

NBER WORKING PAPER SERIES

IDENTIFYING EQUILIBRIUM MODELS OF LABOR MARKET SORTING<sup>^</sup>

Marcus Hagedorn  
Tzuo Hann Law  
Iouri Manovskii

Working Paper 18661  
<http://www.nber.org/papers/w18661>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
December 2012

We would like to thank seminar participants at Arizona State, Indiana, Mannheim, Notre Dame, UPenn, Toulouse, Vienna Institute for Advanced Studies, Bank of France, Yeshiva, Search and Matching Workshop at the Philadelphia Fed, 2012 SED Annual Meeting, 2012 NBER Summer Institute, 2012 Konstanz Workshop on Labor Market Search and the Business Cycle, 2012 Canadian Macro Study Group Meetings, 2012 Sandjberg conference, and 2012 Human Capital Conference at Washington University in St. Louis. Support from the National Science Foundation Grant No. SES-0922406 is gratefully acknowledged. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2012 by Marcus Hagedorn, Tzuo Hann Law, and Iouri Manovskii. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Identifying Equilibrium Models of Labor Market Sorting<sup>^</sup>  
Marcus Hagedorn, Tzuo Hann Law, and Iourii Manovskii  
NBER Working Paper No. 18661  
December 2012  
JEL No. C63,C78,E24,J31

### **ABSTRACT**

Does the market allocate the right workers to the right jobs? Since observable (to economists) variables account for only a small fraction of the wage variance in the data, to answer this question it is essential to study assortative matching between employers and employees based on their unobserved characteristics. This paper enables this line of research. 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 matched employer-employee data on wages and labor market transitions. 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 provide computational algorithms that implement our identification strategy given the limitations of the available data sets. Finally, we extend our identification and implementation strategies to the commonly used class of models of sorting based on comparative advantage and provide a test that discriminates between these models.

Marcus Hagedorn  
Institute for Advanced Studies  
Stumpergasse 56  
A-1060 Vienna  
Austria  
marcus.hagedorn07@gmail.com

Iourii Manovskii  
Department of Economics  
University of Pennsylvania  
160 McNeil Building  
3718 Locust Walk  
Philadelphia, PA 19104  
and NBER  
manovski@econ.upenn.edu

Tzuo Hann Law  
Department of Economics  
University of Pennsylvania  
160 McNeil Building  
3718 Locust Walk  
Philadelphia, PA 19104  
tzuohann@gmail.com

# 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 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.<sup>1</sup> The key problem is that the assumption underlying the fixed effect regression is

---

<sup>1</sup>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 a model of sorting on absolute advantage that allows for globally rankable workers and firms, the worker and firm fixed effects in wages have no relationship to underlying productivities. These theoretical insights have been confirmed quantitatively in a range of assortative matching models in Lopes de

that 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, in a seminal contribution 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.<sup>2</sup> 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 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 and labor market transitions 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).

---

Melo (2009), Lentz (2010), and Lise, Meghir, and Robin (2011), among others.

<sup>2</sup>As we discuss below, Eeckhout and Kircher (2011) show that some model objects can be identified under particular assumptions on the production function.

To make progress with this class of models, we find it beneficial to consider the fully dynamic version 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. Eeckhout and Kircher (2011) considered a simplified two period version of the model in which workers and firms meet randomly in the first period. They can either form a match or pay a search cost and be paired up with their ideal partners in a perfectly competitive labor market in the second period. There is no discounting between the two periods. These specific modeling choices simplify the theoretical analysis substantially, but, unfortunately, they also ultimately prevent the identification of the model. The assumption of no discounting, in particular, delivers that the value of a vacancy is constant, implying the impossibility to rank firms. While this assumption is generally not appealing in quantitative work, it is readily tested in the general version of the model we consider because our identification strategy 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 rankings of candidates submitted by voters in the social choice literature. These problems are extremely computationally complex (they are NP-hard) but, fortunately, the computer science literature has recently made substantial progress in designing computational algorithms that can efficiently approximate their solution. We draw on these advances in algorithm research to develop a method 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 model, wages, which are observed in the data, are a function of the output of the match as well as of two objects that our identification strategy allows to measure - the reservation wage of a worker, and the value of a vacancy. Thus, the wage equation can be solved for output as a function of three measurable variables.

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 substantially 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 standard assumption underlying the existing empirical literature, but perhaps not the best one, 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. Alternatively, 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 be 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, Teulings and Gautier (2004) and Gautier and Teulings (2006, 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 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 our identification and implementation methods to the model of sorting based on comparative advantage in Gautier and Teulings (2012), verify their 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), who add time-consuming search to Becker (1973), with slight generalizations and some modifications. In particular, we do not impose symmetry between the two sides of the market, but have workers on the one side and firms on the other; both sides with potentially different primitives. We also use a linear search technology instead of the quadratic search technology in Shimer and Smith (2000), 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 either 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  $y$ .<sup>3</sup>

Output of a match between worker  $x$  and firm  $y$  is given by the twice differentiable nonnegative production function  $f : [0, 1]^2 \rightarrow \mathbb{R}_+$ . 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  $f_y > 0$ .<sup>4</sup> 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 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 and firm types are  $F$  and  $G$ , respectively, and the “original” production function is  $\hat{f}(\hat{x}, \hat{y})$  then we transform the production function

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

and the distributions are  $F(\hat{x}) = x$ ,  $G(\hat{y}) = y$ .

We place no additional assumptions on the production function (except for mild technical conditions that ensure existence of an equilibrium). 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  $d_w = 1$

---

<sup>3</sup>This model of the firm, as simplistic as it is, represents the current state-of-the-art in this 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.

<sup>4</sup>The assumption 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 paper this assumption is only relevant for identifying rankings of workers and firms when they can be ranked. In Section 5 we show that if some agents cannot be ranked, e.g., firms in the comparative advantage model of Gautier and Teulings (2012), our identification strategy will reveal this and it will continue to recover the production function correctly.



and  $d_f = 1$ . The distribution of producing matches is described by  $d_m : [0, 1]^2 \rightarrow \mathbb{R}_+$ . 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.<sup>5</sup> Table 1 summarizes the relationships between these functions.

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

Integrating the densities from Table 1 gives the time-invariant measures of aggregate employment,  $E = \int d_e(x) dx$ , of unemployment,  $U = \int d_u(x) dx$ , of producing firms,  $P = \int d_p(y) dy$ , and vacant firms,  $V = \int d_v(y) dy$ .

### 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 cost  $c$ . 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, existing matches (including newly formed ones) are destroyed with probability  $\delta$ .<sup>6</sup>

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

<sup>6</sup>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.

## 2.2 Search and matching

Only and all unmatched agents engage in random search.<sup>7</sup> A function  $m : [0, 1] \times [0, 1] \rightarrow [0, \min(U, V)]$  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  $\mathbb{M}_u \frac{\int_Y d_v(y) dy}{V}$ . The probability for a firm to meet any worker  $x \in X \subseteq [0, 1]$  equals  $\mathbb{M}_v \frac{\int_X d_u(x) dx}{U}$ . These probabilities reflect our assumption of a linear search technology. Using the quadratic search technology in Shimer and Smith (2000) these probabilities would be  $\mathbb{M}_u \int_Y d_v(y) dy$  and  $\mathbb{M}_v \int_X d_u(x) dx$ , 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 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 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. Symmetrically for firms, the Borel measurable *acceptance set*  $A^f(y)$  is comprised of 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*  $B^w(x)$  consists of firms which accept worker type  $x$  and are accepted by worker type  $x$ . Similarly, for a firm of type  $y$ ,  $B^f(y)$  consists of workers who accept to match with firm type  $y$  and who are accepted by

---

<sup>7</sup>Random search means that workers and firms do not observe the types of their potential trading partners prior to meeting them, i.e. they have the same information as is available to the econometrician (e.g., age, sex, education, occupation, etc. of a worker and industry, location, etc. of a firm). An alternative assumption is that workers (firms) know the type  $y$  of every firm (type  $x$  of every worker) and can direct their search to specific types (e.g., Moen (1997), Shi (2001), Shimer (2005), Eeckhout and Kircher (2010)), e.g., workers direct their search to firms that are willing to accept them. In the analysis below these informational assumptions matter only for the computation of the job filling probability for firms. These informational assumptions will not affect the analysis at all if the data allows to observe the number of vacancies at individual firms (as in, e.g., the German matched employer-employee dataset provided by the IAB). In this case one can compute the job filling rate directly without the need to make any informational assumptions. Without data on vacancies, the computation of the job filling rate is conditional on the specification of the matching process.

firms of type  $y$ . Specifically,

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

$\overline{B^w}$  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:

$$\begin{aligned} \mathcal{B} &\equiv \{(x, y) : y \in A^w(x) \wedge x \in A^f(y)\} \\ &= \{(x, y) : y \in B^w(x)\} \\ &= \{(x, y) : x \in B^f(y)\}. \end{aligned}$$

## 2.4 Bellman equations and surplus sharing

Let  $V_u(x)$  denote the value of unemployment for a worker of type  $x$ ,  $V_e(x, y)$  the value of worker  $x$  employed 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

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

Shimer and Smith (2000) assume that wages are determined by Nash bargaining over the match surplus  $S(x, y)$  between workers and firms who have equal bargaining powers. We maintain this assumption in this paper, although it is not essential. First, we show below that the assumption of equal bargaining powers can be relaxed and the bargaining power can be identified in the data if the model incorporates either an idiosyncratic or an aggregate stochastic component affecting, say, firm productivity. In terms of notation, we allow for unequal bargaining powers by denoting workers' bargaining power  $\alpha \in (0, 1)$  ( $\alpha = \frac{1}{2}$  corresponds to the model in Shimer and Smith (2000)). Second, our method for identifying the sign and strength of sorting does not use the assumption of Nash bargaining but applies to any bargaining game whose solution implies that payoffs to both parties increase in match surplus. Finally, our method for the non-parametric identification of the production function only relies on specifying the bargaining protocol which yields a wage equation that can be inverted for output.

Generalized Nash bargaining over the match surplus with workers' bargaining power  $\alpha$

implies

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

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

$$\left. \begin{aligned} A^w(x) &= B^w(x) = \{y : S(x, y) \geq 0\}, \\ A^f(y) &= B^f(y) = \{x : S(x, y) \geq 0\}. \end{aligned} \right\} \quad (2)$$

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

$$V_u(x) = b + \beta V_u(x) + \beta \alpha (1 - \delta) \mathbb{M}_u \int_{B^w(x)} \frac{d_v(\tilde{y})}{V} S(x, \tilde{y}) \, d\tilde{y}, \quad (3)$$

$$V_v(y) = -c + \beta V_v(y) + \beta (1 - \alpha) (1 - \delta) \mathbb{M}_v \int_{B^f(y)} \frac{d_u(\tilde{x})}{U} S(\tilde{x}, y) \, d\tilde{x}, \quad (4)$$

$$V_e(x, y) = w(x, y) + \beta V_u(x) + \beta \alpha (1 - \delta) S(x, y), \quad (5)$$

$$V_p(x, y) = f(x, y) - w(x, y) + \beta V_v(y) + \beta (1 - \alpha) (1 - \delta) S(x, y). \quad (6)$$

## 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:

$$\forall (x, y) \in \mathcal{B} \quad \underbrace{\delta d_m(x, y)}_{\text{destruction}} = \underbrace{(1 - \delta) d_u(x) \mathbb{M}_u \frac{d_v(y)}{V}}_{\text{new match formation}}. \quad (7)$$

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.

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

## 2.6 Equilibrium

In a *steady state search equilibrium* (SE) all workers and firms maximize expected payoffs, taking the strategies of all other agents as given.<sup>8</sup> 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 (1) and match formation is optimal given wages  $w$ , i.e. a match is formed whenever the surplus is (weakly) positive (see (2)). The densities  $d_u(x)$  and  $d_v(y)$  ensure that the flow equations in (7) hold.

To prove existence, Shimer and Smith (2000) assume that the production is either globally supermodular or globally submodular.<sup>9</sup> A stronger assumption would be to require that the production function induces either positive assortative matching (PAM) or negative assortative matching (NAM), defined as follows:

**Definition 1.** *Consider worker types  $x_1 < x_2$  and firm types  $y_1 < y_2$ .*

*There is 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  $x_2 \in B^f(y_1)$ .*

*There is 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  $x_2 \in B^f(y_2)$ .*

Whereas this stronger assumption is not necessary for the existence proof it is commonly imposed in the literature as we discuss below. The equilibrium existence proof in Shimer and Smith (2000) also uses their assumption of a quadratic matching function. Nöldeke and Tröger (2009) extend the existence proof to a linear matching technology used in this paper: If  $f$  is either supermodular or submodular then a SE exists.<sup>10</sup> Shimer and Smith (2000) suggest that the assumption of either super or submodularity just avoids a more complicated existence proof and thus can be dispensed with. More specifically, this assumption rules out an atom of zero surplus matches, i.e.

$$\forall x \neq x' : \mu(\{y : S(x, y) = S(x', y) = 0\}) = 0, \quad (8)$$

where  $\mu$  is the Lebesgue measure. Imposing

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

<sup>8</sup>As in Shimer and Smith (2000), we assume that a match is formed if agents are indifferent.

<sup>9</sup>A production function is supermodular if the cross-derivative is positive and it is submodular if the cross-derivative is negative.

<sup>10</sup>They also argue that the proofs extend to our non-symmetric environment.

ensures this property. It thus avoids both the assumption of super or submodularity 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 Shimer and Smith (2000) 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  $x + y + k$  ( $k$  is a constant). However for large enough  $k$ , every worker matches with every firm and thus (8) is trivially satisfied.

**Proposition 1.** *A SE exists.*

### 3 Identification: Theory

In this section we establish the identification of the model from matched employer-employee data on wages and labor market transition rates. We proceed in four steps. First, we show how to identify the ranking of workers. Second, we identify the ranking of firms. Third, having identified the rankings of workers and firms, an investigation of the empirical matching patterns 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 of 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 of 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*

$$w^{av}(x) \equiv \left( 1 - \mathbb{M}_u + \delta \mathbb{M}_u + \mathbb{M}_u(1 - \delta) \int_{\overline{B^w(x)}} \frac{d_v(y)}{V} dy \right) w(x, \tilde{y}(x)) \quad (9)$$

$$+ \mathbb{M}_u(1 - \delta) \int_{B^w(x)} \frac{d_v(y)}{V} w(x, y) dy,$$

*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 worker type follows directly from the assumption that the production function is increasing in worker productivity. This yields for every firm a correct ranking of workers in the matching set of this firm. If one firm were to match with all workers in the economy, the ranking of workers based on wages in this firm would automatically represent a global ranking of all workers in the economy. If no firm matches with all workers, we have to aggregate the partial within-firm rankings to a global one. To illustrate how this works, consider a firm  $A$  which hires workers  $a_1 \prec a_2 \prec \dots \prec a_N$  and another firm  $B$  which hires workers  $b_1 \prec b_2 \prec \dots \prec b_M$  where the ranking within each firm is denoted by " $\prec$ ". Now suppose there is an overlap in the matching set of these two firms so that the best ranked workers in firm  $A$  are lowest ranked workers in firm  $B$ , i.e. for some  $k$ ,  $a_{N-k} = b_1, a_{N-k+1} = b_2, \dots, a_N = b_{k+1}$ . We can then combine the two rankings to rank all workers in the two firms to obtain  $a_1 \prec a_2 \prec \dots, a_{N-k} = b_1 \prec a_{N-k+1} = b_2 \prec \dots \prec a_N = b_{k+1} \prec b_{k+2} \prec \dots b_M$ . Iterating yields a global ranking of workers under the mild assumption that the set of workers can be split into overlapping matching sets.<sup>11</sup> In the implementation part of the paper, we deal with the fact that in small samples two workers, although in the matching set of the same firm, may not be observed to work for that firm during their finite lifetime. Higher worker mobility obviously makes it more likely to observe two workers in some firm, leading to a better aggregation of firm rankings to a global ranking. For realistic amounts of mobility across firms we find that having a small sample

---

<sup>11</sup>Note that the highest, lowest, and adjusted average wages provide global rankings of workers. Thus, these ranking can be used to initialize the rank aggregation procedure and this ensures a resulting global ranking even in cases with non-overlapping matching sets.

is not a relevant limitation and we are able to rank workers very well (despite substantial measurement error in wages that we allow for in our simulation exercises).

Finally, we design the ranking method based on the adjusted average wages because it is easy to prove that the average wage (without) the adjustment is not necessarily increasing in  $x$ .<sup>12</sup> To see this, consider two workers with different productivities. 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 because they 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 his employment history might be lower than that of a less productive worker. The more productive worker still obtains higher utility 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  $y$ .<sup>13</sup> 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,  $S(x, y)$ . Our strategy is as follows. We first establish that the 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

---

<sup>12</sup>Since separation rates are identically  $\delta$  at all firms a worker matches with, a worker's average wage is proportional to  $\int_{B^w(x)} w(x, y) d_v(y) dy$ . Assuming, for simplicity, that  $B^w(x) = [\underline{\varphi}(x), \bar{\varphi}(x)]$ , we get

$$\frac{\partial}{\partial x} \int_{B^w(x)} w(x, y) d_v(y) dy = \int_{B^w(x)} \frac{\partial w(x, y)}{\partial x} d_v(y) dy + \bar{\varphi}'(x) w(x, \bar{\varphi}(x)) d_v(\bar{\varphi}(x)) - \underline{\varphi}'(x) w(x, \underline{\varphi}(x)) d_v(\underline{\varphi}(x)).$$

Clearly, this equation is not necessarily increasing in  $x$ .

<sup>13</sup>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  $y$ .



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  $y$ .

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, given by*

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

*is increasing in  $y$ .*

**Proof of Result 6.** Using equation (1),

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

From (4), it follows that

$$\frac{V_v(y)(1 - \beta) + c}{\beta(1 - \alpha)} = (1 - \delta)\mathbb{M}_v \int_{B^f(y)} \frac{d_u(\tilde{x})}{U} S(\tilde{x}, y) \, d\tilde{x}. \quad (10)$$

From Result 5, both sides of (10) are increasing in  $y$ . Multiplying both sides of (10) by  $\alpha$  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 4), the expected surplus is also increasing in firm type (equation (10)). The next step uses that Nash-bargaining implies that both the worker and the firm benefit from an increase in the surplus. Nash bargaining has an even 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 results extend to 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*

$$\Omega(y) = (1 - \delta)\mathbb{M}_v \int_{B^f(y)} \frac{d_u(x)}{U} (w(x, y) - w(x, \tilde{y}(x))) \, dx, \quad (11)$$

*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 and  $\tilde{y}(x)$  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 (5)),

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

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

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

Using Result 6 completes the proof.<sup>14</sup>

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

---

<sup>14</sup>The strict monotonicity of  $\Omega$  and thus the ranking of firms depends on  $V(y)$  being increasing in  $y$ . Result 5 established that this is an endogenous outcome in the model with type-independent vacancy posting costs. It would also trivially follow under the more natural assumption that costs are increasing in firm type even if we allow for free firm entry. Only if costs are type-independent and there is free-entry  $\Omega$  would be constant. This can be tested in the data using, e.g., the monotonicity test by Hall and Heckman (2000) which we describe below in Section 5.1 to discriminate between models with comparative and absolute advantage. While the data will have the final word on this, we expect the constant  $\Omega$  hypothesis to be rejected in the data in view of the large firm heterogeneity documented by, e.g., Bagger, Christensen, and Mortensen (2011).

to fill a vacancy,  $q(y)$ . The average wage premium equals

$$\Omega^e(y) = \int_{B^f(y)} \frac{\frac{d_u(x)}{U}(w(x, y) - w(x, \tilde{y}(x)))}{\int_{B^f(y)} \frac{d_u(\tilde{x})}{U} d\tilde{x}} dx.$$

The probability that a vacancy of type  $y$  is filled equals

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

It then holds that

$$\Omega(y) = q(y)\Omega^e(y).$$

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

$$\hat{\Omega}^e(j) = \sum_{\{i \text{ employed at } j \text{ at some } t\}} \frac{w(i) - w^{\min}(i)}{E(j)},$$

where  $w(i)$  is the wage worker  $i$  gets at firm  $j$ ,  $E(j)$  is the number of workers employed at firm  $j$  at some  $t$  during the sample, and  $w^{\min}(i)$  is the reservation wage of worker  $i$ . While the reservation wage is clearly conceptually related to the lowest accepted wage by a worker, a more sophisticated measurement procedure is required in small samples and in presence of measurement error. We develop such a procedure in Section 4. From the law of large numbers, we obtain that  $\hat{\Omega}^e(j)$  converges to  $\Omega^e(y(j))$ .<sup>15</sup> We discuss the measurement of the job filling rates  $q(y)$  in Section 3.4.3 below.

### 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. The sign of 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.

---

<sup>15</sup> $\Omega^e(y(j))$  is defined for newly hired workers but since the wage is constant during a job we can equivalently choose the wage at some  $t$  while worker  $i$  is employed at firm  $j$ . We choose this definition for the implementation since it allows to use wage information of workers who are currently working at firm  $j$  but were newly hired before the start of the sample.

### 3.3.1 Relationship to the Literature

Note that  $\Omega(y)$  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. Because of this, Result 7 enables us to identify the sign of sorting. This is in contrast to the recent results of Eeckhout and Kircher (2011) who used 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.

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 (as in Chade (2001)). To see why this is essential for ranking firms, rearrange the Bellman equation (4) of a vacancy in our model:

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

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

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

the Constant Surplus Condition in Theorem 1 in Atakan (2006). If, instead,  $\beta < 1$ , then  $V_v(y)(1 - \beta)$  is increasing in  $y$  and so is the expected surplus. This monotonicity (independent of 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 (11)) 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.

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.<sup>16</sup>

However, introducing discounting is not sufficient to achieve identification in 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 not even constant if  $\beta = 1$  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).

### 3.4 Identifying Remaining Model Parameters

We now show how to identify the remaining objects in the model. Our primary interest is in identifying the production function  $f(x, y)$ . We recover it, at the end of this section, by inverting the wage equation. To accomplish that, we require the estimates of the value of unemployment  $V_u(x)$ , the value of a vacancy  $V_v(y)$ , and the probability to fill a vacancy  $q(y)$ . Alongside with measuring these key objects, we also show how to measure the value of being

---

<sup>16</sup>In terms of the quantitative results reported below, we found that even using the monthly discount factor as high as 0.999 does not measurably affect our ability to identify the objects of interest.

employed  $V_e(x, y)$ , the value of producing for a firm  $V_p(x, y)$ , and the meeting probabilities for unemployed workers and vacant firms  $\mathbb{M}_u$  and  $\mathbb{M}_v$ . Estimated values are denoted with a hat,  $\hat{\cdot}$ .

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

The Bellman Equation (5), implies, using  $V_e(x, \tilde{y}(x)) = V_u(x)$ , that  $V_u(x)(1-\beta) = w(x, \tilde{y}(x))$ . Thus, the reservation wage for workers of type  $x$  can be used to measure the (type-dependent) value of unemployment as<sup>17</sup>

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

To measure  $V_e(x, y)$ , consider a worker of type  $x$ , who starts working at a firm of 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

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

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  $y$  yields the estimate  $\hat{V}_e(x, y)$ .

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

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

Using that  $\alpha = \frac{1}{2}$  in the model of Shimer and Smith (2000), yields an estimate of  $S(x, y)$ . In Appendix I.4 we describe how the parameter  $\alpha$  can be identified from the data in a more general version of the model.

### 3.4.2 Measuring $V_v(y)$ and $V_p(x, y)$

We next turn to the measurement of  $V_v(y)$ , which is related to our estimate  $\Omega(y)$  through

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

Since, as discussed above, we can measure  $\Omega$  in the data, and we can follow the standard approaches in the literature to estimate or calibrate  $\delta$  and  $\beta$ , we obtain a measure of  $V_v(y)(1 -$

---

<sup>17</sup>Measurement of (type-dependent) reservation wages is described in Section 4.

$\hat{\beta}) + c$ . Our estimate is then

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

which differs from the true value by  $\frac{c}{1 - \hat{\beta}}$ , 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.

### 3.4.3 Transition Rates

Data on firm-level vacancies are available. For every worker type we can measure the probability to leave unemployment, which equals

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

From this equation and with firm level vacancy data (i.e. data for  $d_v(y)$ ), we can measure  $\mathbb{M}_u$  since the integral is known and we measure  $\lambda(x)$ . A more robust way is to integrate over all worker types

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

which again allows us to solve for  $\mathbb{M}_u$ .

Similarly the probability to fill a vacancy for firm type  $y$  equals

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

which we can measure directly in the data if we observe vacancy data at the firm level. Next, we show how these objects can be estimated if vacancy data at the firm level are not available.

Data on firm-level vacancies are not available. 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

$$\tilde{q}_y \equiv \int_{B^f(y)} \frac{d_u(\tilde{x})}{U} dx = \frac{\sum_{\hat{x} \in B^f(y)} \hat{u}(\hat{x})}{\sum_{\hat{x}} \hat{u}(\hat{x})}. \quad (12)$$

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 (12) and the law of large numbers imply that

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

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

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

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.

#### 3.4.4 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 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 with a time dimension of 20 years. Most currently available and commonly used matched data sets (e.g., from Brazil, Denmark, Germany, France) have a similar or longer time span (the longer the data set the more precise our method is). We assume that the data include 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)  $f(x, y) = 0.4 + \mathbb{1}_{\{x \leq 0.4\}}(x + 0.6)y + \mathbb{1}_{\{x > 0.4\}}((x - 0.4)^2 + y^2)^{1/2}$ , which induces neither positive nor negative assortative matching (NEITHER). Instead, the pattern of sorting changes over its domain (PAM for  $x \leq 0.4$  and NAM for  $x > 0.4$ ).

The literature has largely restricted attention to identifying sorting assuming that the production function induces either positive or negative assortative matching. This motivates our choice of the first two production functions. Our method, however, does not rely on placing such restrictions on the production function. 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  $F$  and  $G$ ). Common choices in the literature are 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: Parameterizations

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	$\beta$	0.996		
Separation rate	$\delta$	0.01	0.025	
Meeting function scale	$\kappa$	0.4	0.7	
Meeting function elasticity	$\nu$	0.5		
Worker's bargaining weight	$\alpha$	0.5		
Worker's flow utility of unemp.	b	0.0		
Vacancy cost	c	0.0		
Measurement error in wages	$\epsilon$	20% of overall wage variance		

Thus, all combinations of parameters result in 108 distinct parameterizations that we consider. Appendix Figure A-1 summarizes the 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 error in wages we report the rank correlations of the true worker types and types recovered using each of these methods across simulations.

In addition, as discussed above, Result 1 implies that wages within a firm are increasing in worker productivity  $x$ . This provides a way to rank workers according to their productivity within a firm and these partial rankings can be aggregated to construct a more global ranking. However, the presence of measurement error in wages complicates 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 inconsistently with each other. In our application voters correspond to

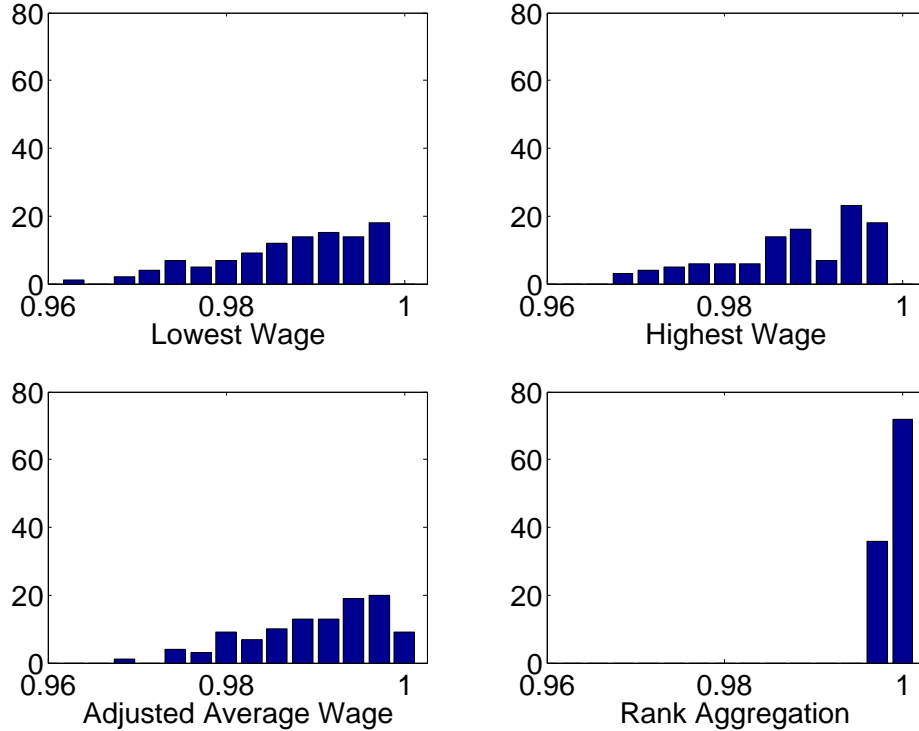


Figure 1: Distribution of the correlation between the estimated and true worker ranks across parameterizations.

firms and workers to voting alternatives. An aggregate ranking then 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 sum of 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. In practice, we found that implementing a portion of their algorithm achieves a high level accuracy while being quite fast. A detailed description is provided in Appendix III.

Figure 1 reports the distribution 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

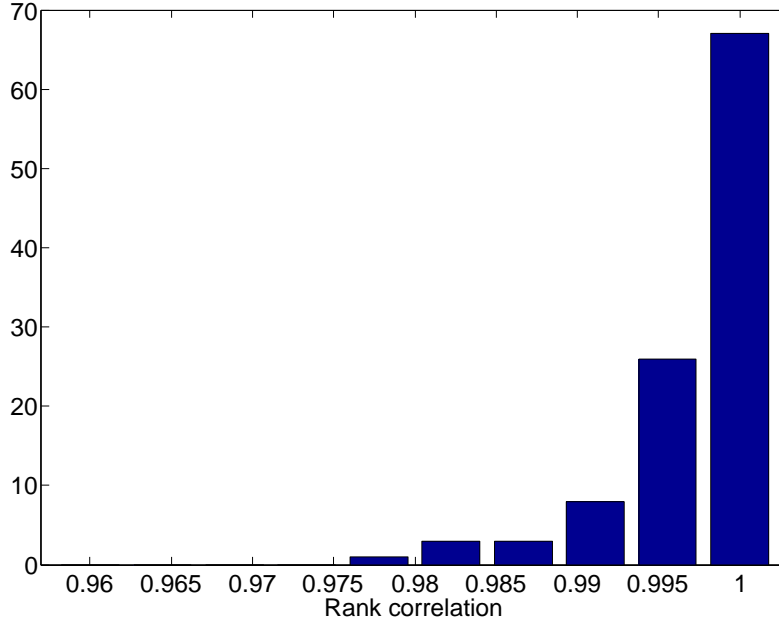


Figure 2: Distribution of the correlation between the estimated and true firm ranks across parameterizations.

performance in ranking workers. The measure based on rank aggregation outperforms any of the 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 Appendix II.

Figure 4.3 plots the distribution 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  $\Omega$  values) into bins or types. The measured value of a vacancy for a firm type is then

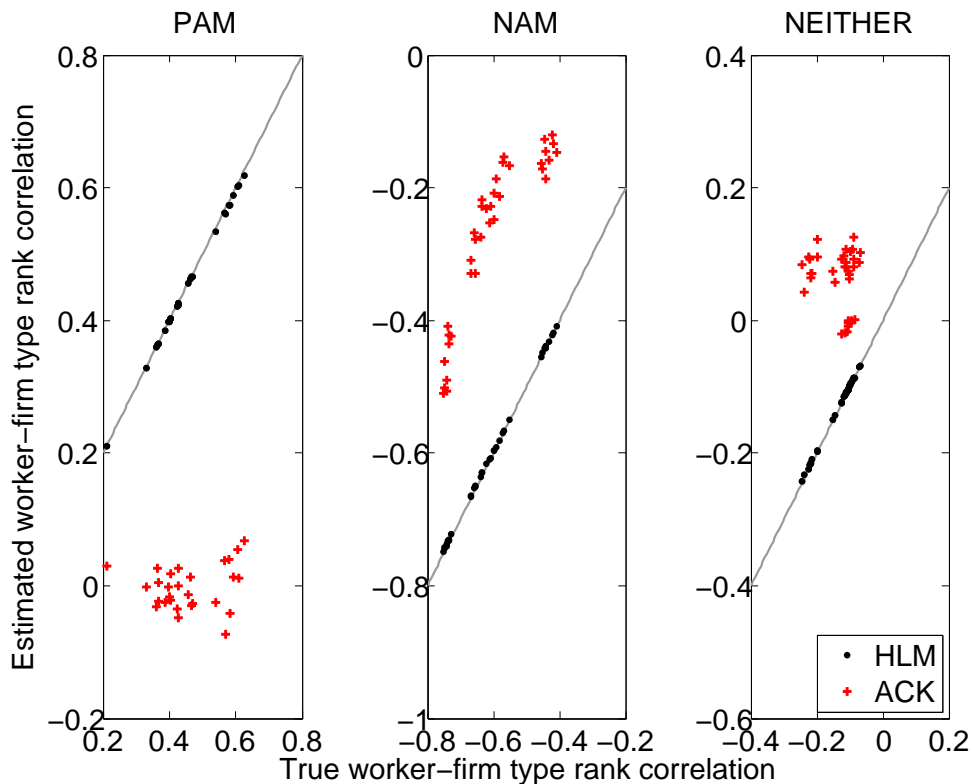


Figure 3: Correlation between identified worker and firm ranks against true correlation.

the average 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 effects of the sampling error in small samples.

#### 4.4 Sign and Strength of Sorting

Figure 3 plots the correlation between identified worker and firm ranks against the true correlation (each black dot corresponds to a different parameterization). Here we separate the three production functions to illustrate that our identification strategy easily identifies the sign of sorting. In all cases this relatively crude measure of the strength of sorting performs quite well (as all the black dots lie on a 45 degree line, the correlation between worker and firm types are identified nearly exactly).

For comparison, the (red, if viewed in color) crosses in the same figures correspond to the correlation between worker and firm fixed effects estimated using the exact least squares formulas provided by Abowd, Creedy, 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.

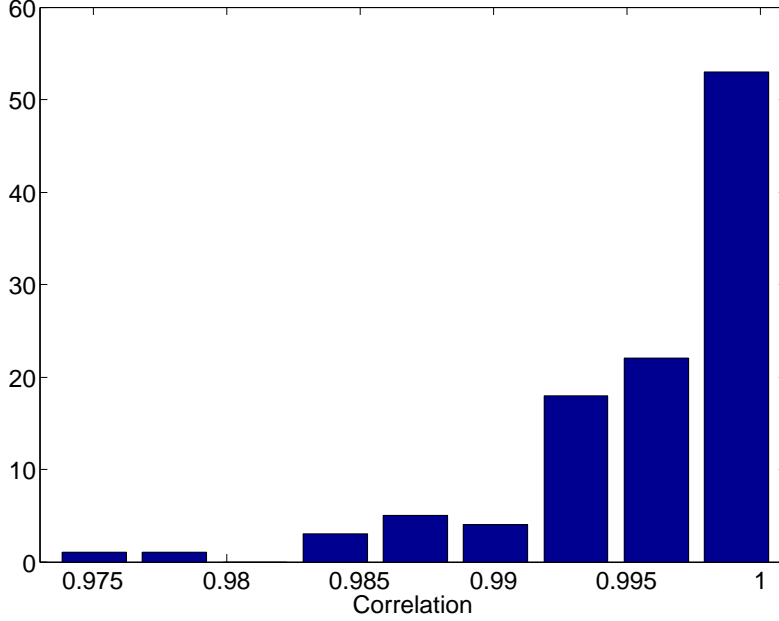


Figure 4: Distribution of correlation between true and estimated production functions across parameterizations.

#### 4.5 Measuring $f(x, y)$

To evaluate how well our method recovers the production function, in Figure 4 we plot the distribution 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 Appendix Figures A-2 - A-4 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  $d_w = d_f = U[0, 1]$ , 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.

## 4.6 Measuring the effect of search frictions on output

Having obtained the nonparametric estimate of the production function and of the rankings of workers and firms, one can address a number of important economic questions, some of which we outline in the Introduction and the Conclusion. One such question is to assess the magnitude of output losses due to mismatch between workers and firms. We now evaluate the ability of our identification and implementation strategies to provide a reliable quantitative answer to this question without imposing functional form assumptions on technology.<sup>18</sup>

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)  $\mu : [0, 1] \rightarrow [0, 1]$  of workers to firms such that the total output  $\sum_x f(x, \mu(x))$  is maximized. 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 frictional matching set is zero.

This assignment problem is a well studied combinatorial optimization problem and there are several existing algorithms that can solve it in polynomial time.<sup>19</sup> 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  $f(\cdot, \mu(\cdot))$ :

$$\mathcal{E}^{no\ fric} = \int_0^1 f(x, \mu(x)) \, dx, \quad (13)$$

where we used that worker ranks are uniformly distributed. In the presence of frictions, let

---

<sup>18</sup>An alternative approach to measuring the cost of mismatch was pursued by Eeckhout and Kircher (2011). They noted (as did Gautier and Teulings 2004, 2006) that differentiating the wage equation (such as Equation (A2) in this paper) the cross-partial derivative of the wage equals the cross-partial derivative of the production function multiplied by the worker bargaining weight. They also show that in their two period model, where a period with search frictions is followed by a frictionless period and there is no discounting, and under the assumption that the production function globally induces PAM or NAM, the average cross-partial of the production function can be theoretically related to the output cost of mismatch. They do not explore the empirical properties of this estimator given the nature of the available data. Even maintaining their restrictive assumptions on the production technology, it is not clear whether it is possible to generalize this method to the fully dynamic model in Shimer and Smith (2000).

<sup>19</sup>See Burkard, DellAmico, and Martello (2009) for a thorough review.



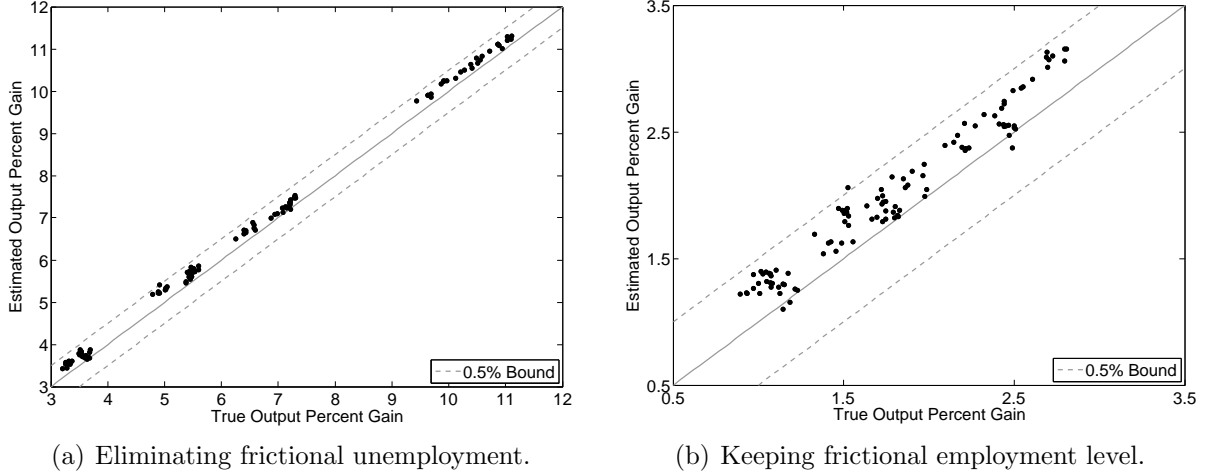


Figure 5: Estimated gains from eliminating frictions.

$\mathcal{E}^{fric}$  be the expectation of  $f$ :

$$\mathcal{E}^{fric} = \int_B f(x, y) d_m(x, y) dx dy. \quad (14)$$

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,  $\mathcal{E}^{fric}$ :

$$\Delta \mathcal{E} = \mathcal{E}^{no\ fric} - \mathcal{E}^{fric}. \quad (15)$$

In Figure 5, 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 horizontal axis while the estimated gain as a percent of frictional output is on the vertical axis. 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 5(a) we plot the estimated gains from eliminating all frictions, including the elimination of frictional unemployment.

In Panel 5(b) we only consider the effect from optimal reassignment of workers employed in the economy with frictions (that is, we keep the employment of each worker type fixed at its level in the economy with frictions).<sup>20</sup> We interpret the results as indicating that the method performs quite well in estimating the gain from the optimal worker reallocation.

<sup>20</sup>Note that the relatively small gains from the reallocation in this experiment are driven by the parsimonious CES production 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.

## 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 can be described in terms of 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 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  $x'$  can produce more than  $x$  in some other firm  $y'$ . An example of such a production function would be  $\zeta - \frac{1}{2}\gamma(x - y)^2$ , where  $\zeta$  and  $\gamma$  are some parameters. 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:<sup>21</sup>

$$\log f(x, y) = x - \frac{1}{2}\gamma(x - y)^2. \quad (16)$$

We show now that the method developed in this paper 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 (16) this is the case if  $\gamma > 1$ ). Our quantitative analysis of the models with the production function (16) will show that our worker ranking procedures continue to perform very well although we will use the relatively high value of  $\gamma = 2$ .

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

---

<sup>21</sup>Gautier and Teulings (2012) use a comparative advantage model with this production function to study sorting on observables taking measurement error into account, while our focus is on measuring sorting on worker and firm characteristics that are not observed in the data.

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.

Thus, our procedure remains unchanged up to this point. We have obtained estimates of all the components needed to invert the wage equation and identify the production function at the level of the individual firm. However, to further reduce the effect of measurement error, it is desirable to be able to group similar firms into bins and to estimate the production function at the bin level. In our analysis of absolute advantage models we could order firms by their estimated  $\Omega$  and assign firms with very similar values of  $\Omega$  to the same bin/type. This was justified as we showed that  $\Omega$  is increasing in firm type  $y$  if the production function is increasing in  $y$ . As this does not hold in the comparative advantage model, we need a different procedure to bin firms. Accordingly, instead of considering firms with very similar  $\Omega$  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 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.

Thus, the firm binning procedure differs depending on whether a production function is monotone or not. In Section 5.1 we will provide a test to determine the monotonicity of the production function in the data. The outcome of this test then determines the appropriate firm binning method.

To quantitatively assess the performance of our method in a comparative advantage model with the production function (16), we conduct a Monte Carlo study similar to the one in Section 4. We set the production function parameter  $\gamma = 2$ , the calibrated value in Gautier and Teulings (2012).<sup>22</sup> We use the same parameterizations and worker and firm distributions as before (see Table 2), resulting in 36 distinct parameterizations. Appendix Figure A-5 summarizes the values that a number of variables of interest take across all simulations. They continue to lie within empirically plausible ranges.

To evaluate the ability of our method to recover the production function, in Figure 6 we plot the distribution of the correlation between the true and the estimated production

---

<sup>22</sup>Strictly speaking, the value of  $\gamma$  in Gautier and Teulings (2012) depends on the scale of  $x - y$  and only the value of  $\gamma * (x - y)^2$  is scale independent. For our scaling of  $x - y$ , taking this account, would imply a lower value of  $\gamma$  although we could not establish the precise mapping. We use the value of  $\gamma = 2$  in this experiment as it makes ranking of workers and the recovery of the production function more difficult. We verified that results become even better when lower values of  $\gamma$  are used.

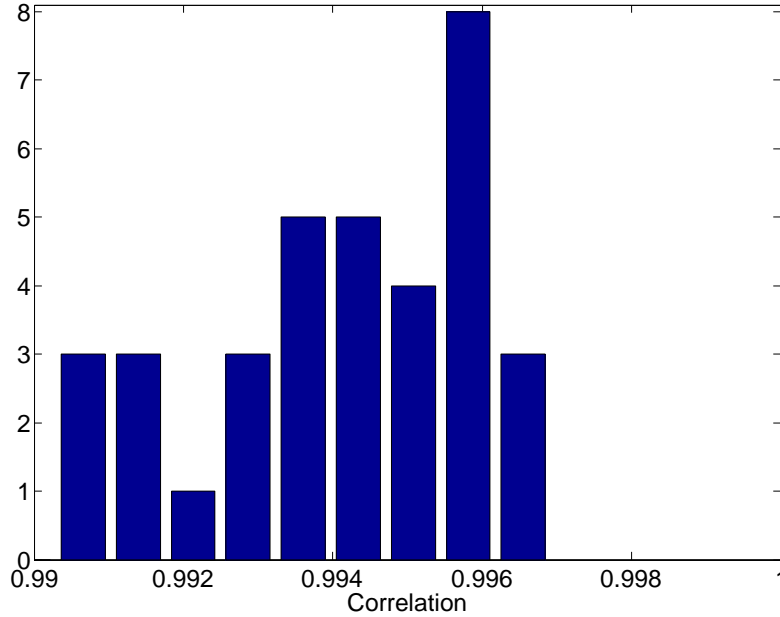


Figure 6: Distribution of correlation between true and estimated production functions across parameterizations in Gautier and Teulings.

functions across all parameterizations. The correlations are generally very high. In Appendix Figure A-6 we plot several views for of the true and estimated production function for the same parameterization that was used to plot similar figures for the monotone production functions. Evidently, the fit is very close.

To assess the reliability of these estimates for measuring the potential gains from eliminating frictions, in Figure 7 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 horizontal axis while the estimated gain as a percent of frictional output is on the vertical axis. As before, the two dotted lines that represents a mistake of plus or minus one half of one percent of output. In Panel 7(a) we plot the estimated gains from eliminating all frictions, including the elimination of frictional unemployment. In Panel 7(b) we only consider the effect from optimal reassignment of workers employed in the economy with frictions (that is, we keep the employment of each worker type fixed at its level in the economy with frictions). We interpret the results as indicating that the method continues to perform well in estimating the gain from the optimal worker reallocation in a comparative advantage model of Gautier and Teulings (2012).

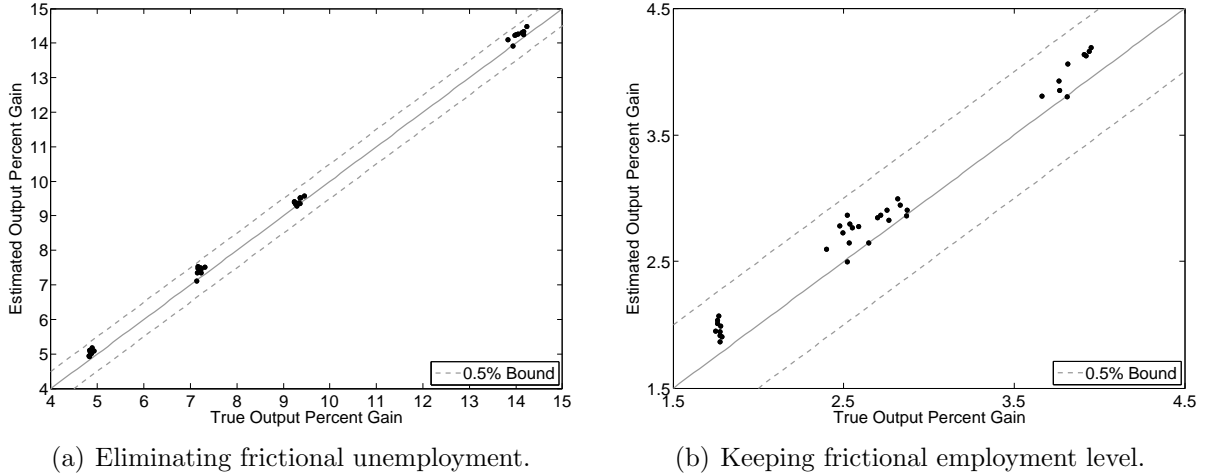


Figure 7: Estimated gains from eliminating frictions in Gautier and Teulings.

## 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 (16). 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  $f(x, \cdot)$ . More concretely, the algorithm selects a set of intervals in the  $y$ -space and on each interval it computes a linear OLS estimate of  $f(x, \cdot)$  on  $y$ . The test rejects that the function is increasing in  $y$  if one of these slopes is not statistically significantly positive, i.e. if the function is decreasing on some interval. For our application, suppose that worker type  $x$  matches with firms  $y_1, \dots, y_n$  and the estimated outputs are  $z_1 := f(x, y_1), \dots, z_n := f(x, y_n)$ . Let the variance of  $z_i$  be  $\sigma^2$ . To test for the monotonicity of a one-dimensional function  $f(x, \cdot)$ , Hall and Heckman (2000) define the following test statistic. Let  $0 \leq r \leq s - 2 \leq n - 2$  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

$$(\hat{a}, \hat{b}) = \operatorname{argmin}_{(a,b)} S(a, b \mid r, s),$$

which is the average slope of the function  $f(x, \cdot)$  on the interval  $[z_{r+1}, z_s]$ , and let

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

The reason to define  $Q$  is that multiplying  $Q$  with  $\hat{b}$  results in a statistic  $\hat{b}(r, s)Q(r, s)$  which has variance  $\sigma^2$  for each pair  $(r, s)$ . The test statistic of Hall and Heckman (2000) is then

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

where the parameter  $m$  puts a lower bound on the size of the intervals. Only intervals with  $s - r \geq m$  are considered ( $r + 1$  is the lower bound of the interval and  $s$  is the upper bound. The distance has to be at least  $m$ ). If  $m$  is too small, say  $m = 1$ , every bump e.g. due to a small numerical error would result in  $f(x, y_k) > f(x, y_{k+1})$  for some  $k$  and in a rejection of monotonicity. if  $m$  is too large, say  $m = n$  then the test only delivers only the average slope. it would for example not detect the decreasing part of a hump-shaped function. We therefore choose  $m = 0.1n$ , where  $n$  is the size of the matching set of type  $x$ .<sup>23</sup>

The monotonicity of  $f(x, \cdot)$  is tested against a constant function. To generate 99% 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  $\sigma^2$ , i.e. we test against a constant function ( $z_i$  are constant) with the same noise as the data.

We implement this procedure for all production functions and all parameterizations considered above. For each production function and each parametrization we rank the firms through  $\Omega$  and 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.

---

<sup>23</sup>We only consider matching with at least 10 elements. We also impose that  $m$  is at least 4.

Table 3: Monotonicity Test I

Production Function	Increasing	Constant	Decreasing
PAM	100.00%	0.00%	0.00%
NAM	99.94%	0.06%	0.00%
NEITHER	99.94%	0.06%	0.00%
Gautier and Teulings	20.94%	61.72%	17.34%

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 (16) we find that we can reject a constant function in favor of an increasing function only in 20.94 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 79.06 percent of cases we either accept the constant function or we even reject in favor of a decreasing function.

While Table 3 refers to the average outcome of the test across all parameterizations of each production function, another informative statistic summarizes the distribution of the test results across parameterizations. To this end, for each parameterization we say that it fails the monotonicity test (or that the increasing function is rejected) if we reject a constant function in favor of an increasing function for less than 95% of worker types  $x$ . Equivalently we “accept” an increasing function if we reject in favor of an increasing function in more than 95% of worker types  $x$ . More concretely, for each parametrization we implement the monotonicity test for all types  $x$ , i.e we test the monotonicity of  $\hat{f}(x, \cdot)$ .<sup>24</sup> We count the percentage where we reject in favor of an increasing, cannot reject, and reject in favor of decreasing. We “accept increasing” for this parametrization if for at least 95% of types  $x$  we rejected in favor of increasing. We then count in percentage terms for how many parameterizations we accept an increasing function. The results summarized in Table 4 imply that in the case of monotone functions we never reject an increasing function. In contrast, the non-increasing function fails this test in 100% of the parameterizations.

<sup>24</sup>Excluding those worker types who match with less than 10 firm types.

Table 4: Monotonicity Test II

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

## 6 Conclusion

In this paper we have developed an empirical methodology that allows to study assortative matching between employers and employees based on their unobserved (to economists) characteristics. In particular, we have shown theoretically that all parameters of the classic Becker model of sorting based on absolute advantage with search frictions, as analyzed in Shimer and Smith (2000), can be identified using only matched employer-employee data on wages and labor market transitions. 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 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 extended our identification and implementation strategies to the model of assortative matching based on comparative advantage in Gautier and Teulings (2006, 2012) and provided a test that allows to discriminate between these models.

Our theoretical analysis was based on a relatively simple version of the model which allowed us to obtain clear and transparent results. We expect that many of these insights will be applicable (or can be adapted) to more sophisticated models that will be taken in the data. Those models will perhaps include on-the-job search, sorting based on both observables and unobservables, and hopefully a more sophisticated theory of the firm.

There are numerous important positive and normative questions that can be answered once the nonparametric identification of this class of models is established, in particular of the production function and of the rankings of workers and firms (if the production function is such that they can be ranked). For example, as we have shown, without any additional



assumptions we can compute the optimal assignment of workers to jobs 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. A similar approach can be used to separately measure the extent of changes in technology and sorting frictions over time (e.g, to understand the reasons for the substantial rise in the college premium in the US in the 1980s) and across, say, countries. The ability to separately identify changes in technology from changes in mismatch seems essential for understanding the effects of changes in many economic policies, e.g., trade liberalizations. 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.

## 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.
- BAGGER, J., B. J. CHRISTENSEN, AND D. T. MORTENSEN (2011): “Wage and Productivity Dispersion: The Roles of Rent Sharing, Labor Quality and Capital Intensity,” mimeo.

- 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.
- BURKARD, R., M. DELLAMICO, AND S. MARTELLO (2009): Assignment Problems. SIAM.
- CHADE, H. (2001): “Two-Sided Search and Perfect Segregation with Fixed Search Costs,” Mathematical Social Sciences, 42, 31–51.
- CONDORCET, J. (1785): “Essai sur l’application de l’analyse à la Probabilité des Décisions Rendues à la pluralité des 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.
- ECKHOUT, J., AND P. KIRCHER (2010): “Sorting and Decentralized Price Competition,” Econometrica, 78(2), 539–574.
- (2011): “Identifying Sorting – In Theory,” The Review of Economic Studies, 78(3), 872–906.
- 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 (2008): “The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited,” American Economic Review, 98(4), 1692–1706.
- (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): “Matching, Sorting, and Wages,” Working paper, University College London.
- LOPES DE MELO, R. (2009): “Sorting in the Labor Market: Theory and Measurement,” Working paper, University of Chicago.
- MOEN, E. R. (1997): “Competitive Search Equilibrium,” Journal of Political Economy, 105(2), 385–411.
- NÖLDEKE, G., AND T. TRÖGER (2009): “Matching Heterogeneous Agents with a Linear Search Technology,” Working paper, Basel and Mannheim.
- SHI, S. (2001): “Frictional Assignment. I. Efficiency,” Journal of Economic Theory, 98(2), 232 – 260.
- SHIMER, R. (2005): “The Assignment of Workers to Jobs in an Economy with Coordination Frictions,” Journal of Political Economy, 113(5), 996–1025.
- SHIMER, R., AND L. SMITH (2000): “Assortative Matching and Search,” Econometrica, 68(2), 343–370.
- TEULINGS, C. N., AND P. A. GAUTIER (2004): “The Right Man for the Job,” Review of Economic Studies, 71(2), 553–80.

## APPENDICES

# 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 receives  $b$ , and moves into employment only if he meets a firm in his acceptance set, and does not experience immediate match destruction. Any breakdown in this sequence leaves the worker unemployed again in the next period.

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

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 (1) to obtain

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

where terms cancel to give (3).

An employed worker receives  $w(x, y)$ , and remains employed next period with probability  $(1 - \delta)$  or becomes unemployed with complementary probability. Minor rearranging and (1)

yield (5):

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

## I.2 Proofs of Results in Section 3.1

**Proof of Result 1.** Adding (5) and (6) 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,

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

so that, using (1), 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

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

Using (5) again, gives us wages<sup>25</sup>

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

---

<sup>25</sup>Wages can also be derived using (6):

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

We now establish that  $V_u(x)$  is increasing in  $x$ . From (3),

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

so that

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

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

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

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

$$\begin{aligned} \frac{\partial V_u(x)}{\partial x} \left( 1 - \beta + \frac{(1 - \beta)\beta\alpha(1 - \delta)\mathbb{M}_u}{1 - \beta(1 - \delta)} \int_{B^w(x)} \frac{d_v(\tilde{y})}{V} \, d\tilde{y} \right) = \\ \frac{\beta\alpha(1 - \delta)\mathbb{M}_u}{1 - \beta(1 - \delta)} \int_{B^w(x)} \frac{d_v(\tilde{y})}{V} \frac{\partial f(x, \tilde{y})}{\partial x} \, d\tilde{y} \end{aligned}$$

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):

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

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 a worker of type  $x$ . Then (5) can be written as

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

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 unions of intervals. For the ease of exposition we assume that there is just one interval:

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

First rewrite the adjusted average wage as

$$w^{av}(x) = w(x, \tilde{y}(x)) + M_u(1 - \delta) \int_{B^w(x)} \frac{d_v(y)}{V} [w(x, y) - w(x, \tilde{y}(x))] dy.$$

Take derivatives with respect to  $x$ :

$$\begin{aligned}\frac{\partial w^{av}(x)}{\partial x} &= \frac{\partial w(x, \tilde{y}(x))}{\partial x} + \mathbb{M}_u(1 - \delta) \int_{B^w(x)} \frac{\partial w(x, y) - w(x, \tilde{y}(x))}{\partial x} \frac{d_v(y)}{V} dy \\ &+ \mathbb{M}_u(1 - \delta) \overline{\varphi}'(x) \frac{d_v(\overline{\varphi}(x))}{V} [w(x, \overline{\varphi}(x)) - w(x, \tilde{y}(x))] \\ &- \mathbb{M}_u(1 - \delta) \underline{\varphi}'(x) \frac{d_v(\underline{\varphi}(x))}{V} [w(x, \underline{\varphi}(x)) - w(x, \tilde{y}(x))].\end{aligned}$$

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

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

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

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

so that

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

and thus that

$$\begin{aligned}\frac{\partial V_v(y)}{\partial y}(1 - \beta + \frac{(1 - \beta)\beta(1 - \alpha)(1 - \delta)\mathbb{M}_v}{1 - \beta(1 - \delta)} \int_{B^f(y)} \frac{d_u(\tilde{x})}{U} d\tilde{x}) &= \\ \beta(1 - \alpha)(1 - \delta)\mathbb{M}_v \int_{B^f(y)} \frac{d_u(\tilde{x})}{U} \frac{\frac{\partial f(\tilde{x}, y)}{\partial y}}{1 - \beta(1 - \delta)} d\tilde{x} &> 0,\end{aligned}$$

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

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

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

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

## I.4 Measuring $\alpha$ in the data

In the model of Shimer and Smith (2000), the value of  $\alpha$  is fixed at  $\frac{1}{2}$ . More generally, one may consider leaving the value of  $\alpha$  unrestricted in the  $(0, 1)$  interval and recovering it from the data. Note that  $\alpha$  governs the responsiveness of wages to changes in match surplus (if  $\alpha \rightarrow 0$ , workers receive  $b$  regardless of the movements in the match surplus, while if  $\alpha \rightarrow 1$ , workers' wages fully co-move with surplus). While this provides a natural source of variation for the identification of this parameter, the fluctuations of surplus are absent from the simple baseline version of the model considered in the main text. In this appendix we extend the model to incorporate two sources of fluctuation in match surplus, an idiosyncratic and an aggregate productivity shock and show how each of these stochastic components allows to identify the bargaining powers.<sup>26</sup>

### I.4.1 Measure $\alpha$ from fluctuation in firm output

To measure the bargaining power  $\alpha$  in the data, we first consider an extended version of the model with i.i.d. shocks to the firm's technology,  $\epsilon_j$ , which changes output from  $f(x, y)$  to  $f(x, y) + \epsilon_j$  for all worker types  $x$  employed at firm  $j$  of type  $y$ . In response to such a shock to the firm's technology, Nash bargaining with worker bargaining power  $\alpha$  implies that profits increase by  $(1 - \alpha)$  and wages increase by  $\alpha$ . To measure  $\alpha$  using this experiment we can

---

<sup>26</sup>Eeckhout and Kircher (2011) have argued that the bargaining power can also be identified in the baseline deterministic version of the model. Unfortunately, their proof appears to contain a mistake (Equation (28) in their paper does not follow from equation (26) since the term  $w_x^*(x)$  is missing in equation (28)).

use any data where the response of wages is observable. This approach for identifying the bargaining powers was pursued in a number of papers in the literature reviewed in Hagedorn and Manovskii (2008).

Adding these shocks to our model is simple and does not change any of our other results and conclusions as we verified in simulated data. The reason is that these shocks are unanticipated and their impact is only to make statistics slightly noisier in the same way as measurement error does (and we have established in the main text that adding a large amount of measurement error does not have a significant impact on our inference). The ranking of workers within a firm is not affected at all since all wages within a firm are shifted by the same amount,  $\alpha\epsilon_j$ , and thus the ranking of workers is preserved. The ranking of firms is based on the statistic  $\Omega$  which is proportional to the value of a vacancy. Since the technology shocks are unanticipated this statistic is not affected either. Neither is the estimation of the production function  $f$ . The only object that is affected is the estimation of the matching set, as now workers may become acceptable only because of a large positive  $\epsilon_j$  whereas they were not acceptable in the absence of any shocks. This makes the model computationally much more burdensome. Given that adding these idiosyncratic shocks to the model obviously allows to identify  $\alpha$  but has no material impact on any of our results, we adopted a simpler model as a benchmark in the main text.

#### I.4.2 Using Business Cycles to Measure $\alpha$

We now show how the bargaining power  $\alpha$  can be measured in the data by considering 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 that

$$V_e(x, y, z) = w(x, y, z) + \beta E(V_u(x, z') | z) + \beta\alpha(1 - \delta)E(S(x, y, z') | z),$$

and for the value of a filled vacancy that

$$V_p(x, y, z) = zf(x, y) - w(x, y, z) + \beta E(V_v(y, z') | z) + \beta(1 - \alpha)(1 - \delta)E(S(x, y, z') | z).$$

Adding up these two Bellman equations yields:

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

and equivalently

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

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

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

so that the surplus equals

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

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

$$w(x, y, z) = \alpha S(x, y, z)(1 - \beta(1 - \delta)) + V_e(x, z) - \beta E(V_u(x, z') | z).$$

Making the same approximation for  $V_u$ ,

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

and using the equation for the surplus  $S$ , we obtain

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

The differences in wages for types  $x$  and  $x'$  equals

$$w(x', y, z) - w(x, y, z) = \alpha z(f(x', y) - f(x, y)) + (1 - \alpha)(1 - \beta)(V_u(x', z) - V_u(x, z)).$$

To figure out  $\alpha$  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  $\kappa$  is then our estimate of  $1 - \alpha$  so that  $\hat{\alpha} = (1 - \kappa)$ .

## II Computation and Implementation

In this section we describe how we compute the model and measure and construct the variables mentioned in the text. We first discretize the continuous type space for both workers and firms with 50 evenly distributed grid points on the type space  $[0, 1]$ . To compute the

model we use an iterative procedure on the match density,  $d_m(x, y)$  and the surplus,  $S(x, y)$ . Let  $d_{m,k}(x, y)$  and  $S_k(x, y)$  be the values in the  $k^{\text{th}}$  iteration. To initialize the iteration we set,  $\forall(x, y)$ , the initial match distribution,  $d_{m,0}(x, y) = 0.5$ , and the initial surplus,  $S_0(x, y) = f(x, y)$ . We obtain a solution by alternatively updating exactly once on either the match density (equation 7) or the flow equation for the surplus (which we get by summing equations 3 - 6). 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 a probability (between 0 and 1) such that the surplus of the match is positive, but very close to zero. Denote iteration  $k$  of the acceptance strategy of workers with  $A_k^w(x, y)$ .  $A_k^w(x, y)$  is the probability worker  $x$  accepts a job at firm  $y$ . We then update the acceptance strategy in the following way.

**Mixed strategy**

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

$$A_{k+1}^w(x, y) = A_k^w(x, y) + 0.001 \cdot \text{rand}() \cdot (1 - A_k^w(x, y))$$

**elseif**  $S_k(x, y) < 5 \times 10^{-7}$  and  $A_k^w(x, y) > 0$

$$A_{k+1}^w(x, y) = A_k^w(x, y) - 0.001 \cdot \text{rand}() \cdot (1 - A_k^w(x, y)),$$

**end**

where  $\text{rand}()$  is a pseudo-random value drawn from the standard uniform distribution on the open interval  $(0, 1)$ . A mixed solution is found if the maximum absolute change between iterations of both the surplus and the match density are less than  $2.5 \times 10^{-6}$ . 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 correspond to higher productivity; e.g., a worker with rank 10 is better than a worker with rank 2, a firm in bin 7 is better than a firm 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 =  $X = \text{\#firm types} = Y = 50$ .
- iii) Worker ID,  $i = 1, \dots, N$ .
- iv) Rank of worker  $i$ ,  $\hat{i} = 1, \dots, N$ .

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

- v) True worker type  $x = 1, \dots, X$ . Each  $x$  has  $N/X$  individual workers.  
E.g., if  $i = 6$  has type 3,  $x(6) = 3$ . For convenience,  $x(i) = 1$  if  $i \in \{1, \dots, N/X\}$ ,  $x(i) = 2$  if  $i \in \{1 + N/X, \dots, 2N/X\}$  and so on. In our estimation of the assignment of individual workers to worker types,  $\hat{x}$ , we use no information on the true assignment  $x$ .
- vi) Estimated worker type (worker bin)  $\hat{x} = 1, \dots, X$ . Each  $\hat{x}$  has  $N/X$  workers.  
E.g., if  $i = 5$  is in bin 45,  $\hat{x}(5) = 45$ .  
For our simulations,  $\hat{x}(i) = 1$  if  $\hat{i} \in \{1, \dots, N/X\}$ ,  $\hat{x}(i) = 2$  if  $\hat{i} \in \{1+N/X, \dots, 2N/X\}$  and so on.
- vii) Firm ID,  $j = 1, \dots, J$ .  $J = N/100$ . Jobs and vacancies sum to 100 at all  $j$ .
- viii) Rank of firm  $j$ ,  $\hat{j} = 1, \dots, J$ .  
E.g., if  $j = 4$  has rank 10,  $\hat{j}(4) = 10$ .
- ix) True firm type,  $y = 1, \dots, Y$ . Each  $y$  has  $N/(100 \cdot Y)$  unique  $j$ 's. Denote  $J_Y \equiv J/Y = N/(100 \cdot Y)$ .  
E.g., if  $j = 4$  has type 10,  $y(4) = 10$ .
- x) Estimated firm type (firm bin)  $\hat{y} = 1, \dots, Y$ . Each  $\hat{y}$  has  $J_Y$  unique  $j$ 's.  
E.g., if  $j = 4$  is in bin 10,  $\hat{y}(4) = 10$ .

First we take simulated matched employer-employee datasets and rank workers using the algorithm described in Appendix III. The algorithm delivers the ranking of workers  $\hat{i}(i)$  and the estimated worker type  $\hat{x}(i)$ . At each firm  $j$ , we observe all workers  $i$  matching with this firm and we have their estimated type  $\hat{x}$ . This gives us an estimate of the set of worker types matching with this firm  $j$ , i.e. we obtain an indicator function  $\hat{\mathcal{B}}(\hat{x}, j)$ , which is one if firm  $j$  hires a worker of type  $\hat{x}$  and is zero otherwise. We now want to refine this estimate of which types match with firm  $j$ . The reason is that whereas we observe whether a worker  $i$  works at a firm  $j$  in the data his type  $\hat{x}(i)$  is just estimated, potentially with error due large measurement error in wages. To take this into account we now provide an algorithm to detect misranked workers. We then exclude the wage histories of these misranked workers. Using **IDNoise** we locate matches that are likely caused by very noisy wage histories. We include all these workers in the set  $\hat{\mathcal{N}}$ .<sup>27</sup> This algorithm also updates the estimate of the set of workers types matching with firm  $j$ ,  $\hat{\mathcal{B}}(\hat{x}, j)$ , by excluding those estimated types of workers who are included in  $\hat{\mathcal{N}}$ .

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

<sup>27</sup>The fraction of workers excluded is small (less than 5%) for most parameterizations.

Construct  $p(\hat{x}, j), \pi(\hat{x}, j)$  and  $N(j)$ .<sup>28</sup>

**for each firm  $j$**

    Compute  $F(p(\hat{x}, j); N(j), \pi(\hat{x}, j))$ .<sup>29</sup>

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

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

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

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

        Return to \*.

**else**

**if**  $(\hat{\mathcal{B}}(\hat{x} + 1, j) = 0 \mid \hat{\mathcal{B}}(\hat{x} - 1, j) = 0)$

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

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

            Return to \*.

**end**

**end**

**end**

**end**

**end**

$i \in \hat{\mathcal{N}}$  if a firm  $j$ , which matches with  $i$ , exists such that  $\hat{\mathcal{B}}(\hat{x}(i), j) = 0$ .

**return**  $[\hat{\mathcal{B}}(\hat{x}, j), \hat{\mathcal{N}}]$

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**  $\hat{x}$

---

<sup>28</sup> $p(\hat{x}, j)$  is the number of workers of estimated type  $\hat{x}$  hired by firm  $j$ .  $N(j) = \sum_{\hat{x}} p(\hat{x}, j)$  is the total number of workers actually hired by firm  $j$  which sums over all types from the matching set of firm  $j$ . The theoretical fraction of workers of type  $\hat{x}$  hired by firm  $j$  over all workers hired by  $j$  is

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

<sup>29</sup>The probability of observing at most  $p(\hat{x}, j)$  given the hiring probability  $\pi(\hat{x}, j)$  from  $N(j)$  trials is

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

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

**foreach**  $j \in J(\hat{x})$ , compute  $\bar{w}(\hat{x}, j) = \text{average wage paid by } j \text{ to workers } i \in \hat{x}$ .

$w_{res}(\hat{x}) = \text{lowest average of } \bar{w}(\hat{x}, j) \text{ possible from pooling } J_Y \text{ firms in } J$ .<sup>30</sup>

**end**

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

Then, for each firm  $j$ , compute the average wage premium as in (12). 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  $\hat{\mathcal{B}}$  of firm  $j$ , which includes all types  $\hat{x}$  for which  $\hat{\mathcal{B}}(\hat{x}, j) = 1$ . Our estimate of  $\hat{q}(j)$  is  $\hat{\mathbb{M}}_v \sum_{\hat{x} \in \hat{\mathcal{B}}} \frac{u(\hat{x})}{u}$ . Multiplying the average wage premium and the acceptance rate gives the statistic  $\hat{\Omega}$  which allows us to rank firms.

We now assign individual firms to firm types  $\hat{y}$ . Using our ranking of firms, we can assign the first  $J_Y$  firms to firm bin 1, the next  $J_Y$  firms to firm bin 2, and so on. The assignment of firms to types allows us to compute statistics for firm types only. For example, statistics for all firms belonging to firm type  $\hat{y} = 1$  will be the firm size (measured by average employment) weighted average of firms with  $\hat{j}(j) \in \{1..J_Y\}$ . This step only serves to aggregate information across firms and yields smoother statistics and better estimates. We could have also proceeded by assigning each individual firm to its own type, i.e.  $\hat{y}(j) = \hat{j}(j)$ .

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  $\hat{\Omega}(\hat{y})$ . 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 aggregated estimated output equals the observed true frictional aggregate output.

To measure output losses due to frictions we optimally assign a sub-sample (5000 workers and 5000 jobs) from the pool of employed workers.<sup>31</sup> 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 search frictions is compared the same procedure repeated using true model generated distributions and production functions.

<sup>30</sup>For  $\hat{x} > 1$ , we additionally impose  $w_{res}(\hat{x}) > w_{res}(\hat{x} - 1)$  which is consistent with theory.

<sup>31</sup>See Section 4.6 for references to the algorithms used.



### 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 among 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, wage data might contain 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. This requires then the specification of how to aggregate disagreements between voters and a method how to find this aggregate ranking.

#### 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.<sup>32</sup> The goal then is to find an aggregate ranking which generates the minimum number of disagreements 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

---

<sup>32</sup>This is first described in Kemeny (1959) and Kemeny and Snell (1963)

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

$$\sum_{i>j} c(i, j)\Pi(i, j) + c(j, i)\Pi(j, i), \quad (\text{A3})$$

where the weighting  $c(i, j)$  is the probability (computed from wage observations) that  $i$  is ranked above  $j$ .

We now construct  $c(i, j)$ . 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  $n_{i,k}$  wage observations and  $n_{j,k}$  from workers  $i$  and  $j$  respectively at firm  $k$ .<sup>33</sup> We know that observed wages follows:

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

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

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

where all of the  $\epsilon$ 's are independent.

We are interested in computing the probability that  $w_{i,k} > w_{j,k}$  given the observations on  $\bar{w}_{i,k}$  and  $\bar{w}_{j,k}$ . 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:

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

The posterior density over  $w_{i,k}$  conditional on knowing  $\sigma^2$  (we explain below how to measure

---

<sup>33</sup> $n_{i,k}$  periods (months, in our case) of observations need not be in one employment spell. Moreover,  $i$  and  $j$  do not need to be employed at the same time.

it in the data) is given by:

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

where

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

and

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

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

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

and

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

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

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

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

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

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

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

$$\begin{aligned} \mathbb{P}(w_{i,k} > w_{j,k}) &= 1 - \Phi\left(\frac{0 - (\bar{w}_{i,k} - \bar{w}_{j,k})}{\frac{\sigma_{i,k}^2}{n_{i,k}} + \frac{\sigma_{j,k}^2}{n_{j,k}}}\right) \\ &= \Phi\left(\frac{\bar{w}_{i,k} - \bar{w}_{j,k}}{\frac{\sigma_{i,k}^2}{n_{i,k}} + \frac{\sigma_{j,k}^2}{n_{j,k}}}\right). \end{aligned}$$

If  $i$  and  $j$  are on the same payroll at only 1 firm,  $\mathbb{P}(w_{i,k} > w_{j,k}) = c(i, j)$ . If more than one firm hires  $i$  and  $j$ , we compute  $\mathbb{P}(w_{i,k} > w_{j,k})$  for all those firms and assign the product of

these probabilities to  $c(i, j)$ , i.e. we consider observations in different firms as independent.

The variance of noise is computed from the variance of wages for all workers within jobs since all variation in wages for a specific job arises from measurement error only.

The solution to the problem of finding the best ranking is then conceptually trivial: (1) Enumerate all possible rankings; (2) Evaluate (A3) for all of them; (3) Select the rank which maximizes the objective function. Unfortunately, the *Kemeny-Young rank aggregation* problem is NP-hard.<sup>34</sup> We therefore approximate the solution to the problem and use the following algorithm:

**Algorithm 3. *Single worker moves***

**Initialize**  $\Pi(i, j)$  that maximizes (A3). Choose ranking from:

- (a) lowest wage,
- (b) highest wage,
- (c) adjusted average wage,

**While** some rearrangement of  $\Pi(i, j)$  improves (A3)

Choose worker  $x$  and rank  $j$ .

Rearrange  $\Pi(i, j)$  so that  $x$  has rank  $j$ , leaving all other relative rankings intact.<sup>35</sup>

**Return**  $\Pi(i, j)$

This algorithm is a simplified version of the algorithm in Kenyon-Mathieu and Schudy (2007) which is capable of approximating the solution arbitrarily well. We choose this algorithm as it provides for us the best compromise between accuracy and computational complexity. Indeed, we show that this simplified algorithm provides a very accurate ranking of workers in our model. It is straightforward to implement the complete algorithms in Kenyon-Mathieu and Schudy (2007) if more precision is required for a particular application.

---

<sup>34</sup>See Bartholdi, Tovey, and Trick (1989). Consider a simple case of a 100 candidates and at least 4 submitted rankings. There are  $100 \times 99 \dots \times 2$  combinations to consider!

<sup>35</sup>Suppose there are workers,  $A, B, C$  and  $D$  ranked alphabetically,  $\{A, B, C, D\}$ . Moving  $C$  to rank 2 would mean rearranging them so that the ranking is now  $\{A, C, B, D\}$ .

## IV Appendix Figures

Figure A-1: Distributions of selected variables of interest across all parameterizations with PAM, NAM, and NEITHER production functions.

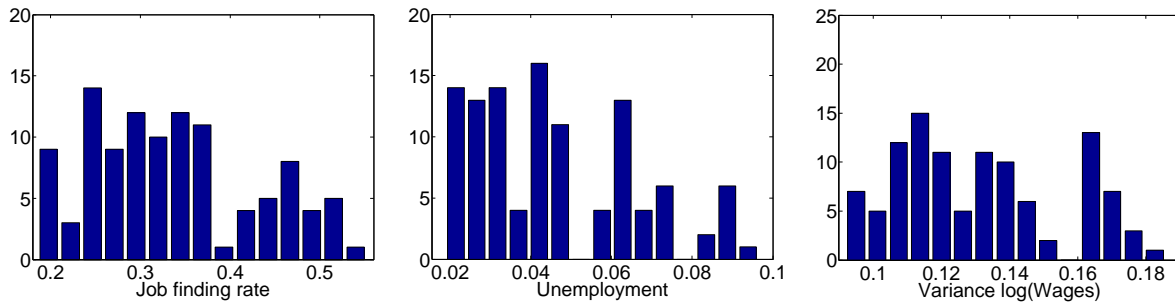


Figure A-2: True and estimated PAM production function.

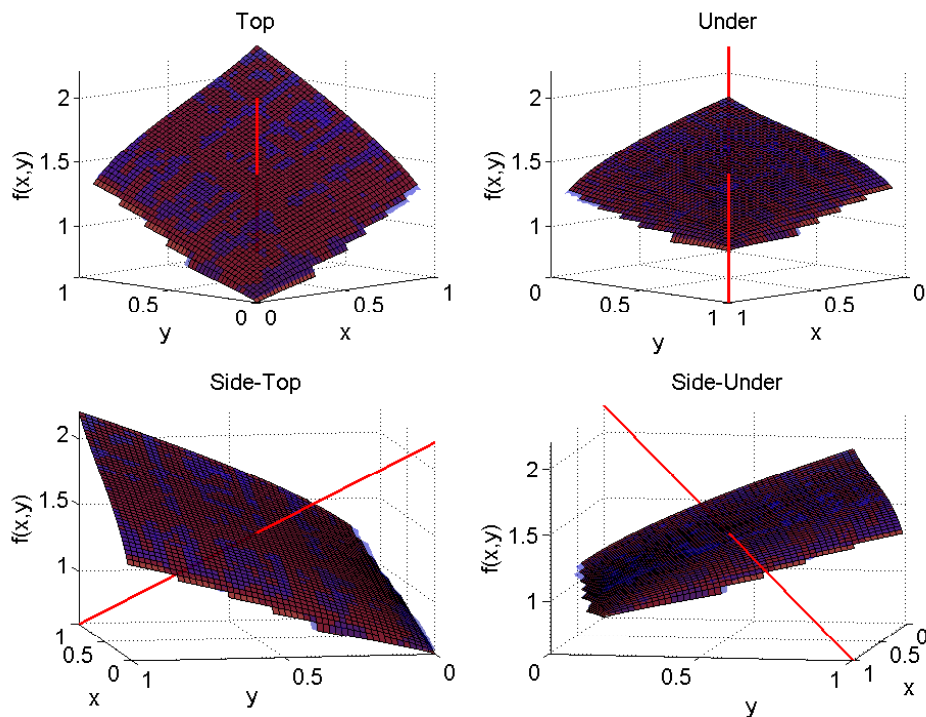


Figure A-3: True and estimated NAM production function.

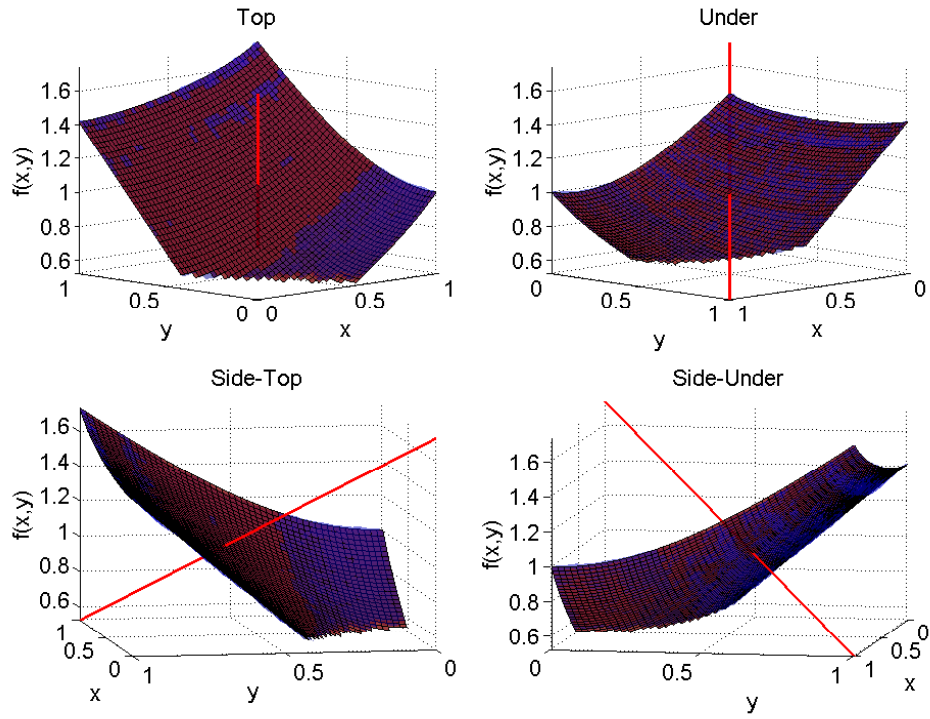


Figure A-4: True and estimated Neither NAM nor PAM production function.

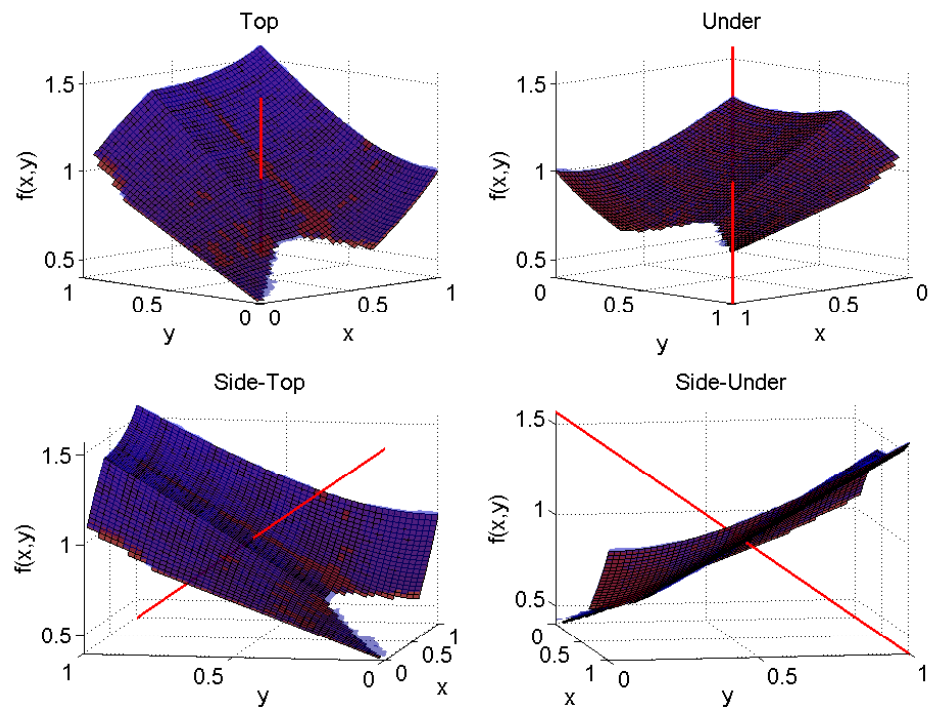


Figure A-5: Distributions of selected variables of interest across all parameterizations with Gautier and Teulings production function.

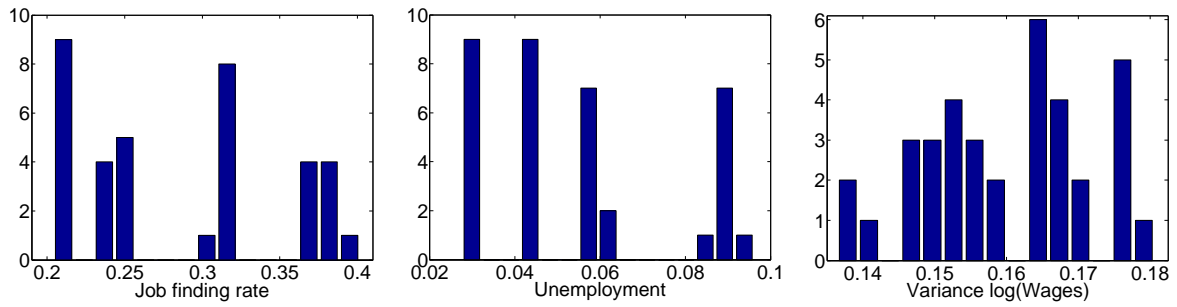


Figure A-6: True and estimated GT production function.

