Identifying Sorting*

Marcus Hagedorn†
Institute for Advanced Studies

Tzuo Hann Law ‡
University of Pennsylvania

Iourii Manovskii§
University of Pennsylvania

PRELIMINARY AND INCOMPLETE

Abstract

Does the market allocate the right workers to the right jobs? What is the allocation of workers to employers that maximizes total output? Since observable (to economists) worker characteristics account for only a small fraction of the wage variance in the data, to answer such questions it is essential to study the patterns of assortative matching between employers and employees based on their unobserved characteristics. This paper enables this line of research. In particular, we show theoretically that all parameters of the classic model of sorting based on absolute advantage in Becker (1973) with search frictions can be identified using only data on wages and labor market transitions rates. In particular, these data are sufficient to assess whether matching between workers and firms is assortative, whether sorting is positive or negative, and to measure the potential effect on output from moving any given worker to any given employer in the economy. We also provide computational algorithms that allow to implement our identification strategy given the limitations (on sample size, frequency of labor market transitions, measurement error, etc.) of the commonly used matched worker-firm data sets and assess their performance. Finally, we extend our identification and implementation strategies to the commonly used class of models of assortative matching based on comparative advantage and provide a test that discriminates between these models.

*This Version: November, 2012. We would like to thank seminar participants at Indiana, Mannheim, Notre Dame, Toulouse, Vienna Institute for Advanced Studies, Bank of France, Search and Matching Workshop at the Philadelphia Fed, 2012 Society for Economic Dynamics Annual Meetings, 2012 NBER Summer Institute, and 2012 Konstanz Workshop on Labor Market Search and the Business Cycle. Support from the National Science Foundation Grant No. SES-0922406 is gratefully acknowledged.

†Institute for Advanced Studies, Stumpergasse 56, A-1060 Vienna, Austria. Email: marcus.hagedorn07@gmail.com.

‡Department of Economics, University of Pennsylvania, 160 McNeil Building, 3718 Locust Walk, Philadelphia, PA, 19104-6297 USA. E-mail: tzuolaw@sas.upenn.edu.

§Department of Economics, University of Pennsylvania, 160 McNeil Building, 3718 Locust Walk, Philadelphia, PA, 19104-6297 USA. E-mail: manovski@econ.upenn.edu.

1 Introduction

Does the market allocate the right workers to the right jobs? Are complementarities between workers and employers important in determining output, productivity, and wages? Do large employers pay higher wages because they employ better workers? What are the sources of inter-industry wage differentials? What is the allocation of workers to employers that maximizes total output? These classic questions are at the heart of current debates in many areas of economics. In business cycle research, there is an ongoing discussion on whether the slow productivity and employment recovery after the Great Recession is due to the mismatch between human capital of unemployed workers and skill requirements of potential employers. In the international trade literature, researchers attempt to determine whether the wage premium of exporting firms is due to them being more productive or having better workers, a question with important implications for understanding the effects of changes in trade regimes. The industry dynamics literature is interested in the role of effective labor input reallocation across producers for productivity dynamics at the micro level. Misallocation at the micro level is relevant for the macro literature as it typically reduces total factor productivity with a potentially important impact on, e.g., income differences across time and across countries. The enhanced focus on this role of resource misallocation represents one of the most important recent developments in the economic growth literature.

It has been long recognized that to make progress in studying these issues it is essential to move the analysis beyond relying on the observable worker and firm attributes that account for only some 30% of the observed variation in wages. This involves expanding the scope of the analysis to include the study of the assortative matching between workers to employers based on their unobservable characteristics, which account for much of the remaining variation. These unobserved characteristics are typically associated, following the lead of Abowd, Kramarz, and Margolis (1999), with worker and firm fixed effects in wages that are estimated using longitudinal matched employer-employee datasets. Unfortunately, the literature has recently established that the key identifying assumptions of this regression approach are inconsistent with the standard equilibrium sorting models and that the worker and firm fixed effects identified using this methodology have no economic interpretation in the context of these models.[1] The key problem is that the fixed effect regression assumes that

[1] Gautier and Teulings (2006) were the first to establish this in a model of sorting based on comparative advantage. This important class of models violates the underlying assumption of the fixed effect regression that workers and firms are globally rankable. Eeckhout and Kircher (2011) later make an even stronger point. They prove that even in the model of sorting on absolute advantage that allows for globally rankable workers and firms, the worker and firm fixed effects in wages have no economic interpretation. These theoretical insights have been confirmed quantitatively in a range of assortative matching models in Lopes de Melo (2009), Lentz (2010), and Lise, Meghir, and Robin (2011),

wages are monotone in firm's productivity (fixed effect). This is inconsistent with an explicit sorting model, where a productive firm may agree to hire a relatively unproductive worker only if that worker accepts a low wage to compensate the firm for the option value of waiting for a more productive potential hire.

Faced with the limitations of the fixed effect regression approach one might hope that an approach more firmly grounded in the theory of sorting models might prove more fruitful. From the perspective of economic theory, a typical starting point for thinking about assignment problems in heterogeneous agent economies is the model of Becker (1973). In labor market applications, the current state-of-the-art formulation is due to Shimer and Smith (2000) who extend the competitive framework in Becker (1973) to allow for time consuming search between heterogeneous workers and firms. This framework is then a natural choice to answer the empirical questions motivating this research agenda. Unfortunately, the hopes of making empirical progress with this class of models have been dashed in the recent literature that argued that their parameters cannot be recovered from the available data. In particular, Eeckhout and Kircher (2011) show that the competitive Becker (1973) model is not identified, and neither is a simplified two period version of the frictional model in Shimer and Smith (2000) with no discounting. It is not even possible to identify which one of two given firms is more productive. As one consequence of the inability to solve this identification problem, existing quantitative work on assortative matching in the labor market has to rely on strong assumptions on technology etc. to be able to take the model to the data. This is problematic as it is these assumptions on technology that determine the patterns and consequences of sorting in the model.

The first contribution of this paper is theoretical. In particular, we provide a solution to this identification problem and establish that the empirical questions motivating this research agenda can be precisely answered using the general model of Shimer and Smith (2000) and using only routinely available matched employer-employee data on wages and labor market transition rates. To put it differently, we establish that all parameters of the model are nonparametrically identified, including the production function. This implies that from wage data alone we can recover the output of any observed employer-employee match and the consequences for output, productivity, and wages of moving any worker to any firm in the economy (subject to some limitations that will be formally spelled out below).

To make progress with this class of models, we find it beneficial to consider the fully dynamic

among others.

versions with discounting. In the general version of the model the value of opening a job vacancy can be established to be monotonically increasing in a firm's productivity. Moreover, we show that the model implies a way to infer the value of a vacancy from data on wages and labor market transitions. This yields a statistic which is monotone in firm productivity and thus can be used to rank firms according to their productivity. The assumption of no discounting used in Eeckhout and Kircher (2011), simplifies the analysis substantially as it delivers that the value of a vacancy is constant. Unfortunately, it is this simplifying assumption that prevents the identification of the model.[2] While the assumption of no discounting is generally not appealing in quantitative work, it is readily tested because the identification strategy proposed here recovers the value of the vacancy for each firm.

The second contribution of this paper is to develop an implementation algorithm for the proposed identification strategy. The first element of the identification strategy is ranking workers. We develop several measures of ranking that are theoretically correct, but whose performance and informational content might differ in small samples and in the presence of measurement error. Our starting point is an equivalence between the problem to rank workers and the problem how to aggregate votes in the social choice literature. These problems are extremely computationally complex (they are NP-hard) but fortunately, spurred by the demand for information aggregation by the Internet search engines, the computer science literature has recently made substantial progress in designing computational algorithms that can efficiently approximate the solution to this problem. We draw on these advances in algorithm research to develop an algorithm that is fast and accurate for the applications we study. The second step of the identification strategy is to rank firms. We show that firms can be ranked based on the expected average difference between the wages they pay to each of their workers and the reservation wages of those workers. This is a simple statistic to compute, but it relies on having an accurate estimate of the reservation wage for each worker, which might be difficult to obtain in short samples. The key insight we use is that once workers are ranked, similarly ranked workers must have similar reservation wages. Thus, we can estimate the reservation wage by considering a group of similar workers, despite the fact that each of those workers is observed for a relatively short period of time.

Being able to rank firms and workers allows us to recover the output of every match. In the

[2] The finding that an assumption made for simplification (here no discounting) impedes identification is reminiscent of a related discussion in the context of hedonic models. In this framework, Brown and Rosen (1982) argue for the non-identifiability of equilibrium hedonic models. Ekeland, Heckman, and Nesheim (2004), however, show that this conclusion is due to a simplifying assumption (linearization of the model) and that the model is identified if the generic nonlinearities of equilibrium hedonic models are used.

model, wages, which are observed in the data, are a function of the output of the match which we are interested in measuring, as well as two objects that our identification strategy allows to measure - the reservation wage of a worker, and the value of a vacancy. Thus, the wage equations can be solved for output as a function of three measurable variables.

Having obtained the full nonparametric identification of the model, in particular of the production function and of the rankings of workers and firms, allows us to answer all questions that motivate our analysis. For example, without any additional assumptions we can compute the optimal assignment on the set where match productivity can be measured. A comparison between the optimal assignment and the observed one reveals the extent of the output loss due to search frictions. We can also determine the importance of complementarities in production and measure the role of frictions and sorting in determining the dispersion of output and productivity across establishments. It is also possible to measure the extent to which sorting on unobservables can account for wage differences across groups of employers (large or small, exporters and no-exporters, belonging to different industries, located in different geographic regions, etc.). Turning to wage dispersion, an application of our method allows to decompose wages into components due to workers, firms, and the assortative matching between them as well as to estimate the role of search frictions and sorting in driving the observed wage dispersion. We discuss other interesting applications of our methodology in the Conclusion.

We assess the performance of the proposed methods in a Monte Carlo study imposing the limitations (on sample size, frequency of labor market transitions, measurement error, etc) of the commonly used matched worker-firm data sets. We find that the identification strategy and the implementation method that we develop are successful at measuring the relevant objects in the model.

Thus, in this paper we develop all the theoretical and computational tools required to enable the empirical analysis using the Becker (1973) model with time consuming search. We limit our analysis to its formulation in Shimer and Smith (2000) because our identification proofs rely on its theoretical properties, such as the existence and properties of equilibria, which have not yet been established more generally. We also think it has considerable pedagogical merit to understand the sources of identification and to tackle the key implementation issues in the simplest possible but relevant model. The desire for analytical clarity also prevents us from attempting to provide substantive answers to the empirical questions motivating this research agenda in this paper. To take the model to the data we think it is necessary to introduce additional features, such as on-the-job search which is prevalent in the data but is absent from the benchmark model we consider. While we expect our identification and implementation strategy to be adaptable to a wider class of related models, including those with on-the-job search, the theoretical analysis becomes tremendously more complicated without yielding additional insights. In addition, the challenge of future empirical analysis will be to disentangle the role of observable and unobservable characteristics in assortative matching. The underlying assumption we make here is that the effects of the observables can be removed using the standard Mincer wage regressions and sorting on the unobservables can then be studied. This is the standard assumption in the literature, but perhaps not the best one. It is possible to adjust the model where sorting occurs on a combination of observable and unobserved characteristics, but the precise formulation will have to by guided by the data. We do not feel that we can do justice to this analysis in this paper.

We do, however, consider one extension that seems substantively important and insightful methodologically. The model in Becker (1973) is based on sorting on absolute advantage. In other words, it is assumed that workers and firms can be ranked in terms of their productivity. In an influential line of research, Gautier and Teulings (e.g., 2012) studied a version of the model in which firms cannot be globally ranked, although there is a well defined notion of the most appropriate firm for each worker. In this model, sorting is based on comparative advantage. We show that our identification strategy can be extended to this version of the model and provide Monte-Carlo evidence that our implementation strategy continues to recover well the objects of interest, including the production function. Finally, we implement a test based on Hall and Heckman (2000) to distinguish between models with and without absolute advantage.

The paper is organized as follows. In Section 2 we describe the standard model with frictional labor market and assortative matching between between workers and firms based on absolute advantage. Section 3 shows theoretically the identification of the model. In Section 4 we develop computational tools needed to implement our identification strategy and evaluate its performance in simulated data sets designed to mimic existing matched employer-employee data sets. In section 5 we extend the analysis to the model of sorting based on comparative advantage in Gautier and Teulings (2012), verify its quantitative performance, and implement a test that allows to distinguish between the two models. Section 6 concludes. Most proofs and details of computations are in the Appendix.

2 The Model

The model description builds on Shimer and Smith (2000), which adds time-consuming search to Becker (1973), with slight generalizations and some modifications. In particular, we do not impose symmetry between both sides of the market. Since the main application of our theory will be the labor market, we instead have workers on the one side of the market and firms on the other side. Both sides with potentially different primitives. Most importantly we impose a linear search technology instead of the quadratic search technology in SS, which seems the better choice for labor market applications. None of our results hinge on this modification.

2.1 Environment

2.1.1 Basics

Time is discrete, all agents are infinitely-lived and maximize the present value of payoffs, discounted with a common discount factor $\beta \in (0,1)$. The unit mass of workers is either employed (e) or unemployed (u) while the unit mass of firms is producing (p) or vacant (v). Workers and firms are heterogeneous with respect to their productivities, denoted by $x \in [0,1]$ and $y \in [0,1]$, respectively. To simplify the exposition we treat each firm as having one job. All the results immediately generalize, however, to each firm having a mass of jobs sharing the same productivity y.[3]

Output of a match between worker x and firm y is given by the twice differentiable nonnegative production function $f : [0,1]^2 \to \mathbb{R}_+$. The existence proof in Shimer and Smith (2000) also requires that f has uniformly bounded first partial derivatives on $[0,1] \times [0,1]$. It is assumed that match output is increasing in worker and firm type, i.e., $f_x > 0$ and $f_y > 0$.[4] This assumption allows x and y to be measured as a worker's or a firm's rank in the corresponding productivity distribution. The rank of a worker (firm) is given by the fraction of workers (firms) who produce weakly less with the

[3] This model of the firm, as simplistic as it is, represents the current state-of-the-art in the literature. As Lentz and Mortensen (2010), pp. 593-594 put it, "all the analyses that we know of assume that output of any given job-worker match is independent of the firm's other matches. Furthermore, firm output is the sum of all the match outputs. Hence, the identification challenge reduces to that of identifying worker and firm contributions over matches and a common match production function. Of course, as the research frontier moves to improve our understanding of multiworker firms, it is likely and appropriately an assumption that will be challenged." We agree with this assessment and hope the identification results established here will continue to be relevant as more sophisticated and empirically implementable theories of the firm are developed.

[4] The assumptions that economic agents can be globally ranked is standard in the models of sorting based on absolute advantage, such as Becker (1973) and Shimer and Smith (2000), and is implicit in the approach of Abowd, Kramarz, and Margolis (1999). In this research project this assumption is only relevant for identifying rankings workers and firms when they can be ranked. If some agents cannot be ranked, e.g., firms in the comparative advantage model of Gautier and Teulings (2012), the proposed identification strategy will reveal this and it will recover the production function correctly.

same firm (worker). In this paper, *productivity*, *rank*, or *type* have identical meanings. Therefore the distributions of worker and firm types are both uniform. If the "original" (non-rank) distributions of worker types are F and of firm types are G and the "original" production function is $f(x, y)$ then we transform the production function

$$\tilde{f}(\tilde{x}, \tilde{y}) = f(\tilde{F}^{-1}(\tilde{x}), \tilde{G}^{-1}(\tilde{y}))$$

and the distributions are

$$\tilde{F}(x) = \tilde{x}$$
$$\tilde{G}(y) = \tilde{y}.$$

We use f instead of $\tilde{f}$ in what follows. We place no additional assumptions on this function. In particular, we do not assume that sorting is either positive or negative but show how to recover this information from the data.

2.1.2 Distributions

The measures characterizing the set of matched and unmatched workers and firms are assumed to be absolutely continuous, implying the existence of a density. Given our identification of types with ranks, the worker and firm time invariant populations are given by $d_w = 1$ and $d_f = 1$. The distribution of producing matches is described by $d_m : [0, 1]^2 \to \mathbb{R}_+$. The functions characterizing the employed and unemployed workers as well as the producing and vacant firms are denoted $d_e(x)$, $d_u(x)$, $d_p(y)$ and $d_v(y)$, respectively.[5] Table 1 summarizes the relationships between these functions.

Table 1: Functions describing distributions

Description	Density Function
Matches	$d_m(x, y)$
Employed workers	$d_e(x) = \int d_m(x, y) \ \mathrm{d}y$
Unemployed workers	$d_u(x) = d_w(x) - d_e(x)$
Producing firms	$d_p(y) = \int d_m(x, y) \ \mathrm{d}x$
Vacant firms	$d_v(y) = d_f(y) - d_p(y)$

[5] Note that these functions do not integrate to one but to the mass of employed, unemployed workers and to the mass of producing and vacant firms, respectively.

Integrating the densities from Table 1 gives the time-invariant measures of aggregate employment, E, of unemployment, U, of producing firms, P, and vacant firms, V:

$$E = \int d_m(x,y) \ \mathrm{d}x\mathrm{d}y = \int d_e(x) \ \mathrm{d}x, \tag{1}$$
$$P = \int d_m(x,y) \ \mathrm{d}x\mathrm{d}y = \int d_p(y) \ \mathrm{d}y, \tag{2}$$
$$U = 1 - E = 1 - \int d_m(x,y) \ \mathrm{d}x\mathrm{d}y = \int d_u(x) \ \mathrm{d}x, \tag{3}$$
$$V = 1 - P = 1 - \int d_m(x,y) \ \mathrm{d}x\mathrm{d}y = \int d_v(y) \ \mathrm{d}y. \tag{4}$$

2.1.3 Timing

It is convenient to think of each period as consisting of two subperiods. In the first subperiod, a worker of type x matched with a firm of type y produces $f(x,y)$. Output of this match is exhausted by payments to the firm, $\pi(x,y)$, and the worker, $w(x,y)$. Vacant firms pay vacancy maintenance costs, c. Unemployed workers obtain flow payoff from non-market activity b. In the second subperiod, new matches are formed when all unmatched workers and firms participate simultaneously in a single labor market subject to search frictions. After matching, each existing match (including a newly formed one) is destroyed with probability δ.[6]

2.2 Search and matching

Only and all unmatched agents engage in random search. A function $m : [0,1] \times [0,1] \to [0, \min(U,V)]$ takes the masses of unemployed workers U and vacant firms V as its inputs and generates meetings. The probability a worker meets a potential employer is given by $\mathbb{M}_u = \frac{m(U,V)}{U}$, while the probability of a vacant firm meeting a potential hire is $\mathbb{M}_v = \frac{m(U,V)}{V}$. These probabilities are time-invariant in the steady-state equilibrium we will consider. The probability for a worker to meet any firm $y \in Y \subseteq [0,1]$ equals $\mathbb{M}_u \frac{\int_Y d_v(y) \ \mathrm{d}y}{V}$. The probability for a firm to meet any worker $x \in X \subseteq [0,1]$ equals $\mathbb{M}_v \frac{\int_X d_v(x) \ \mathrm{d}x}{U}$. These probabilities reflect our assumption of a linear search technology. Using the quadratic search technology in Shimer and Smith (2000) these probabilities would be $\mathbb{M}_u \int_Y d_v(y) \ \mathrm{d}y$ and $\mathbb{M}_v \int_X d_v(x) \ \mathrm{d}x$, respectively. Since we obtain the same search technology by simply setting $U = V = 1$ in the matching process, it will become clear that our results do not depend on the returns to scale of the matching function. Not all meetings necessarily result in

[6] The assumption that newly formed matches are also subject to job destruction shocks enhances the elegance of some expressions below but has no relevance for the substantive results.

matches. Some meetings are between workers and firms who are unwilling to consummate a match and who prefer to continue the search process.

2.3 Strategies and acceptance sets and surplus

The steady-state pure strategy of a worker of type x is to decide with which firms to match with, taking all other strategies as given. This strategy is described by a Borel measurable *acceptance set* $A^w(x)$ of firms that a worker type x is willing to match with (Shimer and Smith (2000) prove that the same type agents use the same strategy). Symmetrically for firms, the Borel measurable *acceptance set* $A^f(y)$ are the workers that a firm of type y is willing to match with. Matching takes place when both the worker and the firm find it mutually acceptable. For a worker of type x, the *matching set* $B^w(x)$ are firms which accept worker type x and are accepted by worker type x. Symmetrically for a firm of type y, $B^f(y)$ are workers who accept to match with firm type y and who are accepted by firms of type y. The matching sets are related to the acceptance sets in the obvious way:

$$B^w(x) \equiv \{\tilde{y} : x \in A^f(\tilde{y}) \wedge \tilde{y} \in A^w(x)\},$$
$$B^f(y) \equiv \{\tilde{x} : y \in A^w(\tilde{x}) \wedge \tilde{x} \in A^f(y)\}.$$

$\overline{B^w}$ and $\overline{B^f}$ denote the complements of B^w and B^f, respectively. Define $\mathcal{B}$ to represent all (x, y) pairs that form in equilibrium:

$$\mathcal{B} \equiv \{(x, y) : y \in A^w(x) \wedge x \in A^f(y)\} \tag{5}$$
$$= \{(x, y) : y \in B^w(x)\} \tag{6}$$
$$= \{(x, y) : x \in B^f(y)\}. \tag{7}$$

2.4 Bellman equations and surplus sharing

Let $V_u(x)$ denote the value of unemployment for a worker of type x, $V_e(x, y)$ worker's x value of employment at a firm of type y, $V_v(y)$ the value of a vacancy for firm y, and $V_p(x, y)$ the value of firm y employing a worker of type x. The surplus of a match between worker x and firm y is defined as

$$S(x, y) \equiv V_p(x, y) - V_v(y) + V_e(x, y) - V_u(x). \tag{8}$$

Generalized Nash bargaining over the match surplus $S(x,y)$ with workers' bargaining power $\alpha = \frac{1}{2}$ implies

$$\left.\begin{aligned} \alpha S(x,y) &= V_e(x,y) - V_u(x), \\ (1-\alpha)S(x,y) &= V_p(x,y) - V_v(y). \end{aligned}\right\} \tag{9}$$

Following this rule, it is clear that by construction, $y \in A^w(x)$ if and only if $x \in A^f(y)$. Hence,

$$\left.\begin{aligned} A^w(x) &= B^w(x) = \{y : S(x,y) \geq 0\}, \\ A^f(y) &= B^f(y) = \{x : S(x,y) \geq 0\}. \end{aligned}\right\} \tag{10}$$

Using the surplus sharing rule (9), we obtain the following steady state value functions. The derivations of these equations are provided in Appendix I.1.

$$V_u(x) = b + \beta V_u(x) + \beta\alpha(1-\delta)\mathbb{M}_u \int\limits_{B^w(x)} \frac{d_v(\tilde{y})}{V} S(x,\tilde{y})\ \mathrm{d}\tilde{y}, \tag{11}$$

$$V_v(y) = -c + \beta V_v(y) + \beta(1-\alpha)(1-\delta)\mathbb{M}_v \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U} S(\tilde{x},y)\ \mathrm{d}\tilde{x}, \tag{12}$$

$$V_e(x,y) = w(x,y) + \beta V_u(x) + \beta\alpha(1-\delta)S(x,y), \tag{13}$$

$$V_p(x,y) = f(x,y) - w(x,y) + \beta V_v(y) + \beta(1-\alpha)(1-\delta)S(x,y). \tag{14}$$

2.5 Stationary distribution of matches

In the stationary match distribution, for all worker and firm type combinations in the matching set the numbers of destroyed and created matches are the same:

$$\forall (x,y) \in \mathcal{B} \qquad \underbrace{\delta d_m(x,y)}_{\text{destruction}} = \underbrace{(1-\delta) d_u(x)\mathbb{M}_u \frac{d_v(y)}{V}}_{\text{new match formation}}. \tag{15}$$

The probability for a worker (of any type) to meet a firm of type y is the product of the probability to meet any firm, $\mathbb{M}_u$, and the probability that this firm is of type y, $\frac{d_v(y)}{V}$. This is multiplied by $(1-\delta)$ because newly formed matches can get destroyed in the same period. Integrating over all matches yields that the total inflow into unemployment equals the total outflow out of unemployment.

$$\underbrace{\int_{\mathcal{B}} \delta d_m(x,y)\ \mathrm{d}x\mathrm{d}y = \delta E}_{\text{inflow}} = \underbrace{(1-\delta)\int_0^1 d_u(x)\mathbb{M}_u \int_{B^w(x)} \frac{d_v(y)}{V}\ \mathrm{d}y\mathrm{d}x}_{\text{outflow}}.$$

2.6 Equilibrium

In a *steady state search equilibrium* (SE) all workers and firms maximize expected payoff, taking the strategies of all other agents as given.[7] The economy is in steady-state. A SE is then characterized by the density $d_u(x)$ of unemployed workers, the density $d_v(y)$ of vacant firms, the density of formed matches $d_m(x, y)$ and wages $w(x, y)$. The density $d_m(x, y)$ implicitly defines the matching sets as it is zero if no match is formed and is strictly positive if a match is consummated. Wages are set to ensure the surplus sharing rule (9) and match formation is optimal given wages w, i.e. a match is formed whenever the surplus is (weakly) positive (see (10)). The densities $d_u(x)$ and $d_v(x)$ ensure that the flow equations in (15) hold.

To prove existence Shimer and Smith (2000) assume that the production is consistent with either positive assortative matching (PAM) or negative assortative matching (NAM), defined as follows:

Definition 1. *Let worker types $x_1 < x_2$ and firm types $y_1 < y_2$.*
The production function f exhibits PAM if $x_1 \in B^f(y_1)$ and $x_2 \in B^f(y_2)$ whenever $x_1 \in B^f(y_2)$ and $x_2 \in B^f(y_1)$.
The production function f exhibits NAM if $x_1 \in B^f(y_2)$ and $x_2 \in B^f(y_1)$ whenever $x_1 \in B^f(y_1)$ and $x_2 \in B^f(y_2)$.

The equilibrium existence proof in Shimer and Smith (2000) uses their assumption of a quadratic matching function. Nöldeke and Tröger (2009) extend the existence proof to linear matching technology used in this paper.[8]

Proposition 1 (**Shimer and Smith (2000) and Nöldeke and Tröger (2009)).** *If f either exhibits PAM or NAM then a SE exists.*

SS claim that the assumption of either PAM or NAM just avoids a more complicated existence proof and thus can be dispensed with. More specifically, the assumption of PAM or NAM rules out an atom of zero surplus matches, i.e.

$$\forall x \neq x' : \mu(\{y : S(x, y) = S(x', y) = 0\}) = 0, \tag{16}$$

where μ is the Lebesgue measure. Imposing

$$\forall x \neq x', \forall y : \mu(\{y' : f(x, y) + f(x', y') = f(x, y') + f(x', y)\}) = 0, \tag{17}$$

[7] As SS we assume that a matched is formed if agents are indifferent.
[8] They also argue that the proofs extend to our non-symmetric environment.

ensures this property. It thus avoids both the assumption of PAM or NAM and also a more complicated existence proof (see the Step 1 of the proof of Lemma 3 in Shimer and Smith (2000)). This property is for example satisfied by the two production functions used in SS as examples which satisfy neither PAM nor NAM: $(x+y)^2$ and $(x+y-1)^2$. It does not hold for modular production functions such as $x+y+k$ (k a constant). However for large enough k, every worker matches with every firm and thus (16) is trivially satisfied.

3 Identification: Theory

The roadmap for identification is as follows. We show how to use matched employer-employee data on wages and labor market transition rates to first identify the ranking of workers and then to identify the ranking of firms. This allows us to identify the presence and sign of sorting. Finally we identify the remaining primitives of the model, in particular, the output of every observed match between any worker and any firm.

3.1 Ranking Workers

We now derive several statistics which are monotonically increasing in worker types. Such statistics naturally provide a way to rank workers.

Result 1. *$V_u(x)$, $V_e(x,y)$ and $w(x,y)$ are increasing in x.*

Let $\tilde{y}(x)$ be the firm that pays the lowest wage accepted by worker type x.

Result 2. *The reservation wage given by $w(x,\tilde{y}(x))$, is increasing in x.*

Let $y^{max}(x)$ be the firm that pays the highest wage to worker type x.

Result 3. *The maximum wage given by $w(x,y^{max}(x))$, is increasing in x.*

Result 4. *The adjusted average wage defined as*

$$w^{av}(x) \equiv \left(1-\mathbb{M}_u+\delta\mathbb{M}_u+\mathbb{M}_u(1-\delta)\int\limits_{\overline{B^w(x)}}\frac{d_v(y)}{V}\,\mathrm{d}y\right)w(x,\tilde{y}(x))+\mathbb{M}_u(1-\delta)\int\limits_{B^w(x)}\frac{d_v(y)}{V}w(x,y)\,\mathrm{d}y \tag{18}$$

is increasing in x.

Formal proofs of these results can be found in Appendix I, although the results themselves are intuitive. The fact that the value of unemployment is increasing in worker's type follows because a more productive worker can always imitate the acceptance strategy of the less productive worker but produce more and consequently receive higher wages. This induces a more productive worker to set a higher reservation wage. The fact that wages within firms are increasing in workers type follows directly from the assumption that the production function is increasing in worker productivity. It is easy to prove that the average wage (without) the adjustment is not necessarily increasing in x.[9] To see this, consider two workers with different productivity. A more productive worker might be matching with a wider set of firms (some of which do not accept the less able worker). However, the more able worker might be only marginally acceptable to those firms who typically match with even better workers. As a consequence, those firms pay low wages to this worker. Thus, the average wage of the worker over the employment history might be lower then that of a less productive worker. The more productive worker still obtains the higher value because he spends a larger fraction of his lifetime employed. Result 4 corrects for this effect by imputing the value of unemployment to unemployed workers and defining the average wage over the lifetime rather than of the portion of lifetime the worker spends employed.

We have derived a number of statistics that allow us to rank workers. In particular, wages within firms, lowest and highest accepted wages, and adjusted average wages provide theoretically valid and equivalent rankings of workers. The performance of these ranking procedures might differ, however, in small samples and in the presence of measurement error in wages. We assess their quantitative performance in Section 4. For the rest of this section we assume that a complete ranking of workers has been constructed.

3.2 Ranking Firms

To rank firms we derive a statistic which is monotonically increasing in firm type y. This is non-trivial since the wage of worker x, $w(x, y)$, is not always increasing in firm productivity. The same problem applies to the surplus of a match, $S(x, y)$. The strategy is as follows. We first establish

[9] Formally, since separation rates are identically δ at all firms a worker matches with, a worker's average wage is proportional to $\int_{B^w(x)} w(x,y) d_v(y)\, \mathrm{d}y$. Assuming, for simplicity, that $B^w(x) = [\underline{\varphi}(x), \overline{\varphi}(x)]$, we get

$$\frac{\partial}{\partial x} \int\limits_{B^w(x)} w(x,y) d_v(y)\, \mathrm{d}y = \int\limits_{B^w(x)} \frac{\partial w(x,y)}{\partial x} d_v(y)\, \mathrm{d}y + \overline{\varphi}'(x) w(x, \overline{\varphi}(x)) d_v(\overline{\varphi}(x)) - \underline{\varphi}'(x) w(x, \underline{\varphi}(x)) d_v(\underline{\varphi}(x)). \tag{19}$$

Clearly, Equation (19) is not necessarily increasing in x.

that value of a vacancy is increasing in y. This implies that the surplus a vacancy is expected to generate is also increasing in y. Any bargaining game where both parties benefit from an increase in the surplus implies that the average surplus of workers employed by firm y is also increasing in y. Finally, we show that the average surplus of workers employed by firm y can be expressed as a function of wages, yielding a simple observable statistic that is increasing in y and thus allows to rank firms. In this Section, we include some of the proofs in the main text as we consider them instructive (and surprisingly simple).

The foundation for our strategy of ranking firms is provided by the following result.

Result 5. *$V_v(y)$ and $V_p(x, y)$ are increasing in y.*

Since the data on the value of the vacancy or on the expected profits from posting a vacancy are not available in standard data sets, our strategy is to relate these monotone statistics to observable statistics from the worker side. The next result is stated only in terms of workers' value functions.

Result 6. *The expected surplus due to newly hired workers is given by*

$$(1-\delta)\mathbb{M}_v \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U} \left(V_e(\tilde{x}, y) - V_u(\tilde{x})\right) \, \mathrm{d}\tilde{x}$$

is increasing in y.

Proof of Result 6. Using equation (9),

$$(1-\delta)\mathbb{M}_v \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U} \left(V_e(\tilde{x}, y) - V_u(\tilde{x})\right) \, \mathrm{d}\tilde{x} = \alpha(1-\delta)\mathbb{M}_v \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U} S(\tilde{x}, y) \, \mathrm{d}\tilde{x}.$$

From (12) it follows that

$$\frac{V_v(y)(1-\beta) + c}{\beta(1-\alpha)} = (1-\delta)\mathbb{M}_v \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U} S(\tilde{x}, y) \, \mathrm{d}\tilde{x}$$

From Result 5 both sides of (20) are increasing in y. Multiplying both sides of (20) by α yields the desired result. ■

The proof used that the value of a vacancy is increasing in firm type y and then involved two steps. First, since the value of a vacancy is related to the expected surplus by an accounting identity (equation 12), the expected surplus is also increasing in firm type (equation (20)). The next step uses that Nash-bargaining implies that both the worker and the firm benefit from an increase in the surplus. Nash bargaining even has a stronger implication as the two parties benefit from an increase in the surplus in fixed proportions, determined by the bargaining power. This strong implication is however not used here and our ranking results will hold for other bargaining games where both parties benefit from an increase in the surplus.

Next, we relate this statistic to wages which are observable in the data.

Result 7. *The expected wage premium over the reservation wage of newly hired workers given by*

$$\Omega(y) = (1-\delta)\mathbb{M}_v \int_{B^f(y)} \frac{d_u(x)}{U}(w(x,y) - w(x,\tilde{y}(x)))\ \mathrm{d}x \tag{20}$$

is increasing in y.

Note that this expectation is taken when the vacancy is still unfilled. The proof uses three simple insights. Let $w(x,\tilde{y}(x))$ be the lowest wage (the reservation wage) that worker x receives where $\tilde{y}(x)$ is the firm type that pays this wage. The first insight is that the lowest wage is equal to the return of being unemployed,

$$w(x,\tilde{y}(x)) = (1-\beta)V_u(x) = (1-\beta)V_e(x,\tilde{y}(x)).$$

Second, the wage of a worker is a premium over the reservation wage (see Equation (13)),

$$\begin{aligned} w(x,y) &= (1-\beta)V_u(x) + (1-\beta(1-\delta))\left(V_e(x,y) - V_u(x)\right) \\ &= w(x,\tilde{y}(x)) + (1-\beta(1-\delta))\left(V_e(x,y) - V_u(x)\right). \end{aligned}$$

Finally, this implies that the worker's surplus is proportional to the difference between the wage and the reservation wage,

$$w(x,y) - w(x,\tilde{y}(x)) = (1-\beta(1-\delta))\left(V_e(x,y) - V_e(x,\tilde{y}(x))\right).$$

Using Result 6 completes the proof.

For the empirical implementation it turns out to be useful to decompose $\Omega(y)$ into two factors that, as we show below, can be easily measured in the data. The first is the average wage premium of newly hired workers at firm y, $\Omega^e(y)$, and the second one is the probability to fill a vacancy, $q(y)$.

The average wage premium equals

$$\Omega^e(y) = \int\limits_{B^f(y)} \frac{\frac{d_u(x)}{U}(w(x,y) - w(x,\tilde{y}(x)))}{\int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U}\,\mathrm{d}\tilde{x}}\,\mathrm{d}x. \tag{21}$$

The probability that a vacancy of type y is filled equals

$$q(y) = (1-\delta)\mathbb{M}_v \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U}\,\mathrm{d}\tilde{x} \tag{22}$$

It then holds that

$$\Omega(y) = q(y)\Omega^e(y). \tag{23}$$

The empirical counterpart of Ω^e is defined for a firm j of type $y(j)$ as

$$\hat{\Omega}^e_t(j) = \sum_{\{i \text{ employed at } j \text{ at } t\}} \frac{(w_t(i) - w^{min}(i))}{E_t(j)},$$

where $E_t(j)$ is the number of workers employed at firm j at time t. From the law of large numbers, we obtain that $\hat{\Omega}^e_t(j)$ converges to $\Omega^e(y(j))$. We discuss the measurement of the (type-dependent) job filling rates $q(y)$ below.

3.2.1 Ranking Firms: Special Cases of PAM and NAM

Result 7 enables us to rank firms in terms of their productivity. Note that Ω is increasing in y regardless of whether the model features positive or negative assortative matching, or indeed neither. In particular, it does not require any assumptions on the production function f, i.e. neither super- nor sub-modularity.[10] If we however assume a special case of PAM or NAM (see Definition 1), we can establish another result that can be used to refine the ranking.

Let $\hat{w}(x)$ be a function that is increasing in worker type, for instance, the reservation wage, maximum wage or adjusted average wage derived above. We then define

$$\Theta(y) \;=\; \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{\int_{B^f(y)} d_u(\tilde{x})\,\mathrm{d}\tilde{x}} \hat{w}(\tilde{x})\,\mathrm{d}\tilde{x}.$$

[10] A production function is supermodular if the cross-derivative is positive and it is submodular if the cross-derivative is negative.

Result 8. *If PAM then* $\frac{\partial \Theta(y)}{\partial y} > 0$*. If NAM then* $\frac{\partial \Theta(y)}{\partial y} < 0$*.*

Result A-1 in the appendix provides the empirical counterpart of this statistic.

Note that this statistic is only useful if the production function is everywhere either PAM or NAM. Moreover, it is increasing if the production function exhibits PAM and is decreasing if the it exhibits NAM and consequently cannot be used to identify the sign of sorting. However, it is monotone, and as a consequence can help refine (if we have PAM or NAM) the ranking based on Result 7 in small samples.

Also note that firms cannot be ordered based on the data on average profits that are available in some datasets. This is because just as average wages do not necessarily increase in x, average profits are not necessarily increasing in y.

3.3 Sign and Strength of Sorting

Having ranked workers and firms, we can compute Spearman's rank correlation between x and y in the data, which is just the Pearson correlation coefficient since both types are already ranked. This correlation is a natural indicator of the sign of sorting. For example, a value of 1 indicates perfect positive assortative matching and a value of -1 indicates perfect negative assortative matching.

3.3.1 Relationship to the Literature

Being able to determine whether we have PAM or NAM seems surprising in view of the recent results of Eeckhout and Kircher (2011). They use a simplified version of Atakan (2006) to show that the sign of sorting cannot be identified from wage data. More precisely, they demonstrate, for every supermodular production function that induces PAM, the existence of a submodular production function that induces NAM and generates identical wages.

In terms of empirical applications, the non-identification is related to the observation that wages are not always increasing in firm productivity. As a result, papers using the methodology pioneered in Abowd, Kramarz, and Margolis (1999), do not identify firm productivity. Abowd, Kramarz, and Margolis (1999) implement a linear regression of wages on person and firm fixed effects. This implicitly assumes that wages are increasing in firm type, contrary to what most search models imply. Our method allows to rank firms taking into account that wages are not monotone in y.

The key difference between the model here (that is Shimer and Smith (2000)) and the models in Atakan (2006) and Eeckhout and Kircher (2011) is that we discount whereas search cost are explicit (and additive) in the latter two papers. To see why this is essential for ranking firms rearrange the Bellman equation (12) of a vacancy in our model:

$$V_v(y)(1-\beta) \quad = -c + \beta(1-\alpha)(1-\delta)\mathbb{M}_v \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U} S(\tilde{x}, y)\ \mathrm{d}\tilde{x}.$$

In the limit $\beta \to 1$ we get that the expected surplus is a constant,

$$c \quad = \quad (1-\alpha)(1-\delta)\mathbb{M}_v \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U} S(\tilde{x}, y)\ \mathrm{d}\tilde{x},$$

the Constant Surplus Condition in Theorem 1 in Atakan (2006). If instead $\beta < 1$ then we have that $V_v(y)(1-\beta)$ is increasing in y and so is the expected surplus. This monotonicity (independent from the production function) of expected surplus is the key step in our ranking of firms. We then measure the surplus as being proportional to the wage premium of a worker resulting in our statistic $\Omega(y)$ (see equation (20)), that is expressed in terms of wages only. Constructing the same statistic in Atakan (2006) does not yield a function that is monotonically increasing in y but is, instead, a constant. The impossibility to rank firms in Atakan (2006) is thus due to the knife-edge assumption of no discounting. As soon as this assumption is relaxed, firms can be ranked. However, this reasoning does not extend to Eeckhout and Kircher (2011) due to the simplifications they make relative to Atakan (2006). They assume essentially a two period model where the first period is a standard labor market with search frictions and the second period is frictionless. As a result, the frictionless second period outcome ($w^*(x)$ for workers and $\pi^*(y)$ for firms) serves as the continuation value, i.e. the value of a vacancy equals

$$V(y) = -c + \underbrace{\int S(x, y) dx}_{\text{expected surplus}} + \beta\pi^*(y),$$

where Eeckhout and Kircher (2011) assume $\beta = 1$. Our statistic $\Omega(y)$ is monotone in y if and only if the expected surplus is. In Shimer and Smith (2000) we show that this is the case because the value of a vacancy is increasing in y. In Eeckhout and Kircher (2011) such a simple relationship between the value of a vacancy and the expected surplus does not exist. Solving the above equation for the expected surplus yields

$$\int S(x, y) dx = V(y) + c - \beta\pi^*(y),$$

which is not necessarily increasing in y since $\pi^*(y)$ is increasing in y and enters with a negative sign. As a result our statistic $\Omega(y)$ which is proportional to expected surplus is not necessarily monotonically increasing.

Expected surplus is even not constant if $\beta = 1$ (as it is in Atakan (2006)) since the continuation value in Eeckhout and Kircher (2011) is the frictionless allocation and not the value of a vacancy as in Shimer and Smith (2000) and in Atakan (2006). Thus, even relaxing the assumption of no discounting, our method will not recover the ranking of firms in Eeckhout and Kircher (2011) due to their modeling of frictionless second period matching.

3.4 Identifying Remaining Model Parameters

We now show how to identify the remaining model features, i.e. the value of being unemployed $V_u(x)$, the value of being employed $V_e(x, y)$, the value of a vacancy $V_v(y)$ and the value of producing for a firm $V_p(x, y)$. We also measure the probability to receive an offer for both unemployed workers and vacant firms $\mathbb{M}_u$ and $\mathbb{M}_v$ and the probability to fill a vacancy $q(y)$. Finally we demonstrate how to measure the output of a match $f(x, y)$. Estimated values are denoted with a hat, $\hat{}$.

3.4.1 Measuring $V_u(x)$ and $V_e(x, y)$

The Bellman Equation (13), implies, using $V_e(x, \tilde{y}(x)) = V_u(x)$, that

$$V_u(x)(1 - \beta) = w(x, \tilde{y}(x)), \tag{24}$$

that is we use the lowest wage to measure the (type-dependent) value of being unemployed as

$$\hat{V}_u(x) = \frac{w(x, \tilde{y}(x))}{1 - \beta}. \tag{25}$$

We next turn to $V_e(x, y)$. Consider a worker of type x, who starts working at firm type y at time $t = 0$, becomes unemployed at time t_U, and receives wage $w_t = w(x, y)$ for all t between $t = 0$ and $t = t_U - 1$. We then define

$$\sum_{t=0}^{t_U - 1} \beta^t w_t + \beta^{t_U} \hat{V}_u(x), \tag{26}$$

where, of course, we use the estimated value for $\hat{V}_u(x)$. Averaging across all these sums for all types x starting at firm y gives us our estimate $\hat{V}_e(x, y)$.

We then also have an estimate of surplus multiplied by the bargaining power

$$\hat{\alpha}\hat{S}(x,y) = \hat{V}_e(x,y) - \hat{V}_u(x). \tag{27}$$

The next step to recover the surplus is to measure the bargaining power.

3.4.2 Measuring α

Shimer and Smith (2000) impose $\alpha = 0.5$. Thus, in their model this is not a free parameter the value of which needs to be identified. Nevertheless, if one is willing to relax this assumption, in Appendix I.4 we describe how this parameter can be identified from the data. From now we assume that we have an estimate for α, which we denote $\hat{\alpha}$.

3.4.3 Measuring $V_v(y)$ and $V_p(x,y)$

Since we already have an estimate of $\alpha S(x,y)$, we now also have an estimate of $S(x,y)$. We next turn to the measurement of $V_v(y)$, which equals

$$V_v(y) = -c + \beta V_v(y) + \beta(1-\alpha)(1-\delta)\mathbb{M}_v \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U} S(\tilde{x},y)\ \mathrm{d}\tilde{x}.$$

The value of a vacancy is related to our estimate $\Omega(y)$ through

$$V_v(y)(1-\beta) = -c + \beta\frac{1-\alpha}{\alpha}(1-\delta)\Omega(y).$$

Since we have obtained measures of α and Ω in the data, and we can also measure δ and β, we obtain a measure of

$$V_v(y)(1-\beta) + c.$$

Our estimate is then

$$\hat{V}(y) = \frac{\beta}{1-\beta}\frac{1-\hat{\alpha}}{\hat{\alpha}}(1-\delta)\hat{\Omega}(y),$$

which differs from the true value by $\frac{c}{1-\beta}$, a constant that does not depend on y.

Using this our estimate of $V_p(x,y)$ then equals

$$\hat{V}_p(x,y) = \hat{V}_v(y) + (1-\hat{\alpha})\hat{S}(x,y),$$

which also differs from the true value by a constant.

3.4.4 Transition Rates

For every worker type we can measure the probability to move out of unemployment, which in the model equals

$$\hat{\lambda}(x) = (1-\delta)\mathbb{M}_u \int\limits_{B^w(x)} \frac{d_v(\tilde{y})}{V}\,\mathrm{d}\tilde{y}.$$

From this equation and with firm level vacancy data, we can already measure $\mathbb{M}_u$ since the integral is known and we measure the LHS. A more robust way is to integrate over all worker types

$$\int_0^1 \hat{\lambda}(x)\,\mathrm{d}x = (1-\delta)\mathbb{M}_u \int_0^1 \int\limits_{B^w(x)} \frac{d_v(\tilde{y})}{V}\,\mathrm{d}\tilde{y}\mathrm{d}x,$$

which again allows us to solve for $\mathbb{M}_u$.
Similarly the probability to fill a vacancy for firm type y equals

$$q(y) = (1-\delta)\mathbb{M}_v \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U}\,\mathrm{d}\tilde{x},$$

which we can measure directly in the data if we observe vacancy data at the firm level. If vacancy data at the firm level are not available, $q(y)$ can still be easily estimated using only the aggregate number of vacancies or the aggregate number of unemployed workers, as we now show.

The probability, $\tilde{q}_y$ that a vacancy posted by firm j of type $y(j)$ is filled conditional on meeting a worker is simply the share of unemployed workers belonging to this firm's matching set in total unemployment. We now index workers by their estimated rank $\hat{x}$ and $\hat{u}(\hat{x})$ denotes this type's lifetime unemployment rate. Using the law of large numbers it holds that

$$\tilde{q}_y \equiv \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U}\,\mathrm{d}x = \frac{\sum_{\hat{x}\in B^f(y)} \hat{u}(\hat{x})}{\sum_{\hat{x}} \hat{u}(\hat{x})}. \tag{28}$$

Note that $\tilde{q}_y$ can be computed without any vacancy data. Denote by $H_t(y)$ the observed number of new hires in firms of type y at time t, and by $V_t(y)$ the unobserved number of vacancies posted by these firms. Equation (28) and the law of large numbers imply that

$$\frac{H_t(y)}{(1-\delta)\tilde{q}_y} = \mathbb{M}_v V_t(y). \tag{29}$$

Adding up across all firms and time periods, and rearranging yields an estimate for $\mathbb{M}_v$ (and $\mathbb{M}_u$ as $U = V$):

$$\hat{\mathbb{M}}_v = \frac{1}{1-\delta}\frac{\sum_{y,t}\frac{H_t(y)}{\tilde{q}_y}}{\sum_{y,t} V_t(y)}. \tag{30}$$

Note that in the model the aggregate number of vacancies is equal to aggregate unemployment ($\sum_{y,t} V_t(y) = \sum_t U_t$), which can be readily computed in the data.

3.4.5 Measuring output $f(x, y)$

Using the equation for wages (A2), our estimate of the production function $f(x, y)$ on the set of matches that actually form, then equals

$$\hat{f}(x,y) = \frac{w(x,y) - \hat{\alpha}(\beta-1)\hat{V}_v(y) - (1-\hat{\alpha})(1-\beta)\hat{V}_u(x)}{\hat{\alpha}}.$$

The output of a match is determined by inverting the wage equation, expressing the output $f(x, y)$ as a function of the observed wage $w(x, y)$ and the two measured outside options $V_v(y)$ and $V_u(x)$. For this to be feasible the researcher has to know the exact wage equation. In the model of Shimer and Smith (2000) this is the case since Nash bargaining is imposed. Other wage determination mechanisms which imply an invertible wage equation would also allow for an identification of output.

4 Implementation and Quantitative Evaluation

In this section we develop the key implementation steps of the proposed identification strategy and evaluate their performance in a Monte Carlo study over a wide range of parameter values that are likely to be encountered in empirical work. The detailed implementation algorithm is described in Appendix II.

4.1 Parameterization

We assume that a researcher has access to a matched employer-employee panel data set. We conduct the analysis with a data set that has a time dimension of 20 years. This is a conservative choice because the longer the data set the more precise our method is. Most currently available and commonly used matched data sets (e.g., from Brazil, Denmark, Germany, France) have a longer time span. We assume that the data includes the information on wages, all employment and unemployment spells of the worker over the duration of the sample, and all hires and separations at the firm level. We simulate the model at a monthly frequency. The production functions commonly used in the literature belong to the constant elasticity of substitution (CES) family. We consider three such function:

i) $f(x, y) = 0.6 + 0.4(x^{1/2} + y^{1/2})^2$, which induces positive assortative matching (PAM),

ii) $f(x, y) = (x^2 + 2y^2)^{1/2}$, which induces negative assortative matching (NAM), and

iii) $f(x, y) = \mathbb{1}_{\{x \leq 1/2\}} \left(\frac{1}{2} + xy\right) + \mathbb{1}_{\{x > 1/2\}} \left(\frac{1}{2} + \left((x - \frac{1}{2})^2 + \left(\frac{y}{2}\right)^2\right)^{1/2}\right)$, which induces neither positive nor negative assortative matching (NEITHER). Instead, the pattern of sorting changes over its domain (PAM for $x \leq 1/2$ and NAM for $x > 1/2$).

The literature has largely restricted attention to identifying sorting assuming that the production function features either positive or negative assortative matching. This motivates our choice of the first two production functions. However, our method does not place any restrictions on the production function, except for it being increasing in each argument. The choice of the third production function is designed to illustrate this point. The production functions are scaled to generate a realistic amount of wage dispersion.

We also consider three distributions of workers and firms (these are the "original" non-rank distributions). The literature largely restricts to either a uniform or normal distributions. We consider both and for the normal distribution we choose the mean of 0.5 and the variance of 0.25 (the distribution is truncated and normalized on $[0, 1]$ interval). We also consider a bimodal distribution constructed as the sum of two normals: $N(0.2, 0.25) + N(0.8, 0.25)$ truncated and normalized to integrate to one on $[0, 1]$. The distributions are discretized into 50 values on an evenly spaced grid. We simulate a small sample of 30,000 workers. There is the same number of jobs in the economy. Jobs of the same productivity level are assigned to firms with an upper bound of 100 jobs per firm. As not all these jobs are filled at a point in time, the actual size of employment at each firm varies across parameterizations but is not more than 100 workers. We set the discount factor to 0.996 at monthly frequency to be consistent with the annual interest rate of 4%.

We assume the standard Cobb-Douglas form of the meeting function $m(s, v) = \kappa s^{\nu} v^{(1-\nu)}$. We set the elasticity parameter $\nu = 0.5$ as this parameter plays no interesting role in our stationary model. We consider a wide range for the scale parameter $\kappa = \{0.4, 0.7\}$ to generate the job finding probabilities ranging between a high of about 50% a month in the US and a low of about 15% in some European countries. Similarly, we choose two values for the separation rate $\delta = \{0.01, 0.025\}$, roughly spanning the US and European evidence.

We set the flow utility of unemployed workers and the vacancy cost to zero to ensure that all workers and firms weakly prefer to be match with at least some partners to remaining always vacant or unemployed. We also consider symmetric bargaining weights of 0.5 for workers and firms.

Finally, we allow for measurement error in wages. Hagedorn and Manovskii (2012) estimate that measurement error accounts for approximately 20% of the variance of residual wages in the US NLSY data. This is likely an upper bound on the matched employer-employee data sets as these data are typically based on administrative sources with highly reliable wage information. Nevertheless, to make the test of the proposed method more stringent, we add iid noise to every wage observation with the variance of 20% of the correctly measured wage variance. The error is simulated as draws from a normal distribution truncated at three standard deviations around the mean of zero.

The values of parameters used in simulations are summarized in Table 2.

Table 2: Parametrization

Parameter	Symbol	Option 1	Option 2	Option 3
Production function	$f(x, y)$	PAM	NAM	NEITHER
Worker distribution	d_w	Uniform	Normal	Bi-Modal
Firm distribution	d_f	Uniform	Normal	Bi-Modal
Discount factor	β	0.996		
Separation rate	δ	0.01	0.025	
Meeting function scale	κ	0.4	0.7	
Meeting function elasticity	ν	0.5		
Worker's bargaining weight	α	0.5		
Worker's flow utility of unemp.	b	0.0		
Vacancy cost	c	0.0		
Measurement error in wages	ϵ	20% of overall wage variance		

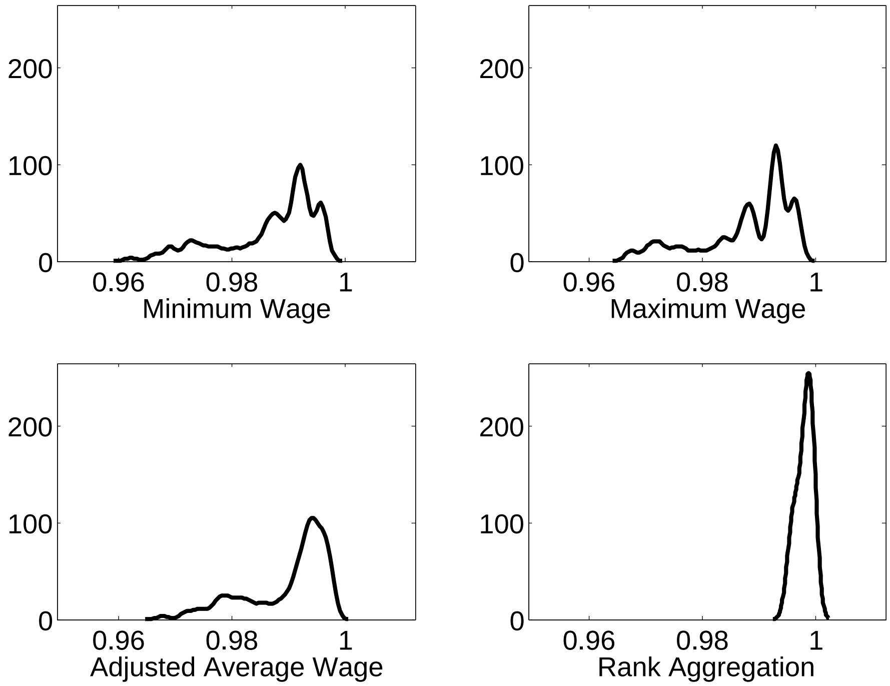
Thus, all combinations of parameters result in 108 distinct parameterizations that we consider. Appendix Figure A-1 summarizes the range of values that a number of variables of interest take across all simulations. Most tend to lie within empirically plausible ranges.

4.2 Ranking of Workers

Results 2, 3 and 4 have established that in large samples workers can be ranked based on the lowest accepted wage, the highest accepted wage, or the adjusted average wage. To assess the performance of each of these methods in small samples and in the presence of measurement errors in wages we report the rank correlations of the true worker types and types recovered using each of these methods across simulations.

In addition, Result 1 implies that wages within a firm are increasing in worker productivity x, which provides another way to rank workers according to their productivity. However, the presence of measurement error in wages dilutes this way of ranking. Within one firm one worker could be ranked higher than another worker not because he is more productive (actually he is less productive) but just because of the measurement error. And the ranking between these two workers may not be consistent with the ranking from other firms. To solve this problem, we build on the insights from social choice theory, where voters rank candidates, potentially inconsistent with each other. In our application voters correspond to firms and workers to voting alternatives. An aggregate ranking then

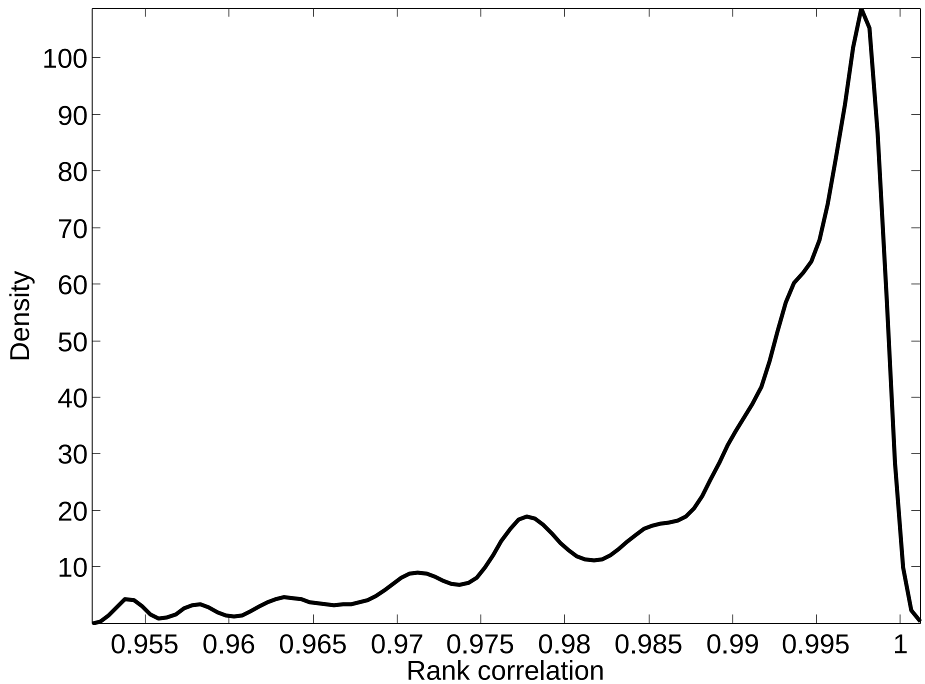
Figure 1: Density of the correlation between the estimated and true worker ranks across parameterizations.



minimizes the number of disagreements between individual votes which defines the *Kemeny-Young rank aggregation* problem first described in Kemeny (1959) and Kemeny and Snell (1963). We refine this procedure. We not only count the number of disagreements but instead assign weights to the ranking of all worker pairs, determining how likely it is that the observed wages (with measurement error) indicate the true ranking. If for example the wage of worker i is much higher than the wage of worker j, we assign a high weight whereas the weight is small if the wages are very similar. We use a Bayesian approach to compute these weights. The goal is then to find a ranking that maximizes the total weights in favor of a proposed ranking. To deal with the computational complexity of this problem, we build on insights from Kenyon-Mathieu and Schudy (2007) who provide a polynomial time algorithm that approximates the solution to this problem with arbitrary degree of accuracy. We use a portion of their algorithm and find that it is reasonably fast while serving our purposes. A detailed description is provided in Appendix III.

Figure 1 reports kernel density estimates of the correlation between the estimated and true worker ranks across all simulations. The correlations are relatively high for all measures. In particular, the adjusted average wage dominates both the minimum and the maximum wage in its performance in ranking workers. The measure based on rank aggregation outperforms any of the

Figure 2: Density of the correlation between the estimated and true firm ranks across parameterizations.



individual measures.

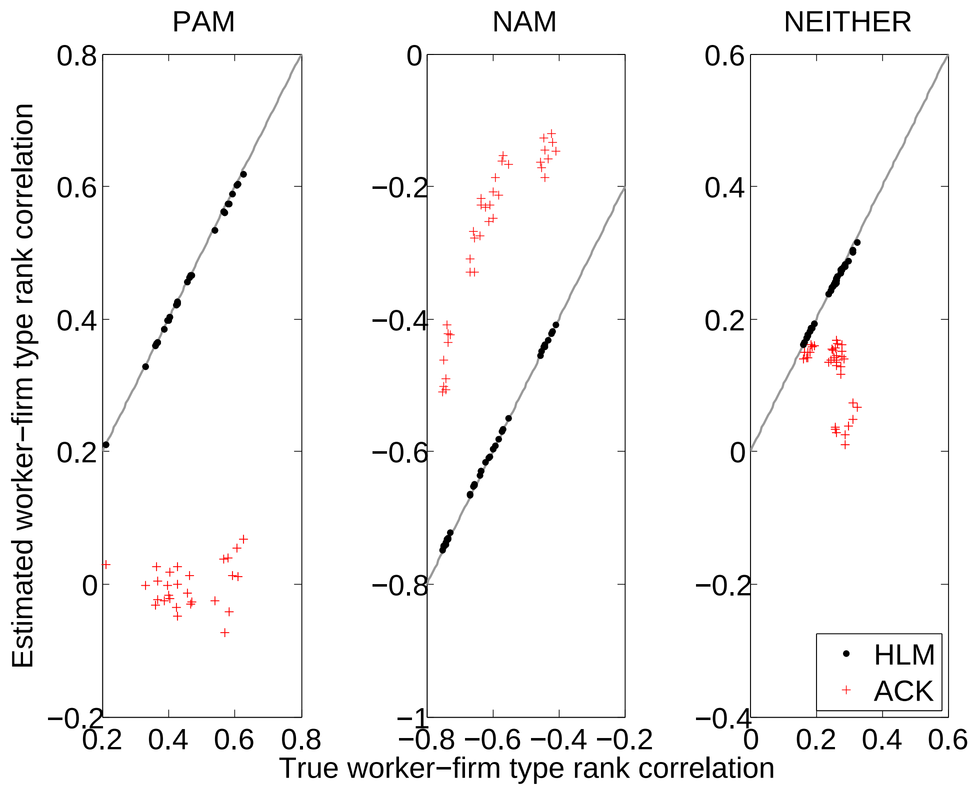
4.3 Ranking of Firms

To rank firms one simply needs to compute the expected average difference between the wages a firm pays to each of its workers and the reservation wage of those workers. The only challenge is to obtain an accurate estimate of the reservation wage for each worker, despite the limited time dimension of the available data. The key insight we use is that once workers are ranked, similarly ranked workers must have similar reservation wages. Thus, we can estimate the reservation wage by considering a group of similar workers, despite the fact that each of those workers is observed for a relatively short period of time. It is also straightforward to correct for the presence of the measurement error in wages as described in the Appendix.

Figure 4.3 plots the density of the correlation between the ranking of firms based on Result 7 and true ranking across the parameterizations we consider. In all parameterizations that we consider the ranking of firms is identified quite precisely.

After firms are ranked, we can also group similarly ranked firms (i.e., firms with very similar Ω values) into bins or types. The measured value of a vacancy for a firm type is then the average

Figure 3: Correlation between identified worker and firm ranks against true correlation.



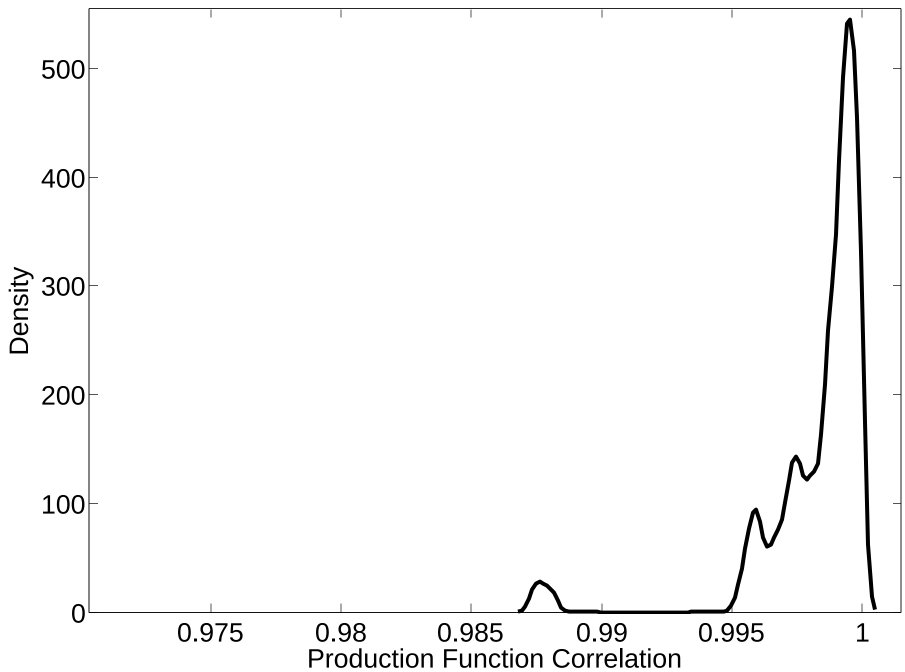
value of a vacancy for all the firms assigned to this type. While the production function can be estimated at the level of an individual firm, estimating it on worker and firm bins helps eliminate the sampling error in small samples.

4.4 Strength of Sorting

Figure 3 plots the correlation between identified worker and firm ranks against true correlation. Here we separate the three production functions to illustrate that or identification strategy easily identifies the sign of sorting. In all cases this crude measure of the strength of sorting performs quite well.

For comparison, in the same figures we report the correlation between worker and firm fixed effects estimated using the exact least squares formulas provided by Abowd, Creecy, and Kramarz (2002). The results confirm the findings in the literature that such reduced form estimates of the strength of sorting are severely biased towards zero.

Figure 4: Correlation between true and estimated production functions.



4.5 Measuring $f(x, y)$

To evaluate how well our method recovers the production function, in Figure 4 we plot the density of correlation between the true and estimated production functions across all parameterizations (we have grouped similarly ranked firms and workers into types x and y as discussed above). The correlations are generally very high.

To provide a better sense of the ability of our identification and implementation strategy to recover the production function, in Figures 5 - 7 we plot the true and estimated production functions for three particular examples from the set of the parameterizations we considered. The three production functions induce positive, negative, or neither positive nor negative assortative matching and we use the same set of parameters for all cases: $\beta = 0.996$, $\kappa = 0.3$, $\delta = 0.01$, $\alpha = 0.5$, $b = c = 0$ and $\mathcal{d}_w = \mathcal{d}_f = U[0, 1]$, with the measurement error equal to 20% of the variance of wages. Each figure contains the true production function (dark red with black lines if viewed in color) defined on the equilibrium matching set and the estimated one (transparent blue). As the functions are essentially on top of each other, to help appreciate the closeness of the fit, for each production function we provide four views with the red line representing the axis of rotation. The estimated production functions are presented without any smoothing or filtering.

Figure 5: True and estimated PAM production function.

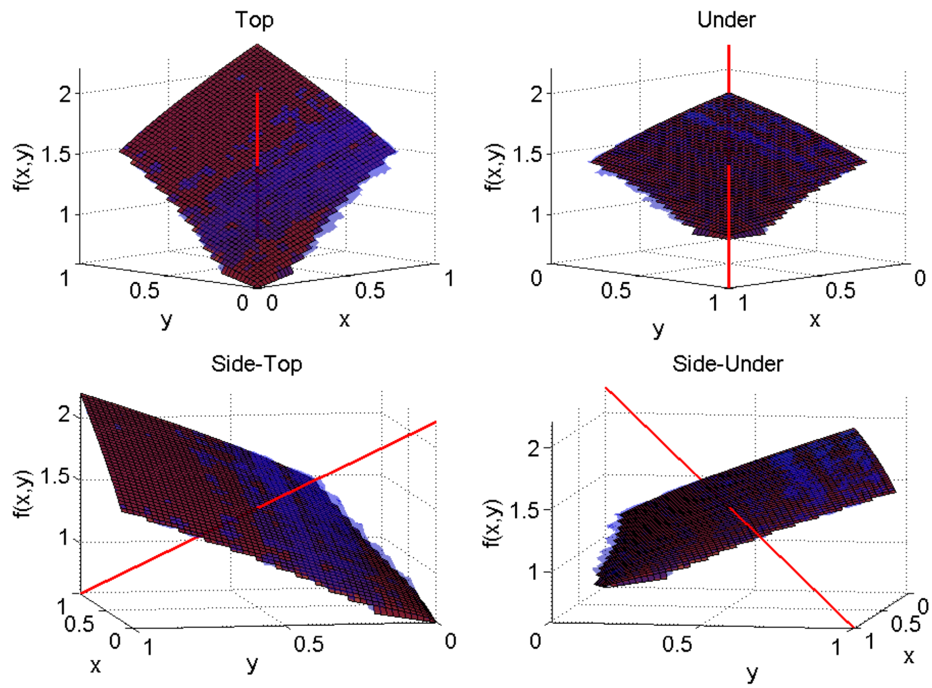


Figure 6: True and estimated NAM production function.

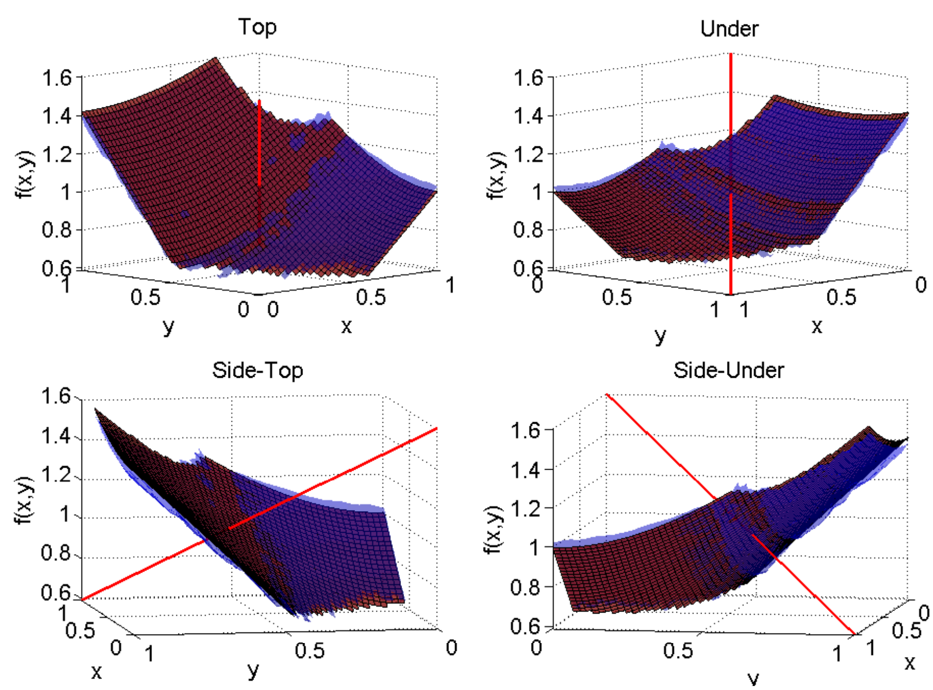
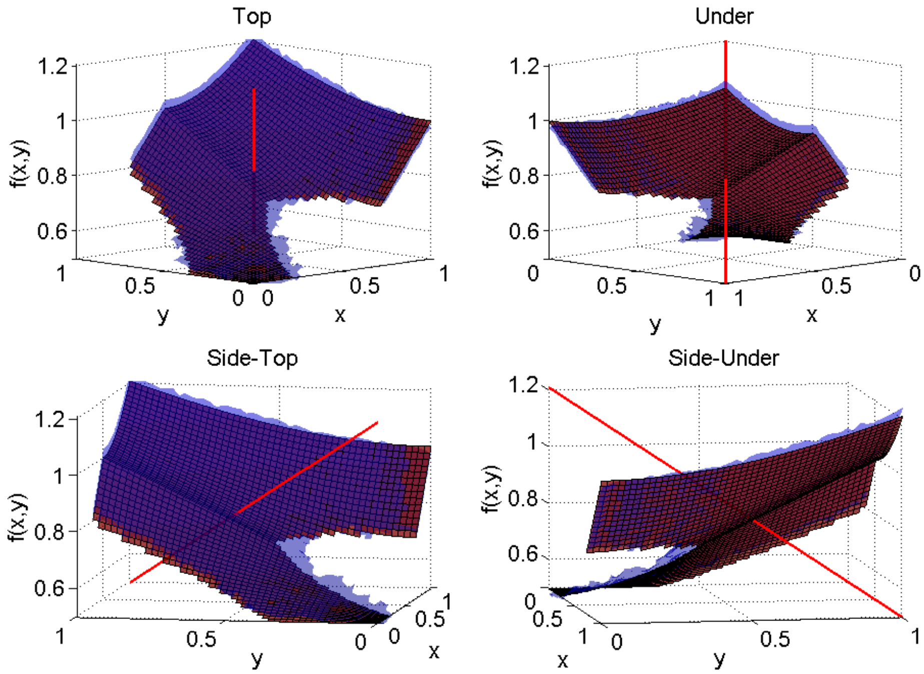


Figure 7: True and estimated Neither NAM nor PAM production function.



4.6 Measuring the effect of search frictions on output

A classic question in the literature is to assess the magnitude of output losses due to mismatch between workers and firms. We now evaluate the ability of our identification strategy to provide a reliable quantitative answer to this question.[11]

To do so we first derive the (counterfactual) allocation in a world without frictions. To solve for the frictionless assignment we need to find a one-to-one assignment (bijection) $\mu : [0,1] \to [0,1]$ of workers to firms such that the total output $\sum_x f(x, \mu(x))$ is maximized. This assignment problem is a classic and well studied combinatorial optimization problem. Our identification strategy identifies the production function only on the set of (x, y) matches observed in the data. Since our objective is to find an optimal assignment on this set, we assume that the output outside of the observed

[11] An alternative approach to measuring the cost of mismatch is pursued by Teulings and Gautier (2004), and Gautier and Teulings (2006, 2012). These authors analyze related but fundamentally different models from Shimer and Smith (2000). In particular, jobs can not be ranked according to productivity in the work of Gautier and Teulings. They also assume a functional form for output which features complementarities between jobs and workers. As a result the questions of how to rank firms, determine the sign of sorting, and identify the production function in the data do not come up. We assess the ability of our identification strategy to identify the production function and to provide a quantitative assessment of the cost of mismatch in their model below. Eeckhout and Kircher (2011) measure the cost of mismatch in a finite horizon version of the model in Shimer and Smith (2000), where a period with search frictions is followed by a frictionless period. In this simplified framework they measure the cost of mismatch and are able to measure the cross-derivative of the production function on average. Their method however does not extend to Shimer and Smith (2000) as it uses their assumption that a period with frictions is followed by a frictionless world.

frictional matching set is zero.

There are several existing algorithms that can solve this problem in polynomial time.[12] However, a complete solution is not required to approximate the effect of the elimination of search frictions on output. Instead, a much smaller scale assignment problem can be solved on a random sample of workers and firms. We choose the size of the sample so that the maximum total output of the sample scaled to the size of the total population of workers and firms becomes invariant to the sample size. Across our simulations, we found that a sample of about 5000 workers and 5000 jobs is sufficient. On a sample of that size we can solve the problem in minutes using the Jonker and Volgenant (1987) algorithm without special hardware.

Denote by $\mathcal{E}^{no\,fric}$ the expectation of frictionless output $f(\cdot, \mu(\cdot))$:

$$\mathcal{E}^{no\,fric} = \int_0^1 f(x, \mu(x)) \, \mathrm{d}x, \tag{31}$$

where we used that worker ranks are uniformly distributed. In the presence of frictions, let $\mathcal{E}^{fric}$ be the expectation of f:

$$\mathcal{E}^{fric} = \int_{\mathcal{B}} f(x, y) d_m(x, y) \, \mathrm{d}x\mathrm{d}y. \tag{32}$$

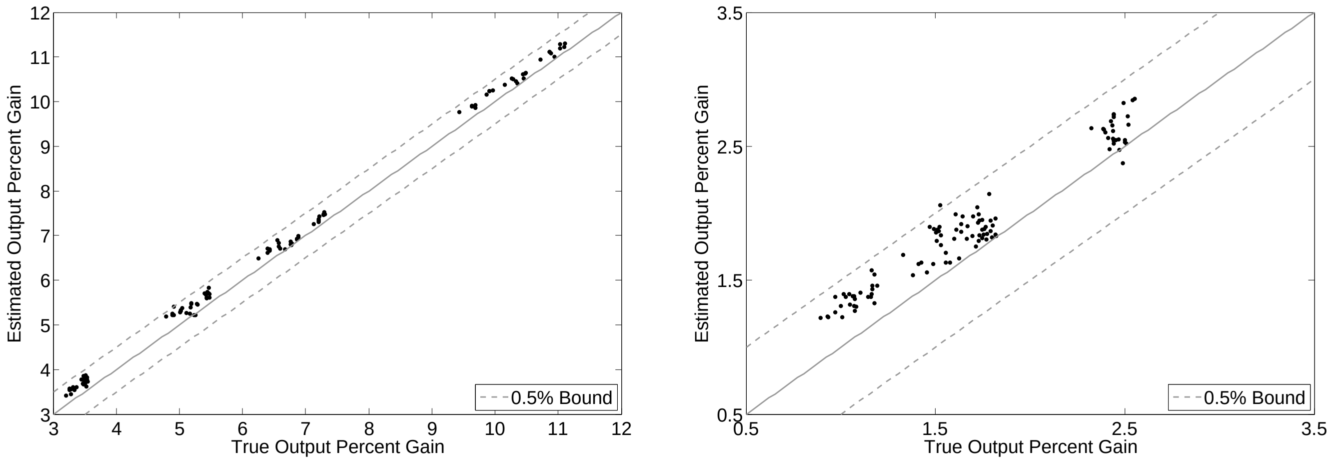
Then, the output loss due to misallocation is the difference between the expected output without frictions, $\mathcal{E}^{no\,fric}$ and the expected output with frictions, $\mathcal{E}^{fric}$:

$$\Delta^{\mathcal{E}} = \mathcal{E}^{no\,fric} - \mathcal{E}^{fric}. \tag{33}$$

In Figure 8, we plot the percent output gain from the optimal reallocation of workers $100 \cdot \frac{\Delta^{\mathcal{E}}}{\mathcal{E}^{fric}}$. The true output gain as a percent of frictional output is on the x-axis while the estimated gain as a percent of frictional output is on the y-axis. When estimating the gain from the reassignment, we use the estimated agent type as well as the estimated production function. To help visually interpret the results, the figure also includes two dotted lines that represents a mistake of plus or minus one half of one percent of output. In Panel 8(a) we plot the estimated gains from eliminating all frictions, including the elimination of frictional unemployment. In Panel 8(b) we keep the employment level of each worker type fixed and only consider the effect from optimal reassignment of workers employed in the economy with frictions.[13] We interpret the results as indicating that the method performs

[12]See Burkard, DellAmico, and Martello (2009) for a thorough review.

[13]Note that the relatively small gains from the reallocation in this experiment are driven by the CES production



(a) Eliminating frictional unemployment.

(b) Keeping frictional employment level.

Figure 8: Estimated gains from eliminating frictions.

quite well in estimating the gain from the optimal worker reallocation.

4.7 The Role of Discounting

Recall that Eeckhout and Kircher (2011) prove that in a special case of the search model with two periods and without discounting the ranking of firms cannot be established. Eeckhout and Kircher (2011) discuss their non-identification results in the presence of discounting. They find that their theoretical results do not exactly apply in this case but note (correctly) that it is very difficult to detect any ranking of firms from individual wages. Our approach does not suffer from this problem because it departs from considering individual wages in two important ways, both consistent with the theoretical model. First, for every individual we consider the difference between his actual wage and his reservation wage. Second we aggregate this difference across all workers employed in a firm. We have repeated our complete Monte Carlo analysis using the monthly discount factor as high as 0.999 and found that this does not measurably affect our ability to identify the objects of interest.

5 Identification in Gautier and Teulings (2006, 2012)

So far we have considered only production functions $f(x, y)$ which are increasing in both arguments. With this assumption both workers and firms have absolute advantage: A firm y either produces more output with all workers than some firm y' or it produces less with all workers. And similarly

function. Larger gains from eliminating mismatch can be obtained with more cumbersome specifications of the production functions. We have explored a number of such specifications and found that our method continues to perform equally well.

for workers. If a worker x produces more output than a worker x' at some firm y then this worker x also produces more output than x' in all other firms. If the production function features pure comparative advantage instead, such a global ranking of workers and firms does not exist and the ranking of workers in terms of output can be different in different firms. That is a worker x can produce more output than worker x' in a firm y but worker x' produces more than x in some other firm y'. An example of such a production function would be $c - \frac{1}{2}\gamma(x-y)^2$. The idea of sorting on comparative advantage is built upon in, e.g., the work of Gautier and Teulings (2012). They use a production function which combines elements of absolute and comparative advantage:

$$\log f(x,y) = x - \frac{1}{2}\gamma(x-y)^2. \tag{34}$$

We show now that we can also recover this production function. It will become clear that our identification strategy, with one modification, generalizes to production functions which are not monotonically increasing.

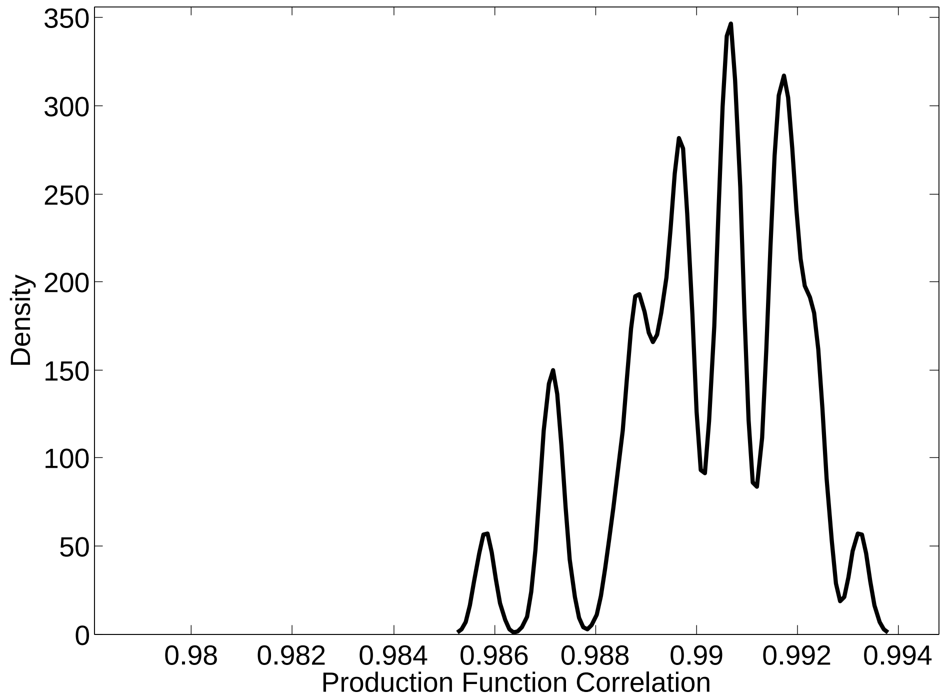
The procedure to rank workers is unchanged. It's implementation is however a greater challenge since a higher ranked worker does not obtain higher wages in all firms but just in most of them (For the production function (34) this is the case if $\gamma > 1$). Our implementation for the production function (34) will show that this works still very well although we use $\gamma = 2$, the calibrated value in Gautier and Teulings (2012). Ranking of firms is not possible with this production function since what matters here is just how close x and y are, i.e. y enters only in the $(x-y)^2$ term. As a result, a firm's expected profit is not necessarily increasing in y and therefore $\Omega(y)$ is not necessarily increasing in y. Nevertheless we can compute $\Omega(y)$ as before and use it as an input to recover the production function, since it still holds that

$$V_v(y)(1-\beta) = -c + \beta\frac{1-\alpha}{\alpha}(1-\delta)\Omega(y).$$

The other value functions $V_e(x,y), V_p(x,y)$ and $V_u(x)$ can also be computed as before.

In the benchmark we assigned firms to bins to measure a (monotone) production function $f(x,y)$. Concretely, we ordered firms by their estimated Ω and assigned firms with very similar values of Ω to the same bin/type. This was justified as we showed that Ω is increasing in firm type y if the production function is increasing in y. As this does not hold here anymore we use a different procedure to bin firms. Instead of considering firms with very similar Ω to be in the same bin, we now consider firms with identical matching sets to be in the same bin. That is all firms which match

Figure 9: Correlation between true and estimated production functions in Gautier and Teulings.



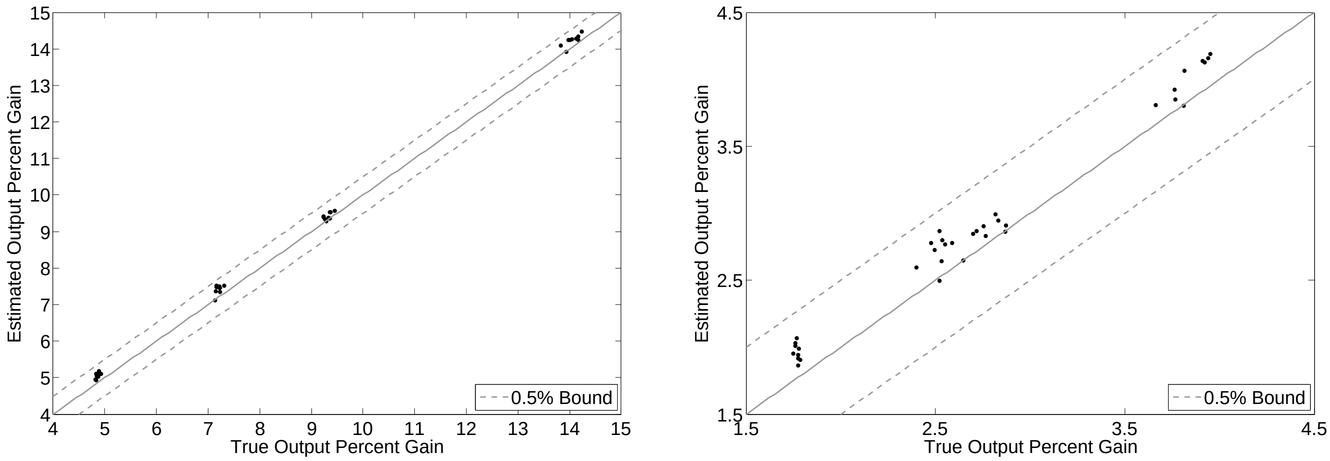
with the same set of worker types are now considered to be in the same bin. We then compute the value of a vacancy and output for a firm as the average value for all firms in this bin.

This leaves open the question how we can determine from the data whether a production function is monotone or not. We tackle this issue in Section 5.1.

We use the same parameterizations and worker and firm distributions as before, resulting in 36 distinct parameterizations. Appendix Figure A-2 summarizes the range of values that a number of variables of interest take across all simulations. They continue to lie within empirically plausible ranges.

To evaluate the ability our method to recover the production function, in Figure 9 we plot the density of the correlation between the true and the estimated production functions across all parameterizations. The correlations are generally very high.

In Figure 10, we plot the percent output gain from the optimal reallocation of workers. The true output gain as a percent of frictional output is on the x-axis while the estimated gain as a percent of frictional output is on the y-axis. When estimating the gain from the reassignment, we use the estimated agent type as well as the estimated production function. To help interpret the results, the figure also includes two dotted lines that represents a mistake of plus or minus one half of one percent of output. We interpret the results as indicating that the method performs quite well in



(a) Eliminating frictional unemployment. (b) Keeping frictional employment level.

Figure 10: Estimated gains from eliminating frictions in Gautier and Teulings.

estimating the gain from the optimal worker reallocation.

5.1 Testing for Absolute vs. Comparative Advantage

In this section we develop a test to distinguish between production functions that are monotonically increasing in firm type and those that are not monotone such as the one in Equation (34). We conduct this test for all worker types x, i.e. we test for the monotonicity of $f(x,\cdot)$ in y for all x. To this aim we use the method developed in Hall and Heckman (2000), which is based on the slopes of local linear estimates of $f(x,\cdot)$. Suppose that type x matches with firms $y_1, \ldots y_n$ and the estimated outputs at this n levels are $z_1 := f(x, y_1), \ldots, z_n := f(x, y_n)$. Let the variance of z_i be σ^2. To test for the monotonicity of a one-dimensional function $g := f(x,\cdot)$ Hall and Heckman (2000) define the following test statistic. Let $0 \le r \le s-2 \le n-2$ be integers, let a, b be constants and put

$$S(a, b \mid r, s) = \sum_{i=r+1}^{s} \{z_i - (a + by_i)\}^2.$$

For each choice of (r, s) define the OLS-estimates $\hat{a} = \hat{a}(r, s)$ and $\hat{b} = \hat{b}(r, s)$ by

$$(\hat{a}, \hat{b}) = argmin_{(a,b)} S(a, b \mid r, s),$$

and let

$$Q(r, s)^2 = \sum_{i=r+1}^{s} \{y_i - (s-r)^{-1} \sum_{j=r+1}^{s} y_j\}^2.$$

Then $\hat{b}(r,s)Q(r,s)$ has variance σ^2 for each pair (r,s). The test statistic is then

$$T_m = \max\{-\hat{b}(r,s)Q(r,s) : 0 \leq r \leq s-m \leq n-m\}$$

The monotonicity of $f(x,\cdot)$ is tested against a constant function. To generate confidence bands, the statistic T_m is generated 500 times for artificial data generated by drawing n times from a Normal distribution with mean zero and variance σ^2.

We implement this procedure for all production functions and all parameterizations considered above. For each production function and each parametrization we conduct the test for all worker types. Three outcomes are possible for each of these test: reject a constant function in favor of an increasing function, cannot reject a constant function, and reject a constant function in favor of a decreasing function. For each test we count in percentages how often we observe each of these three possible outcomes. For each production function we then take the average of these values and report them in Table 3.

Table 3: Monotonicity Test

Production Function	Increasing	Constant	Decreasing
PAM	100.00%	0.00%	0.00%
NAM	99.94%	0.06%	0.00%
NEITHER	98.06%	1.94%	0.00%
Gautier and Teulings	20.94%	61.72%	17.33%

We find that for all three monotone functions we nearly always reject the constant function in favor of an increasing function. In the case of the non-monotone function (34) we find that we can reject a constant function in favor of an increasing function only in 22.39 percent of the cases (note that this is not the evidence of a low power of the test; the production function is indeed increasing in y for some of the worker types x with this production function). In the remaining 77.61 percent of cases we either accept the constant function or we even reject in favor of a decreasing function.

We also provide a second way to assess the results. For each test we say that we reject an increasing function if we reject in favor of an increasing function in less than 95% of all tests. In case of monotone functions we never reject an increasing function. For the non-increasing function we reject in 100% of all tests. Table 4 summarizes the results.

Table 4: Monotonicity Test

Production Function	Reject Increasing	Accept Increasing
PAM	0	100%
NAM	0	100%
NEITHER	0	100%
Gautier and Teulings	100%	0%

6 Conclusion

We have shown theoretically that all the parameters of the assortative matching model with search frictions analyzed in Shimer and Smith (2000) can be identified using only data on wages and labor market transitions rates. In particular, these data are sufficient to assess whether matching between workers and firms is assortative and whether sorting is positive or negative. We have also provided computational algorithms that allow to implement our identification strategy given the limitations (on sample size, frequency of labor market transitions, measurement error, etc.) of the commonly used matched worker-firm data sets, and found that they perform well in a Monte Carlo study. Finally, we extend our identification and implementation strategies to the model of assortative matching based on comparative advantage in Gautier and Teulings (2006, 2012) and provide a test that allows to discriminate between these models.

The key contribution of the paper is that it provides a way to empirically study assortative matching between employers and employees based on their unobserved characteristics. We hope that our methodology would be applied in future work to address many important questions. We already outlined several applications in the Introduction. Some further applications are listed below.

Excess output dispersion due to misallocation. As discussed in Section 4.6 our identification strategy allows to measure output of each worker-firm match and the observed allocation of workers to firms can be compared to the frictionless one, yielding, e.g, an estimate of the output loss due to misallocation. Similar arguments allow to measure the effect of misallocation.

The dispersion of output in the frictionless world is just the variance of the frictionless output computed above $f(x, \mu(x))$, i.e.

$$\mathcal{V}^{no\,fric} = \int_0^1 (f(x, \mu(x)) - \mathcal{E}^{no\,fric})^2 dx, \tag{35}$$

where $\mathcal{E}^{no\,fric}$ is the expectation of $f(\cdot, \mu(\cdot))$.

In the presence of frictions the variance of output equals

$$\mathcal{V}^{fric} = \int_{\mathcal{B}} (f(x,y) - \mathcal{E}^{fric})^2 d_m(x,y)dxdy, \tag{36}$$

where $\mathcal{E}^{fric}$ is the expectation of f.

The excess variance in the world with search frictions (which could be negative) then equals

$$\Delta^{\mathcal{V}} = \mathcal{V}^{fric} - \mathcal{V}^{no\,fric}. \tag{37}$$

Decomposing Wage Differences between Groups of Firms. Many of the questions motivating this research agenda - e.g, the persistent employer size-wage differences, interindustry wage differentials, spatial wage differences, exporter wage premium - relate to differences in wages paid by various groups of firms. Once productivity rankings of workers and firms are established and the production function is identified, decomposing these differences is fairly straightforward in the context of the model.

Wage Variance Decomposition. At the most basic level, once the rankings of workers and firms have been identified, the variance of wages can be decomposed into the contribution of workers, firms and complementarity/sorting.

Let w^0 be the mean wage, $w^0 = \int_{X\times Y} w(x,y)d(x,y)$, where $d(x,y)$ is the density on the set $X \times Y$ (can be a subset). To decompose the variance define the following functions: $w_X(x) = \int_Y w(x,y)d(x,y)dy - w^0$, $w_Y(x) = \int_X w(x,y)d(x,y)dx - w^0$, $w_{XY}(x,y) = w(x,y) - w_X(x) - w_Y(y) - w^0$. Clearly, $w(x,y) = w^0 + w_Y(y) + w_X(x) + w_{XY}(x,y)$.

Sobol (1993) shows that w_X, w_Y and w_{XY} are mutually orthogonal, so that the following variance decomposition holds:

$$Var(w) = Var(w_X) + Var(w_Y) + Var(w_{XY}), \tag{38}$$

with the interpretation that $Var(w_X)$ is the variance due to worker effects, $Var(w_Y)$ is the variance due to firm effects and $Var(w_{XY})$ is the variance due to complementarity and assortative matching between X and Y.

Mismatch and Wages. It is also possible to separately identify the role of assortative matching and search frictions in determining the extent of the observed dispersion in wages. The methodology parallels the analysis of the role of sorting and misallocation in determining the variance of output discussed above.

Wage Discrimination. Interestingly, the assortative matching model combined with the proposed identification strategy provides a simple way to measure discrimination in wages. Consider two distinct groups of workers, say, men and women. Our identification strategy provides two ways of determining how similar the workers in these two groups are. One is - using wages - by looking at how workers are ranked. The second one is to look at matching sets. Identical or similar workers types have identical or similar matching sets. In the benchmark model, both ways of comparison will yield the same result. This is because in the benchmark model workers are paid according to productivity and there is no discrimination in wages. If there is discrimination in wages then the first way of comparison provides an incorrect ranking of productivity. The hiring is however still according to productivity. In the presence of discrimination, the discrepancy between the two ways of ranking provides information about how large discrimination in wages is.

References

ABOWD, J. M., R. H. CREECY, AND F. KRAMARZ (2002): "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," LEHD Program Technical Paper TP-2002-06, U.S. Census Bureau.

ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): "High Wage Workers and High Wage Firms," Econometrica, 67(2), 251–334.

AILON, N., M. CHARIKAR, AND A. NEWMAN (2008): "Aggregating Inconsistent Information: Ranking and Clustering," Journal of the Association for Computing Machinery, 55(5), 23:1–23:27.

ATAKAN, A. E. (2006): "Assortative Matching with Explicit Search Costs," Econometrica, 74(3), 667–680.

BARTHOLDI, J., C. A. TOVEY, AND M. A. TRICK (1989): "Voting Schemes for Which It Can Be Difficult to Tell Who Won the Election," Social Choice and Welfare, 6, 157–165, 10.1007/BF00303169.

BECKER, G. (1973): "A Theory of Marriage: Part I," Journal of Political Economy, 81(4), 813–846.

BLANCHFLOWER, D. G., A. J. OSWALD, AND P. SANFEY (1996): "Wages, Profits and Rent-Sharing," Quarterly Journal of Economics, 111(1), 227–251.

BROWN, J. N., AND H. ROSEN (1982): "On the Estimation of Structural Hedonic Price Models," Econometrics, 50(3), 765–769.

BURKARD, R., M. DELLAMICO, AND S. MARTELLO (2009): Assignement Problems. SIAM.

CONDORCET, J. (1785): "Essai sur l'application de l'analyse la probabilitdes dcisions rendues la p lu ra litdes voix," in American Mathematical Society Bookstore.

DE BORDA, J. C. (1781): Memoire sur les Elections au Scrutin. Histoire de l'Academie Royale des Sciences, Paris.

DRISSI-BAKHKHAT, M., AND M. TRUCHON (2004): "Maximum likelihood approach to vote aggregation with variable probabilities," Social Choice and Welfare, 23(2), 161–185.

Eeckhout, J., and P. Kircher (2011): "Identifying Sorting – In Theory," The Review of Economic Studies, 78(3), 872–906.

Ekeland, I., J. J. Heckman, and L. Nesheim (2004): "Identification and Estimation of Hedonic Models.," Journal of Political Economy, 112(1), S60 – 109.

Gautier, P. A., and C. N. Teulings (2006): "How Large Are Search Frictions?," Journal of the European Economic Association, 4(6), 1193–1225.

——— (2012): "Sorting and the Output Loss due to Search Frictions," Discussion Paper TI 2011-010/3, Tinbergen Institute.

Hagedorn, M., and I. Manovskii (2012): "Search Frictions and Wage Dispersion," mimeo, University of Pennsylvania.

Hall, P., and N. E. Heckman (2000): "Testing For Monotonicity of a Regression Mean by Calibrating for Linear Functions," The Annals of Statistics, 28(1), 20–39.

Jonker, R., and A. Volgenant (1987): "A Shortest Augmenting Path Algorithm for Dense and Sparse Linear Assignment Problems," Computing, 38, 325–340.

Kemeny, J., and J. Snell (1963): Mathematical Models in the Social Sciences. Blaisdell Publ., New York [u.a.], 2. print edn.

Kemeny, J. G. (1959): "Mathematics without Numbers," Daedalus, 88(4), pp. 577–591.

Kenyon-Mathieu, C., and W. Schudy (2007): "How to Rank with Few Errors," in Proceedings of the thirty-ninth annual ACM symposium on theory of computing, STOC '07, pp. 95–103, New York, NY, USA. ACM.

Lentz, R. (2010): "Sorting by Search Intensity," Journal of Economic Theory, 145(4), 1436–1452.

Lentz, R., and D. T. Mortensen (2010): "Labor Market Models of Worker and Firm Heterogeneity," Annual Review of Economics, 2(1), 577–602.

Lise, J., C. Meghir, and J.-M. Robin (2011): "Matcing, Sorting, and Wages," Working paper, University Colege London.

Lopes de Melo, R. (2009): "Sorting in the Labor Market: Theory and Measurement," Working paper, University of Chicago.

Nöldeke, G., and T. Tröger (2009): "Matching Heterogeneous Agents with a Linear Search Technology," Working paper, Basel and Mannheim.

Shimer, R., and L. Smith (2000): "Assortative Matching and Search," Econometrica, 68(2), 343–370.

Sobol, I. M. (1993): "Sensitivity Analysis for Non-Linear Mathematical Models," Mathematical Modeling & Computational Experiment (Engl. Transl.), 1, 407–414.

Teulings, C. N., and P. A. Gautier (2004): "The Right Man for the Job," Review of Economic Studies, 71(2), 553–80.

I Proofs and Derivations

I.1 Derivation of value functions

We derive workers' value functions only since the functions for firms follow by symmetry.

An unemployed worker consumes b, and moves into employment only if he obtains a meeting with a firm in his acceptance set, and does not face immediate match destruction. Any breakdown in this sequence leaves the worker unemployed again in the next period.

$$
\begin{aligned}
V_u(x) = b + &\underbrace{\beta(1-\delta)\mathbb{M}_u \int\limits_{B^w(x)} \frac{d_v(\tilde{y})}{V} V_e(x,\tilde{y})\ \mathrm{d}\tilde{y}}_{\text{successful matching}} \\
&+ \underbrace{\beta\delta V_u(x)}_{\text{destruction}} + \underbrace{\beta(1-\delta)(1-\mathbb{M}_u)V_u(x)}_{\text{no meeting}} \\
&+ \underbrace{\beta(1-\delta)\mathbb{M}_u V_u(x) \int\limits_{\overline{B^w(x)}} \frac{d_v(\tilde{y})}{V}\ \mathrm{d}\tilde{y}}_{\text{meet unacceptable firm}}
\end{aligned}
$$

To express the continuation value from successful matching in terms of surplus, subtract $V_u(x)$ from the integrand and add it back to rebalance the equation. Then, use (9) obtain

$$
\begin{aligned}
V_u(x) = b + &\beta\alpha(1-\delta)\mathbb{M}_u \int\limits_{B^w(x)} \frac{d_v(\tilde{y})}{V} S(x,\tilde{y})\ \mathrm{d}\tilde{y} \\
&+ \beta\delta V_u(x) + \beta(1-\delta)(1-\mathbb{M}_u)V_u(x) \\
&+ \beta(1-\delta)\mathbb{M}_u V_u(x) \left[\int\limits_{\overline{B^w(x)}} \frac{d_v(\tilde{y})}{V}\ \mathrm{d}\tilde{y} + \int\limits_{B^w(x)} \frac{d_v(\tilde{y})}{V}\ \mathrm{d}\tilde{y} \right]
\end{aligned}
$$

where terms cancel to give (11).

A worker who is employed consumes $w(x,y)$, and stays employed if he escapes destruction and otherwise moves into unemployment. Minor rearranging and (9) gives us (13).

$$\begin{aligned} V_e(x,y) &= w(x,y) + \beta\delta V_u(x) + \beta(1-\delta)V_e(x,y) \\ &= w(x,y) + \beta\delta V_u(x) + \beta\alpha(1-\delta)S(x,y) + \beta(1-\delta)V_u(x) \\ &= w(x,y) + \beta V_u(x) + \beta\alpha(1-\delta)S(x,y) \end{aligned}$$

I.2 Proofs of Results in Section 3.1

Proof of Result 1. Adding (13) and (14) yields:

$$V_e(x,y) + V_p(x,y) = f(x,y) + \beta V_v(y) + \beta V_u(x) + \beta(1-\delta)S(x,y),$$

and equivalently

$$V_e(x,y) - V_u(x) + V_p(x,y) - V_v(y) = f(x,y) + (\beta-1)V_v(y) + (\beta-1)V_u(x) + \beta(1-\delta)S(x,y),$$

so that, using (9), gives

$$S(x,y)(1-\beta(1-\delta)) = f(x,y) + (\beta-1)V_v(y) + (\beta-1)V_u(x)$$

and thus surplus equals

$$S(x,y) = \frac{f(x,y) + (\beta-1)V_v(y) + (\beta-1)V_u(x)}{1-\beta(1-\delta)}. \tag{A1}$$

Using (13) again gives us wages[14]

$$\begin{aligned} w(x,y) &= S(x,y)\alpha(1-\beta(1-\delta)) + (1-\beta)V_u(x) \\ &= \alpha f(x,y) + \alpha(\beta-1)V_v(y) + (1-\alpha)(1-\beta)V_u(x). \end{aligned} \tag{A2}$$

[14] Using (14) gives the wage as

$$\begin{aligned} w(x,y) &= f(x,y) - S(x,y)(1-\alpha)(1-\beta(1-\delta)) + (\beta-1)V_v(y) \\ &= f(x,y) - (1-\alpha)f(x,y) - (1-\alpha)(\beta-1)V_u(x) + \alpha(\beta-1)V_v(y) \\ &= \alpha f(x,y) + (1-\alpha)(1-\beta)V_u(x) + \alpha(\beta-1)V_v(y), \end{aligned}$$

which is unsurprising.

We now establish that $V_u(x)$ is increasing in x. From (11),

$$V_u(x)(1-\beta) \;=\; b+\beta\alpha(1-\delta)\mathbb{M}_u \int\limits_{B^w(x)} \frac{d_v(\tilde{y})}{V} S(x,\tilde{y})\ \mathrm{d}\tilde{y}$$

so that

$$\frac{\partial V_u(x)}{\partial x}(1-\beta) \;=\; \beta\alpha(1-\delta)\mathbb{M}_u \int\limits_{B^w(x)} \frac{d_v(\tilde{y})}{V} \frac{\partial S(x,\tilde{y})}{\partial x}\ \mathrm{d}\tilde{y}$$

keeping in mind that $S(x,y)=0$ at the boundaries. As a result, we have using (A1), that

$$\frac{\partial V_u(x)}{\partial x}(1-\beta) \;=\; \frac{\beta\alpha(1-\delta)\mathbb{M}_u}{1-\beta(1-\delta)} \int\limits_{B^w(x)} \frac{d_v(\tilde{y})}{V} \frac{\partial f(x,\tilde{y})+(\beta-1)V_u(x)}{\partial x}\ \mathrm{d}\tilde{y}$$

Solving for $\frac{\partial V_u(x)}{\partial x}$ yields

$$\frac{\partial V_u(x)}{\partial x}\left(1-\beta+\frac{(1-\beta)\beta\alpha(1-\delta)\mathbb{M}_u}{1-\beta(1-\delta)}\int\limits_{B^w(x)} \frac{d_v(\tilde{y})}{V}\ \mathrm{d}\tilde{y}\right) = \frac{\beta\alpha(1-\delta)\mathbb{M}_u}{1-\beta(1-\delta)}\int\limits_{B^w(x)} \frac{d_v(\tilde{y})}{V}\frac{\partial f(x,\tilde{y})}{\partial x}\ \mathrm{d}\tilde{y} \tag{A3}$$

and thus $\frac{\partial V_u(x)}{\partial x}>0$ since $\frac{\partial f(x,y)}{\partial x}>0$.

To show that $w(x,y)$ is increasing x, we differentiate (A2):

$$\frac{\partial w(x,y)}{\partial x} = \alpha\frac{\partial f(x,y)}{\partial x} + (1-\alpha)(1-\beta)\frac{\partial V_u(x)}{\partial x}, \tag{A4}$$

which is positive because $\frac{\partial f(x,y)}{\partial x}>0$ and $\frac{\partial V_u(x)}{\partial x}>0$.

Finally, we show that $V_e(x,y)$ is increasing in x as well. We have

$$V_e(x,y) = w(x,y) + \beta\delta V_u(x) + \beta(1-\delta)V_e(x,y)$$

and thus that

$$V_e(x,y)(1-\beta(1-\delta)) = w(x,y) + \beta\delta V_u(x),$$

which is increasing in x since $\frac{\partial w(x,y)}{\partial x}>0$ and $\frac{\partial V_u(x)}{\partial x}>0$. ■

Proof of Result 2. Let $\tilde{y}(x)$ be a firm type such that worker x is indifferent between matching with this firm and staying unemployed,

$$V_e(x, \tilde{y}(x)) = V_u(x).$$

$\tilde{y}(x)$ is the firm that pays the reservation wage to worker of type x. Then (13) can be written as

$$V_e(x, \tilde{y}(x)) \quad = \quad w(x, \tilde{y}(x)) + \beta V_u(x)$$

so that

$$w(x, \tilde{y}(x)) = V_e(x, \tilde{y}(x)) - \beta V_u(x) = (1-\beta)V_u(x).$$

which from Result 1 is increasing in x. ■

Proof of Result 3. The maximum wage given by $w(x, y^{max}(x))$. Taking derivatives w.r.t. x yields

$$\frac{\partial w(x, y^{max}(x))}{\partial x} = w_x(x, y^{max}(x)) + w_y(x, y^{max}(x))y_x^{max}(x) = w_x(x, y^{max}(x)) > 0.$$

Proof of Result 4. Assume that the matching sets are intervals[15]

$$B^w(x) = [\underline{\varphi}(x), \overline{\varphi}(x)].$$

First rewrite the adjusted average wage as

$$w^{av}(x) = w(x, \tilde{y}(x)) + \mathbb{M}_u(1-\delta) \int\limits_{B^w(x)} \frac{d_v(y)}{V} \left[w(x,y) - w(x, \tilde{y}(x))\right] \, \mathrm{d}y.$$

Take derivatives with respect to x:

$$\begin{aligned}
\frac{\partial w^{av}(x)}{\partial x} \quad &= \quad \frac{\partial w(x, \tilde{y}(x))}{\partial x} + \mathbb{M}_u(1-\delta) \int\limits_{B^w(x)} \frac{\partial w(x,y) - w(x, \tilde{y}(x))}{\partial x} \frac{d_v(y)}{V} \, \mathrm{d}y \\
&+ \quad \mathbb{M}_u(1-\delta)\overline{\varphi}'(x) \frac{d_v(\overline{\varphi}(x))}{V} \left[w(x, \overline{\varphi}(x)) - w(x, \tilde{y}(x))\right] \\
&- \quad \mathbb{M}_u(1-\delta)\underline{\varphi}'(x) \frac{d_v(\underline{\varphi}(x))}{V} \left[w(x, \underline{\varphi}(x)) - w(x, \tilde{y}(x))\right]
\end{aligned}$$

[15]This is a result in a symmetric environment of Shimer and Smith (2000). Our proofs go through without this assumption and we make it occasionally throughout this paper only for ease of exposition.

The last two terms go to zero as $w(x, \overline{\varphi}(x)) = w(x, \underline{\varphi}(x)) = w(x, \tilde{y}(x))$. Now simply rewrite

$$\begin{aligned}\frac{\partial w^{av}(x)}{\partial x} &= \frac{\partial w(x, \tilde{y}(x))}{\partial x}\left[1 - \mathbb{M}_u + \delta\mathbb{M}_u + \mathbb{M}_u(1-\delta)\int\limits_{\overline{B^w(x)}} \frac{d_v(y)}{V}\,\mathrm{d}y\right] \\ &+ \mathbb{M}_u(1-\delta)\int\limits_{B^w(x)} \frac{\partial w(x,y)}{\partial x}\frac{d_v(y)}{V}\,\mathrm{d}y\end{aligned}$$

to see that $\frac{\partial w^{av}(x)}{\partial x} > 0$. ■

I.3 Proofs of Results in Section 3.2

Proof of Result 5. For the value of a vacancy we have that

$$V_v(y)(1-\beta) = -c + \beta(1-\alpha)(1-\delta)\mathbb{M}_v \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U} S(\tilde{x}, y)\,\mathrm{d}\tilde{x},$$

so that

$$\frac{\partial V_v(y)}{\partial y}(1-\beta) = \beta(1-\alpha)(1-\delta)\mathbb{M}_v \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U} \frac{\frac{\partial f(\tilde{x},y) + (\beta-1)V_v(y)}{\partial y}}{1-\beta(1-\delta)}\,\mathrm{d}\tilde{x}$$

and thus that

$$\frac{\partial V_v(y)}{\partial y}\Big(1 - \beta + \frac{(1-\beta)\beta(1-\alpha)(1-\delta)\mathbb{M}_v}{1-\beta(1-\delta)}\int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U}\,\mathrm{d}\tilde{x}\Big) = \beta(1-\alpha)(1-\delta)\mathbb{M}_v \int\limits_{B^f(y)} \frac{d_u(\tilde{x})}{U}\frac{\frac{\partial f(\tilde{x},y)}{\partial y}}{1-\beta(1-\delta)}\,\mathrm{d}\tilde{x} > 0,$$

so that $\frac{\partial V_v(y)}{\partial y} > 0$ since the coefficient multiplying it is positive. Finally we show that the value of a filled job for a firm is increasing in y. We have that

$$\begin{aligned} V_p(x,y) &= f(x,y) - w(x,y) + \beta V_v(y) + \beta(1-\alpha)(1-\delta)S(x,y) \\ &= f(x,y)(1-\alpha) - (1-\alpha)(1-\beta)V_u(x) + \alpha(1-\beta)V_v(y) + \beta V_v(y) + \beta(1-\delta)(V_p(x,y) - V_v(y)),\end{aligned}$$

so that

$$V_p(x,y)(1-\beta(1-\delta)) = f(x,y)(1-\alpha) - (1-\alpha)(1-\beta)V_u(x) + V_v(y)(\beta\delta + \alpha(1-\beta)).$$

and

$$\frac{\partial V_p(x,y)}{\partial y}(1-\beta(1-\delta)) \quad = \quad \frac{\partial f(x,y)}{\partial y}(1-\alpha) + \frac{\partial V_v(y)}{\partial y}(\beta\delta + \alpha(1-\beta)) > 0. \ \blacksquare$$

Proof of Result 8. For now we assume (can be generalized) that the matching sets are intervals

$$B_u(y) = [\underline{\varphi}(y), \overline{\varphi}(y)].$$

If we have PAM then both $\underline{\varphi}(y)$ and $\overline{\varphi}(y)$ are increasing and if we have NAM then both $\underline{\varphi}(y)$ and $\overline{\varphi}(y)$ are decreasing. We then have

$$\Theta(y) \quad = \int\limits_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} \frac{d_u(\tilde{x})}{\int_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} d_u(\tilde{x}) \, \mathrm{d}\tilde{x}} \hat{w}(\tilde{x}) \, \mathrm{d}\tilde{x}$$

and thus that

$$\begin{aligned}
\frac{\partial \Theta(y)}{\partial y} \quad &= \frac{d_u(\overline{\varphi}(y))\hat{w}(\overline{\varphi}(y))\overline{\varphi}'(y) - d_u(\underline{\varphi}(y))\hat{w}(\underline{\varphi}(y))\underline{\varphi}'(y)}{\int_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} d_u(\tilde{x}) \, \mathrm{d}\tilde{x}} \\
&- \frac{\int_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} d_u(\tilde{x})\hat{w}(\tilde{x}) \, \mathrm{d}\tilde{x}[d_u(\overline{\varphi}(y))\overline{\varphi}'(y) - d_u(\underline{\varphi}(y))\underline{\varphi}'(y)]}{(\int_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} d_u(\tilde{x}) \, \mathrm{d}\tilde{x})^2} \\
&= \frac{d_u(\overline{\varphi}(y))\overline{\varphi}'(y)[\hat{w}(\overline{\varphi}(y)) - \int_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} \frac{d_u(\tilde{x})}{\int_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} d_u(\tilde{x}) \, \mathrm{d}\tilde{x}} \hat{w}(\tilde{x}) \, \mathrm{d}\tilde{x}]}{(\int_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} d_u(\tilde{x}) \, \mathrm{d}\tilde{x})} \\
&- \frac{d_u(\underline{\varphi}(y))\underline{\varphi}'(y)[\hat{w}(\underline{\varphi}(y)) - \int_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} \frac{d_u(\tilde{x})}{\int_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} d_u(\tilde{x}) \, \mathrm{d}\tilde{x}} \hat{w}(\tilde{x}) \, \mathrm{d}\tilde{x}]}{(\int_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} d_u(\tilde{x}) \, \mathrm{d}\tilde{x})}
\end{aligned}$$

Since $\underline{\varphi}'(y), \overline{\varphi}'(y) > 0$ if PAM, $\underline{\varphi}'(y), \overline{\varphi}'(y) < 0$ if NAM, and since $\hat{w}$ is increasing,

$$\hat{w}(\overline{\varphi}(y)) - \int_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} \frac{d_u(\tilde{x})}{\int_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} d_u(\tilde{x}) \, \mathrm{d}\tilde{x}} \hat{w}(\tilde{x}) \, \mathrm{d}\tilde{x} > 0$$

and

$$\hat{w}(\underline{\varphi}(y)) - \int_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} \frac{d_u(\tilde{x})}{\int_{\underline{\varphi}(y)}^{\overline{\varphi}(y)} d_u(\tilde{x}) \, \mathrm{d}\tilde{x}} \hat{w}(\tilde{x}) \, \mathrm{d}\tilde{x} < 0 \quad \blacksquare$$

Result A-1. *Our estimate of* $\Theta(j)$ *is*

$$\hat{\Theta}_t(j) = \sum_{\{i \text{ employed at } j \text{ at } t\}} \frac{\hat{w}(i)}{E_t(j)},$$

where $E_t(j)$ *is the number of workers employed at firm* j *at time* t. $\hat{w}$ *is the function increasing in the type of worker* i *(if worker* i *is of type* x *then* $\hat{w}(i) = \hat{w}(x)$.

Proof of Result A-1.

From the law of large numbers, we obtain the equivalent of this weighted average as

$$\hat{\Theta}_t(j) = \int_{B_u(y(j))} \frac{d_u(x)}{\int_{B_u(y(j))} d_u(\tilde{x})\,\mathrm{d}\tilde{x}} \hat{w}(x)\,\mathrm{d}x.$$

From the definition of $q(j)$, Result A-1 follows. ■

Proof of Result 7. (13) written for worker x matched with firm y or $\tilde{y}(x)$ becomes

$$V_e(x,y) = w(x,y) + \beta V_u(x) + \beta(1-\delta)\left(V_e(x,y) - V_u(x)\right)$$
$$V_e(x,\tilde{y}(x)) = w(x,\tilde{y}(x)) + \beta V_u(x) + \beta(1-\delta)\left(V_e(x,\tilde{y}(x)) - V_u(x)\right)$$

Differencing, we get that wages follow

$$w(x,y) - w(x,\tilde{y}(x)) \quad = \quad (1-\beta(1-\delta))\left((V_e(x,y) - V_e(x,\tilde{y}(x)))\right).$$

Finally, by integrating and multiplying both sides by $(1-\delta)\mathbb{M}_v$ we obtain

$$(1-\delta)\mathbb{M}_v \int_{B^f(y)} \frac{d_u(x)}{U}(w(x,y) - w(x,\tilde{y}(x)))\,\mathrm{d}x$$
$$= (1-\beta(1-\delta))(1-\delta)\mathbb{M}_v \int_{B^f(y)} \frac{d_u(x)}{U}(V_e(x,y) - V_u(x))\,\mathrm{d}x$$

which by Result 6 is increasing in y. ■

I.4 Measuring α in the data

I.4.1 Using Business Cycles to Measure α

To measure the bargaining power α in the data we consider now an extended version of the model with business cycles, i.e. exogenous changes in aggregate productivity z. The output of a pair (x, y) is then $zf(x, y)$. Consider two worker types x and x' (have to be different types, working at firm y when productivity is z and when it is $\hat{z}$. The wages of worker x in the two business cycle states are $w(x, y, z)$ and $w(x, y, \hat{z})$, respectively. For worker x' the corresponding wages are $w(x', y, z)$ and $w(x', y, \hat{z})$. These wages are observed. The equation for wages with business cycles is straightforward and follows the same arguments as the one without business cycles. For the value of a job it holds with the obvious notation

$$V_e(x, y, z) = w(x, y, z) + \beta E(V_u(x, z') \mid z) + \beta\alpha(1-\delta)E(S(x, y, z') \mid z) \tag{A5}$$

and for the value of a filled vacancy,

$$V_p(x, y, z) = zf(x, y) - w(x, y, z) + \beta E(V_v(y, z') \mid z) + \beta(1-\alpha)(1-\delta)E(S(x, y, z') \mid z) \tag{A6}$$

Adding up these two Bellman equations yields:

$$\begin{aligned} V_e(x, y, z) + V_p(x, y, z) &= zf(x, y) + \beta E(V_v(y, z') \mid z) + \beta E(V_u(x, z') \mid z) \\ &\quad +\beta(1-\delta)E(S(x, y, z') \mid z), \end{aligned} \tag{A7}$$

and equivalently

$$\begin{aligned} S(x, y, z) &= V_e(x, y, z) - V_u(x, z) + V_p(x, y, z) - V_v(y, z) \\ &= zf(x, y) - V_v(y, z) - V_u(x, z) + \beta E(V_v(y, z') \mid z) + \beta E(V_u(x, z') \mid z) \\ &\quad +\beta(1-\delta)E(S(x, y, z') \mid z), \end{aligned} \tag{A8}$$

Motivated by the observation that productivity basically follows a random walk, we now make the approximation that

$$E(S(x, y, z') \mid z) = S(x, y, z) + \text{expectational error}, \tag{A9}$$

so that the surplus equals

$$
\begin{aligned}
S(x,y,z)(1-\beta(1-\delta)) &= zf(x,y) - V_v(y,z) - V_u(x,z) + \beta E(V_v(y,z') \mid z) \\
&\quad + \beta E(V_u(x,z') \mid z).
\end{aligned} \tag{A10}
$$

Using the Bellman equation for V_e and the approximation we can solve for wages:

$$
w(x,y,z) = \alpha S(x,y,p)(1-\beta(1-\delta)) + V_e(x,z) - \beta E(V_u(x,z') \mid z) \tag{A11}
$$

Making the same approximation for V_u,

$$
E(V_u(x,z') \mid z) = V_u(x,z) + \text{expectational error}, \tag{A12}
$$

and using the equation for the surplus S,

$$
\begin{aligned}
w(x,y,z) &= \alpha(zf(x,y) - V_v(y,z) - (1-\beta)V_u(x,z) + \beta E(V_v(y,z') \mid z)) \\
&\quad + V_u(x,z)(1-\beta) \\
&= \alpha z f(x,y) + \alpha(\beta E(V_v(y,z') \mid z) - V_v(y,z)) + (1-\alpha)(1-\beta)V_u(x,z).
\end{aligned} \tag{A13}
$$

The differences in wages for types x and x' equals

$$
\begin{aligned}
& w(x',y,z) - w(x,y,z) \\
= \; & \alpha z(f(x',y) - f(x,y)) + (1-\alpha)(1-\beta)(V_u(x',z) - V_u(x,z))
\end{aligned}
$$

To figure out α we have to measure $V_u(x,z)$ and $V_u(x',z)$ in the data. For this we use the Bellman equation for V_e and the approximation for the expected surplus

$$
\begin{aligned}
V_u(x,z) = V_e(x,\underline{y}(x,z),z) &= w(x,\underline{y}(x,z),z) + \beta E(V_u(x,z') \mid z) + \beta\alpha(1-\delta)E(S(x,\underline{y}(x,z),z') \mid z) \\
&= w(x,\underline{y}(x,z),z) + \beta V_u(x,z) + \beta\alpha(1-\delta)S(x,\underline{y}(x,z),z) \\
&= w(x,\underline{y}(x,z),z) + \beta V_u(x,z),
\end{aligned}
$$

so that

$$
V_u(x,z)(1-\beta) = w(x,\underline{y}(x,z),z),
$$

i.e. we measure the value of employment at the lowest firm at productivity level z through the lowest wage accepted by type x at level z. Using this expression for the reservation wage in the wage equation to substitute for the value of unemployment, yields

$$\begin{aligned} & w(x', y, z) - w(x, y, z) \\ = \; & \alpha z(f(x', y) - f(x, y)) + (1-\alpha)(1-\beta)(V_u(x', z) - V_u(x, z)) \\ = \; & \alpha z(f(x', y) - f(x, y)) + (1-\alpha)(w(x', \underline{y}(x', z), z) - w(x, \underline{y}(x, z), z)) \end{aligned}$$

For the empirical implementation define then dummies $\delta_{x,y}$ which is one if worker type x works at firm type y and zero otherwise. We then regress

$$w_t(x') - w_t(x) = z_t(\delta_{x',y} - \delta_{x,y}) + \kappa(w(x', \underline{y}(x', z), z) - w(x, \underline{y}(x, z), z)).$$

The estimated value of κ is then our estimate of $1-\alpha$ so that $\hat{\alpha} = (1-\kappa)$.

I.4.2 Measure α from fluctuation in profits

To measure the bargaining power α in the data we consider now an extended version of the model with i.i.d. shocks to the firm's technology. In response to such a shock to the firm's technology than profits increase by $(1-\alpha)$ and wages increase by α. To measure α using this experiment we use the results in Blanchflower, Oswald, and Sanfey (1996) who estimate the response of wages to a change in profits. The value of the bargaining power α is then chosen to match the estimated response of wages and profit.

II Computation and Implementation

We discretize the type space for both workers and firms with 50 evenly distributed grid points on $[0, 1]$. From an initial match distribution, $d_{m,0}(x, y) = 0.5 \;\; \forall(x, y)$, and initial surplus, $S_0(x, y) = f(x, y) \;\; \forall(x, y)$,we obtain a solution by alternatively updating exactly once on either match density (equation 15) or surplus (summing equations 11 - 14). When $d_{m,k}(x, y) < 10^{-6}$, we set $d_{m,k}(x, y) = 0$. A solution is found if the maximum absolute difference between iterations of both surplus and match density is less than 10^{-12}.

If no solution admitting a pure acceptance strategy is found (due to discretization), we solve for a mixed strategy; i.e. unemployed agents accept matches with an interior probability such that the surplus of the match is positive, but very close to zero. Denote iteration k of the acceptance strategy of workers with $A^w_k(x,y)$. $A^w_k(x,y)$ is the probability worker x accepts a job from firm y. We then update the acceptance strategy in the following way.

Mixed strategy

if $S_k(x,y) > 5 \times 10^{-7}$ and $A^w_k(x,y) < 1$

$A^w_{k+1}(x,y) = A^w_k(x,y) + 0.001 \cdot rand() \cdot (1 - A^w_k(x,y))$

elseif $S_k(x,y) < 0.50$ and $A^w_k(x,y) > 0$

$A^w_{k+1}(x,y) = A^w_k(x,y) - 0.001 \cdot rand() \cdot (1 - A^w_k(x,y))$

end

A mixed solution is found if the maximum absolute change between iterations of both the surplus and the match density are less than 2.5×10^{-7}. We find a mixed strategy solution in all parameterizations that we use.

With the computed solution, we simulate 600 workers and 600 jobs for each grid point giving 60000 agents (30000 workers and 30000 jobs) over a period of 240 months with an initial burn-in of 100 months. This corresponds to 20 years of monthly data.

Where order is meaningful (e.g. ranks, types or bins), higher numbers represent better workers; e.g. a worker with rank 10 is better than a worker with rank 0, a firm in bin 7 is estimated to be better than a worker in bin 3.

Here, we define quantities that we will use to sketch the procedures we use.

i) #workers = #jobs = $N = 30000$.

ii) #worker types = X = #firm types = $Y = 50$.

iii) Worker ID, $i = 1..N$.

iv) Rank of worker i, $\hat{i} = 1..N$.

E.g. $i = 4$ has rank 10; $\hat{i}(4) = 10$.

v) True worker type $x = 1..X$. Each x has N/X individual workers.

E.g. $i = 6$ has type 3; $x(6) = 3$. For convenience, $x(i) = 1$ if $i \in \{1..600\}$, $x(i) = 2$ if $i \in \{601..1200\}$ and so on. We do not use this information anywhere in the code.

vi) Estimated worker type (worker bin) $\hat{x} = 1..X$. Each $\hat{x}$ has N/X workers.

E.g. $i = 5$ is in bin 45; $\hat{x}(5) = 45$.

For our simulations, $\hat{x}(i) = 1$ if $\hat{i} \in \{1..600\}$, $\hat{x}(i) = 2$ if $\hat{i} \in \{601..1200\}$ and so on.

vii) Firm ID, $j = 1..J$. $J = N/100$. Jobs and vacancies sum to 100 at all j.

viii) Rank of firm j, $\hat{j} = 1..J$.

E.g. $j = 4$ has rank 10; $\hat{j}(4) = 10$.

ix) True firm type, $y = 1..Y$. Each y has $N/(100 \cdot Y)$ unique j's.

E.g. $j = 4$ has type 10; $y(4) = 10$.

x) Estimated firm type (firm bin) $\hat{y} = 1..Y$. Each $\hat{y}$ has $N/(100 \cdot Y)$ unique j's.

E.g. $j = 4$ is in bin 10; $\hat{y}(4) = 10$.

First we take simulated matched employer-employee datasets and rank workers using the algorithm described in Section III. The algorithm deliver $\hat{i}(i)$ and $\hat{x}(i)$. At all firms j, we observe i's and hence have an estimate of the matching set of firm j, $\hat{B}^f(j)$. Using **IDNoise** we locate matches that are likely caused by very noisy wage histories, and exclude the wage histories of workers belong to such matches from all estimations. The fraction of workers excluded is small (less than 5%) which we denote $\hat{\mathcal{N}}$. This algorithm also gives us the estimate of the matching set $\hat{\mathcal{B}}$.

Algorithm 1. ***IDNoise***$[\hat{x}(i)] \implies [\hat{\mathcal{B}}, \hat{\mathcal{N}}]$

Construct $p(\hat{x}, j), \pi(\hat{x}, j)$ *and* $N(j)$.[16]

for each j

Compute $F(p(\hat{x}, j); N(j), \pi(\hat{x}, j))$.[17]

$\forall \hat{x}$, *Initialize* $Acc(\hat{x}, j) = 1$ *iff* $p(\hat{x}, j) > 0$.

****for*** $\hat{x}$ *with* $Acc(\hat{x}, j) = 1$

if $\hat{x} \in \{1, X\}$ *and* $F(p(\hat{x}, j); N(j), \pi(\hat{x}, j)) < 0.1$

Set $Acc(\hat{x}, j) = 0$.

Return to $*$.

else

if $(Acc(\hat{x} + 1, j) = 0 \quad | \quad Acc(\hat{x} - 1, j) = 0)$

if $F(p(\hat{x}, j); N(j), \pi(\hat{x}, j)) < 0.1$

Set $Acc(\hat{x}, j) = 0$.

Return to $*$.

[16] $p(\hat{x}, j)$ is the number of employment spells from bin $\hat{x}$ by j. $N(j) = \sum_{\hat{x}} p(\hat{x}, j)$. j's probability of accepting workers in $\hat{x}$

$$\pi(\hat{x}, j) = \frac{u(\hat{x})\mathbb{1}\{p(\hat{x}, j) > 0\}}{\sum_{\hat{x}} u(\hat{x}) \cdot \mathbb{1}\{p(\hat{x}, j) > 0\}}.$$

[17] The probability of observing at most $p(\hat{x}, j)$ given the hiring probability $\pi(\hat{x}, j)$ from $N(j)$ trials is

$$F(p(\hat{x}, j); N(j), \pi(\hat{x}, j)) = \sum_{i=0}^{p(\hat{x}, j)} \binom{N(j)}{i} \pi(\hat{x}, j)^i (1 - \pi(\hat{x}, j))^{N(j)-i}. \tag{A14}$$

__end__

__end__

__end__

__end__

__end__

Construct $\hat{\mathcal{B}} = 1$ *iff* $Acc(\hat{x}, j) = 1$

$i \in \hat{\mathcal{N}} = 1$ *if* i *belongs to any match with* $Acc(\hat{x}, j) = 0$.

return $[\hat{\mathcal{B}}, \hat{\mathcal{N}}]$

The next crucial statistic to estimate is reservation wages for each worker $\hat{w}_{res}(i)$. To this end, we implement **ResWage**.

Algorithm 2. *ResWage*$[w(i,j), \hat{x}(i), \hat{\mathcal{N}}] \implies \hat{w}_{res}(i)$

Consider wages histories of $i \notin \hat{\mathcal{N}}$.

for $\hat{x}$

Construct $J = \{j$ *s.t.* j *hires any* $i \in \hat{x}\}$.

Compute $\bar{w}(\hat{x}, j) =$ *average wage paid by* j *to* $i \in \hat{x}$ *for all* j.

$w_{res}(\hat{x}) =$ *lowest average of* $\bar{w}(\hat{x}, j)$ *possible from pooling N/(100*Y) firms in J.*[18]

end

return $\hat{w}_{res}(i) = w_{res}(\hat{x}(i))$

Then, for each firm j, compute the average wage premium as in (21). We next estimate job filling rates $\hat{q}(j)$ using information from all workers whether or not they belong to $\hat{\mathcal{N}}$ over the acceptance set $\hat{\mathcal{B}}$. We can now rank firms.

We now aggregate the individual firm data into $\hat{y}$ level. For example, statistics for $\hat{y} = 1$ will be the firm size (measured by average employment) weighted average of firms with $\hat{j}(j) = \{1..6\}$. This step only serves to aggregate information across firms. In principle, the production function for each firm can be estimated in isolation, $\hat{y}(j) = \hat{j}(j)$ especially if the firm is large.

Taking present values of estimated minimum wages for each bin yields $V_u(\hat{x})$. Compute the average wages each bin $\hat{x}$ receives with all firms of bin $\hat{y}$. This is $w_{av}(\hat{x}, \hat{y})$. Compute the corresponding value of employment, $V_e(\hat{x}, \hat{y})$ and $V_v(\hat{y})$ from $\Omega(\hat{y})$. The estimate of the production function $\hat{f}(\hat{x}, \hat{y})$ follows.

Using unemployment rates at the $\hat{x}$ level and estimated firm size at the $\hat{j}$ level, we can estimate frictional output with the estimated production function. We add a constant to $\hat{f}(\hat{x}, \hat{y})$ so that the estimated output equals the frictional output from solving the model.

[18]For $\hat{x} > 1$, we additionally impose $w_{res}(\hat{x}) > w_{res}(\hat{x} - 1)$ which is consistent with theory.

To measure output losses due to frictions we optimally assign a sub-sample (5000 workers and 5000 jobs) from the pool of employed workers. The sub-sample reflects the estimated type distributions of employed workers and producing firms. To evaluate the accuracy of our method, our estimated gains from eliminating sorting is compared the same procedure repeated using true model generated distributions and production functions.

III Rank aggregation

Our goal is to rank workers according to their productivity. We know that wages within a firm are increasing in worker productivity x. Thus, in the absence of measurement error, considering the workers within one specific firm gives us a correct ranking between these workers. Repeating this ranking for every firm yields a globally consistent and, if workers are sufficiently mobile between firms, a complete ranking of workers since worker rankings are transitive across firms.

However, data are usually ridden by substantial measurement error. Consequently, within one firm, a less productive worker could be ranked above a truly more productive worker because of measurement error. Furthermore, the ranking between these two workers may not be transitive across firms where they happen to be co-workers. Thus, the rankings from all firms are not consistent and thus do not yield an aggregate ranking. To solve this problem, we build on the insights from social choice theory, which considers a equivalent problem in the context of voting.

Imagine that voters were asked to rank candidates from the most to the least preferred one. Voters will rank candidates according to their own preferences but when the need to have a single ranking of candidates comes up, a disagreement is likely to arise. Unless every voter ranks all candidates identically, there will not be an aggregate ranking that all voters agree with completely. How then does one accomplish two important tasks? Firstly, we need some notion of how “good” the aggregate ranking is. Secondly, we need a method to find the ranking that is “best” given our criteria of what is “good”.

III.1 Kemeny-Young rank aggregation

Given many (perhaps) inconsistent rankings of candidates, how does one aggregate the ranks to determine who the best candidate is? This problem is ancient, and first studied by de Borda (1781) and Condorcet (1785). One natural starting point to use as a metric for evaluating the posited aggregate ranking is the number of disagreements generated in the voter submitted ranks as done in the Kemeny-Young formulation of this classic problem. The goal then is to find an aggregate ranking which generates the minimum number of disagreements with the data, or equivalently, agrees the most with the data. Drissi-Bakhkhat and Truchon (2004) argue in a context of a social choice model that the disagreements in the ranking of two alternatives should be weighted by the probability that the voters compare them correctly. Similarly, in our labor market application weighting means that the disagreements are weighted by the probabilities that a worker is ranked higher than another worker (which are computed from wage data). Fortunately the computer science literature provides algorithms to handle these weighted ranking problems as well since they can be cast as a special case of a weighted feedback arc set problem on tournaments (see for example Ailon, Charikar, and Newman (2008)).

For a candidate ranking Π, $\Pi_{ij} = 1$ if i is ranked higher than j and $\Pi_{ij} = 0$. There are no ties. The objective is to rank Π which maximizes

$$\sum_{i>j} c(i,j)\Pi(i,j) + c(j,i)\Pi(j,i),$$

where the weighting c_{ij} is the probability (computed from wage observations) that i is ranked above j.

We now construct c_{ij}. First, we use head-to-head wage information at all firms to calculate the probability that worker i is ranked higher than worker j. Note, that we can only use this ranking when we observe worker i and worker j at the same firm. We first discuss the simple case where we only observe i and j at one firm.

Suppose we observe n_{i,f_k} wage observations and n_{j,f_k} from workers i and j respectively at firm k. We know that observed wages follows:

$$\hat{w}_{i,f_k,t} = w_{i,f_k} + \epsilon_t$$

contains noise with variance σ^2. We can compute the average wages $\bar{w}_{i,f_k,t}$ and $\bar{w}_{j,f_k,t}$, which can be written as:

$$\begin{aligned}
\bar{w}_{i,f_k,t} - \bar{w}_{j,f_k,t} &= \frac{1}{n_{i,f_k}} \sum_{t=1}^{n_{i,f_k}} \hat{w}_{i,f_k,t} - \frac{1}{n_{j,f_k}} \sum_{t=1}^{n_{j,f_k}} \hat{w}_{j,f_k,t} \\
&= w_{i,f_k} - w_{j,f_k} + \frac{1}{n_{i,f_k}} \sum_{t=1}^{n_{i,f_k}} \epsilon_{i,f_k,t} - \frac{1}{n_{j,f_k}} \sum_{t=1}^{n_{j,f_k}} \epsilon_{j,f_k,t},
\end{aligned}$$

where all of the ϵ's are independent.

We are interested in computing the probability that $w_{i,f_k} > w_{j,f_k}$ given the observations on $\bar{w}_{i,f_k,t}$ and $\bar{w}_{j,f_k,t}$. A Bayesian approach seems a natural one to follow to accomplish this. Suppose that we had a normal prior distribution over the wages, that is we assume that:

$$w_{i,f_k} \sim \mathcal{N}(\mu_0, \tau_0^2).$$

The posterior density over w_{i,f_k} conditional on knowing σ^2 (we explain below how to measure it in the data) is given by:

$$p(w_{i,f_k}|\hat{w}_{i,f_k,1}, \cdots, \hat{w}_{i,f_k,n_{i,f_k}}) = p(w_{i,f_k}|\bar{w}_{i,f_k}) = \mathcal{N}(\mu_n, \tau_n^2),$$

where

$$\mu_n = \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{n_{i,f_k}}{\sigma^2}\bar{w}_{i,f_k}}{\frac{1}{\tau_0^2} + \frac{n_{i,f_k}}{\sigma^2}}$$

and

$$\frac{1}{\tau_n^2} = \frac{1}{\tau_0^2} + \frac{n_{i,f_k}}{\sigma^2}.$$

If in the baseline case we assume an uninformative prior, that is, we take $\tau_0^2 \to \infty$, this simplifies to:

$$\mu_n = \bar{w}_{i,f_k}$$

and

$$\frac{1}{\tau_n^2} = \frac{n_{i,f_k}}{\sigma^2}.$$

Then the posterior densities for w_{i,f_k}, w_{j,f_k} given the data would just be given by:

$$p(w_{i,f_k}|\bar{w}_{i,f_k}) = \mathcal{N}\left(\bar{w}_{i,f_k}, \frac{\sigma^2}{n_{i,f_k}}\right),$$
$$p(w_{j,f_k}|\bar{w}_{j,f_k}) = \mathcal{N}\left(\bar{w}_{j,f_k}, \frac{\sigma^2}{n_{j,f_k}}\right).$$

Since these posteriors are independent normals, we know that the distribution over the difference $p(w_{i,f_k} - w_{j,f_k}|\bar{w}_{i,f_k}, \bar{w}_{j,f_k})$ is simply:

$$p\left(w_{i,f_k} - w_{j,f_k}|\bar{w}_{i,f_k}, \bar{w}_{j,f_k}\right) = \mathcal{N}\left(\bar{w}_{i,f_k} - \bar{w}_{j,f_k}, \frac{\sigma_{i,f_k}^2}{n_{i,f_k}} + \frac{\sigma_{j,f_k}^2}{n_{j,f_k}}\right).$$

Thus, the posterior probability that $w_{i,f_k} > w_{j,f_k}$ can simply be computed as:

$$\mathbb{P}(w_{i,f_k} > w_{j,f_k}) = 1 - \Phi\left(\frac{0 - (\bar{w}_{i,f_k} - \bar{w}_{j,f_k})}{\frac{\sigma^2_{i,f_k}}{n_{i,f_k}} + \frac{\sigma^2_{j,f_k}}{n_{j,f_k}}}\right)$$
$$= \Phi\left(\frac{\bar{w}_{i,f_k} - \bar{w}_{j,f_k}}{\frac{\sigma^2_{i,f_k}}{n_{i,f_k}} + \frac{\sigma^2_{j,f_k}}{n_{j,f_k}}}\right).$$

If more than one firm hires both both workers i and j, we compute $\mathbb{P}(w_{i,f_k} > w_{j,f_k})$ for all those firms and assign the product of these probabilities to c_{ij}.

The variance of noise is computed from the variance of wages from all workers since we assumed that wages for any given type (x, y) is subject to the same measurement error. Again, this is easily generalized to other settings.

The solution to the problem of finding the best ranking is then conceptually trivial: (1) Enumerate all possible rankings; (2) Evaluate the weighted number of disagreements; (3) Select the ranking that gives the lowest cost. Unfortunately, the *Kemeny-Young rank aggregation* problem is NP-hard.[19] We therefore use the results in Kenyon-Mathieu and Schudy (2007). They provide a polynomial time algorithm capable of approximating the correct ranking to any degree of accuracy in theory. While a polynomial time approximation scheme is a huge improvement, this is still a very difficult problem. In our implementation, we appeal to the theoretical guarantees in Kenyon-Mathieu and Schudy (2007) but implement a simplified version of their algorithm. In particular, we implement **Step 1: Single vertex moves** from their algorithm which we reproduce below. We show using calibrated examples that our implementation works well in practice.

Algorithm 3. ***Deterministic polynomial time approximation scheme Kenyon-Mathieu and Schudy (2007)***

Given: *Fixed parameters* $\epsilon > 0$ *and* $b \in (0, 1]$.

Input: *A weighted tournament.*

Round weights to integer multiples of ϵ/n^2.

$\pi \leftarrow$ *constant factor approximation eg* ***KwikSort*** *from Ailon, Charikar, and Newman (2008).*[20]

[19]See Bartholdi, Tovey, and Trick (1989). Consider a simple case of a 1000 candidates and at least 4 submitted rankings. There are $1000 \times 999... \times 2$ combinations to consider! This is an extremely large number. Ranking millions of workers in our application is simply not feasible. We need methods to map the problem into something more manageable and consider approximation techniques.

[20]Instead of a constant factor approximation as an initialization, we use the ranking from (a) minimum wage, (b) maximum wage, (c) adjusted average wage, and, (d) average wage, which yields the least weighted disagreements.

While *some move decreases cost, do it. Types of moves:*

1. Single vertex moves.

Choose vertex x and rank j.

Take x out of the ordering and put it back in so that its rank is j.

2. Additive approximation.

Choose integers $i < j$.

Let $U = \{vertices\ with\ rank\ in[i, j]\}$.

$\pi_U \leftarrow AddApprox(U)$ with parameter $\beta(\epsilon)$.

Replace the restriction π_U of π to U by π_U.

Output: *π.*

IV Appendix Figures

Figure A-1: Non-parametric plots of selected variables of interest across all parameterizations with PAN, NAM, and NEITHER production functions.

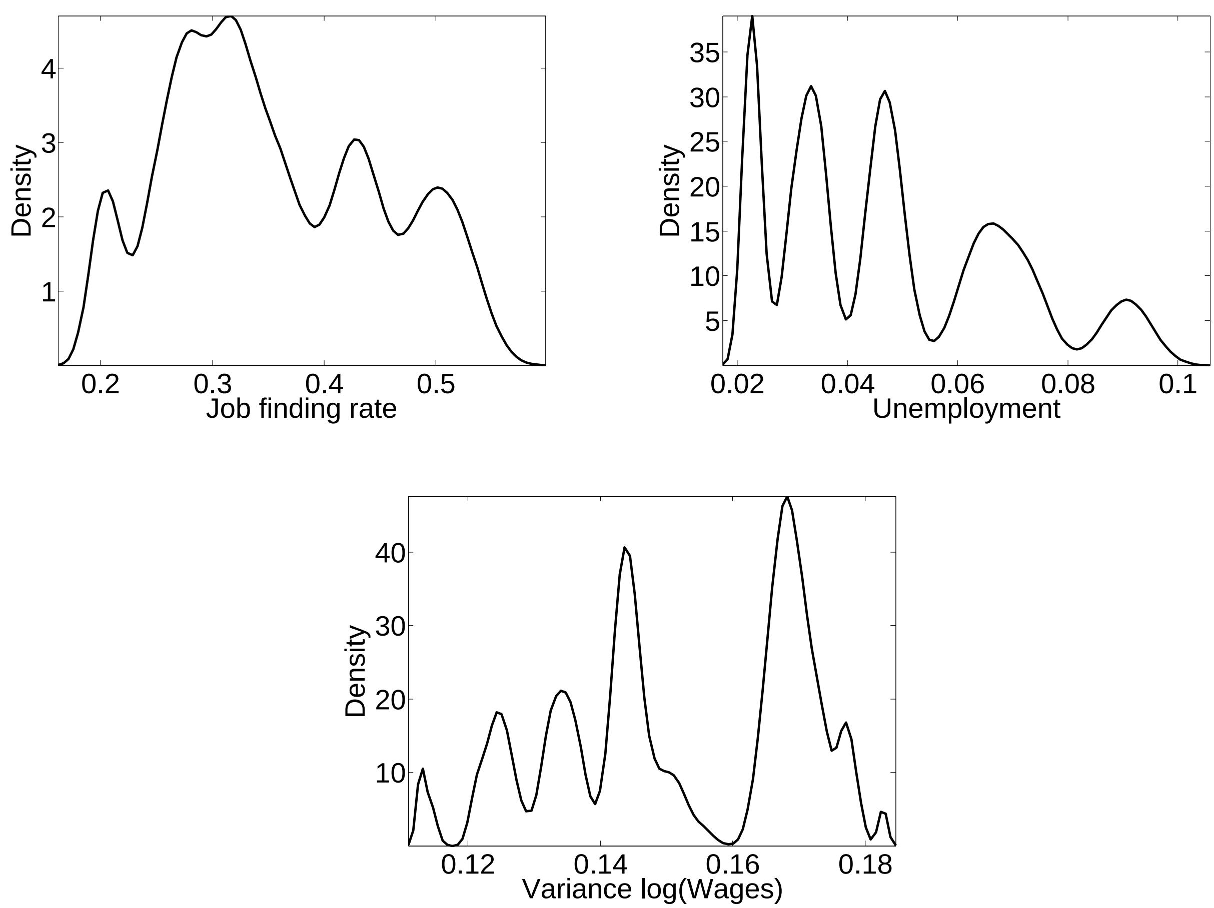


Figure A-2: Non-parametric plots of selected variables of interest across all parameterizations with Gautier and Teulings production function.