and η remain, whereby the latter controls the degree of complementarity in the production function. Without frictions, sorting is perfect and positive, as long as $\eta < 1$. In the presence of frictions, workers and firms tolerate mismatch, the extent of which will depend on the degree of complementarities. As discussed in Section V, the moment $\mathrm{Corr}(\theta, \tilde{\theta})$ captures well the intensity of sorting among workers. Thus, the value of this moment must be set as a means of fixing η. I choose a target value of .4 for $\mathrm{Corr}(\theta, \tilde{\theta})$, which corresponds to the case of Denmark and tends toward the top of the range found across data sets (.17 in Italy and ~.5 in Brazil and Germany). In the extended calibration presented in Section E of the appendix, I also test an alternative value of .2 for each education group. Finally, I need to set the parameter value of home production, b. I follow the approach used by Shimer (2005) and assume that b is such that the average value of $b(x)$ is 40 percent of the average wage, a

---

[12] I adopt her results using Survey of Income and Program Participation (SIPP) data because they pertain to time aggregation. To obtain the implied EErates, I use her reported shares of job-to-job transitions as a fraction of all separations (49.7 percent for high school and 58.5 percent for college).

TABLE 3
CALIBRATION TARGETS, MODEL FIT, AND PARAMETERS

| | A. CALIBRATION TARGETS | | | | | B. CALIBRATED PARAMETER | |
| | High School | | College | | | | |
TARGET	Data	Model	Data	Model		High School	College
UErate	.412	.412	.373	.367	λ^U	.495	.404
EUrate	.019	.017	.008	.008	δ	.016	.008
EErate	.018	.020	.012	.013	λ^E	.031	.034
Var(w)	.22	.224	.303	.302	σ^W	.151	.163
Corr($\theta, \tilde{\theta}$)	.4	.394	.4	.400	η	$-.986$	.587
$\int_{\underline{x}}^{\bar{x}} \frac{b(x')}{E(w|x')} u(x')\,dx'$	.4	.405	.4	.400	b	.048	.043
Var(Δw|JJT)[a]	.096	.092	.090	.086	σ^F	.296	.252
Pr($\Delta w < 0$|JJT)	.27	.375	.260	.289	λ^R	.010	.004
$E(\Delta w$|JJT)	.039	.042	.047	.050	β	.147	.167

[a] The Δ refers to yearly change, and JJT refers to workers who were continuously employed during that year but switched jobs at least one time.

number derived from the average replacement rate from the US unemployment insurance system. However, since the pool of unemployed workers may differ from the overall pool of workers, I compute $b(x)/E(w|x)$ for each worker type and then average over the distribution of unemployed workers, obtaining the target

$$\int_{\underline{x}}^{\bar{x}} \frac{b(x')}{E(w|x')} \, u(x')\,dx'.$$

I choose the combination of parameters that minimize the distance between the target moments and the model moments.[13] The calibration results are shown in table 3, panels A and B. As can be seen from table 3, the model is successful in matching the proposed moments, especially for the sample of college-educated workers (there are too many wage cuts in job switches in the sample comprising workers with a high school education). When considering the estimated parameters, it is worth mentioning that there is a substantial difference in the estimated degree of complementarities pertaining to the high school versus the college sample. While both η values (-0.99 and 0.59, respectively) reflect supermodularity, the degree of complementarities in the market populated by workers with a high school diploma is stronger. The intuition for that result is that there are "more frictions" in the market serving high school graduates ($\lambda^U/\delta \approx 30$ in high school and $\lambda^U/\delta \approx 50$ in college), which leads to less sorting.

[13] I use the global optimizer "covariance matrix adaptation evolution strategy," which combines evolutionary algorithms with traditional gradient methods and is well suited for non-smooth problems with potentially many local optima (see Hansen and Ostermeier [2001] for details). I use the code provided by the authors, with the default parameters, and at least three (possibly manual) restarts.

TABLE 4
AKM MOMENTS ON CALIBRATED ECONOMY

AKM MOMENTS	DATA	MODEL		MODEL (Large Sample)[a]	
		High School	College	High School	College
$\mathrm{Corr}(\theta, \psi)$	$-.01$	.175	.338	.173	.335
$\mathrm{Corr}(\theta, \tilde{\theta})$	.40	.394	.400	.397	.395
$\mathrm{Var}(\psi)/\mathrm{Var}(\theta + \psi)$	.22	.023	.042	.023	.044
R^2	.89	.892	.943	.890	.942
$\mathrm{Corr}(x, y)$	. . .	.486	.473	.488	.478
$\mathrm{Corr}(x, \tilde{y})$	. . .	.415	.411	.418	.409
$E[\mathrm{Var}(w|x)]/\mathrm{Var}(w)$	. . .	.198	.122	.200	.121

[a] The larger simulated sample has 300,000 workers.

However, since I am imposing the same degree of sorting on both markets, $\mathrm{Corr}(\theta, \tilde{\theta}) = .4$, I need stronger complementarities to achieve that.

B.   *AKM Regression*

Given that the model has been calibrated to match several features of the US economy, its ability to reproduce the AKM empirical moments needs to be verified. Table 4 summarizes the results. First, as expected, the worker-coworker correlation $\mathrm{Corr}(\theta, \tilde{\theta})$ (a targeted moment) captures the "true" degree of sorting (the value of $\mathrm{Corr}(x, \tilde{x})$) in the economy quite well. Second, while the $\mathrm{Corr}(\theta, \psi)$ value is much lower than its theoretical counterpart, this value is still well above that measured in any of the data sets described in Section III. Third, and perhaps more significant, the dispersion in firm fixed effects, $\mathrm{Var}(\psi)/\mathrm{Var}(\theta + \psi)$, is substantially smaller than that observed in the data. The intuition for these results derives from the nonmonotonicities explained in Section IV.C. Since they distort the mapping between true firm types, y, and firm fixed effects, ψ, the model is consistent with a low value for $\mathrm{Corr}(\theta, \psi)$ despite incorporating positive sorting. However, the same nonmonotonicities reduce the dispersion in ψ, as firms determine the wages of different workers in different ways. In addition, I included results of a simulation based on a sample of 300,000 workers (instead of 10,000) to establish whether sample size affects the results. As can be seen, the results based on the larger sample are very similar to the baseline results.

VI.   **Conclusion and Future Directions**

My objective in this paper was to provide a structural interpretation for the facts pertinent to firm-wage differentials and worker-firm sorting obtained by adopting the AKM methodology. First, I reviewed reduced-form evidence from the AKM regressions, including my own, focusing on three

moments: a significant variance of firm fixed effects, worker and firms fixed effects that are uncorrelated, and a significant correlation between the fixed effects of workers and their coworkers. I demonstrated that these three moments are inconsistent with the "face value" interpretation provided by the piece-rate model. This is the case because the degree of coworker segregation that I documented violates the implication that the worker skill distribution is independent across firms. I extended the analyses further and verified that this fact is not driven by composition using Brazilian data.

In the next step of the investigation, I used a natural alternative to explain those facts: a frictional model that allows for assortative matching. A clear challenge for the model was to generate a zero correlation between fixed effects while allowing for sorting patterns. I showed that the model incorporates a mechanism that allows for the fact that the wage function is nonmonotone in a firm's productivity, which distorts the mapping between firm types and fixed effects, biasing the correlation. This nonmonotonicity does not affect the mapping between worker types and their wages; hence, the worker-coworker measure is not contaminated and accurately captures the degree of sorting in the economy. However, the same mechanism can generate problems in matching with respect to the first fact. As I have shown, because wages are nonmonotone, a firm will pay some workers more and other workers less, which depresses the dispersion in firm fixed effects. Thus, it is a quantitative matter if the model can succeed or not in matching the data. To address that issue, I calibrated the model to US data, confirming that the model can explain a number of labor market transition characteristics, wage dynamics, and the wage distribution. The model also performs well in explaining the correlations of the AKM regression. However, the model generates too little dispersion in firm fixed effects relative to the data, which can be explained by the aforementioned nonmonotonicities. The result remains valid after robustness checks with a much less parsimonious calibration.

These findings lead to the conclusion that neither model fully succeeds in matching the data, indicating that features need to be changed or added. I discuss a few possibilities. One possible extension would be to allow for types of workers or jobs to be stochastic. It is clear from the frictionless matching model that this has strong implications for the variance of the fixed effects. The argument is as follows. The model described in Section IV with $J = 1$, no frictions, and a production function such that $F_{hp} > 0$ can be applied. This is a version of the Becker (1973) marriage model, which features perfect assortative matching (top worker with top firm, and vice versa) under such a parameterization. Now, the same economy can be replicated from one period to the next, allowing workers to switch types while keeping the distribution unchanged. In such a scenario, firms will hire new workers, but the new hires will have the same level of skill

and will be paid the same wage. Therefore, a set of firm dummies would explain 100 percent of wage variation (over both periods), whereas the same is not true for worker dummies. Naturally, one could reverse the argument and fix worker types while allowing firm types to change. In that case, worker dummies would explain 100 percent of wage variation. It is not well known how these conclusions change when both types change simultaneously and how these interact with frictions.[14] Second, since the coworker measure points to sorting between workers, another extension would be to consider explicitly models with team production and cross-complementarities in production. This is a challenging and promising area of research.

Another promising direction for future investigations is the one I pursued in the companion paper (Lopes de Melo 2015). In that work, I extend the frictional sorting model to include firm attributes that affect workers' preferences, in line with the theory of compensating differentials (e.g., Rosen 1986). That makes the model analysis more complex because it requires identifying two distinct dimensions of firm heterogeneity (firm productivity and the amenities). In that model, there are still complementarities in production between worker skill level and firm productivity, which induces assortative matching. Moreover, there is the compensating differential that shifts the wages of all workers in the same direction, which resembles the firm fixed effects in the AKM methodology. Consequently, the model can explain the correlations in AKM, as well as generate a fair amount of dispersion in firm fixed effects.

Appendix

A. Data Description and Sample Selection

I use the labor market census RAIS (Relacao Anual de Informacoes Sociais), an administrative data set collected annually by the Brazilian Labor Ministry, which includes all firms in the Brazilian formal sector and provides information for all their workers. The ministry collects demographic information of all workers, such as age, education, and sex, along with some information about establishments, such as sector and location. In addition, it provides information about the job, such as the average wage earned during that year (the measure I use), the wage in December, the average number of hours worked, occupation, dates of admission and resignation, type of contract, and causes for the termination of employment.[15]

I use data pertaining to an 11-year period, ranging from 1995 to 2005. Although data from previous years were available, I opted for this time frame because, until

---

[14] Lise et al. (2016) allow for stochastic job types but consider only the implications for the correlation between fixed effects, $\mathrm{Corr}(\theta, \psi)$.

[15] The remaining variables are race, nationality, a measure of disability, and the juridic nature of the firm.

1994, Brazil suffered from extremely high inflation, which caused serious measurement problems in variables such as wages and also had structural implications for the macroeconomy. I restrict the sample to the state of Sao Paulo, which is the richest state of the country, contributing to the GDP and industrial production by over 13 percent and 30 percent, respectively. Sao Paulo is characterized by a much smaller level of informal employment relative to all other regions of the country. Work informality is an important feature in Brazil, just as it is in many developing countries, and is excluded in the analysis. Furthermore, RAIS is an enormous data set, and reducing the number of states makes the data employed in analyses more manageable.

I start with a sample of 25,856,195 observations pertaining to workers aged between 20 and 60, with at least a high school diploma, employed in the state of Sao Paulo, in the 1995–2005 period, for whom wage observations are available and who worked for a particular firm during the full year. I then apply a series of inclusion criteria on this sample. First, I select individuals who worked at least an average of 25 hours per week. Then I collapse observations to individual, establishment, and year levels. If data for a particular worker are given for the same establishment multiple times, I select the observation with the highest wage. Next, I restrict the focus to workers with at least a moderate degree of formal labor market attachment, whose data are included in the sample for at least 5 of the 11 years, and I eliminate observations pertaining to individuals who have worked in more than three firms per year. I also restrict the sample to firms with at least two workers in order to compute coworker measures. These restrictions yield a final sample comprising 16,253,899 observations. Finally, I use the algorithm described by Abowd et al. (2002) for computing the connected groups in the sample and select the largest group. As discussed previously, applying the AKM methodology allows for identifying only worker and firm fixed effects within groups of connected workers. The largest group contains 16,027,426 observations, or 98.6 percent of the sample. Table A1 provides some statistics related to this sample.

TABLE A1
SAMPLE STATISTICS

	Firms			Workers	
			Whole Sample		
Observations	25.9 million	Number	579,000	Number	6.1 million
% college	36.2%	Average observations	44.59	Average observations	4.18
% female	49%	Average firm size	11.7	Average no. firms	1.36
Average age	35.5	Average no. years	3.81	Average no. years	4.01
			Selected Sample		
Observations	16.0 million	Number	137,000	Number	2.0 million
% college	38%	Average observations	116.27	Average observations	7.97
% female	49%	Average no. workers	22.65	Average no. firms	1.72
Average age	36.7	Average no. years	5.13	Average no. years	7.71

B. Derivation of the Value Functions and Surplus

Vacancies can meet three kinds of workers: unemployed, employed from reallocation shocks, and unemployed from job-to-job transitions. The first two kinds of workers have as their outside option unemployment, so the firm is able to extract a share $(1 - \beta)$ of the surplus of the newly formed match. The latter kind has as an outside option the total surplus generated on her current match, which leads the firm to gain a share $1 - \beta$ of the excess surplus over the previous match. This implies the following value equation:

$$
\begin{aligned}
rV(p) = {}& \lambda^{FE}(1 - \beta) \int_{\underline{y}}^{\bar{y}} \int_{\underline{x}}^{\bar{x}} \alpha^{E}(x', y', y)[S(x', y) - S(x', y')]e(x', y') \, dx' \, dy' \\
& + \lambda^{F}(1 - \beta) \int_{\underline{x}}^{\bar{x}} \alpha^{U}(x', y) S(x', y) u(x') \, dx' \\
& + \lambda^{FR}(1 - \beta) \int_{\underline{y}}^{\bar{y}} \int_{\underline{x}}^{\bar{x}} \alpha^{U}(x', y) S(x', y) e(x', y') \, dx' \, dy'.
\end{aligned}
\tag{A1}
$$

Next I compute the net present discounted value of a worker employed at wage $w(x, y, q)$:[16]

$$
\begin{aligned}
& \left\{ r + \delta + \lambda^{R} + \lambda^{E} \int_{\underline{y}}^{\bar{y}} [S(x, y') > S(x, q)] v(y') \, dy' \right\} \\
& \times [W(x, y, w(x, y, q)) - U(x)] = w(x, y, q) - rU(x) \\
& \qquad + \lambda^{R} \beta \int_{\underline{y}}^{\bar{y}} [S(x, y') > 0] S(x, y') v(y') \, dy' \\
& + \lambda^{E} \int_{\underline{y}}^{\bar{y}} [S(x, y') > S(x, y)][\beta \, \max\{S(x, y'), S(x, y)\} \\
& \qquad + (1 - \beta) \min\{S(x, y'), S(x, y)\}] v(y') \, + dy'.
\end{aligned}
\tag{A2}
$$

The firm analogous to this equation is given by

$$
\begin{aligned}
& \left\{ r + \delta + \lambda^{R} + \lambda^{E} \int_{\underline{y}}^{\bar{y}} [S(x, y') > S(x, q)] v(y') \, dy' \right\} [J(x, y, w(x, y, q)) - V(y)] \\
& = F(x, y) - w(x, y, q) - rV(y) + \lambda^{E} \int_{\underline{y}}^{\bar{y}} [S(x, y') > S(x, q)][(1 - \beta)[S(x, y) \\
& \quad - \min\{S(x, y'), S(x, y)\}] v(y') \, dy'.
\end{aligned}
\tag{A3}
$$

If we add equations (A2) and (A3) and do some algebra, we obtain equation (3). Note that this argument applies to matches that pay equilibrium wages. Follow-

---

[16] One can use a very similar equation to compute the value for the worker x of matching with firm y at an arbitrary wage w, $W(x, y, w)$. The only modification would be that job-to-job transitions affect the worker value whenever the surplus of the new match exceeds the value of the current contract for the worker: $S(x, y') > W(x, y, w) - U(x)$.

ing footnote 16, one can easily adapt these equations to compute values at arbitrary wages, which yields the same surplus function.

C.   Steady-State Flows

I now describe the equilibrium equations that jointly determine the stationary distributions $e(x, y)$, $u(x)$, and $v(y)$, the idleness rates u and v, and the necessary restrictions on λ^F, λ^{FE}, and λ^{FR}. The first equilibrium equation described is between the flows in and out of employed matches of type (x, y). If $S(x, y) \leq 0$, then $e(x, y) = 0$. Otherwise, $e(x, y)$ is determined by equating the inflows to the outflows:

$$
\left\{ \delta + \lambda^R + \lambda^E \int_{\underline{y}}^{\bar{y}} [S(x, y') > S(x, y)] v(y') \, dy' \right\} (1 - u) e(x, y)
$$

$$
= u \lambda^U u(x) v(y) + \lambda^R (1 - u) v(y) \int_{\underline{y}}^{\bar{y}} e(x, y') \, dy' \tag{A4}
$$

$$
+ \lambda^E (1 - u) v(y) \int_{\underline{y}}^{\bar{y}} [S(x, y) > S(x, y')] e(x, y') \, dy'.
$$

On the left-hand side of the equation, matches of type (x, y) can be ended in three ways: exogenous destruction shocks, reallocation shocks, and successful job-to-job transitions. On the other hand, the inflows into those matches come from workers hired from unemployment and employed workers of type x who switch to jobs of type y via either reallocation shocks or job-to-job transitions.

In addition, I use two consistency conditions to determine the densities of unemployed workers and vacancies, $u(x)$ and $v(y)$. The mass of employed workers of type x has to equal the total mass of workers of that type minus the unemployed ones:

$$
1 - u(x) u = (1 - u) \int_{\underline{y}}^{\bar{y}} e(x, y') \, dy'. \tag{A5}
$$

A similar condition needs to hold for jobs of type y:

$$
J - v(y) v = (J - v) \int_{\underline{x}}^{\bar{x}} e(x', y) \, dx'. \tag{A6}
$$

Next, the unemployment rate is determined by integrating (A4) over the full support of x and y:

$$
(1 - u) \left\{ \delta + \lambda^R \int_{\underline{y}}^{\bar{y}} \int_{\underline{y}}^{\bar{y}} \int_{\underline{x}}^{\bar{x}} [S(x', y'') \leq 0] v(y'') e(x', y') \, dx' \, dy' \, dy'' \right\}
$$

$$
= u \lambda^U \int_{\underline{y}}^{\bar{y}} \int_{\underline{x}}^{\bar{x}} [S(x', y') > 0] u(x') v(y') \, dx' \, dy'. \tag{A7}
$$

Another equilibrium requirement for equilibrium is that the total number of employed workers has to equal the number of filled jobs:

$$1 - u = J - v. \tag{A8}$$

Finally, the contact rates of workers and jobs must be such that the total number of contacts is the same in equilibrium:

$$u\lambda^{U} = v\lambda^{F},$$

$$(1 - u)\lambda^{E} = v\lambda^{FE}, \tag{A9}$$

$$(1 - u)\lambda^{R} = v\lambda^{FR}.$$

D.   Computation Algorithm

One can use the following method to solve the model. First, select N grid points on the supports of x and y in which we compute our endogenous objects.[17] I begin with initial guesses for $e(x, y)$, $S(x, y)$, and u on the selected node points. I then adopt the following interactive procedure:

1. Given the old value of γ^{O}, S^{O}, and u^{O}, we can update these using equilibrium equations (2)–(A9), thus obtaining γ^{N}, S^{N}, and u^{N}. When computing integrals we use numerical integration using as node points the N grid points previously selected.
2. We compute the distance between the old value and the update. We first compute the maximum absolute value of $(X^{N} - X^{O})/(1 + X^{O})$ on the grid for each of the three objects and then take the maximum value between the three.
3. Repeat steps 1 and 2 until the norm is sufficiently small.

Because we do not have results showing that these mappings are contractions, we cannot anticipate that this algorithm works. Adding on-the-job search increases the computational burden of the problem substantially, as it requires us to compute double integrals. Because of that I use the solution to the problem without on-the-job search as the initial condition for the model with on-the-job search. I set N to 30 when solving the model with on-the-job search, which I found to be a good compromise between speed and accuracy of the approximation. Also, because of discretization, it may happen sometimes that the algorithm does not converge because matching decisions change discontinuously with minimal changes to the surplus. In order to smooth that, assume that the decision rule takes a logistic form $1/(1 + e^{-100S})$ instead of indicator functions.[18] Furthermore, I have found that "slowing" the updates $\gamma^{O} = \alpha\gamma^{N} + (1 - \alpha)\gamma^{O}$ improves the performance of the code.

E.   Expanded Calibration

In Section V, I showed that the baseline calibration fails to match jointly all the moments from the AKM regression. One possibility is that this was due to some of the parameter restrictions that were imposed in the calibration. In order to ad-

---

[17] I select evenly spaced grid points (the midpoints of N evenly spaced intervals) instead of Chebychev nodes because the endogenous matching sets imply that the regions of integration of the program are worker and firm specific.

[18] This influences only very small surplus values, below 0.1 in modulus. Surplus values can take the value of hundreds or thousands in typical calibrations.

TABLE A2
EXPANDED CALIBRATION

| | $\mathrm{Corr}(\theta, \tilde{\theta}) = .4$ | | | | $\mathrm{Corr}(\theta, \tilde{\theta}) = .2$ | | | |
| | High School | | College | | High School | | College | |
TARGET	Data	Model	Data	Model	Data	Model	Data	Model
UErate	.412	.392	.373	.371	.412	.337	.373	.363
EUrate	.019	.015	.008	.007	.019	.016	.008	.007
EErate	.018	.022	.012	.015	.018	.023	.012	.015
Var(w)	.220	.170	.303	.240	.220	.133	.303	.257
Corr$(\theta, \tilde{\theta})$	.400	.303	.400	.347	.200	.195	.200	.192
Var$(\Delta w|\mathrm{JJT})$	.096	.099	.090	.106	.096	.110	.090	.099
Pr$(\Delta w < 0|\mathrm{JJT})$	.270	.347	.260	.290	.270	.368	.260	.285
$E(\Delta w|\mathrm{JJT})$	.039	.046	.047	.053	.039	.044	.047	.053
Skew(w)	−.028	−.028	−.487	−.492	−.028	−.027	−.487	−.491
Corr(θ, ψ)	−.013	.171	−.013	.154	−.013	−.047	−.013	.097
Var$(\psi)/\mathrm{Var}(\theta + \psi)$	.217	.165	.217	.156	.217	.160	.217	.166
R^2_{AKM}	.889	.879	.889	.922	.889	.794	.889	.944

dress that, I perform a second calibration, this time with more flexibility and many more free parameters in order to see if the model can match the AKM moments jointly with the moments that I had previously targeted for some configuration of parameters.

I make the following modifications to the exercise. First, I allow the discount rate, ρ, the measure of jobs in the economy, J, and the weight of the CES function, φ, to vary as part of the calibration. I cap the discount rate at 20 percent a year and J at 1.1, which would imply a vacancy rate around 14.5 percent and a monthly vacancy-filling rate around 0.25 for the high school market.[19] This rate is well below the values documented in Davis, Faberman, and Haltiwanger (2010) using JOLTS data (reducing J increases the job-filling rate). Second, I use beta distributions instead of lognormal, which allows for very flexible shapes. This adds four shape parameters (in addition to the two variance parameters) to the calibration, leaving us with 16 parameters on the new formulation. Finally, I allow for negative complementarities in production by allowing for $\eta > 1$.

I choose to match the same target moments as before, as well as the additional moments from the AKM regression, $\mathrm{Corr}(\theta, \psi)$, $\mathrm{Var}(\psi)/\mathrm{Var}(\theta + \psi)$, and R^2_{AKM} and the skewness of wages, to assure that the wage distribution has a reasonable shape. Another modification is that I also consider a value of .2 for $\mathrm{Corr}(\theta, \tilde{\theta})$, in line with the lower values for that statistic presented in table 1. Finally, I drop

$$\int_0^1 \frac{b(x')}{E(w|x')} u(x') \, dx'$$

as one of the target moments from the calibration because there is a debate in the related literature about the appropriate value for the flow value of leisure (see,

---

[19] The vacancy-filling rate is

$$\frac{u\mathrm{UErate} + (1 - u)\mathrm{EErate}}{J - 1 - u}.$$

TABLE A3
EXTENDED CALIBRATION: PARAMETER VALUES

	Corr$(\theta, \tilde{\theta}) = .4$		Corr$(\theta, \tilde{\theta}) = .2$	
	High School	College	High School	College
λ^U	.535	.468	.388	.379
δ	.013	.006	.015	.007
λ^E	.050	.052	.042	.047
λ^R	.012	.005	.012	.005
σ^W	2.94	3.62	7.04	1.458
σ^F	52.01	19.66	9.78	63.83
η	.448	.565	2.16	.635
β	.244	.234	.172	.276
b	.000	.000	.138	.000
ρ	.009	.015	.016	.016
J	1.050	1.016	1.072	1.055
φ	.224	.234	.455	.312
α_1^W	48.94	51.57	58.71	12.88
α_2^W	35.03	7.18	43.26	13.75
α_1^F	57.16	3.52	50.46	12.65
α_2^F	3.41	.50	16.86	.54

e.g., Hall and Milgrom 2008). This leaves us with an overparameterized specification with 12 target moments and 16 parameters.

The main results are displayed in tables A2 and A3. As we can see, the expanded calibration allows the model to match the AKM moments better, especially $\mathrm{Var}(\psi)/\mathrm{Var}(\theta + \psi)$. However, the model still falls short in some dimensions, such as a too high value of $\mathrm{Corr}(\theta, \psi)$ relative to the data. I see this as consistent with the tension to match $\mathrm{Var}(\psi)/\mathrm{Var}(\theta + \psi)$ and $\mathrm{Corr}(\theta, \psi)$ jointly because of the nonmonotonicities explained in Section III. Using the lower value for $\mathrm{Corr}(\theta, \tilde{\theta})$ of .2, I obtain the closest fit of the AKM moments in the high school sample. However, that calibration falls short with respect to other moments (e.g., too little dispersion in wages and a too low UErate) and features a discount rate of 17.5 percent a year.

References

Abowd, J., R. Creecy, and F. Kramarz. 2002. "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data." Manuscript, http://www2.census.gov/ces/tp/tp-2002-06.pdf.

Abowd, J., F. Kramarz, P. Lengerman, and S. Perez-Duarte. 2004. "Are Good Workers Employed by Good Firms? A Test of a Simple Assortative Matching Model for France and the United States." Working Paper no. 15546, NBER, Cambridge, MA.

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

Andrews, M., L. Gill, T. Schank, and R. Upward. 2008. "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?" *J. Royal Statis. Soc., Ser. A, Statis. in Society* 171 (3): 673–97.

Bagger, J., F. Fontaine, F. Postel-Vinay, and J.-M. Robin. 2014. "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics." *A.E.R.* 104 (6): 1551–96.

Bagger, J., and R. Lentz. 2014. "An Empirical Model of Wage Dispersion with Sorting." Working paper, NBER, Cambridge, MA.

Barlevy, G. 2008. "Identification of Search Models Using Record Statistics." *Rev. Econ. Studies* 75 (1): 29–64.

Becker, G. 1973. "A Theory of Marriage: Part I." *J.P.E.* 81 (4): 813–46.

Burdett, K., and D. Mortensen. 1998. "Wage Differentials, Employer Size, and Unemployment." *Internat. Econ. Rev.* 39 (2): 257–73.

Cahuc, P., F. Postel-Vinay, and J. Robin. 2006. "Wage Bargaining with On-the-Job Search: Theory and Evidence." *Econometrica* 74 (2): 323–64.

Card, D., J. Heining, and P. Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Q. J.E.* 128 (3): 967–1015.

Christensen, B., R. Lentz, D. Mortensen, G. Neumann, and A. Werwatz. 2005. "On-the-Job Search and the Wage Distribution." *J. Labor Econ.* 23:31–58.

Davis, S. J., R. J. Faberman, and J. C. Haltiwanger. 2010. "The Establishment-Level Behavior of Vacancies and Hiring." Technical report, NBER, Cambridge, MA.

Davis, S. J., and J. C. Haltiwanger. 1991. "Wage Dispersion Between and Within U.S. Manufacturing Plants, 1963–86." *Brookings Papers Econ. Activity: Microeconomics,* 115–200.

Eeckhout, J., and P. Kircher. 2010. "Sorting and Decentralized Price Competition." *Econometrica* 78 (2): 539–74.

———. 2011. "Identifying Sorting-In Theory." *Rev. Econ. Studies* 78 (3): 872–906.

Elsby, M. W., B. Hobijn, and A. Şahin. 2010. "The Labor Market in the Great Recession." *Brookings Papers Econ. Activity* (Spring): 1–69.

Fallick, B., and C. Fleischman. 2004. "The Importance of Employer-to-Employer Flows in the US Labor Market." Discussion Paper no. 200434, Fed. Reserve Board, Washington, DC.

Hagedorn, M., T. H. Law, and I. Manovskii. 2012. "Identifying Equilibrium Models of Labor Market Sorting." Working Paper no. 18661, NBER, Cambridge, MA.

Hall, R. E., and P. R. Milgrom. 2008. "The Limited Influence of Unemployment on the Wage Bargain." *A.E.R.* 98 (4): 1653–74.

Hansen, N., and A. Ostermeier. 2001. "Completely Derandomized Self-Adaptation in Evolution Strategies." *Evolutionary Computation* 9 (2): 159–95.

Hornstein, A., P. Krusell, and G. L. Violante. 2011. "Frictional Wage Dispersion in Search Models: A Quantitative Assessment." *A.E.R.* 101 (7): 2873–98.

Iranzo, S., F. Schivardi, and E. Tosetti. 2008. "Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data." *J. Labor Econ.* 26 (2): 247–85.

Jolivet, G., F. Postel-Vinay, and J.-M. Robin. 2006. "The Empirical Content of the Job Search Model: Labor Mobility and Wage Distributions in Europe and the US." *European Econ. Rev.* 50 (4): 877–907.

Kremer, M., and E. Maskin. 1996. "Wage Inequality and Segregation by Skill." Working paper, NBER, Cambridge, MA.

Krueger, A. B., and L. H. Summers. 1988. "Efficiency Wages and the Inter-industry Wage Structure." *Econometrica* 56:259–93.

Lise, J., C. Meghir, and J.-M. Robin. 2016. "Matching, Sorting and Wages." *Rev. Econ. Dynamics* 19:63–87.

Lopes de Melo, R. 2007. "The Implications of Search Models for Wage Dynamics: An Empirical Assessment." Manuscript, https:goo.gl/9KS61A.

———. 2015. "How Firms Affect Wages: A Structural Decomposition." Manuscript, Univ. Edinburgh.

Lucas, R., and E. Prescott. 1974. "Equilibrium Search and Unemployment." *J. Econ. Theory* 7 (2): 188–209.

Mortensen, D. 2003. *Wage Dispersion: Why Are Similar Workers Paid Differently?* Cambridge, MA: MIT Press.

Nagypál, É. 2008. "Worker Reallocation over the Business Cycle: The Importance of Employer-to-Employer Transitions." Manuscript, Northwestern Univ.

Oi, W. Y., and T. L. Idson. 1999. "Firm Size and Wages." In *Handbook of Labor Economics*, vol. 3B, edited by Orley Ashenfelter and David Card, 2165–2214. Amsterdam: Elsevier.

Postel-Vinay, F., and J. Robin. 2002. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica* 70 (6): 2295–2350.

Rosen, S. 1986. "The Theory of Equalizing Differences." In *Handbook of Labor Economics*, vol. 1, edited by Orley Ashenfelter and Richard Layard, 641–92. Amsterdam: Elsevier.

Shi, S. 2001. "Frictional Assignment. I. Efficiency." *J. Econ. Theory* 98 (2): 232–60.

Shimer, R. 2005. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies." *A.E.R.* 95:25–49.

———. 2012. "Reassessing the Ins and Outs of Unemployment." *Rev. Econ. Dynamics* 15 (2): 127–48.

Shimer, R., and L. Smith. 2000. "Assortative Matching and Search." *Econometrica* 68 (2): 343–69.

Woodcock, S. D. 2015. "Match Effects." *Res. Econ.* 69 (1): 100–121.