



European  
University  
Institute

MAX WEBER  
PROGRAMME FOR  
POSTDOCTORAL  
STUDIES

# WORKING PAPERS

MWP 2017/09  
Max Weber Programme

Sorting When Firms Have Size

Maxim Goryunov



European University Institute  
**Max Weber Programme**

## **Sorting When Firms Have Size**

Maxim Goryunov

EUI Working Paper **MWP** 2017/09

This text may be downloaded for personal research purposes only. Any additional reproduction for other purposes, whether in hard copy or electronically, requires the consent of the author(s), editor(s). If cited or quoted, reference should be made to the full name of the author(s), editor(s), the title, the working paper or other series, the year, and the publisher.

ISSN 1830-7728

© Maxim Goryunov, 2017

Printed in Italy  
European University Institute  
Badia Fiesolana  
I – 50014 San Domenico di Fiesole (FI)  
Italy  
[www.eui.eu](http://www.eui.eu)  
[cadmus.eui.eu](http://cadmus.eui.eu)

## **Abstract**

In this paper, I study the sorting of workers to firms, when firm size is explicitly taken into account. I develop a method to *non-parametrically* identify the match production function from data on workers' wages and firms' revenues and posted job vacancies. Under the proposed identification procedure, ordering of workers and firms is identified independently, and can therefore be achieved using potentially different data sets. The model sheds light on the question of the exporter wage premium: exporters pay higher wages because they are larger, and higher wages are required to support a larger firm size.

## **Keywords**

matching, sorting, wage determination, firm size

**JEL classification:** C78, E24, J31, L11

*I wish to thank Michal Fabinger, Philipp Kircher, Filip Matějka, Kenneth Mirkin, Jakub Steiner and Ludo Visschers for insightful discussions, suggestions and comments. I also benefited from comments and questions from seminar participants in Yekaterinburg, Edinburgh, Saint-Petersburg and the CMSSE Summer Conference. Part of this research was carried out during my stay at the School of Economics of the University of Edinburgh. I am grateful to the School for its enormous hospitality and CERGE-EI for making my research stay feasible.*

*Maxim Goryunov*

*Max Weber Fellow, 2016-2017*

*European University Institute, Florence, Italy, and National Research University Higher School of Economics, Moscow, Russian Federation. Email: maxim.goryunov@eui.eu*



# 1 Introduction

The increasing availability of detailed micro-level data sets has made us well informed about the large amount of heterogeneity on both sides of the labour market. Recent research has shown that firms differ in size, productivity, capital intensity and many other characteristics. More importantly, the differences are enormous even within narrowly defined industries (Crozet and Trionfetti, 2013). This suggests that unobservable firm characteristics play an important role in any explanation of firm behaviour and outcomes. A similar observation holds for workers. Studies of wage determination have revealed that wage inequality has been growing in recent decades in virtually all countries, and that most of the inequality and its growth cannot be attributed to observable characteristics of workers, even in narrowly defined occupation-sector cells (see, for example, Helpman, Itskhoki, Muendler and Redding, 2012). Therefore, a study of the labour market cannot disregard the heterogeneity in, and interplay between, the unobserved characteristics of both firms and workers.

How does this enormous heterogeneity in the labour market play out in the interaction between its two sides — firms and workers? Does the market allocate workers to employers in an optimal fashion? If it does not, how large are the output losses from the misallocation? The answer to these and related questions is important for the resolution of numerous economic debates. Study of misallocation at the micro level is important for the macroeconomics literature, as it has a direct impact on aggregate productivity fluctuations and long-run income dynamics. Understanding the allocation of workers to jobs is particularly vital in international trade, where researchers are interested in whether exporters pay higher wages due to their own higher productivity or because they hire better workers.

Given the importance of these questions, there is no shortage of attempts at addressing them in the economic literature. However, the implications of the size choice made by firms for the answers has been largely ignored so far. In order to provide an insight into the role of firm size in answering these questions, I develop a model that features heterogeneous price-making monopolistically competitive firms, heterogeneous workers and a frictional labour market. The model addresses the implications of the creation of vacancies within firms and their choice of size for labour market sorting outcomes. In a nutshell, it introduces the random search model developed by Shimer and Smith (2000) into the monopolistic competition framework à la Melitz (2003). Workers searching for jobs are randomly matched with vacancies posted by firms. In contrast to the most standard search models, there is no free entry of firms; however, each firm can post as many vacancies as it needs. One can view this as a free entry of vacancies within a firm.

Production is linear in the *quantity* of labour, i.e. firm output is the sum of outputs of its workers. Nevertheless, even the most productive firms do not grow indefinitely, due to decreasing demand for their products and the local market power they enjoy.

The contribution of the paper is twofold. First, I show that match output, and hence a firm's production function, can be identified non-parametrically. The identification strategy I develop utilizes firm-level data. One of its advantages is that identification of firms' unobserved characteristics is achieved independently of workers' characteristics in a straightforward and intuitive way. In other words, the ranking of firms is identified from firms' variables only, such as revenue and vacancies, and does not depend on the wages they pay. Identification of the match output function is essential for any counterfactual analysis that includes shifting different workers between different jobs. In such a context, non-parametric identification is especially important because of the lack of microfoundations behind the structure of the match production function. Moreover, the arguments of the match production function, i.e. workers' and firms' unobserved heterogeneity drivers, are not well understood themselves.

In addition, I show that, along with the match output, vacancy creation costs are identified non-parametrically. Understanding the shape of vacancy creation cost is important because it has potential implications for business cycle models: the faster the marginal cost increases with vacancy creation, the more incentives a firm has to smooth the hiring process. Thus, a convex vacancy creation cost may be one explanation for slower recoveries.

Second, I extend the model into an international trade setting to show how taking the firm size explicitly into consideration can dramatically change predictions about sorting. Exporters allow for a wider range of quality among their workers and pay them higher wages than non-exporters. This is in dramatic contrast to the result of Bombardini, Orefice and Tito (2014), who show that in a one-to-one matching framework, exporters tend to choose, on average, better workers and have less skill dispersion in their workforce. The different result is driven by the firm size effect. In the present model, exporters are larger, and the cost of supporting a larger firm size requires a larger equilibrium match surplus. Therefore, the set of acceptable workers, i.e. those who generate positive match surplus, expands. In addition, exporters pay a higher wage to their workers because the wage is positively related to the match surplus.

Understanding predictions about the matching sets of exporters relative to non-exporters is important because it has implications for the effects of trade liberalization on wage inequality. Indeed, if on average matching sets become tighter when more firms engage in exporting activity, wage inequality will also increase, for two reasons. First, the share of workers who can enjoy the exporter wage premium increases more slowly than

the share of exporting firms. Second, the tighter matching sets indicate that matches are closer to perfect and workers' wages are closer to their maximal attainable wages (given the aggregate environment). Conversely, expanding matching sets put downward pressure on wage inequality.

This paper fits into the extensive line of research on the estimation of models of sorting on the labour market based on unobserved characteristics of workers and firms, firmly grounded in theory, such as Lopes de Melo (2013) and Lise, Meghir and Robin (2013). However, the paper addresses two shortcomings of the current literature. First, a large part of the literature imposes a good deal of structure on the model and, in particular, does not allow for varying degrees of complementarity — Hagedorn, Law and Manovskii (2017) being a notable exception. Second, virtually no research in the area so far has made any distinction between 'a firm' and 'a job'. In other words, in these models, firm boundaries are arbitrary, and a firm with  $n$  workers is equivalent to  $n$  firms with one worker each. Although theoretically convenient, this approach has a major shortcoming when applied to data. Since no data set has information on the profitability or output of a particular worker or workplace in a firm, the identification of firm characteristics is based on information about wages and labour flows. Therefore, identification strategies tend to be indirect and computationally intensive.

Apart from this empirical reason for considering firm size in labour market sorting models, there is also a purely theoretical reason that deserves attention. As Bagger and Lentz (2015) note, in a one-to-one matching framework the decision to accept or reject a match relies heavily on a fundamental scarcity. In such a world, the decision to agree upon a match is equivalent to a decision to discontinue searching. However, the relevance of this assumption is not so obvious, since workers can continue to search for opportunities while employed and firms can have many workers. There is a large literature that relaxes the scarcity assumption on the worker side of the model via on-the-job search. This paper can be viewed as a mirror image of that literature. Although I retain the scarcity on the worker side, I relax it on the firm side of the model via explicit introduction of the firm's choice of size.

The rest of the paper is organized as follows. The next section provides a literature overview. Section 3 sets up the model of the frictional labor market with heterogeneous workers and heterogeneous firms choosing their size. Section 4 characterizes the equilibrium and theoretically develops the identification procedure. Section 5 extends the model into the international trade setting, and Section 6 concludes.

## 2 Literature Review

Since the seminal study by Abowd, Kramarz and Margolis (1999), it has been believed that one can grasp unobserved characteristics by first running Mincerian regressions of wages on the observable characteristics of firms and workers and their respective fixed effects using longitudinal linked employer-employee data sets. Second, examining these fixed effects in particular, correlation between them conditional on being matched has been considered a rough measure of sorting. With the increasing availability of linked employer-employee data sets, this approach has been widely adopted and applied to data sets from a number of countries, with a general conclusion that the correlation coefficient between fixed effects in worker-firm matches is not very large. Moreover, most studies have found it to be either insignificant or even negative<sup>1</sup>. Although in their review of early literature, Abowd and Kramarz (1999) cautioned that “it is important to keep in mind that it is not always possible to make a direct interpretation of the statistical parameters (for individuals or firms) in terms of simple economic model” (p. 2671), the lack of a significant positive correlation between worker and firm fixed effects has been widely interpreted as an absence or unimportance of sorting.

Recent research has shown that the identifying assumptions of this reduced form approach are inconsistent with virtually every equilibrium model of sorting, and that the estimated fixed effects do not contain information on underlying unobserved characteristics. In other words, applied to data generated by an equilibrium sorting model, this regression approach yields fixed effect estimates that have no interpretation within the original model. The intuition behind this, uncovered by Eeckhout and Kircher (2011), is that wages are potentially non-monotone in a firm’s type: a better firm has to be compensated for hiring workers who are worse than the firm desires and therefore, a linear model is fundamentally misspecified. In addition to the purely theoretical argument against interpreting the absence of a correlation between workers’ and firms’ fixed effects as an absence of sorting, Lopes de Melo (2013) has shown that the correlation between the fixed effects of a worker and her coworkers is strong, suggesting that similar workers do indeed sort together.

Understanding the limitations of the reduced form approach of two-way fixed effects regressions has led to the development of a literature on the identification of sorting grounded in theory. The starting point for understanding the assortative matching is Becker’s (1973) assignment model with transferable utility. The main insight from this

---

<sup>1</sup>See Abowd and Kramarz (2009) for a review of early literature, Gruetter and Lalive (2009) for Austria, Abowd, Kramarz, Pérez-Duarte and Schmutte (2009) for the U.S., and Card, Heining and Kline (2013) for Germany, among others.

model is the crucial dependence of the sign of sorting on the complementarities between two sides of the market: positive assortative matching (PAM) — mating of likes — arises when the production function is supermodular, i.e. the marginal product of an agent in a match increases with the quality of her partner. Shimer and Smith (2000), Atakan (2006) and Eekhout and Kircher (2010) build on Becker’s insight and develop the assignment model to introduce search frictions. They show that in the presence of search frictions the interplay between the degree of complementarities and the level of search friction is decisive in determining the degree and sign of sorting. The Shimer and Smith (2000) model, being the most natural approach, has become a cornerstone of the literature on sorting in the labour market. However, one of the limitations of the theoretical literature on sorting is its focus on one-to-one matching, and its resulting disregard of the role of firm size. This paper aims to overcome this limitation by introducing firm size into what is essentially the Shimer and Smith (2000) framework.

Lopes de Melo (2013) and Lise, Meghir and Robin (2013) develop structural models of sorting and wage dynamics. Estimation of these models suggests that PAM between workers and firms is present in the data. However, the main limitation of their approach is the strong assumptions imposed on the functional form of the production function, that do not allow the sign and strength of sorting to differ within the domain of worker and firm types.

Hagedorn, Law and Manovskii (2017) take a step further. Building on Shimer and Smith (2000), they develop an identification procedure that allows for non-parametric identification of the production function. Applying their framework to German linked employer-employee data, they show that although complementarities between worker and firm productivity (and hence, PAM) prevail on average, there are regions of local substitutability, and that the market exploits this feature of the production function: reassigning workers to firms in a perfectly assortative fashion would reduce total output by 1.43%. However, the empirical literature has inherited the limitation of its theoretical predecessor, namely, the focus on one-to-one matching. The identification procedure I develop in this paper borrows heavily from Hagedorn et al. (2017), but overcomes the limitations of one-to-one matching.

The literature on macroeconomic dynamics has recognized that the abstraction from the firm inevitably misses a potentially important intensive margin of employment adjustment over the business cycle: expansion (as opposed to entry) of firms during booms and their contraction (as opposed to exit) during recessions. Intuitively, a firm that can adjust its labour force size has additional room for manoeuvre when faced with shocks, and models taking this into account can produce different amplification mechanisms. Hence, the growing literature on the role of firm size in labour market dynamics in macroeconomics.

Kaas and Kircher (2014), Elsby and Michaels (2003) and Moscarini and Postel-Vinay (2014) develop macroeconomic models that capture sluggish labour market dynamics, job flows and the evolution of the firm size distribution over the business cycle. However, this paper, to the best of my knowledge, is the first attempt to explicitly take into account the role of firm size in the outcome of sorting on the frictional labor market.

Few papers have studied sorting on the labor market in a one-to-many matching framework. Eeckhout and Kircher (2012) expand a frictionless Beckerian approach and show that if firms can choose not only the type (quality), but also the quantity of production factors, the necessary conditions on the primitives for PAM become stricter. Intuitively, switching to worse workers is not as detrimental for a firm as in a one-to-one matching world, since lower quality can be compensated for with a larger number of workers. Therefore, for this not to happen, the loss in match value must be very high in the best firms, i.e. marginal match output should change very sharply with the type of firm. The interesting question of how the presence of search frictions affects the conditions for PAM is outside the scope of this paper. Instead, I take an agnostic stand on the strength of complementarities and a data-driven approach: Given the observed labour market outcomes, I uncover the shape of the primitives.

The only attempt to utilize firm-level data in an empirical study of sorting that I am aware of is the study by Bartolucci, Devicienti and Monzon (2015). They use a number of definitions of firm profit to rank firms, and exploit the patterns of movement of workers between firms to deduce aggregate measures of the degree and sign of sorting on the labour market. Their methodology remains valid in the model I develop here. Therefore, this paper can be viewed as providing theoretical support, in terms of a general equilibrium model, to their empirical procedure.

### 3 Model

The economy consists of two sectors: one producing differentiated intermediate inputs using labour, and the other assembling the final good from intermediate inputs. The final good is produced under perfect competition. The intermediate good sector is the crucial building block of the model. Its structure integrates Shimer and Smith's (2000) model of a time-consuming job search and Melitz's (2003) approach to firm heterogeneity.

Firms in the intermediate sector require labour for production. Both firms and workers are heterogeneous, yet the production function is linear in the *quantity* of labour. However, the market for intermediate inputs is monopolistically competitive, and local market power constrains optimal firm size. Unemployed workers search for jobs in the intermediate sector and firms post vacancies to hire labour. The labour market is frictional:

it takes time to fill a vacancy and to find a job.

### 3.1 Final Good Production

The final or consumption good is assembled from varieties of differentiated inputs under Constant Elasticity of Substitution (CES) technology:

$$Y = \left[ \int_{j \in \Omega} q(j)^{\frac{\sigma-1}{\sigma}} dj \right]^{\frac{\sigma}{\sigma-1}}, \quad \sigma > 1, \quad (1)$$

where  $\Omega$  denotes the set of varieties available for production of the final good, and  $\sigma$  is the elasticity of substitution between varieties. In what follows, I assume that there is a measure one continuum of firms, each producing distinct variety  $j$ . Thus,  $\Omega$  is a set of measure one. Given the constant returns to scale technology and perfect competition in the final good sector, I can focus on a representative firm and its demand for inputs. The price of the final good is normalized to one.

The cost minimization problem in the final good sector is standard. It implies the following inverse demand for intermediate inputs:

$$p_\varphi = Y^{1/\sigma} q_\varphi^{-1/\sigma}, \quad (2)$$

is taken as given by the producers of the intermediate inputs.

### 3.2 Differentiated Sector

Diverging from Melitz (2003), I assume that the mass of varieties of intermediate inputs is fixed<sup>2</sup>. Firms are heterogeneous, they differ in their type  $\varphi$ , which is uniformly distributed on a  $[0, 1]$  interval. Production of a good requires labour, which is also differentiated. I assume that the economy is populated by a unit mass of workers of type  $a$  which is uniformly distributed on a  $[0, 1]$  interval<sup>3</sup>. I assume the following production technology: given any cumulative labour force distribution  $L_\varphi(a)$  within a firm of type  $\varphi$ , the output is

$$q(L_\varphi(a), \varphi) = \int \psi(a, \varphi) dL_\varphi(a). \quad (3)$$

---

<sup>2</sup>One can introduce an entry game similar to the original Melitz model. However, since the main focus of the paper is on the labour market, I leave the entry stage out. Conversely, one can always find the fixed production cost and entry cost levels, such that the measure of stayers is one.

<sup>3</sup>Observe that uniform distribution of types is not a restriction. Any other continuous distribution can be transformed to uniform by effectively “renaming” workers aided by corresponding changes in the match production function. In other words, types are ordinal; they do not measure productivity or skills per se, but only by their effectiveness in production.

Here  $\psi(a, \varphi)$  can be understood as the efficiency units of labour worker  $a$  provides to a firm  $\varphi$ , or as the standard match output of a firm-worker pair. Then, the aggregate firm output is the sum of the output of all the individual pairs. Although this assumption disregards potential complementarities or spillovers between different workers<sup>4</sup>, it is in line with most of the current literature, which treats aggregate output as a sum of match outputs. I apply a similar logic to within-firm production, to facilitate comparison with existing models of frictional labour market sorting, and to highlight the role of firm size not confounded with intrafirm technological spillovers.

I assume that the match output  $\psi(a, \varphi)$  is an increasing function, with the underlying structural assumption that ordering of types  $a$  and  $\varphi$  is meaningful — a higher level implies a more productive type — and global — being more productive does not depend on the match partner. In other words, this is a model of absolute advantage in the labour market. Although restrictive, the last assumption is prevalent in the matching literature. Importantly, I do not put any restrictions on the cross derivative of the match output function, since the main focus of the paper is on its identification.

Intermediate good producers take the demand (2) for their goods as given. Thus, the revenue of a firm producing a differentiated variety as a function of the production volume is given by

$$R(q_\varphi) = Y^{1/\sigma} q_\varphi^{\frac{\sigma-1}{\sigma}}. \quad (4)$$

Observe that revenue is a concave function of firm output. This feature, stemming from the demand structure, limits firm size in this model. Alternatively, one can say that a firm faces decreasing monetary returns to its production scale, and therefore one can easily construct an isomorphic version of the model — with the production function concave in total effective labour, and perfect competition between the firms in the intermediate sector.

### 3.3 Labour Market

Now, I turn to the core of the paper — labor market structure. Firms' and workers' behaviour on the labor market is crucial to identification of the match output function. Furthermore, the labour market is frictional, and frictions are the only source of movement of workers between different firms, and therefore are the source of identification.

---

<sup>4</sup>There are a few exceptions addressing intrafirm worker productivity interdependence. Bombardini, Gallipoli, and Pupato (2012) study the intrafirm complementarities level as a source of industry comparative advantage. Helpman, Itskhoki, and Redding (2012) introduce congestion into production technology. However, since the production function generally depends on the whole labour distribution within a firm, every tractable approach to it is bound to impose restrictive structural assumptions.

Time is discrete. In every period workers can be either employed or unemployed. Employed workers receive wage income, and unemployed workers enjoy a utility equivalent to flow income  $b(a)$ . Firms post vacancies, and unemployed workers search for a job. The chances of finding a job and filling a vacancy are governed by the labour market tightness  $\theta$ , which is defined as the vacancy-to-unemployment ratio. The meeting rate is given by  $m(\theta)$  for a firm and  $\theta m(\theta)$  for a worker, the latter representing the matching function. This indirectly implies a standard assumption of a constant returns to scale matching function. I additionally assume that  $m(\theta)$  is decreasing in  $\theta$  and  $\theta m(\theta)$  is increasing in  $\theta$ , which is equivalent to the assumption that the number of matches increases both with the number of vacancies and with the number of unemployed workers.

The meeting is random: neither firms nor workers can target a potential partner's type. The match is consummated voluntarily upon a meeting, when the types of both partners are perfectly observable. There is no on-the-job search in the baseline model, and a worker stays in the match until its separation. The matches are dissolved exogenously with probability  $\delta$ .

Households are assumed to be risk-neutral suppliers of labour of a given skill  $a$ . They maximize the expected lifetime income flow, discounted at an interest rate  $r$ . Denote as  $U(a)$  the value function of unemployed worker of type  $a$ , and as  $V(a; \varphi)$  the value function of a worker  $a$  employed by firm  $\varphi$ . I impose symmetry across firms of a given type, which allows me to ignore the potential dependence of value functions and wages on firm employment. The value functions of a worker obey the following two Bellman equations<sup>5</sup>:

$$rU(a) = b(a) + \theta m(\theta) \int \gamma(\varphi) \max\{V(a; \varphi) - U(a), 0\} d\varphi, \quad \forall a \quad (5)$$

$$rV(a; \varphi) = w(a, \varphi) + \delta(U(a) - V(a; \varphi)) \quad \forall(a, \varphi) \quad (6)$$

Here  $\gamma(\varphi) = \frac{v(\varphi)}{\int v(\varphi) d\varphi}$  is the distribution of vacancies across firm types. It governs the chances of meeting a firm of any given type. It is straightforward that a worker engages in a match if the value of being employed exceeds the value of being unemployed. The interpretation of these equations is somewhat standard. Equation (5) states that the flow value of unemployment consists of income in unemployment and the expected gain in value from meeting a firm. Equation (6) represents the flow value of employment by firm  $\varphi$  as the wage  $w(a, \varphi)$  at this firm and the potential loss in value from separation. Generally,

---

<sup>5</sup>Due to monopolistic competition in the intermediate goods sector, firms there obtain positive profits that workers can potentially have claims on. However, unless these claims depend on the state of employment, they do not affect employment decisions. For this reason, I omit them from the Bellman equations for workers' monetary flow to ease notation, effectively assuming that a separate class of entrepreneurs enjoys all the profits from the intermediate goods sector.

wages depend not only on the type of worker and firm but also on the composition of the labour force within a particular firm. I ignore this dependence because I focus on the symmetric steady state equilibrium, in which all firms of the same type have the same labour force structure.

Thus, the behaviour of workers is straightforward: They look for a job and take any that brings them a higher income flow. As I show later, this is equivalent to a simple reservation wage rule. Denote as  $\mathcal{A}(a)$  the set of acceptable matches for a worker of type  $a$ , i.e.  $\mathcal{A}(a)$  is a subset of firm types a worker  $a$  is willing to work for given the equilibrium wage:

$$\mathcal{A}(a) = \{\varphi : V(a, \varphi) - U(a) > 0\}.$$

Now, I turn to firms' behaviour on the labor market. Denote as  $J(L_\varphi, \varphi)$  the value of a firm of type  $\varphi$  and labour force  $L_\varphi(a)$ . Firms maximize their present value, which is equal to the discounted stream of profits. In order to hire workers, firms choose a measure of vacancies  $v$  to post. The Bellman equation for the firm problem is:

$$J(L_\varphi, \varphi) = \max_{v, L'} \frac{1}{1+r} \left\{ R(q(L_\varphi, \varphi)) - \int w(a, \varphi) dL_\varphi(a) - c(v) + J(L', \varphi) \right\}, \quad (7)$$

subject to the definition of revenue (4) and the law of motion of within-firm employment

$$0 \leq L'(a) \leq (1 - \delta)L(a) + m(\theta)v \int_0^a \frac{u(a')}{u} \mathcal{I}\{\varphi \in \mathcal{A}(a')\} da', \quad (8)$$

where  $u(a)$  is a measure of unemployed workers of particular type  $a$ , and  $u = \int_0^1 u(a') da'$  is total unemployment.  $c(v)$  is the vacancy posting flow cost, with  $c'(v) > 0$ ,  $c''(v) \geq 0$ . The change in within-firm employment consists of labour attrition due to separation, and new hires obtained from filled vacancies.  $\mathcal{I}\{\cdot\}$  is the indicator function, and the term  $\mathcal{I}\{\varphi \in \mathcal{A}(a')\}$  ensures that a worker takes a job if offered. The inequality in the labour law of motion implies that a firm can shadow any of its labour force without cost if necessary, creating an asymmetry in the labour adjustment cost. Although artificial, this assumption is standard in the search literature. It is not crucial for my results, since in a steady state there is no voluntary match destruction.

Lastly, because hires are not interchangeable with workers outside the firm, workers have bargaining power. I assume that wages are determined through the generalized Nash bargaining solution in the spirit of Stole and Zwiebel (1996). Thus

$$(1 - \beta)(V(a; \varphi) - U(a)) = \beta \frac{dJ(L, \varphi)}{dl(a)}, \quad (9)$$

where  $\beta$  is the bargaining power of a worker. Put briefly, a worker and firm receive fixed shares of the match surplus, with  $\beta$  being the worker's share. Due to the risk-neutrality on both sides, the total match surplus can be viewed as a monetary gain of size  $S = \frac{dJ(L,\varphi)}{dl(a)} + V(a; \varphi) - U(a)$ . Here, the meaning of the derivative  $\frac{dJ(L,\varphi)}{dl(a)}$  is the marginal benefit for a firm from hiring a worker of a particular type. Effectively, the firm enjoys an increase in its value from a match with a worker relative to non-consummation of this particular match. A worker's gain from the match,  $V(a; \varphi) - U(a)$ , is the share  $\beta$  (his bargaining power) of the match surplus. This is the intuition behind (9).

Note that in general the solution to the bargaining problem would depend not only on the firm's and worker's types, but also on the whole distribution of employment in the firm, and, in particular, on the firm's size. However, since I focus on the steady state of the economy, firm type  $\varphi$  captures all the latter, and therefore the equilibrium wage only depends on the pair of types.

Finally, labour balance identities should hold, i.e. the sum of employed and unemployed workers should be equal to the total population:

$$M \int L_\varphi(a) d\varphi + \int_0^a u(a') da' = a. \quad (10)$$

Recall that the distribution of workers by type in the population is uniform, and therefore the mass of people with a type weakly below  $a$  is equal to  $a$ .

## 4 Labour Market Equilibrium and Identification of Sorting

In this section, I provide a partial characterization of the firms' optimal behaviour. The established properties of the firms' behaviour will be central to the development of a strategy for the identification of the model primitives. As is standard in models of frictional sorting, I use a properly defined match surplus function, as a gain from consummating the match relative to the outside option. Then, I describe how the match surplus feeds into wages and hiring decisions. The established interdependence of surpluses, wages and vacancy creation allows for the identification of the production function from data on wages and firm revenue.

I focus on the steady state equilibrium of the economy. First, define a surplus function

$$s(a, \varphi) = \frac{\sigma}{\sigma - \beta} \frac{dR(\varphi)}{dl(a)} - rU(a). \quad (11)$$

The next proposition establishes that this is a proper definition of the surplus function in the sense that matches are consummated whenever it is positive. This can be understood

intuitively: an additional worker brings to the negotiation table the value of his marginal product,  $\frac{dR(\varphi)}{dl(a)}$  in flow terms. At the same time, his presence increases output, damping marginal revenue. This in turn leads to a decrease in the value of the marginal product of other workers. Hence, the firm improves its position in bargains with other workers it employs in the period. Thus, the multiplier  $\frac{\sigma}{\sigma-\beta} > 1$  takes care of this pecuniary externality in the negotiation process. A smaller  $\beta$ , i.e. a weaker bargaining position of the workers, naturally leads to less importance of this externality. At the limit, if the workers have no bargaining power, the match gain is equal to the marginal revenue. The negative part of the surplus is straightforward: the worker forgoes her value in unemployment and the firm has nothing to lose, since the vacancy cost is sunk at the point of wage negotiation.

**Proposition 1.** (i) *The hiring decision is governed by the surplus function with matches being consummated whenever  $s(a, \varphi) \geq 0$ .*

(ii) *The outcome of wage bargaining yields the following wage rule:*

$$w(a; \varphi) = \beta s(a, \varphi) + rU(a). \quad (12)$$

(iii) *The vacancy creation policy of a firm is indirectly defined by*

$$c'(v) = \frac{m(\theta)}{r + \delta} \int (1 - \beta) s^+(a, \varphi) \frac{u(a)}{u} da, \quad (13)$$

where  $s^+(\cdot) = \max\{s(\cdot), 0\}$ .

Moreover, with this result at hand, the Bellman equation for an unemployed worker (5) can be reformulated as follows:

$$rU(a) = b(a) + \beta \frac{\theta m(\theta)}{r + \delta} \int s^+(a, \varphi) \gamma(\varphi) d\varphi. \quad (14)$$

The intuitive interpretation of the results presented in Proposition 1 is clear in the light of the definition of the surplus function. The wage is nothing but a standard Nash-bargaining-type surplus sharing rule that assigns  $\beta$  share of the surplus to a worker. Although one should bear in mind that the parties bargain over the surplus over the whole period of the relationship, due to the focus on the steady state, this is equivalent to sharing flow surplus every period.

The last point of Proposition 1 requires that the optimal policy of a firm equates the marginal cost of a vacancy on the left-hand side to its expected marginal benefits. Indeed, on the right-hand side of equation (13) the chance of meeting a worker,  $m(\theta)$ , is multiplied by a firm's average share of the surplus resulting from a meeting (whether a

match follows or not), represented as an integral over worker types, with the distribution of unemployed workers being the relevant one. In addition, multiplier  $\frac{1}{r+\delta}$  accounts for the total discounted flow over the expected length of the relationship. Equation (14) has a similar meaning, only from the worker's viewpoint.

It is worth emphasizing the two assumptions of the model that allow for extension of the job search model from a one-to-one to a one-to-many matching framework and neat equilibrium characterization. First, as stressed earlier, there are no direct complementarities between the workers in the production function, and output is the sum of the marginal products of the workers. Second, the CES aggregation of the intermediate goods into the final good in (1) guarantees that the marginal value-product of a worker is proportional to her marginal product, with the proportionality coefficient only depending on the firm's type. These two assumptions taken together allow for a neat characterization of the solution to the firm's problem.

The identification procedure I develop hinges on the established properties of the optimal behaviour of firms and workers. I now discuss how they allow for identification of the worker ranking, value function in unemployment, vacancy creation cost and match output function.

## 4.1 Identifying Worker Types and Unemployment Value

The worker side of the model is similar to the standard one-to-one frictional search setup of Shimer and Smith (2000). Thus, the identification of worker types developed by Hagedorn et al. (2017) carries over to the model presented in this paper. The validity of this identification strategy is warranted by the following proposition:

**Proposition 2.** *Let the value in unemployment,  $b(a)$ , be non-decreasing in worker type  $a$ . Then,*

- *value in unemployment  $U(a)$  is increasing in worker type  $a$ , and hence, wage  $w(a, \varphi)$  and value in employment  $V(a; \varphi)$  are increasing in  $a$ ;*
- *minimum and maximum wages attainable by a worker of a given type  $a$  are increasing in  $a$ .*

*In addition, if for a worker of type  $a$  there is a firm type  $\varphi$  that does not hire her in equilibrium, then the minimal wage she attains is equal to the flow value of unemployment:  $\min_{\varphi} w(a, \varphi) = rU(a)$ .*

The intuition behind Proposition 2 is quite straightforward. For a given firm, a better worker has a larger value of marginal product  $\frac{dR(\varphi)}{da} = \frac{\sigma-1}{\sigma} p_{\varphi} \psi(a, \varphi)$ , which is reflected

in the wage ranking within a firm. Given the uniformity of ranking within a firm and the non-decreasing flow value in unemployment, better workers have better prospects when they are unemployed. The last part of the proposition states that if workers do not accept wage offers from some firms they follow the simple reservation wage rule in accepting offers with the reservation wage being equal to the flow value in unemployment.

The identification of worker ranking from wage data is based on the derived monotonicity properties. These properties also hold in a one-to-one matching model as in Hagedorn et al. (2017) and their identification strategy applies. For the sake of completeness I reiterate their argument in the remainder of this subsection.

If a firm matches with all worker types, the wages it pays would provide a global ranking of the workers. However, this can hardly be expected, especially if the search frictions are sufficiently small and complementarities in the production function are sufficiently strong. Nevertheless, wage ranking within a firm provides a ranking of workers within that firm. With sufficient mobility of workers across firms, one can employ a transitivity argument. Consider a worker  $a$  moving from firm  $\varphi_1$  to firm  $\varphi_2$ . Then, any worker  $\tilde{a}$  with a higher wage in firm  $\varphi_2$ , i.e.  $w(\tilde{a}, \varphi_2) > w(a, \varphi_2)$ , should have a higher rank than any worker  $\tilde{a}$  in firm  $\varphi_1$  with wages  $w(\tilde{a}, \varphi_1) < w(a, \varphi_1)$ , and vice versa. In other words, if, according to wage ranking within firm  $\varphi_2$ ,  $\tilde{a} > a$  and, according to wage ranking within firm  $\varphi_1$ ,  $a > \tilde{a}$ , transitivity implies that  $\tilde{a} > \tilde{a}$ . With sufficient mobility of workers across firms, aggregation of interfirm ranking, aided by this transitivity argument, identifies the global ranking of workers in a linked employer-employee dataset.

Two complications might arise in empirical implementation of this procedure. First, the measurement error in wages can distort the observed workers' ranking within the firm. To address this problem, Hagedorn et al. (2017) augment the aggregation procedure by assigning weights to worker pairs within the firm. The particular structure of weights depends on the distributional assumption about the measurement error. Hagedorn et al. (2017) work with a normal distribution, imposing independence across workers and firms. Under this assumption, weights have an intuitive structure: a higher wage difference in an observed worker pair leads to a higher incremental value in the aggregation objective function (effectively Bayesian probability) if they are ranked according to the wage differential.

Second, the exact aggregation of ranks within a firm is computationally complex. The results of Proposition 2 on global ranking in maximal and minimal attainable wages help to improve the procedure by providing an initial ranking that should be close to the exact ranking (and is not exact only due to the measurement error). Hagedorn et al. (2017) show that one can initialize the algorithm with a global ranking by maximal or minimal wage, and use single worker moves for improvement. This procedure yields an accurate

solution without being as computationally demanding as the original problem.

The last part of Proposition 2 allows for identification of the unemployment value as a minimal attainable wage. However, given the relatively short time span of linked employer-employee data sets usually used in a sorting estimation, straightforward empirical implementation could be problematic. Hagerdorn et al. (2014) put forward a solution based on the fact that the ranking of workers is identified. By the continuity argument, similarly ranked workers must have similar reservation wages. Thus, one can group together similarly ranked workers and consider them as being of the same type. This approach dramatically expands the number of observations available for a given worker type and yields sufficiently precise estimates of reservation wages and unemployment values.

## 4.2 Identification of Vacancy Posting Cost

Insights from Proposition 1 allow for identification of the vacancy posting cost function when a researcher has data on the number of vacancies within a firm, in addition to the wage data. Observe that with a Nash bargaining result (12), the surplus can be identified from wage data:

$$s(a, \varphi) = [w(a, \varphi) - \min_{\varphi} w(a, \varphi)] / \beta. \quad (15)$$

In other words, the surplus is proportional to the wage premium over the reservation wage of a worker. Here, I have used the fact that the flow value in unemployment is identified by the minimal attainable wage using the procedure developed in the previous subsection. Now, I rewrite the vacancy creation policy (13) in the following way:

$$(r + \delta)c'(v) = m(\theta) \int \mathcal{I}\{s(a, \varphi) > 0\} \frac{u(a)}{u} da \times \int \frac{(1 - \beta)s^+(a, \varphi) \frac{u(a)}{u}}{\int \mathcal{I}\{s(a', \varphi) > 0\} \frac{u(a')}{u} da'} da, \quad (16)$$

with  $\mathcal{I}$  being an indicator function. The two terms on the right-hand side of (16) have direct empirical counterparts. The first term represents the chances of a vacancy meeting a worker multiplied by the share of acceptable workers in the unemployment pool. Together, this constitutes a probability that the vacancy is filled at the end of the period. Thus, the empirical counterpart of the first term is the ratio of the number of new hires to the number of posted vacancies. The second term is the firm's share of the surplus from a match averaged across new matches. With the surplus identified with (15), it is proportional to the average wage premium of new hires over their respective reservation wages.

With the marginal posting cost identified, one can test for convexity of the vacancy creation cost, i.e. for increasing marginal cost. The degree of vacancy posting cost

convexity has important implications for macroeconomic models. Although the model does not feature business cycle fluctuations, the role of the shape of the vacancy creation cost function in firm dynamics can be understood intuitively. Indeed, if the cost function were found to be convex, it would imply that firms have incentives to smooth vacancy creation over recoveries, i.e. to distribute vacancy creation over a longer period of time, leading to a slower recovery process. On the other hand, a constant marginal cost would imply that firms immediately adjust their labour force to the optimal level.

Identification of the vacancy posting cost function depends crucially on the linearity of the relationship between the surplus and the wage premium, which is the result of the particular assumption about the bargaining process. However, in any model where the wage depends positively on the match surplus, the relationship between the wage premium and the surplus would be monotone. This assumption seems a natural outcome of a wage setting process. Therefore, the proposed procedure for vacancy cost identification is robust to alternative specifications of the bargaining arrangement. Although this identification strategy would not correctly identify the exact functional form of the vacancy posting cost, with wages monotone in the match surplus, it identifies a monotone transformation of the marginal cost of vacancy posting. Therefore, the test for the constant marginal cost would not be misspecified, and would still discriminate correctly between linear and convex cost functions as long as the vacancy creation technology is the same for all firms.

### 4.3 Identifying Firm Types and the Production Function

The main focus of the literature is on the identification of sorting. In this subsection, I show how the model structure allows one to identify the firm ranking and production function with the help of the firm-level data. Since most of the previous literature has equated firms and jobs, the identification procedures developed so far can rarely make use of firm-level data. The only exception I am aware of is Bartolucci et al. (2015), who use firm data to recover aggregate measures of the strength and sign of sorting on the labour market. The explicit introduction of the firm into the model allows for much simpler identification of the details of the sorting outcome from an additional source of information.

I start with the following proposition:

**Proposition 3.** *In equilibrium, better firms enjoy higher profits, i.e. firm value  $J(\varphi)$  is increasing in  $\varphi$ . Additionally,*

(i) *if  $c''(v) > 0$  then  $\varphi_i > \varphi_j$  implies that either  $R_i > R_j$ , or  $v_i > v_j$ , or both;*

(ii) if the marginal vacancy posting cost is constant,  $c'(v) = c$ ,  $\varphi_i > \varphi_j$  implies  $R_i > R_j$ .

Turned on its head, the proposition implies that the profit ordering pins down type ordering. However, due to potential measurement error in or misreporting of profits, the ranking of firm types from profits alone might not be measured efficiently. The second part of Proposition 3 provides an additional source of identification that might be useful in practical applications. It states that the ordering of revenues and vacancies identifies firms' ranking as well as profits. The more aligned profits, revenue and vacancies rankings are, the easier it is for the model to reconstruct the firms' ranking confidently.

One of the advantages of this identification strategy lies in the fact that two rankings are identified using different sources of information: the workers' ranking is identified from individual wage data, whereas the firms' ranking is from firm-level data. In addition to making identification of the firms' ranking more straightforward and intuitive, relative to the current literature, it avoids potential biases in the sorting estimation, stemming from the fact that the firm type is identified from wages, and hence the types of its workers.

If the correlation between revenue and vacancies is not perfect, but sufficiently high, the researcher can use the analogous tactics that were applied in the identification of flow unemployment value. Firms with similar revenue and vacancies, yet an opposing ranking of the two, can be grouped together as firms of the same type.

The last step is identification of the production function. Observe that using surplus definition (11) we can find the match production function:

$$\psi(a, \varphi) = \frac{\sigma - \beta}{\sigma - 1} Y^{\frac{1}{\sigma-1}} R_\varphi^{\frac{1}{\sigma-1}} [s(a, \varphi) + rU(a)].$$

From this, together with the identification results developed above, it follows that, up to a constant multiplier, the production function can be identified from wages and revenues with the following equation:

$$\tilde{\psi}(a, \varphi) = R_\varphi^{1/(\sigma-1)} \left[ w(a, \varphi) - (1 - \beta) \min_\varphi w(a, \varphi) \right] \quad (17)$$

Note that the production function is identified non-parametrically. Thus, it allows for flexibility in the sign and degree of complementarity on the domain of the function. With the production function identified, one can investigate the degree of complementarities locally and globally. In addition, one can ask how much of the total output can be gained by worker reallocation between firms or jobs reallocation (changes in firm sizes), i.e. how detrimental search frictions are.

The procedure developed for identification of the production function relied on knowledge of the elasticity of the substitution parameter  $\sigma$ . An alternative approach would be to augment the wage equation and to use a linear-regression technique to estimate

the augmented version of it. This would allow a researcher to identify the elasticity of substitution simultaneously with the production function. Therefore, it would provide, in addition, an indirect check of the model’s validity: the estimated value of the elasticity of substitution should lie in the region agreeable with the literature<sup>6</sup>.

The idea behind the alternative production function identification is somewhat straightforward. Rather than invert the wage equation (12) for the match output, one can write it as follows:

$$\ln(w(a, \varphi) - (1 - \beta) \min_{\varphi} w(a, \varphi)) = \chi + \frac{1}{\sigma - 1} \ln R_{\varphi} + \ln \psi(a, \varphi). \quad (18)$$

This is a standard log-linear equation that can be estimated with ordinary least squares methodology. The disadvantage of this identification technique is that it requires variation in the firm revenue at the level of type. In empirical implementation any data set provides two sources of such variation. First, recall that for identification of types we grouped similar workers and similar firms together. Consider a worker-firm pair  $(i, j)$  and denote by  $a(i)$  and  $\varphi(j)$  respectively the worker and the firm type assigned to them during the identification. Then the wage equation for econometric estimation can be written as

$$\ln(w_{ijt} - (1 - \beta) \min_{\varphi} w(a(i), \varphi)) = \chi + \frac{1}{\sigma - 1} \ln R_{jt} + \ln \psi(a(i), \varphi(j)) + \varepsilon_{ijt}, \quad (19)$$

where  $\varepsilon_{ij}$  keeps track of the measurement error in wages. Thus, grouping similar firms would allow identification of the firm production function and the elasticity of substitution. However, the variation in revenue at the firm level must be small by construction, leading to very imprecise estimates

Arguably more importantly, there is inevitable time variation in firm revenue stemming from the business cycle. Although this sort of fluctuation is likely to be the main source of identification of  $\sigma$  in practice, the model so far does not account for productivity fluctuation. More work is needed to understand the extent to which aggregate productivity movement would alter the identification procedure developed. However, if they are small relative to the labour market adjustment velocity, one can conclude that business cycle fluctuations should not distort the identification too much.

Observe that the structure of the model suggests a specification of the wage equation similar to that of Hagedorn et al. (2017), but different from the one usually considered in the literature. Rather than decomposing log-wages into two-way fixed effects components,

---

<sup>6</sup>However, the power of this test is rather low, since consensus on the acceptable values of the elasticity of substitution has not yet been achieved. A survey by Hillbery and Hummels (2013) reports elasticity values in the range from 0.9 to 34.4.

it suggests looking at the firm-specific component of the wage premium over a worker's reservation wage, and not of at the wage itself. This is an implication of strategic wage bargaining at the firm level.

#### 4.4 Discussion of the Size Effect

What difference does firm size make to production function identification? There are two channels through which it plays its role. First, the interaction between the market power and the firm size effectively creates disparity between match output and its marginal value. The same effect would be experienced in the presence of a concave production function, which would lead to disparity between marginal and average products. Second, vacancy creation within the firm might lead to a different marginal posting cost between firms. Indeed, if the vacancy creation cost function is convex, in equilibrium different firm size unequivocally leads to different marginal vacancy costs for different firms. However, most models of one-to-one matching assume a constant entry cost for vacancies, independently of their type. Thus, this unaccounted variation in vacancy cost can become a cause of misidentification of the production function.

To illustrate how these effects may play out, consider the identification of the production function in Hagedorn et al. (2017). Their model is particularly close to the one developed in this paper, lacking only the firm size component. The separate existence of vacancies outside firms gives a value to an unfilled vacancy, and the outcome of wage bargaining accounts for this value. Hence, the production function is identified from the following equilibrium condition:

$$f(a, \varphi) = \frac{w(a, \varphi) - \beta r V_v(\varphi) - (1 - \beta)rU(a)}{\beta}, \quad (20)$$

with  $V_v(\varphi)$  being the value of an unfilled vacancy of type  $\varphi$ . In contrast to a free entry of vacancies into the economy, voluntary creation of vacancies within firms leads to the value of a vacancy being equal to its marginal cost within a firm, but not at the aggregate level.

Therefore, if the data generating process is described by the model developed here, the Hagedorn et al. (2017) identification procedure will identify the following transformation of the production function:

$$\tilde{f}(a, \varphi) = \xi R_\varphi^{-\frac{1}{\sigma-1}} \psi(a, \varphi) + (r + s)c'(v_\varphi), \quad (21)$$

with  $\xi$  being a multiplier reflecting the size of the economy.

The second summand in (21) accounts for the unaccounted variation in the marginal vacancy cost across firms described above. The intuition behind its appearance is the

following. With the match surplus identification coming from the worker side of the model, and therefore being unaltered by the presence of the outside option for firms, the value of the vacancy shifts the identified match output up. Although the vacancy cost does not enter the bargaining procedure in my model, it would be premature to claim that vacancy creation costs do not play a role in wage bargaining, since this result depends on the intricacies of the nature and timing of the vacancy cost, as well as on the bargaining protocol. However, the importance of this effect should not be overstated. Since this effect does not influence the cross derivative of the production function, it is irrelevant to an analysis of complementarities, and to the results of counterfactuals that do not substantially change the distribution of the firm sizes. Even in the latter case, for this effect to be important, the vacancy posting cost function should have a high degree of convexity.

The firm size effect comes from the multiplier  $R_\varphi^{-\frac{1}{\sigma-1}}$  in (21), and exactly accounts for the disparity between match output and its value (or between marginal and average products in the alternative specification). In other words, the one-to-one matching model equates the marginal product and its value, but it identifies the latter. Importantly, this disparity affects the cross derivative of the production function. One would expect that the multiplier is decreasing in firm size. Therefore, conclusions from models that do not take it into account might understate the degree of complementarity between workers' skills and firms' productivity in the economy. This effect might be an important driver behind the modest gains from re-sorting workers found in the literature so far.

Importantly, the one-to-one matching model, and hence the identification procedure, can be considered a limiting case of the model and the identification procedure I develop in this paper. In particular, when  $\sigma \rightarrow \infty$ , aggregation of the intermediate goods into the final goods (1) becomes linear, i.e. intermediate sector firms produce perfect substitutes, and firm boundaries effectively disappear. In particular, as can be seen from (21), both identification strategies identify the same production function (up to an additive shifter). In this empirical sense, the model I develop can replicate a one-to-one matching model.

## 5 Extension: Exporter Wage Premium

In this section, I show how the model can be useful above and beyond addressing the identification of sorting. In particular, I develop an extended version of the model that is relevant to international trade.

Since the seminal work by Melitz (2003), firm heterogeneity and firm size distribution have become important explanatory aspects of new trade models and applications. However, worker heterogeneity is rarely addressed in these models. Grossman, Helpman and

Kircher (2017) study sorting of workers into firms in a different framework. They employ the Heckscher-Ohlin trade model and focus mostly on sorting between, rather than within, industries. In addition, search frictions in their model do not alter the sorting pattern. The closest work in spirit to mine is the model of Helpman et al. (2010). Their model generates positive assortative matching in a similar set-up, due to the functional forms they utilize. Furthermore, in their model better workers are paid more only because they are employed by more productive firms, i.e. personal productivity affects the wage only through the chance of being hired by a better firm, and the exporters pay higher wages solely due to their higher productivity.

My model easily allows for an extension into international trade because it combines a workhorse model of the matching of heterogeneous types with the standard model of heterogeneous firm sizes. Therefore, it is natural to think about such an extension and the effect international trade has on wages in this framework. As I show later, there is room for an exporter wage premium even if the exporters do not differ from non-exporters in their own productivity (type).

I briefly outline the extended model here, relegating a more detailed description to the Appendix. Consider a world of two symmetric countries, with economies consisting of final and intermediate goods sectors, as described in Section 2. The intermediate inputs can now be traded across the border with impediments a la Melitz (2003). The exports involve a fixed cost of numeraire  $f_x$ , which is idiosyncratic to a firm, and an iceberg cost  $\tau$  common to all firms, i.e. to ship one unit of the good into the foreign country,  $\tau$  units of it must be shipped out of the country of origin. Under these assumptions about trade costs, symmetric countries and with the demand structure introduced, the revenue of a firm becomes:

$$R_\varphi = [Y(1 + I\tau^{1-\sigma})]^{1/\sigma} q_\varphi^{\frac{\sigma-1}{\sigma}}. \quad (22)$$

Here,  $I$  is an indicator of exporting activity, i.e.  $I = 1$  if a firm decides to export and  $I = 0$  otherwise. Although in this new environment, the distribution of unemployment and firm sizes would be different from that in the closed economy, Proposition 1 continues to hold. However, although it is hard to track the general equilibrium effects of trade opening on unemployment and production, one could make an interfirm comparison in an open economy. Due to the idiosyncratic cost of exporting, there might be two firms of the same type (production function) one of which is exporting while the other is not. The following proposition summarizes the differences between such two hypothetical firms.

**Proposition 4.** *Consider two firms,  $i$  and  $j$ , such that  $\varphi_i = \varphi_j$ . Assume that in equilibrium firm  $j$  exports and firm  $i$  does not, i.e.  $I(\varphi_j) = 1$  and  $I(\varphi_i) = 0$ . Then*

(i)  $q_j > q_i$  and  $R_j > R_i$ , i.e. the exporting firm is larger measured by output and

revenue;

(ii) if type  $a$  of workers is hired by firm  $i$ , it is hired by firm  $j$ , i.e. firm  $j$  has a weakly larger matching set.

In addition, if the vacancy posting cost function is convex,  $c''(\cdot) > 0$ , then

(iii)  $v_j > v_i$ , i.e. the exporting firm posts more vacancies;

(iv)  $w_j(a, \varphi) > w_i(a, \varphi)$ , i.e. the exporting firm pays a higher wage to any given worker type.

Parts (i) and (iii) of the proposition are quite standard in the trade literature. Since the seminal work by Bernard and Jensen (1999), it has been confirmed both empirically and theoretically that exporting firms are larger than non-exporting ones. Result (ii) is less straightforward, and deserves special explanation. Start with a hypothetical situation in which these two firms have the same output. Due to the availability of a foreign market, the conversion of output into revenue is higher for the exporting firm. i.e. its total and marginal revenue are higher. This drives up the surplus from a match with any given worker, and the exporting firm has incentives to expand, both in terms of vacancy creation and type acceptance. Expansion puts a downward pressure on the surplus, but does it in a uniform fashion across workers. However, since a larger firm has to create more vacancies, in equilibrium the average surplus from its new hires has to be larger as well. Together with the fact that a firm cannot change the surplus from matches with different workers differently, the surplus from any given match should be higher, implying a larger matching set for an exporting firm. From this the last part of the proposition follows immediately, as wages are tightly connected to the match surplus. Thus, the model highlights the different foundation of an exporter wage premium: the cost of supporting the firm size. In this model, large firms have to create more vacancies at a higher marginal cost. The results of the proposition, especially of its second part, rely heavily on this assumption. However, as noted in the previous subsection, this particular assumption can be tested in the future.

The result on comparison of the matching sets is drastically different from that of the model with one-to-one matching in Bombardini et al. (2014). In a world where a firm can match with only one worker, exporting increases the importance of a good match, shifting up and narrowing down the acceptance set of the exporting firm relative to a non-exporting firm. This difference highlights how explicit incorporation of firm size into the search models of labour markets can substantially change the predictions of the models.

## 6 Conclusion

This paper has developed an equilibrium model of matching between workers and firms where firms, as opposed to jobs, have size. In other words, firms make decisions not only about the extensive margin — what types of workers to hire — but also about the intensive margin — how many workers to hire. I have also shown theoretically how equilibrium conditions resulting from the optimizing behaviour of workers and firms allow for identification of the model primitives such as match output and vacancy creation cost functions. I have shown that for the identification, one needs data on workers' wages and firms' revenues and vacancies, which are usually observable in modern linked employer-employee data sets. Importantly, identification of the production function, which is the cornerstone in addressing the question of sorting, is performed non-parametrically.

The proposed identification procedure permits the quantification of the role of search friction in its interplay with the complementarities in production and with firm size. In order to quantify the role of frictions on the extensive margin of hiring, one can compute the change in total output resulting from optimal reassignment of workers to firms, conditional on the observed number of jobs within each firm. Additionally, to assess the role of search frictions with regard to firm size, one can look at the loss in the aggregate output relative to the globally optimal assignment of workers to firms. I believe the empirical quantification of these effects will be an important step in the further advancement of this line of research.

Next, I have extended the model to allow for international trade. This exercise shows the importance of considering firm size in predictions about equilibrium sorting. In particular, in this model exporters have larger matching sets than non-exporters; i.e. they hire more types of workers, in contrast to the prediction of the one-to-one matching model of Bombardini et al. (2014). In addition, this formulation sheds new light on the exporter wage premium: I have shown that the necessity of supporting a larger firm size forces exporters to pay higher wages.

This paper is a first step in the study of the role of firm size in sorting on the labour market. Further advancement of this line of research requires empirical assessment of the model and quantification of the role of the firm size. However, prior to that, the model should be enriched to include prominent features of the data, such as job-to-job transitions and on-the-job search. First, allowing workers to search for a job while employed would relax the scarcity assumption imposed on the workers' side of the labour market in the same way this paper has relaxed the scarcity assumption on the firms' side. Second, I expect that this extension will dramatically improve the performance of the model when faced with data.

## References

- Abowd, J.M., Kramarz, F. (1999) The analysis of labor markets using matched employer-employee data. In Ashenfelter O. and Card D. (Eds.). *Handbook of labor economics 3B* (pp. 2629-2710). Amsterdam: Elsevier North Holland.
- Abowd, J.M., Kramarz, F., Margolis D.N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2), 251-333.
- Abowd, J.M., Kramarz, F., Pérez-Duarte, S., Schmutte, I. (2009). *A formal test of assortative matching in the labor market*. NBER Working Paper No. 15546.
- Atakan, A. (2006). Assortative matching with explicit search cost. *Econometrica*, 74(3), 667-680.
- Bagger, J., Lentz, R. (2015). *An Empirical Model of Wage Dispersion with Sorting*. Unpublished manuscript.
- Bartolucci, C., Devicienti, F., Monzon, I. (2015). *Identifying Sorting in Practice*. Unpublished manuscript.
- Becker, G.S. (1973). A theory of marriage: Part I. *Journal of Political Economy*, 81(4), 813-846.
- Bernard, A. B., Jensen, J. B. (1999). Exceptional exporter performance: cause, effect, or both?. *Journal of international economics*, 47(1), 1-25.
- Bombardini, M., Gallipoli, G., Pupato, G. (2012). Skill Dispersion and Trade Flows. *American Economic Review*, 102(5), 2327-48.
- Bombardini, M., Orefice, G., Tito, M. D. (2014). *Does Exporting Improve Matching? Evidence from French Employer-Employee Data*. Unpublished manuscript.
- Card, D., Heining, J., Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, 128(3), 967-1015.
- Crozet, M., Trionfetti, F. (2013). Firm-level comparative advantage. *Journal of International Economics*, 91(2), 321-328.
- Eeckhout, J., Kircher, F. (2010). Sorting and Decentralized Price Competition. *Econometrica*, 78(2), 539-574.
- Eeckhout, J., Kircher, F. (2011). Identifying sorting—in theory. *Review of Economic Studies*, 78, 872-906.

- Eeckhout, J., Kircher, F. (2012). *Assortative Matching with Large Firms: Span of Control over More versus Better Workers*. Unpublished manuscript.
- Elsby, M. W., Michaels, R. (2013). Marginal jobs, heterogeneous firms, and unemployment flows. *American Economic Journal: Macroeconomics*, 5(1), 1-48.
- Grossman, G. M., Helpman, E., Kircher, P. (2017). Matching, Sorting, and the Distributional Impacts of International Trade. *Journal of Political Economy*, 124(1), 224-264.
- Gruetter, M., Lalive R. (2009). The importance of firms in wage determination. *Labor Economics*, 16(2), 149-160.
- Hagedorn, M., Law, T. H., Manovskii, I. (2017). Identifying Equilibrium Models of Labor Market Sorting *Econometrica*, 85(1), 29-65..
- Helpman, E., Itskhoki, O., Redding, S. (2010). Inequality and unemployment in a global economy. *Econometrica*, 78(4), 1239-1283.
- Helpman, E., Itskhoki, O., Muendler, M.A., Redding, S. (2012). *Trade and inequality: From theory to estimation*. NBER Working Paper No. 17991.
- Hillbery, R., Hummels, D. (2013). Trade Elasticity Parameters for a Computable General Equilibrium Model. In Dixon P. and Jorgenson D. (Eds.). *Handbook of Computable General Equilibrium Modeling 1B* (pp. 1213-1269). Amsterdam: Elsevier North Holland.
- Kaas, L., Kircher, P. (2014). *Efficient firm dynamics in a frictional labor market*. Unpublished manuscript.
- Lise, J., Meghir, C., Robin, J. M. (2013). *Mismatch, sorting and wage dynamics*. NBER Working Paper No. w18719.
- Lopes de Melo, R. (2013). *Firm wage differentials and labor market sorting: Reconciling theory and evidence*. Unpublished Manuscript.
- Melitz, M. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6), 1695-1725.
- Moscarini, G., Postel-Vinay, F. (2014). *Wage Posting and Business Cycles: a Quantitative Exploration*. Unpublished manuscript.
- Shimer, R., Smith, L. (2000). Assortative matching and search. *Econometrica*, 68(2), 343-369.

## A Proof of Proposition 1

I prove the proposition by verifying the following guess: the equilibrium wage is given by:

$$w(a; L_\varphi, \varphi) = \xi_0 r U(a) + \xi_1 \frac{dR_\varphi(q)}{dl(a)}, \quad (23)$$

i.e. the wage is a linear combination of the unemployment value of a worker and his marginal value product. This leads to a firm's wage bill:

$$\int w(a; L_\varphi, \varphi) dL_\varphi(a) = \xi_0 r \int U(a) dL_\varphi(a) + \xi_1 \frac{\sigma - 1}{\sigma} R_\varphi(q). \quad (24)$$

With this at hand, we can move to the firm's problem defined by (7) and (8). The conditional maximization can be written as follows:

$$J(L, \varphi) = \max_{v, L'} \left\{ \frac{1}{1+r} \left[ (1 - \xi_1 \frac{\sigma - 1}{\sigma}) R_\varphi(q) - \xi_0 r \int U(a) dL(a) - c(v) + J(L', \varphi) \right] \right. \\ \left. \int \lambda(a) \left( (1-s)l(a) + m(\theta)v \frac{u(a)}{u} - l'(a) \right) da + \int \mu(a) l'(a) da \right\}.$$

Taking the first order conditions, we obtain:

$$\frac{c'(v)}{1+r} = m(\theta) \int \lambda(a) \frac{u(a)}{u} da \quad (25)$$

$$\frac{1}{1+r} \frac{dJ(L', \varphi)}{dl'(a)} = \lambda(a) - \mu(a). \quad (26)$$

Since for every  $a$  such that  $l'(a) > 0$   $\mu(a) = 0$ ,  $\lambda(a)$  defines the current marginal value of an additional worker of a given type. Thus, the first order condition with respect to  $v$  requires that the cost of a vacancy be equal to the marginal gains from it.

Now we can employ the envelope theorem:

$$\frac{dJ(L, \varphi)}{dl(a)} = \frac{1}{1+r} \left[ (1 - \xi_1 \frac{\sigma - 1}{\sigma}) \frac{dR_\varphi(q)}{dl(a)} - \xi_0 r U(a) \right] + \lambda(a)(1-s). \quad (27)$$

Fix on steady state and an  $a$  so that  $l(a) = l'(a) > 0$ . For these values we can rewrite (27) as

$$\frac{dJ(L, \varphi)}{dl(a)} = \frac{1}{r+s} \left[ (1 - \xi_1 \frac{\sigma - 1}{\sigma}) \frac{dR_\varphi(q)}{dl(a)} - \xi_0 r U(a) \right]. \quad (28)$$

To uncover the left-hand side of (9), observe that from (6) it follows that:

$$V(a; L, \varphi) - U(a) = (w(a; L, \varphi) - rU(a))/(r+s). \quad (29)$$

We can combine the last equation with (28) and (9), obtaining

$$(1-\beta) \left[ \xi_0 r U(a) + \xi_1 \frac{dR}{dl(a)} - rU(a) \right] = \beta \left[ (1 - \xi_1 \frac{\sigma - 1}{\sigma}) \frac{dR}{dl(a)} - \xi_0 r U(a) \right], \quad (30)$$

with a method of indeterminate coefficients yielding

$$\xi_0 = (1 - \beta) \quad \text{and} \quad \xi_1 = \frac{\sigma\beta}{\sigma - \beta}. \quad (31)$$

Hence, the first claim of the proposition.

With this result we can go further. Observe that (5) can be rewritten as

$$rU(a) = b(a) + \frac{\theta m(\theta)}{r+s} \int \gamma(\varphi) \beta \left( \frac{\sigma}{\sigma - \beta} \frac{dR}{dl(a)} - rU(a) \right)^+ d\varphi. \quad (32)$$

Analogously, (25) can be rewritten as

$$c'(v) = \frac{m(\theta)}{r+s} \int \frac{u(a)}{u} (1 - \beta) \left( \frac{\sigma}{\sigma - \beta} \frac{dR}{dl(a)} - rU(a) \right)^+ da. \quad (33)$$

## B Proof of Propositions 2 and 3

By contradiction, suppose that for some  $a' > a$   $U(a') < U(a)$ . This implies that for any  $\varphi$

$$s(a, \varphi) = \frac{\sigma\beta}{\sigma - \beta} B(\varphi) \psi(a, \varphi) - rU(a) < s(a', \varphi),$$

with  $B(\varphi) = \frac{\sigma-1}{\sigma} Y^{1/\sigma} q_\varphi^{-1/\sigma}$ .

From this it immediately follows that

$$rU(a) - b(a) = \beta \frac{\theta m(\theta)}{r+s} \int s^+(a, \varphi) \gamma(\varphi) d\varphi < rU(a') - b(a'),$$

which, together with the assumption of a non-decreasing income flow in unemployment, contradicts the assertion. Thus,  $U(a)$  is increasing in its argument. The part about wage and value in employment for a given  $\varphi$  directly follows from the equilibrium wage rule (12) and value in employment (6). The minimal attainable wage

$$\underline{w} = \min_{\varphi: s(a, \varphi) \geq 0} w(a, \varphi) = rU(a),$$

and thus is increasing in  $a$  as well. The maximum attainable wage

$$\bar{w} = \max_{\varphi: s(a, \varphi) \geq 0} w(a, \varphi) = \max_{\varphi: s(a, \varphi) \geq 0} \left\{ \frac{\sigma\beta}{\sigma - \beta} B(\varphi) \psi(a, \varphi) + (1 - \beta) rU(a) \right\}$$

is increasing by the envelope theorem.

Analogously, assume by contradiction that for some  $\varphi_i > \varphi_j$   $q(\varphi_i) < q(\varphi_j)$ . Then, for any worker  $a$ :

$$s(a, \varphi_j) = \frac{\sigma - 1}{\sigma - \beta} [YM]^{1/\sigma} q^{-1/\sigma}(\varphi_j) \psi(a, \varphi_j) - rU(a) < s(a, \varphi_i).$$

Again, after integration this yields:

$$\frac{(r+s)c'(v_j)}{(1-\beta)m(\theta)} = \int s^+(a, \varphi_j) \frac{u(a)}{u} da < \int s^+(a, \varphi_i) \frac{u(a)}{u} da = \frac{(r+s)c'(v_i)}{(1-\beta)m(\theta)}.$$

If the posting cost function is linear, this is a contradiction. The assumption of convex cost implies the first part of the result. Then, immediately  $R(\varphi) = Y^{1/\sigma} q^{\frac{1-\sigma}{\sigma}}(\varphi)$  behaves analogously to  $q(\varphi)$ .

The firm value  $J(\varphi)$  is trivially increasing in the firm's type. This follows from the simple argument, akin to the revealed preference. Consider two firms of types  $\varphi' > \varphi$ . First, firm  $\varphi'$  can choose to produce an amount of output equal to the equilibrium output of firm  $\varphi$ , and hence have the same revenue, using exactly the same combination of types. However, since the workers in firm  $\varphi'$  are more productive than those in  $\varphi$ , it will require fewer workers and therefore will need to post fewer vacancies. The last step is to show that although firm  $\varphi'$  will pay higher wages to individual workers, the total wage bill will still be smaller than that of firm  $\varphi$ . Given the wage rule (12), the total wage bill is

$$\frac{\beta(\sigma-1)}{\sigma-\beta} R_\varphi + (1-\beta)r \int U(a) dL_\varphi(a),$$

which is straightforwardly smaller for firm  $\varphi'$  under the described scenario. Thus, I have shown that firm  $\varphi'$  can generate the same revenue as firm  $\varphi$  with lower costs. This implies that its equilibrium profit, and value  $J(\varphi')$ , is larger.

## C Extended Model Structure and Proof of Proposition 4

I now allow the world to have two identical countries. Each country is the same as the country described in Section 2. The final good and labour markets are country specific, whereas intermediate goods can be traded across the border with impediments. Now, both home- and foreign-produced varieties of an intermediate good can be used in production of the final good. The production function becomes:

$$Y = \left[ \int_{j \in \Omega^H} q(j)^{\frac{\sigma-1}{\sigma}} dj + \int_{j \in \Omega^F} q(j)^{\frac{\sigma-1}{\sigma}} dj \right]^{\frac{\sigma}{\sigma-1}}, \quad \sigma > 1, \quad (34)$$

where  $\Omega^H$  is the set of intermediate goods produced in the home country and  $\Omega^F$  is the set of intermediate goods imported from the foreign country. This production function generates demand for the intermediate good of the same structure as before:  $p_j = Y^{1/\sigma} q_j^{-1/\sigma}$ . In a symmetric equilibrium, output of the final good is the same in both countries.

An intermediate goods producer can choose whether to sell all quantity produced on the home market, or to export some of it to the foreign country. Markets are assumed to be segregated, so that firms can charge different prices in different countries. To enter the foreign market, a firm must pay a fixed cost  $f_j$  per period in the market, where  $f_j$  is drawn from some distribution  $F(\cdot)$  independently across firms. Additionally, shipping the good to the foreign country involves an iceberg cost  $\tau$ , i.e. to deliver and sell a unit of the good to the foreign country, the firm must ship  $\tau$  units from the home country.

I break the firm's problem down into two steps. First, I describe the optimal way to distribute sales of a given amount  $q$  of the good between two countries and how much revenue it will generate. The answer to this question is the solution to the following net revenue maximization problem:

$$\begin{aligned} \max_{q_H, q_F} Y^{1/\sigma} q_H^{\frac{\sigma-1}{\sigma}} + Y^{1/\sigma} q_F^{\frac{\sigma-1}{\sigma}} \\ \text{s.t.} \quad q_H + \tau q_F = q \end{aligned}$$

The straightforward solution is to distribute the output so that  $q_F = q_H \tau^{-\sigma}$ , and therefore the revenue that can be generated from the given amount  $q$  of the output is

$$R(q) = [Y(1 + I\tau^{1-\sigma})]^{1/\sigma} q^{\frac{\sigma-1}{\sigma}},$$

where  $I$  stands for the indicator of exporting. Now, taking into account the revenue generating function, the firm must decide whether to export, how much output to produce and what type of workers to employ in the production. The slightly modified firm's objective function (7) becomes:

$$J(L_\varphi, \varphi) = \max_{v, L', I} \frac{1}{1+r} \left\{ R(q(L_\varphi, \varphi), I) - \int w(a, \varphi) dL_\varphi(a) - c(v) - If_j + J(L', \varphi) \right\}, \quad (35)$$

subject to the hiring constraint (8) and  $I \in \{0, 1\}$ . The Bellman equations for workers on the labour market do not change. Since exporting does not alter the structure of the firm problem, i.e. it can be solved for each  $I$  with the revenue function scaled up proportionally and then the maximum value can be chosen, the result of Proposition 1 applies, and the firm's vacancy creation policy remains the same. Now I prove Proposition 4.

Consider two firms  $i$  and  $j$  of the same type  $\phi_i = \phi_j$  but due to different fixed exporting cost draws, only firm  $j$  is exporting. Start by contradiction. Suppose that  $q_i > q_j$ . Then,  $\frac{dR_j}{dl(a)} > \frac{dR_i}{dl(a)}$  and  $s(a, \varphi_j) > s(a, \varphi_i)$  for all types  $a$ . Following the optimal vacancy posting rule (13) would imply that firm  $j$  accepts more types and posts (weakly) more vacancies. Given that the production functions of two firms are the same and the interfirm labour force size is proportional to the number of vacancies posted, we arrive at a contradiction. Hence, the exporting firm has a larger output and larger revenue.

However, the output advantage of firm  $j$  cannot be too large. If  $q_j > q_i$  to the extent that  $\frac{dR_j}{dl(a)} < \frac{dR_i}{dl(a)}$ , by analogous reasoning we arrive at a contradiction. Therefore, the exporting firm has a (weakly) higher marginal revenue from and match surplus with every worker. This guarantees that the exporting firm has a larger matching set and posts more vacancies. The last assertion also follows because wages depend on the firm type through the worker's share of the surplus only.