Identifying Sorting*

Marcus Hagedorn†
Institute for Advanced Studies

Tzuo Hann Law ‡
University of Pennsylvania

Iourii Manovskii§
University of Pennsylvania

August 24, 2012

Abstract

We develop a methodology to quantitatively study sorting between workers and firms based on their unobserved (to economists) productive characteristics. It is often hypothesized that sorting on unobservables is an important determinant of output, productivity and wage differences at the micro and macro levels. Sorting on unobservables is thought to be substantially more important than sorting on observable attributes that are typically found to account for only a relatively small amount of variation in variables of interest. For example, sorting on unobserved productivity differences is offered as a possible explanation for the persistent differences in wages across employers of different size, industry affiliation, or exporting status. Key questions in macro involve measuring the degree of mismatch between worker skills and technology and the evolution of this mismatch over the business cycle. Restricting the analysis to observed skills provides at best only an incomplete set of answers. In the growth literature the distribution of observed and unobserved skills is an important determinant of technology adoption and the extent of misallocation is thought to be an important determinant of aggregate TFP. Yet, despite the manifest importance of these questions, the literature currently lacks a theoretically consistent way to measure the unobserved productivities of workers and firms and the patterns of sorting between them. We fill this gap in the literature. This allows to provide answers to the questions mentioned above and many others, discussed in the main text.

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

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

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

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

1 Introduction

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

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

[1] Gautier and Teulings (2006) were the first to establish this in a model of sorting based on comparative advantage. This important class of models violates the underlying assumption of the fixed effect regression that workers and firms are globally rankable. Eeckhout and Kircher (2011) later make an even stronger point. They prove that even in the model of sorting on absolute advantage that allows for globally rankable workers and firms, the worker and firm

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

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

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

fixed effects in wages have no economic interpretation. These theoretical insights have been confirmed quantitatively in a range of assortative matching models in Lopes de Melo (2009), Lentz (2010), and Lise, Meghir, and Robin (2011), among others.

spelled out below).

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

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

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

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

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

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

The paper is organized as follows. In Section 2 we describe the standard model with frictional labor market and assortative matching between between workers and firms. 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. Section 5 concludes. Most proofs and details of computations are in the Appendix.

2 The Model

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

2.1 Environment

2.1.1 Basics

Time is discrete, all agents are infinitely-lived and maximize the present value of payoffs, discounted with a common discount factor $\beta \in (0, 1)$. The unit mass of workers is either employed (e) or unemployed (u) while the unit mass of firms is producing (p) or vacant (v). Workers and firms are heterogeneous with respect to their productivities, denoted by $x \in [0,1]$ and $y \in [0,1]$, respectively. To simplify the exposition we treat each firm as having one job. All the results immediately generalize, however, to each firm having a mass of jobs sharing the same productivity y.[2]

Output of a match between worker x and firm y is given by the twice differentiable nonnegative production function $f : [0,1]^2 \to \mathbb{R}_+$. The existence proof in Shimer and Smith (2000) also requires that f has uniformly bounded first partial derivatives on $[0,1] \times [0,1]$. It is assumed that match output is increasing in worker and firm type, i.e., $f_x > 0$ and $f_y > 0$.[3] 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 types are F and of firm types are G and the "original" production function is $f(x, y)$ then we transform the production function

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

and the distributions are

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

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

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

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

2.1.2 Distributions

The measures characterizing the set of matched and unmatched workers and firms are assumed to be absolutely continuous, implying the existence of a density. Given our identification of types with ranks, the worker and firm time invariant populations are given by $d_w = 1$ and $d_f = 1$. The distribution of producing matches is described by $d_m : [0,1]^2 \to \mathbb{R}_+$. The densities 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.[4] Table 1 summarizes the relationships between these functions.

Table 1: Functions describing distributions

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

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

$$E = \int d_m(x,y) \ \mathrm{d}x\mathrm{d}y = \int d_e(x) \ \mathrm{d}x, \tag{1}$$

$$P = \int d_m(x,y) \ \mathrm{d}x\mathrm{d}y = \int d_p(y) \ \mathrm{d}y, \tag{2}$$

$$U = 1 - E = 1 - \int d_m(x,y) \ \mathrm{d}x\mathrm{d}y = \int d_u(x) \ \mathrm{d}x, \tag{3}$$

$$V = 1 - P = 1 - \int d_m(x,y) \ \mathrm{d}x\mathrm{d}y = \int d_v(y) \ \mathrm{d}y. \tag{4}$$

2.1.3 Timing

It is convenient to think of each period as consisting of two subperiods. In the first subperiod, a worker of type x matched with a firm of type y produces $f(x,y)$. Output of this match is exhausted by payments to the firm, $\pi(x,y)$, and the worker, $w(x,y)$. Vacant firms pay vacancy maintenance costs, c. Unemployed workers obtain flow payoff from non-market activity b. In the second subperiod,

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

new matches are formed when all unmatched workers and firms participate simultaneously in a single labor market subject to search frictions. After matching, each existing match (including a newly formed one) is destroyed with probability δ.[5]

2.2 Search and matching

Only and all unmatched agents engage in random search. A function $m : [0,1] \times [0,1] \to [0, \min(U,V)]$ takes the masses of unemployed workers U and vacant firms V as its inputs and generates meetings. The probability a worker meets a potential employer is given by $\mathbb{M}_u = \frac{m(U,V)}{U}$, while the probability of a vacant firm meeting a potential hire is $\mathbb{M}_v = \frac{m(U,V)}{V}$. These probabilities are time-invariant in the steady-state equilibrium we will consider. The probability for a worker to meet a firm $y \in Y \subset [0,1]$ equals $\mathbb{M}_u \frac{\int_Y d_v(y)dy}{V}$. The probability for a firm to meet a worker $x \in X \subset [0,1]$ equals $\mathbb{M}_v \frac{\int_X d_v(x)dx}{U}$. These probabilities reflect our assumption of a linear search technology. In SS's quadratic search technology these probabilities would be $\mathbb{M}_u \int_Y d_v(y)dy$ and $\mathbb{M}_v \int_X d_v(x)dx$, respectively. Since we obtain the SS 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 with which firm to match, 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 (see SS that the same type agents use the same strategy). Symmetrically for firms, the Borel measurable *acceptance set* $A^f(y)$ are the workers that a firm of type y is willing to match with. Matching takes place when both the worker and the firm find it mutually acceptable. For a worker of type x, the *matching set* $B^w(x)$ are firms which accept worker type x and are accepted by worker type x. Symmetrically for a firm of type y, $B^f(y)$ are workers who accept to match with firm type y and who are accepted by firms of type y. The

[5]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.

matching sets are related to the acceptance sets in the obvious way:

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

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

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

2.4 Bellman equations and surplus sharing

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

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

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

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

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

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

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

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

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

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

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

2.5 Stationary distribution of matches

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

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

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

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

2.6 Equilibrium

In a *steady state search equilibrium* (SE) all workers and firms maximize expected payoff, taking the strategies of all other agents as given.[6] The economy is in steady-state. A SE is then characterized by the density $d_u(x)$ of unemployed workers, the density $d_v(y)$ of vacant firms, the density of formed matches $d_m(x,y)$ and wages $w(x,y)$. The density $d_m(x,y)$ implicitly defines the matching sets as it is zero if no match is formed and is strictly positive if a match is consummated. Wages are set to ensure the surplus sharing rule (9) and match formation is optimal given wages w, i.e. a match is formed whenever the surplus is (weakly) positive (see (10)). The densities $d_u(x)$ and $d_v(x)$ ensure that the flow equations in (15) hold.

[6] As SS we assume that a matched is formed if agents are indifferent.

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

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

The equilibrium existence proof of SS uses their assumption of a quadratic matching function, whereas our matching technology is linear. Fortunately, Nöldeke and Tröger (2009) extend the existence proof of SS to our linear matching technology:[7]

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

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

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

where μ is the Lebesgue measure. Imposing

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

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

3 Identification: Theory

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.

[7] They also argue that the proofs extend to our non-symmetric environment.

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

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

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

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

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

Result 4. *The adjusted average wage defined as*

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

is increasing in x.

Formal proofs of these results can be found in Appendix I, although the results themselves are intuitive. The fact that the value of unemployment is increasing in worker's type follows because a more productive worker can always imitate the acceptance strategy of the less productive worker but produce more and consequently receive higher wages. This induces a more productive worker to set a higher reservation wage. The fact that wages within firms are increasing in workers type follows directly from the assumption that the production function is increasing in worker productivity. It is easy to prove that the average wage (without) the adjustment is not necessarily increasing in x.[8] To see this, consider two workers with different productivity. A more productive worker might be matching with a wider set of firms (some of which do not accept the less able worker). However, the more able worker might be only marginally acceptable to those firms who typically match with even better workers. As a consequence, those firms pay low wages to this worker. Thus, the average wage of the worker over the employment history might be lower then that of a less productive worker. The more productive worker still obtains the higher value because he spends a larger fraction of his lifetime employed. Result 4 corrects for this effect by imputing the value of unemployment to

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

$$\frac{\partial}{\partial x} \int\limits_{B(x)} w(x, y)\text{đ}_v(y) \, \mathrm{d}y = \int\limits_{B(x)} \frac{\partial w(x, y)}{\partial x} \text{đ}_v(y) \, \mathrm{d}y + \overline{\varphi}'(x) w(x, \overline{\varphi}(x)) \text{đ}_v(\overline{\varphi}(x)) - \underline{\varphi}'(x) w(x, \underline{\varphi}(x)) \text{đ}_v(\underline{\varphi}(x)). \quad (19)$$

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

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 accepted wages, and adjusted average wages provide theoretically valid and equivalent rankings of workers. Their performance of these ranking procedures might differ, however, in small samples and in presence of measurement error in wages. We assess their quantitative performance in Section 4 and construct a single measure that aggregates the information implied by all the ranking procedures developed here. 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 statistics which are monotonically increasing in firm type y. This is non-trivial since the wage of worker x, $w(x, y)$, is not always increasing in firm productivity. The same problem applies to the surplus of a match, $S(x, y)$. The strategy is as follows. We first establish that value of a vacancy is increasing in y. This implies that the surplus a vacancy is expected to generate is also increasing in y. Any bargaining game where both parties benefit from an increase in the surplus implies that the average surplus of workers employed by firm y is also increasing in y. Finally, we show that the average surplus of workers employed by firm y can be expressed as a function of wages, yielding a simple observable statistic that is increasing in y and thus allows to rank firms. In this Section, we include some of the proofs in the main text as we consider them instructive (and surprisingly simple).

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

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

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

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

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

is increasing in y.

Proof of Result 6. Using equation (9),

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

From (12) it follows that

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

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

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

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

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

$$\Omega(y) = (1-\delta)\mathbb{M}_v \int\limits_{B(y)} \frac{d_u(x)}{U} \left(w(x,y) - w(x,\tilde{y}(x))\right) \, \mathrm{d}x \tag{20}$$

is increasing in y.

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

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

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

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

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

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

Using Result 6 completes the proof.

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

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

The probability that a vacancy of type y is filled equals

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

It then holds that

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

The empirical counterpart of Ω^e is defined as

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

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

3.2.1 Ranking Firms: Special Cases of PAM and NAM

Result 7 enables us to rank firms in terms of their productivity. Note that Ω is increasing in y regardless of whether the model features positive or negative assortative matching, or indeed neither. In particular, it does not require any assumptions on the production function f, i.e. neither super- nor sub-modularity.[9] If we however assume a special case of PAM or NAM (see Definition 1), we can establish another result that can be used to refine the ranking.

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

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

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

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

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

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

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

3.3 Sign and Strength of Sorting

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

3.3.1 Relationship to the Literature

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

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

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

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

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

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

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

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

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

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

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

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

3.4 Identifying Remaining Model Parameters

Estimated values are denoted with a hat, ˆ.

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

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

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

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

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

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

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

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

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

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

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

3.4.2 Measuring α

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

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

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

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

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

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

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

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

Our estimate is then

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

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

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

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

which also differs from the true value by a constant.

3.4.4 Transition Rates

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

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

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

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

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

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

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

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

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

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

Denote by $H_t(y)$ the observed number of new hires in firms of type y at time t, and by $V_t(y)$ the unobserved number of vacancies posted by these firms. Equation (28) and the law of large numbers imply that

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

Adding up across all firms and time periods, and rearranging yields an estimate for $\mathbb{M}_v$:

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

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

3.4.5 Measuring output $f(x, y)$

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

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

and, as a result, profits (per period) equal

$$\hat{\pi}(x, y) = \hat{f}(x, y) - w(x, y).$$

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

4 Implementation and Quantitative Evaluation

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

4.1 Parameterization

We assume that a researcher has access to a matched employer-employee panel data set. We conduct the analysis with a data set that has a time dimension of 15 years. This is a conservative choice because the longer the data set the more precise our method is. Most currently available and commonly used matched data sets (e.g., from Brazil, Denmark, Germany, France) have a longer time span. We assume that the data includes the information on wages, all employment and unemployment spells of the worker over the duration of the sample, and all hires and separations at the firm level. We simulate the model at a monthly frequency.

We consider three production functions:

i) $f(x, y) = 5xy + 1.1$, which induces positive assortative matching,

ii) $f(x, y) = 5((x+1)(1-y/2) + y(2-x/2)) - 6.5$, which induces negative assortative matching, and

iii) $f(x, y) = ((1 - |x - 0.1|)(1 - |y - 0.2|) + x + y) - 0.7$, which induces neither positive nor negative assortative matching. Instead, the pattern of sorting changes over its domain.

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

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

We set the discount factor to 0.996 at monthly frequency to be consistent with the annual interest rate of 4%. We also conduct a complete set of simulations assuming a much larger discount factor of 0.999, which implies an annual interest rate of about 1%. This choice is motivated not as much by its empirical relevance but by the observation that identification is more difficult when the discount factor is closer to one.

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 $\kappa = \{0.5, 0.8\}$ to generate the job finding probabilities ranging between a high of about 45% a month in the US and a low of about 10% in some European countries. Similarly, we choose two values for the separation rate $\delta = \{0.01, 0.03\}$, 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 two standard deviations around the mean of zero.

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

Table 2: Parametrization

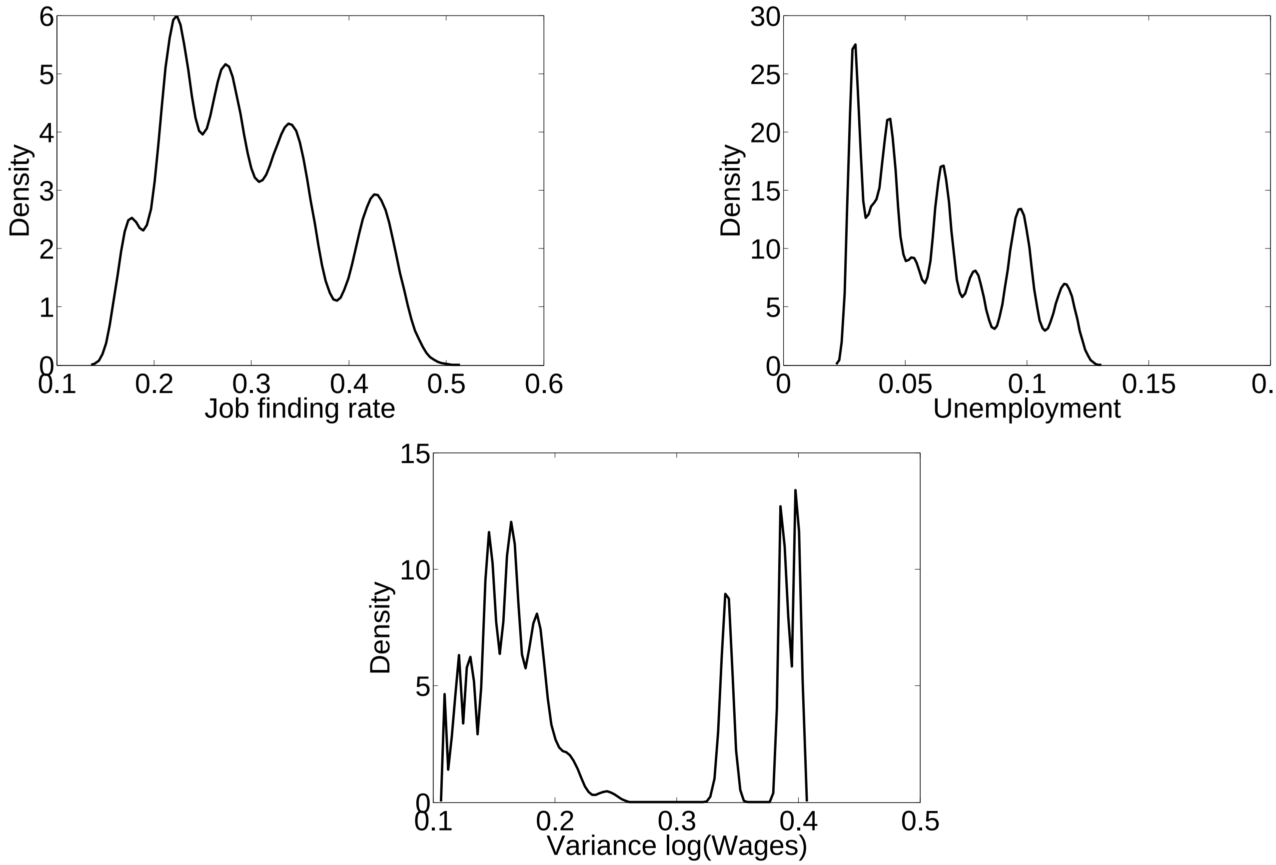
Parameter	Symbol	Option 1	Option 2	Option 3
Production function	$f(x, y)$	PAM	NAM	Neither
Worker distribution	d_w	Uniform	Normal	Bi-Modal
Firm distribution	d_f	Uniform	Normal	Bi-Modal
Discount factor	β	0.996	0.999	
Separation rate	δ	0.01	0.03	
Meeting function scale	κ	0.5	0.9	
Meeting function elasticity	ν	0.5		
Worker's bargaining weight	α	0.5		
Worker's flow utility of unemp.	b	0.0		
Vacancy cost	c	0.0		
Measurement error in wages	ϵ	20% of overall wage variance		

Thus, all combinations of parameters result in 216 distinct parameterizations that we consider. Figure 1 summarizes the range of values that a number of variables of interest take across all simulations. Most tend to lie within empirically plausible ranges.

4.2 Ranking of Workers

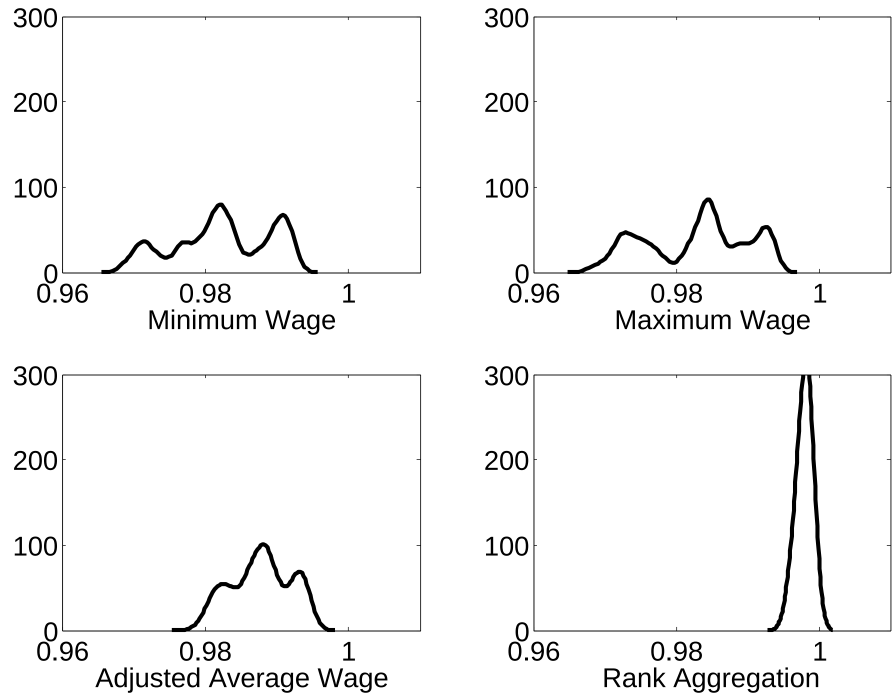
Results 2, 3 and 4 have established that in large samples workers can be ranked based on the lowest accepted wage, the highest accepted wage, or the adjusted average wage. To assess the performance of each of these methods in small samples and in the presence of measurement errors in wages we report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

Figure 1: Non-parametric plots of selected variables of interest across all parameterizations.



In addition, Result 1 implies that wages within a firm are increasing in worker productivity x, which provides another way to rank workers according to their productivity. However, the presence of measurement error in wages dilutes this way of ranking. Within one firm one worker could be ranked higher than another worker not because he is more productive (actually he is less productive) but just because of the measurement error. And the ranking between these two workers may not be consistent with the ranking from other firms. To solve this problem, we build on the insights from social choice theory, where voters rank candidates, potentially inconsistent with each other. In our application voters correspond to firms and workers to voting alternatives. A disagreement arises if a ranking of two workers is the opposite of each other in two firms. The goal is then to find a ranking that minimizes the number of disagreements which defines the *Kemeny-Young rank aggregation* problem first described in Kemeny (1959) and Kemeny and Snell (1963). We refine this problem by weighting the disagreement between two alternatives of each voter by the likelihood that this ranking is due to measurement error. Rank aggregation is a classic problem in social and

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



natural sciences, particularly due to its computational complexity. We use use recent advances in computer science and algorithm research to tackle this problem. An overview of the problem and available methods is provided in Appendix III. Here we provide a brief summary.

We represent the comparisons of workers based on the within-firm wage rankings as a directed graph. As this graph is not necessarily connected, we complete the ranking, i.e., turn the problem into a tournament, by using the ranking based on minimum, maximum, and adjusted average wages. We search for a ranking that has the minimum amount of disagreement with each of the input rankings. This problem is equivalent to removing the minimum feedback arc set of the tournament. Kenyon-Mathieu and Schudy (2007) provide a polynomial time algorithm that can approximate the solution to this problem with arbitrary degree of accuracy. We use a version of their algorithm that is reasonably fast and achieves a high level of accuracy.

Figure 2 reports kernel density of the correlation between the estimated and true worker ranks across all simulations. The correlations are high for all measures. In particular, the adjusted average wage dominates both the minimum and the maximum wage in its performance in ranking workers. The combined measure outperforms any of the individual measures. Its performance advantage appears small because the correlations between true and estimated ranks is very high to begin with.

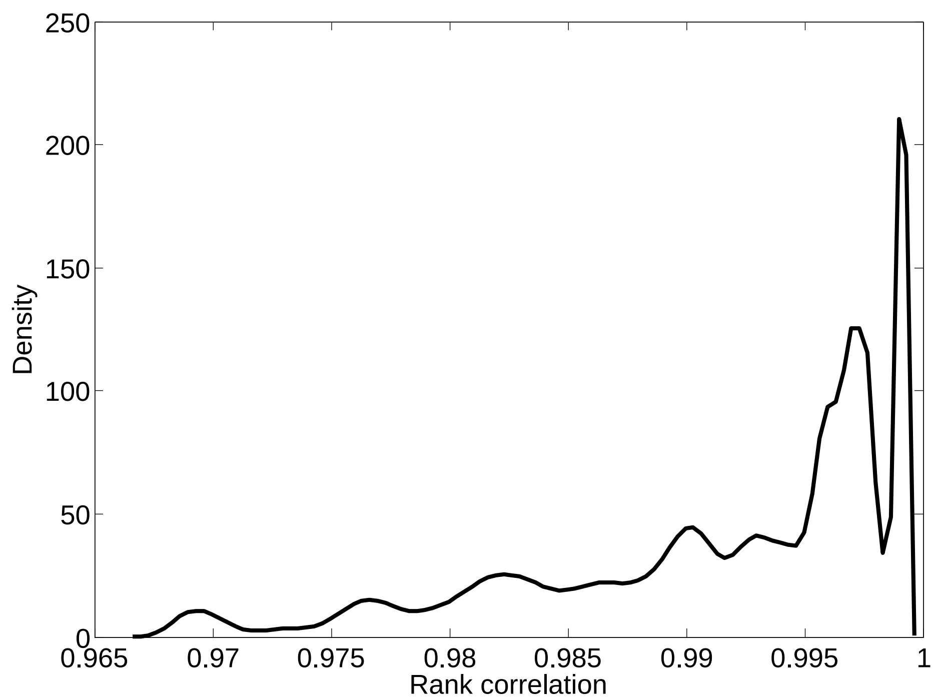
However, its performance advantage becomes clearer when we use it in estimating the production function.

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 the variance of which can be estimated on the sample of job stayers.

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

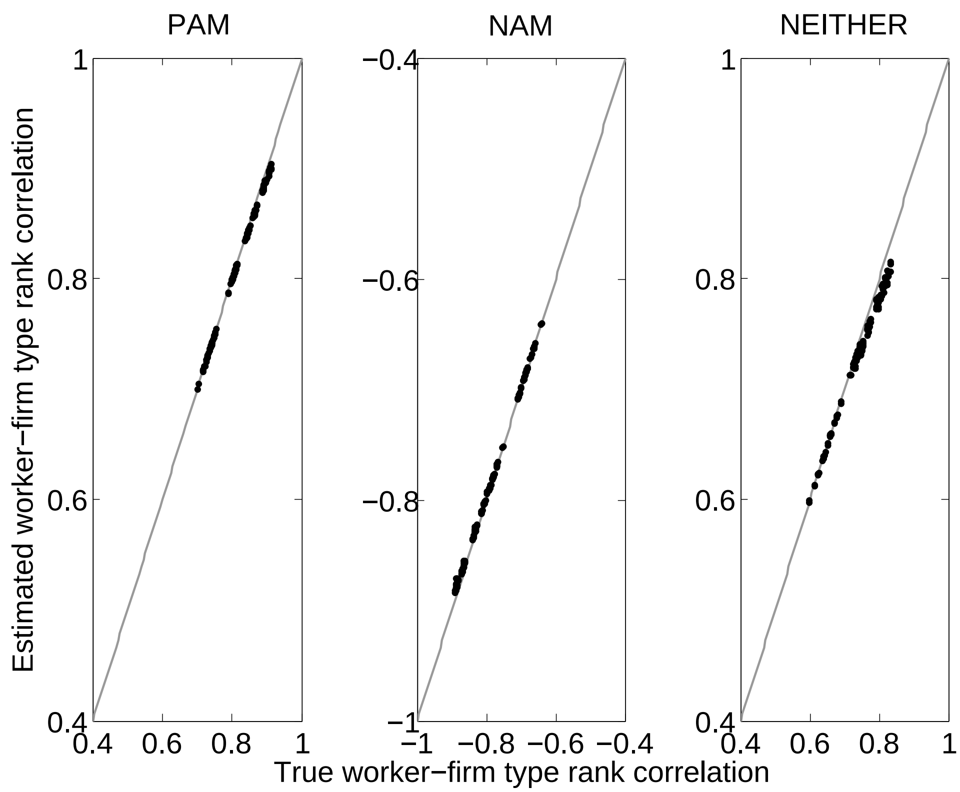
Figure 3: Density of the correlation between the estimated and true firm ranks across parameterizations.



4.4 Strength of Sorting

Figure 4 plots the correlation between identified worker and firm ranks against true correlation. Here we separate the three production functions to illustrate that or identification strategy easily identifies the sign of sorting. In all cases this crude measure of the strength of sorting performs quite well.

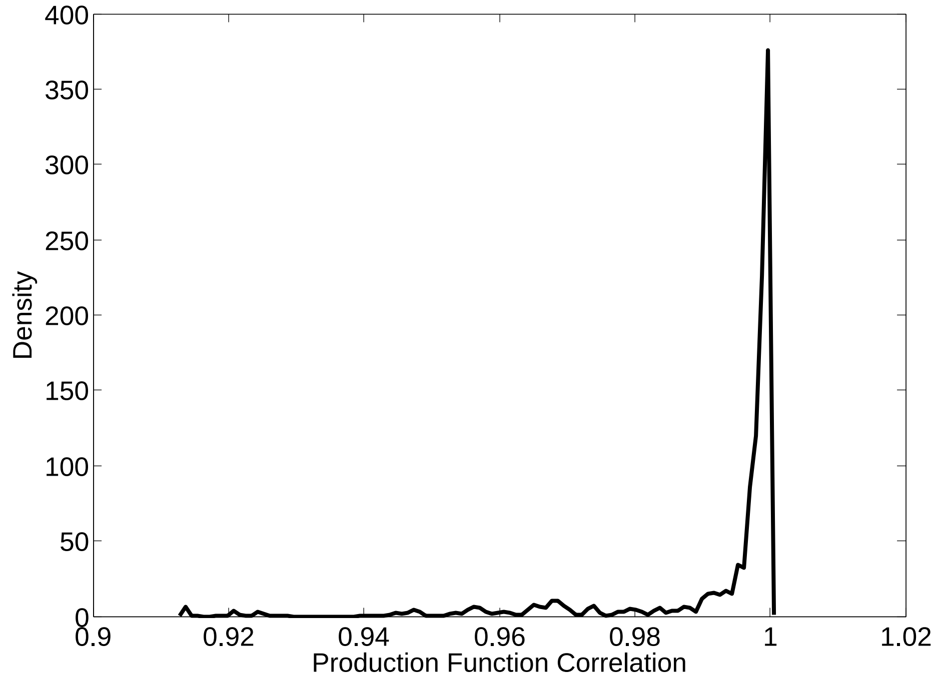
Figure 4: Correlation between identified worker and firm ranks against true correlation.



4.5 Measuring $f(x, y)$

To evaluate how well our method recovers the production function, in Figure 5 we plot the density of correlation between the true and estimated production functions across all parameterizations. The correlations are generally very high.

Figure 5: Correlation between true and estimated production functions.



To provide a better sense of the ability of our identification and implementation strategy to recover the production function, in Figures 6 - 8 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: $\beta = 0.996$, $\kappa = 0.5$, $\delta = 0.03$, $\alpha = 0.5$, $b = c = 0$ and $\mathcal{d}_w = \mathcal{d}_f = U[0, 1]$, with the measurement error equal to 20% of the variance of wages. Each figure contains the true production function (dark red if viewed in color) and the estimated one (with the color transitioning from blue to yellow over the domain). 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.

Figure 6: True and estimated PAM production function.

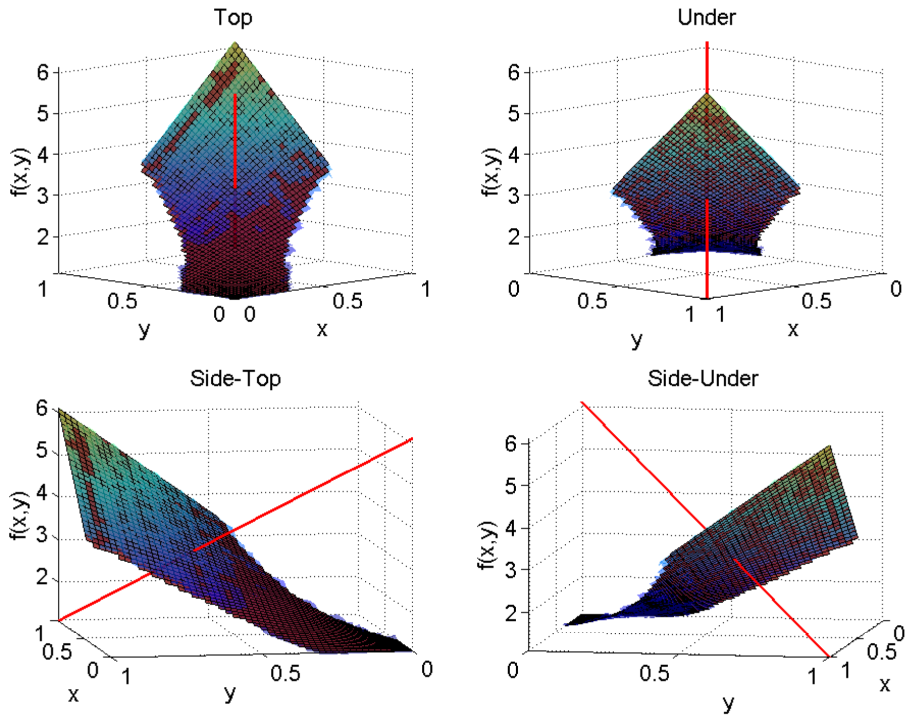


Figure 7: True and estimated NAM production function.

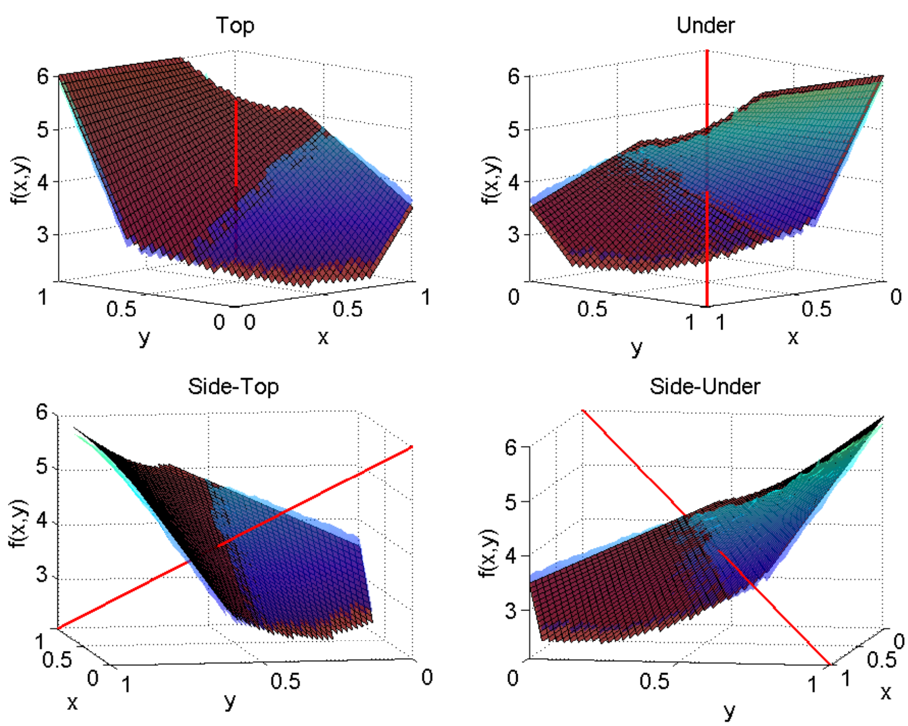
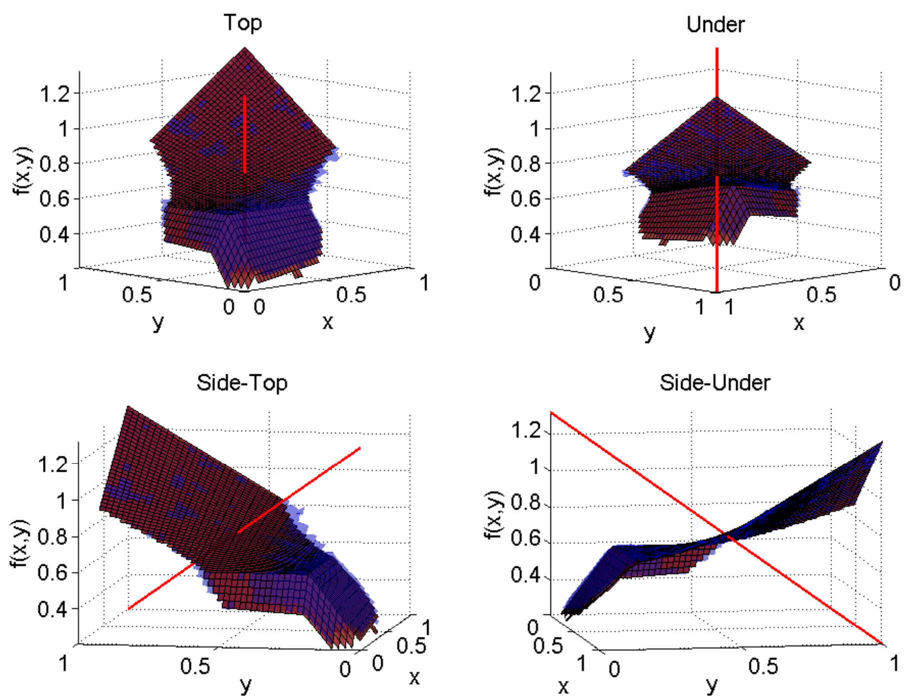


Figure 8: True and estimated Neither NAM nor PAM production function.



4.6 Measuring the effect of search frictions on output

A classic question in this literature is to assess the magnitude of output losses due to mismatch between workers and firms. We now evaluate the ability of our identification strategy to provide a reliable quantitative answer to this question.[10]

To do so we first derive the (counterfactual) allocation in a world without frictions. To solve for the frictionless assignment we need to find a one-to-one assignment (bijection) $\mu : [0, 1] \to [0, 1]$ of workers to firms such that the total output $\sum_x f(x, \mu(x))$ is maximized. This assignment problem is

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

a classic and well studied combinatorial optimization problem. Our identification strategy identifies the production function only on the set of (x, y) matches observed in the data. Since our objective is to find an optimal assignment on this set, we assume that the output outside of the observed frictional matching set is zero.

There are several existing algorithms that can solve this problem in polynomial time.[11] 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 1000 workers and 1000 jobs is sufficient. On a sample of that size we can solve the problem in seconds using the Jonker and Volgenant (1987) algorithm.

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

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

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

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

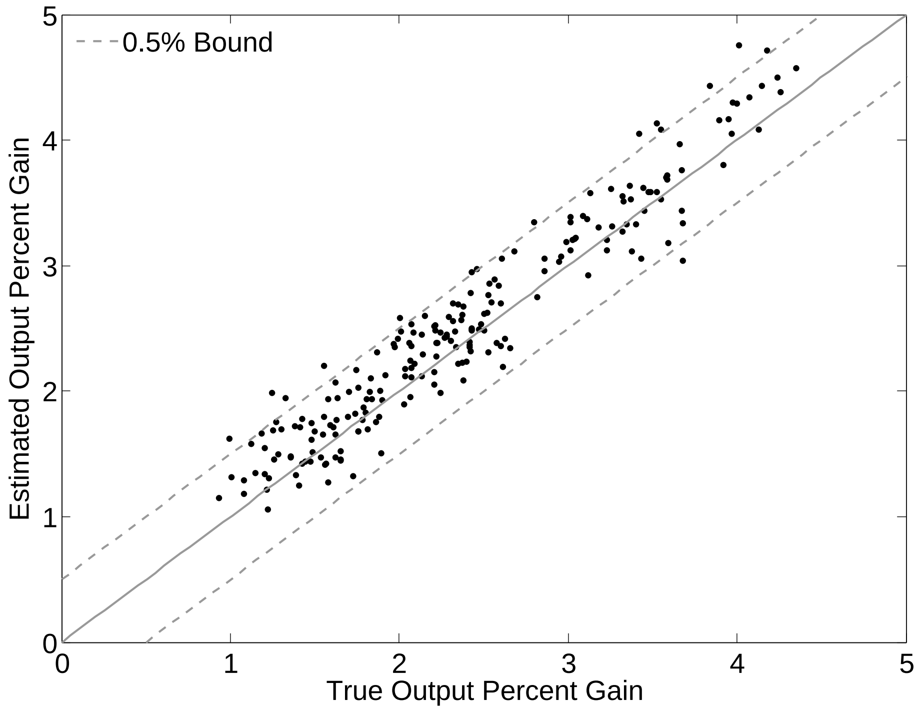
Then, the output loss due to misallocation is the difference between the expected output without frictions, $\mathcal{E}^{no\,fric}$ and the expected output with frictions, $\mathcal{E}^{fric}$:

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

In Figure 9, we plot the percent output gain from the optimal reallocation of workers $100 \cdot \frac{\Delta^{\mathcal{E}}}{\mathcal{E}^{fric}}$. The true output gain as a percent of frictional output is on the x-axis while the estimated gain as a percent of frictional output is on the y-axis. When estimating the gain from the reassignment, we use the estimated agent type as well as the estimated production function. To help interpret the results, the figure also includes two dotted lines that represents a mistake of plus or minus one half of one percent of output. We interpret the results as indicating that the method performs quite well in estimating the gain from the optimal worker reallocation.

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

Figure 9: Estimated gains from eliminating frictions.



4.7 The Role of Discounting

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

5 Conclusion

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

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

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

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

$$\mathcal{V}^{no\,fric} = \int_0^1 (f(x, \mu(x)) - \mathcal{E}^{no\,fric})^2 dx, \tag{34}$$

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

In the presence of frictions the variance of output equals

$$\mathcal{V}^{fric} = \int_{\mathcal{B}} (f(x, y) - \mathcal{E}^{fric})^2 d_m(x, y) dx dy, \tag{35}$$

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

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

$$\Delta^{\mathcal{V}} = \mathcal{V}^{fric} - \mathcal{V}^{no\,fric}. \tag{36}$$

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

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

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

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

$$Var(w) = Var(w_X) + Var(w_Y) + Var(w_{XY}), \tag{37}$$

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

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

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

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High Wage Workers and High Wage Firms," Econometrica, 67(2), 251–334.

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

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

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

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

Blanchflower, D. G., A. J. Oswald, and P. Sanfey (1996): "Wages, Profits and Rent-Sharing," Quarterly Journal of Economics, 111(1), 227–251.

Burkard, R., M. DellAmico, and S. Martello (2009): Assignement Problems. SIAM.

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

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

Eeckhout, J., and P. Kircher (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 (2012): "Search Frictions and Wage Dispersion," mimeo, University of Pennsylvania.

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

Karp, R. M. (2010): "Reducibility Among Combinatorial Problems," in 50 Years of Integer Programming 1958-2008, ed. by M. Janger, T. M. Liebling, D. Naddef, G. L. Nemhauser, W. R. Pulleyblank, G. Reinelt, G. Rinaldi, and L. A. Wolsey, pp. 219–241. Springer Berlin Heidelberg.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

where terms cancel to give (11).

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

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

I.2 Proofs of Results in Section 3.1

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

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

and equivalently

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

so that, using (9), gives

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

and thus surplus equals

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

Using (13) again gives us wages[12]

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

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

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

so that

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

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

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

[12] Using (14) gives the wage as

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

which is unsurprising.

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

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

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

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

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

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

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

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

and thus that

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

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

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

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

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

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

so that

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

which from Result 1 is increasing in x. ■

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

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

Proof of Result 4. Assume that the matching sets are intervals[13]

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

First rewrite the adjusted average wage as

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

Take derivatives with respect to x:

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

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

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

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

I.3 Proofs of Results in Section 3.2

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

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

so that

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

and thus that

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

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

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

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

so that

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

and

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

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

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

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

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

and thus that

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

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

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

and

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

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

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

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

Proof of Result A-1.

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

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

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

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

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

Differencing, we get that wages follow

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

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

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

which by Result 6 is increasing in y. ■

I.4 Measuring α in the data

I.4.1 Using Business Cycles to Measure α

To measure the bargaining power α in the data we consider now an extended version of the model with business cycles, i.e. exogenous changes in aggregate productivity z. The output of a pair (x, y)

is then $zf(x,y)$. Consider two worker types x and x' (have to be different types, working at firm y when productivity is z and when it is $\hat{z}$. The wages of worker x in the two business cycle states are $w(x,y,z)$ and $w(x,y,\hat{z})$, respectively. For worker x' the corresponding wages are $w(x',y,z)$ and $w(x',y,\hat{z})$. These wages are observed. The equation for wages with business cycles is straightforward and follows the same arguments as the one without business cycles. For the value of a job it holds with the obvious notation

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

and for the value of a filled vacancy,

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

Adding up these two Bellman equations yields:

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

and equivalently

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

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

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

so that the surplus equals

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

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

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

Making the same approximation for V_u,

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

and using the equation for the surplus S,

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

The differences in wages for types x and x' equals

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

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

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

so that

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

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

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

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

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

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

I.4.2 Measure α from fluctuation in profits

We can also use Blanchflower, Oswald, and Sanfey (1996) types of results who estimate the response of wages to a change in profits. If these changes in profits are caused by i.i.d. shocks to the firm's technology than profits increase by $(1-\alpha)$ and wages increase by α. Using the above results would provide us with an estimate of α.

II Computation and Implementation Details

The model is solved by iterating on the match density and surplus. If no pure acceptance strategy solution is found (because of the discretization), a mixed strategy is used such that unemployed agents accept matches with an interior probability such that the surplus of the match is positive, but very close to zero. The productivity space is discretized into 50 grid points and there are 1200 agents per type (600 workers, 600 jobs). Firms are simply collections of 100 jobs of the same type. There are 6 firms per type. The model is simulated for 240 months with an initial burn-in period of 100 months starting from an initial match distribution. This corresponds to 20 years of monthly data.

We use the following notation where a $\hat{}$ denotes as estimated value and in its absence, the true value which might not be observable (to the economist).

i) Number of workers = number of jobs = $N = 30000$.

ii) Number of worker types = $X = 50$.

iii) Name of the worker, $i = 1..N$.

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

v) Worker true type $x = 1..X$.

vi) Estimated worker type (worker bin) $\hat{x} = 1..X$. E.g. Workers with $\hat{i} = 1..600$ all have $\hat{x} = 1$

vii) Number of firm types = $Y = 50$.

viii) Firm name, $j = 1..N/100 = 300$.

ix) Rank of firm j, $\hat{j} = 1..N/100$.

x) Firm true type, $y = 1..Y$.

xi) Estimated firm type (firm bin) $\hat{y} = 1..Y$. E.g. Firms with $\hat{j} = 1..6$ all have $\hat{y} = 1$

We use the method outlined in Appendix III to rank workers. An an initial guess of the worker rankings, we compute the cost (from within firm wages) associated with ranking workers using minimum wage, maximum wage, adjusted average wage, or, expected wages, and use the lowest cost ranking. This gives us an estimate of the worker's rankings $\hat{i}$ and by binning them, $\hat{x}$.

Next, we compute the

$$u(\hat{x}) = \frac{1}{|\hat{x}|} \sum_{\hat{i} \in \hat{x}} u(\hat{i}),$$

where $|\hat{x}|$ is the number of workers in bin $\hat{x}$.

For each firm j, sum up the number of new hires from each bin. This is the number of employment spells recorded (workers may have more than 1 spell). Call this $p(\hat{x}, j)$. Additionally, we need $N(j) = \sum_X p(\hat{x}, j)$.

Next, we estimate the firm's probability of accepting workers from any given bin

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

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

Then, we compute the binomial cdf to identify bins where the firm is observed to have hired too few workers from given the total number of workers it hired. The probability of observing at most $p(\hat{x}, j)$ given the hiring probability $\pi(\hat{x}, j)$ from $N(j)$ trials is simply

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

Using this statistic, we drop workers belonging to $j - \hat{x}$ matches that yield too few observed matches. To this end, we create a statistic $D(i) = 1$ if workers are dropped and 0 otherwise. And a statistic called $D_f(x, j) = 1$ if the matches are dropped and 0 otherwise. If worker $\hat{i} \in \hat{x}$ st of one the following three conditions hold for any j,

i) $\hat{x} \in \{1, X\}$ AND $F(p(\hat{x}, j); N(j), \pi(\hat{x}, j)) < \zeta_1$

ii) $\hat{x} \in \{2, X-1\}$ AND ($p_j(\hat{x}+1) == 0$ OR $p_j(\hat{x}-1) == 0$) AND $F(p(\hat{x}, j); N(j), \pi(\hat{x}, j)) < \zeta_1$

iii) $\hat{x} \in \{2, X-1\}$ AND $(p_j(\hat{x}+1) > 0$ AND $p_j(\hat{x}-1) > 0)$ AND $F(p(\hat{x}, j); N(j), \pi(\hat{x}, j)) < \zeta_2$

we set $D(i) = 1$ and $D_f(\hat{x}, j) = 1$ to mark them as dropped matches. ζ_1 and ζ_2 are cut-offs for dropping $j - \hat{x}$ matches that result in too few employment spells given the unemployment rates.

Next, recalculate $u(\hat{x})$ for $D(i) == 0$ bearing in mind that the number of workers in each bin may have been reduced. Also calculate the minimum wage,$w_{min}(\hat{x})$ for each bin.

i) for $\hat{x} = 1..X$

(a) Compute average wages that each firm pays collectively to all workers in the bin. Call this $w_{av}(\hat{x}, j)$. I.E. If a grand total of 4 firms are observed to have hired at least one worker each from a bin, we have 4 observations.

(b) If $\hat{x} > 1$ Drop all $w_{av}(\hat{x}, j) < w_{min}(\hat{x} - 1)$.

(c) Take z_3 fraction of the remaining $w_{av}(\hat{x}, j)$ and if after doing so, fewer than z_4 remain, take the minimum of z_4 or all wages $w_{av}(\hat{x}, j)$. Average these weighted by number of wages observations going into each $w_{av}(\hat{x}, j)$. This gives $w_{min}(\hat{x})$. For all $\hat{i} \in \hat{x}, w_{min}(\hat{i}) \equiv w_{min}(\hat{x})$.

Now, for each firm j, compute the average $w(i, j) - w_{min}(\hat{i})$ for $D(i) = 0$ only over all wages that the firm pays out. This gives the average premium over minimum wages that the firm pays out, $w_{diff}(j)$. Also, compute each firm's job acceptance rate $\hat{q}_{hire}(j)$ using only $j - \hat{x}$ where $D_f(x, j) \neq 1$ as described in the Appendix. Rank firms j using $q_{ac}(j) \cdot w_{diff}(j)$ to obtain ranks $\hat{j}$ and bins $\hat{y}$.

Now we recover $\mathbb{M}_v$ and the job filling rate, $q_{fill}(j)$ using the method described in the Appendix. Instead of using the number of new hires $E_t(j)$, we recognize that the average number of new hires must be equal to $\delta \cdot |j|$, where $|j|$ is the average firm size.

To obtain smoother estimates of the production function, we now aggregate the individual firm data. The $w_{diff}(\hat{y}) = mean(w_{diff}(j))$ weighted by firm size, $|j|$. The job filling rate of bin $q_{fill}(\hat{y})$ is the similarly weighted acceptance rate of each firm.

With the job filling rate and average wage difference, our statistic $\Omega(\hat{y})$ can be readily computed. 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 $\hat{y}$ using only $D(i) == 0$. This is $w_{av}(\hat{x}, \hat{y})$. Compute the corresponding value of employment, $V_e(\hat{x})$ using the equation for value of unemployment which can be used to obtain the flow value of a vacancy $\hat{b}$ using the equation 11.

Next, compute the value of vacancy from $\Omega(\hat{y})$ with equation 20. From here, the production function $f(\hat{x}, \hat{y})$ is easily computed using the wage equation.

The true frictional production is computed from the model. Using unemployment rates at the $\hat{x}$ level and estimated firm size at the $\hat{j}$ level, compute frictional output with the estimated production function. Then, shift the estimated production function point-wise by a constant so these two outputs are exactly the same. This is our final estimate of the production function. Percent difference in the production function is the difference between the estimated and true function on the true acceptance set. Same goes for the correlation and variance.

Finally, we measure output losses due to frictions. Use aggregate employment, $E * 1000$ nodes for the linear assignment problem. Take the model generated employment rate, $u(x)$ of each bin of workers, and assign employed workers to each bin by random rounding. Same goes for the number of jobs of each type y. Compute the frictionless output using the true production function.

Do the same for the estimated frictionless output using $u(\hat{x})$, and firm sizes of each bin, $|\hat{y}$— and the estimated production function. Comparing the output of the linear assignment problem gives us the estimated output losses due to frictions.

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. As mentioned above, maybe surprisingly, workers' average wages are not increasing. But considering the workers within one specific firm gives us a correct ranking between these workers. Repeating this ranking for every firm yields a globally consistent and, if workers are somewhat mobile between firms, a complete ranking of workers. However, data are usually ridden by substantial measurement error. Taking into account this measurement error dilutes this way of ranking. Within one firm one worker could be ranked better than another work 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. Thus the rankings from all firms is not consistent and thus does not yield an aggregate ranking. How can we solve this problem? To this end, it is beneficially to 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 top to bottom. Voters will rank candidates according to their own preferences but when the need to have a single ranking of candidates comes up, one can imagine the disagreement that would arise. Unless every voter ranks all candidates identically, there will not be an aggregate ranking that all voters agree with completely.

How then does one accomplish two important tasks? Firstly, we need some notion of how "good" the aggregate ranking is. Secondly, we need a method to find the ranking that is "best" given our criteria of what is "good".

III.1 Kemeny-Young rank aggregation

Given many (perhaps) inconsistent rankings of candidates, how does one aggregate the ranks to determine who the best candidate this? This problem is ancient, and first studied by de Borda (1781) and Condorcet (1785). A natural metric for evaluating the posited aggregate rank is the number of disagreements generated in the voter submitted ranks. Given voter submitted ranks $\pi_1, \pi_2, ..., \pi_K$ and a posited aggregate rank Π, the total number of disagreements is given by

$$\sum_{k \in K} \sum_{i,j} d_k(i,j) \quad \forall i,j : \pi_k(i,j) \text{ disagrees with } \Pi(i,j)$$

$d_k(i,j)$ is the weight that the econometrician places on ranking k. A disagreement between two ranks is simply an instance where candidate A is ranked higher than candidate B in one ranking and the other way around in the other. The problem of finding the ranking Π^* that minimizes the number of disagreements defines the *Kemeny-Young rank aggregation* problem first described in Kemeny (1959) and Kemeny and Snell (1963). The solution is conceptually trivial:

i) Enumerate all possible rankings possible;

ii) Evaluate the weighted number of disagreements;

iii) Select the ranking that gives the lowest cost.

Unfortunately, the *Kemeny-Young rank aggregation* problem is NP-hard.[14] Consider a simple case of a 1000 candidates and at least 4 submitted rankings. There are $1000 \times 999... \times 2$ combinations to consider! This is an extremely large number. Ranking millions of workers in our application is simply not feasible. We need methods to map the problem into something more manageable and consider approximation techniques.

III.2 Reducing Kemeny-Young

To proceed, we need a few definitions which we use to construct a simple example to illustrate.

[14] See Bartholdi, Tovey, and Trick (1989).

III.2.1 Preliminaries on graphs

Definition 2. ***Directed graph (G)***
A directed graph $G = (V, E)$ *is comprised of a set of vertices*[15], $V \equiv \{1...n\}$, *and, a set of ordered pairs called edges,*[16] $E \equiv \{(i, j) : i, j \in V\}$.

Definition 3. ***Directed path (P)***
P *of length* p *in graph* G *is an ordered sequence of* p *edges* $(i, j)_1, (i, j)_2, ..., (i, j)_p$ *such that for any* $(k, k-1), k \in \{2, ..., p\}$, $(\cdot, u)_{k-1}, (u, \cdot)_k$ *holds for some* $u \in V$.

We work only with *weakly connected graphs*. That is, if we are allowed to violate the order on the edges, then there exists some path connecting any two $u, v \in V$. In short, there are no orphaned vertices.

Definition 4. ***Directed acyclic graph(DAG)***
A DAG is a directed graph G *containing no path* P *which connects any vertex to itself. That is,* $\forall k \in V, \nexists P : (k, \cdot)_1, ..., (\cdot, k)_p$.

We need to use special graphs called tournaments. Here, we abuse notation slightly, and use T for all of them.

Definition 5. ***Tournament (T)***
A tournament $T = (V, E)$ *is a directed graph where any two vertices are connected by at least one edge. That is,* $\forall u, v \in V, (u, v)$ *and/or* $(v, u) \in E$

Here, we define two types of tournaments pertinent to us. There are others.

Definition 6. ***Weighted Tournament with Probability Constraints (T)***
A weighted tournament with probability constraints is a tournament where the sum of weighted edges between any two vertices add up to wlog 1. Denote the weighted edge between vertices (u,v) as $w_{i,j} \in E$. $T = (V, E)$ *is a weighted tournament with probability constraints if,* $\forall u, v \in V, w_{i,j} + w_{j,i} = 1$.

Definition 7. ***Generalized weighted tournament (T)***
A generalized weighted tournament $T = (V, E)$ *is set of vertices,* $V \equiv \{1, 2, ..., n\}$ *and weighted edges* $E \equiv \{w_{i,j}$ *for any ordered-pair* $(i, j) \in V\}$ *such that* $b \leq w_{i,j} + w_{j,i} \leq 1$ *for* $b \in (0, 1]$.

Definition 8. ***Feedback arc set (FAS)***
A FAS, $E' \subset E$ *are edges such that for any graph* $G = (V, E)$, $G' = (V, E \backslash E')$ *is a DAG.*

[15] Also called edges. These will represent candidates or in our case, workers.

[16] Also called arcs. These will represent the ranking between two vertices. In our case, these will indicate the ordering of wages for any two workers in a firm at the same time.

Ailon, Charikar, and Newman (2008) show that the Kemeny-Young rank aggregation problem is equivalent to finding the minimum FAS on a weighted tournament with probability constraints. If each voter ranks *all* candidates, this condition is obviously met. In our case, voters do not submit complete rankings. To tackle the problem, we require more recent advances in the field. We use a simple example to illustrate.

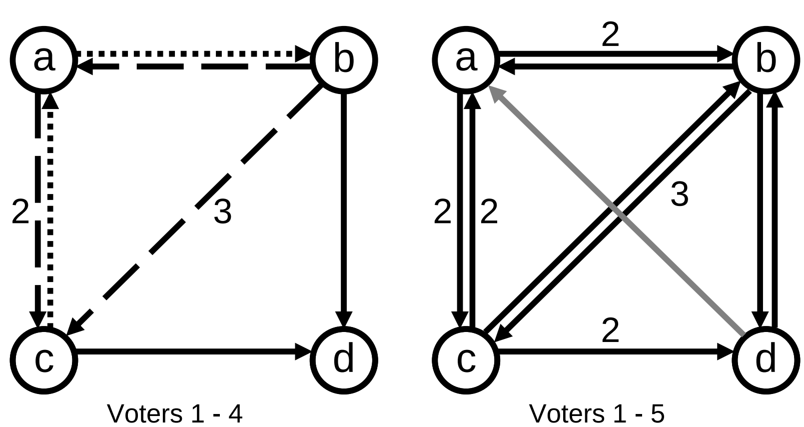
Consider the following votes on 4 candidates, a, b, c and d. The first four voters submit partial rankings only.

Table A-1: Submitted rankings

Voter 1	Voter 2	Voter 3	Voter 4	Voter 5
a	b	c	b	c
b	a	a	c	d
c	c		d	a
				b

Using only the ranks provided by the first four voters for now, we construct a graph in Figure III.2.1 by first creating a vertex for every observed worker. Next, we add a unit weight to the corresponding edge for each instance where a worker is observed to be ranked above another.[17] The different patterns for edges are just for highlighting the edges that we refer to next. The numbers indicate weights.

Figure A-1: Graphs from submitted rankings



[17]For example in Figure III.2.1, $b - c$ is seen 3 times; directly for Voter 1 and Voter 4, and, indirectly for Voter 2 with choice a in between.

Figure III.2.1 depicts a directed graph. Tracing the arrows gives us paths. One such path connects $a \to b \to c \to a$. This makes the graph cyclic. To turn it into a DAG, one option is to remove all $b - c$ edges, the $b - a$ edge, and, the $a - c$ edges. This removes a FAS of weight 6. Once the graph is a DAG, it becomes very easy to determine the ranking of workers. The rank that violates none of the edges is uniquely $c - a - b - d$.[18] The cost of this particular solution is the weights of the number of edges removed. In this case, it is 6.

A smallest FAS set are edges $c-a$ and $a-b$ yielding a cost of 2. The resulting ranking is $b-a-c-d$. As mentioned, this problem is called finding the minimum FAS on a general directed graph. Unfortunately, this problem is still very difficult, even to approximate. In fact, any approximation scheme to guarantee a cost that is at most 1.36 that of the true minimum cost is an NP-hard problem as well (see Karp (2010)).

A simpler version of the problem would be one where every two vertices have at least one edge connecting them. The minimum FAS problem on tournaments is easier than the FAS problem on general directed graphs.

Our graph above is not yet a tournament because there are no edges connecting $a - d$. If we had included Voter 5's rank, the result would be a tournament. Particularly, it is a *generalized weighted tournament* because the sum of edges in both directions do not normalize to a constant. In this particular case, the results in Kenyon-Mathieu and Schudy (2007) apply. They provide a polynomial time algorithm capable of approximating the FAS to any degree of accuracy in theory. While a polynomial time approximation scheme is a huge improvement, this is still a difficult problem. In our implementation, we appeal to the theoretical guarantees in Kenyon-Mathieu and Schudy (2007) but only implement a portion of the entire algorithm. We show using calibrated examples that our implementation works well in practice.

III.2.2 Fully utilizing information

Firms' reports of spot wages in each period are analogous to the partial ranks provided by Voters 1 to 4 in our example above. The cycles represent measurement and reporting error. Kenyon-Mathieu and Schudy (2007) algorithm's need for at least one complete rank are fulfilled by minimum wages, maximum wages and adjusted average wages. We hence use all available information from wages to rank workers.

[18] Uniqueness of a ranking is not guaranteed. This is a property of the solution to the Kemeny-Young problem.