Discussion Paper Series – CRC TR 224

Discussion Paper No. 181
Project A 03

Firm Dynamics with Labor Market Sorting

Andreas Gulyas*

June 2020

* University of Mannheim; andreas.gulyas@uni-mannheim.de

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Collaborative Research Center Transregio 224 - www.crctr224.de
Rheinische Friedrich-Wilhelms-Universität Bonn - Universität Mannheim

Firm Dynamics with Labor Market Sorting*

Andreas Gulyas†

This version: May 2020
First version: November 2016

Abstract

I develop a multi-worker firm model with search frictions, job-to-job transitions, firm dynamics and worker-firm complementarities to study the employment dynamics at the establishment level. Due to the complementarities in production, the ideal worker type changes after productivity shocks, which leads firms to adjust the skill composition of their workforce. Hence, the relationship between changes in workforce quality and firm growth rates in the data informs the strength of complementarities. Using German social security data, I document how firms reorganize the skill composition of their workforce. The estimated model matches many salient facts of establishment level employment dynamics by firm growth rates such as poaching rates, firm size distributions, and the characteristic hockey-stick patterns of the establishment level hire and separation rates by firm growth rates. I decompose the output costs of search frictions and show that the misallocation of jobs and workers across firms generate significant output losses. I conclude that assortative labor market matching is key to understand establishment level employment dynamics.

*This paper is based on a revised version of the first chapter of my Ph.D. thesis. It was previously circulated with the title "Identifying Labor Market Sorting with Firm Dynamics". I thank Lee Ohanian, Pablo Fajgelbaum, Till von Wachter and Pierre-Olivier Weill for invaluable support and advice. I also thank Andy Atkeson, Ariel Burstein, Antonio Ciccone, Kyle Herkenhoff, Andrea Di Miceli, and Ioannis Kospentaris and the seminar participants of UCLA, Vienna Institute for Advanced Studies, UC Louvain, FRB Atlanta, CEU, IIES, University of Mannheim, Bank of Italy, Cambridge University, University of Oslo, NYU Global Macro Workshop, SED 2018, T2M 2018, Search and Matching Conference 2018, Universitat Autonoma de Barcelona, Econometric Society North American and European Summer Meetings 2018 and the NOeG Winter Workshop for useful comments. Financial support from the Austrian Academy of Sciences and the German Research Foundation (DFG) through CRC TR 224 (Project A3) is gratefully acknowledged. All remaining errors are mine.

†University of Mannheim; andreas.gulyas@uni-mannheim.de

1 Introduction

One of the central roles of the labor market is to reallocate workers and jobs across firms. Firms typically hire new workers while at the same time separating from some of their existing workforce. Thus, a significant part of worker turnover does not serve the purpose of reallocating jobs across firms, but rather workers across firms. Indeed, starting with Abowd, Kramarz, and Margolis (1999), a growing literature documents that specific workers tend to work for different types of firms, giving rise to assortative labor market sorting.[1]

In this paper I show that the behavior of worker flows at the establishment level is tightly linked to assortative matching in the labor market. I develop a tractable multi-worker firm model with search frictions, firm dynamics originating from idiosyncratic firm-level shocks and sorting between firm and worker types. The typical identification strategy for assortative labor market sorting in the literature relies on fixed firm types. The introduction of firm dynamics therefore requires a novel identification strategy for the complementarities between unobserved worker and firm attributes that drive the pattern of sorting. The key idea is that firms' reorganization of their workforce in response to productivity shocks will identify the complementarities in production. Intuitively, in a world with positive (negative) assortative matching, firm growth is associated with worker quality upgrading, whereas shrinking firms will reorganize towards lower (higher) skilled workers. The stronger the sorting motive, the more intensively firms are adjusting the skill composition of their workforce in response to shocks. There is surprisingly little evidence on these adjustments. To fill this gap, I use German social security data to document how firms adjust the skill composition of their workforce as they grow and contract. The estimated model replicates a number of salient empirical facts about firm dynamics: (1) symmetric bell shaped firm growth distribution, (2) empirically consistent firm size distribution, (3) realistic relationship between firm size and growth, (4) fast-growing firms hire more workers per vacancy, fill their vacancies faster and disproportionately through poaching from other firms, and (5) "hockey-stick" pattern of hire and separation rates over the firm gowth rate distribution (Davis et al., 2006). The last fact is due to complementarities in production. The ideal worker type changes after productivity shocks, which leads firms to adjust the skill composition of their workforce. Therefore, firms separate from some of the existing workforce while hiring new workers, which generates excess worker turnover after firm productivity shocks. This excess turnover is puzzling from the view of standard firm dynamics models, where firms reduce excess worker turnover to

[1] Since the seminal paper of Abowd, Kramarz, and Margolis (1999), Card et al. (2013), Song et al. (2018) among many others document assortative labor market sorting.

quickly achieve the desired change in size. In contrast, my model provides a theory for the excess worker turnover in growing and contracting firms.

In my model, workers and firms are heterogeneous in their productive capacity and complementarities in production induce sorting in equilibrium, as in Becker (1973). As in Shimer and Smith (2000), search frictions impede the reallocation of workers across firms, so equilibrium sorting is imperfect. To account for the significant fraction of labor reallocation through job-to-job transitions, my model features on-the-job search. I depart from most of the sorting literature in two dimensions. First, I assume that firms face idiosyncratic productivity shocks.[2] Given that a majority of firms change size every quarter (Davis, Faberman, and Haltiwanger, 2006, 2012), it is highly implausible to assume fixed firm types even over short time periods. Second, I relax the assumption of one-dimensional firm productivity. Firm productivity has many dimensions ranging from factors that are fixed over time (or are adjusted at a very low frequency) such as the quality of the capital stock, environmental factors such as access to infrastructure to more transient ones such as managerial talent.[3] It is also easily conceivable that different productivity components interact differently with worker quality in production. For example, it could be that worker skills are complements to the quality of the capital stock, but substitutes to managerial quality. To account for these possibilities, I depart from most of the literature by assuming that firms are heterogeneous in two components. On the one hand, firms are endowed with a fixed firm productivity component, on the other hand they are also characterized by time-varying productivity. In addition, identification is complicated by the fact that worker quality is allowed to interact differently with the two firm productivity dimensions.

The key identification idea is to study how worker skills are distributed across the firm size distribution in the cross section as well as how firms adjust the skill composition of their workforce as they expand and contract. In my setup, more productive firms operate at a larger scale.[4] Because firms face convex job creation costs, firms with positive productivity shocks tend to expand in size whereas firms with negative ones shrink. This allows me to map changes in unobserved productivity to observable changes in firm size in the German social security dataset. In addition, firms adjust the skill composition of their workforce in response to shocks to their time-varying productivity. This reorganization is linked to

[2] To my best knowledge, Lise, Meghir, and Robin (2016) is the only other study with assortative matching considering firm-level shocks.

[3] Syverson (2011)

[4] I follow an extensive literature in economics explaining differences in firm size by productivity differences, see e.g. Hopenhayn (1992), Melitz (2003), Luttmer (2007) and Lentz and Mortensen (2008)

complementarities in production between the time-varying firm productivity and worker types. As in Becker (1973), if the two are complements, high type workers have a relatively higher marginal productivity at high type firms. This implies that high type workers become more valuable to firms with positive productivity shocks, and these reorganize the workforce towards higher skilled workers. With negative sorting, the exact opposite occurs. Low type workers are more valued by firms with currently high productivity and therefore firms downgrade the skill composition of their workforce after positive shocks and upgrade it after negative ones. Therefore, the relationship between changes in average workforce quality and firm growth rates recovers the sign and degree of the complementarities between the time-varying firm productivity component and worker skills. On the other hand, the sorting pattern along the fixed firm productivity component is identified through the cross-sectional relationship between firm size and worker skills. I measure worker types by average annual earnings controlling for observable wage determinants. I show that this metric provides an accurate measure of worker types in my model.[5]

I estimate the model with German matched employer-employee data and find that even though there is positive sorting in the cross section, establishments upgrade their worker skills when they contract and downgrade them when they expand. The multi-dimensional sorting setup is crucial to explain this pattern. Worker productivity is estimated to be a complement to time-invariant firm productivity, which generates the positive cross-sectional sorting. On the other hand, worker skills are identified to be substitutes to time-varying firm productivity. Thus, establishments separate from low type workers when they shrink and hire less skilled workers as they grow. In the data as well as in the model, growing firms replace existing workers with less skilled ones, whereas contracting firms use worker flows to upgrade worker skills. These adjustments generate hiring and separations at the same time, which leads to excess worker turnover after productivity shocks and generates the "hockey-stick" hire and separation rate dynamics documented in Davis et al. (2006). I therefore argue that labor market sorting is key to understand the establishment level behavior of employment dynamics.

I then use the model to quantify the output cost of search frictions, which is generated by three distinct forces. First, lower frictions lead to higher employment rates, which mechanically increases output. Second, they lead more productive firms to gain employment share. Third, the allocation of worker types across firm types improves, which raises output due to

[5]Lopes de Melo (2018) shows in a similar model that worker fixed effects capture the corresponding true worker types closely.

complementarities in production. The decomposition yields that 79% of the output gain of lower search frictions originate from increased employment, whereas 16% from misallocation of jobs across firms and another 5% from the misallocation of worker types across firm types.

My paper contributes to several strands of the literature. First, it relates to the large empirical literature studying employment dynamics at the establishment level. Many papers study how firms adjust their hiring and separation decisions in response to shocks (e.g. Davis et al. (2006, 2012)), but there is surprisingly little evidence on how firms reorganize their workforce in response to shocks. Caliendo, Monte, and Rossi-Hansberg (2015) find that French manufacturing firms grow by adding layers of management and expand preexisting layers with lower skilled workers. Although not directly comparable, I find that German establishments grow by adding lower skilled workers. Traiberman (2019) studies how firms reorganize occupations in response to trade liberalization, whereas I focus on reorganizations based on unobservable worker characteristics in response to idiosyncratic shocks.

My paper is closest to two recent working papers that explain the establishment level behavior of worker flows. In Borovicková (2016) firm productivity shocks are correlated with idiosyncratic match level productivity shocks. This generates excess worker turnover in response to productivity shocks. In contrast, Bachmann et al. (2019) lay out a model where firm dynamics are driven by persistent idiosyncratic separation rate shocks to which firms cannot react immediately due to a time-to-hire friction. This implies that after receiving a surprise separation rate shock, firms decline in size, but start to re-hire workers, generating excess worker turnover. Kaas and Kircher (2015), Schaal (2017), Bilal et al. (2019), among others develop search models with firm dynamics, but do not match or study the pattern of excess worker turnover.

My structural model builds on earlier papers studying wage inequality with search models without sorting. I build upon Postel-Vinay and Robin (2002), Dey and Flinn (2005) and Cahuc, Postel-Vinay, and Robin (2006) to incorporate job-to-job transitions into a search and matching model. I draw upon Postel-Vinay and Turon (2010) to incorporate wage renegotiations after productivity shocks. In addition, my paper is joining a growing literature studying the sorting patterns in labor markets using structural search models. Abowd, Kramarz, and Margolis (1999) (henceforth AKM) pioneered the identification of sorting by correlating worker and firm fixed effects from wage panel data. The fixed effect approach has recently been called into question, Eeckhout and Kircher (2011), Hagedorn et al. (2017), Bagger and Lentz (2018), and Lopes de Melo (2018) show that the AKM fixed effects analysis does not necessarily recover the true sorting pattern.

My paper provides an alternative identification strategy that does not rely on fixed firm types. It also joins Lindenlaub (2017) and Postel-Vinay and Lindenlaub (2017) to study multi-dimensional sorting. Lopes de Melo (2018), Lise and Robin (2017) and Lise, Meghir, and Robin (2016) study sorting in structural models of the labor market. Bonhomme, Lamadon, and Manresa (2019) provide a semi-structural approach to study how firms and workers sort together. Engbom and Moser (2018) provide a structural framework that maps into the AKM framework.

This paper is structured as follows: In the next section, I present the full model. Section 3 discusses the identification of all parameters, and section 4 provides the estimation results. The last section concludes.

2 Model

This section presents the search model with multi-worker firms and firm dynamics that is used to understand the empirical patterns reported in the previous section. The model builds on Shimer and Smith (2000) to study sorting in a frictional environment. I borrow from Postel-Vinay and Robin (2002), Dey and Flinn (2005) and Cahuc et al. (2006) to incorporate job-to-job transitions. Wages are renegotiated after productivity shocks according to the mechanism in Postel-Vinay and Turon (2010).

Time is discrete and the economy is populated with a unit mass of heterogeneous workers and firms. They meet in a frictional labor market to form matches for production. Workers are heterogeneous with respect to a one dimensional productivity type, denoted by x. This summarizes any productive capacity of the worker not observed by the researcher. Most of the literature studying sorting between workers and firms assume that firm productivity is one dimensional and is fixed over time.[6] The notion of fixed firm productivity seems to be at odds with the vast amount of firm dynamics observed in the data. It has long been recognized that firm productivity changes are a driver of firm growth. In addition, firm productivity has many dimensions, ranging from factors that are fixed over time (or are adjusted at a very low frequency) such as the quality of the capital stock, environmental factors such as access to infrastructure to more transient ones such as managerial talent.[7] It is also easily conceivable that different productivity components interact differently with worker quality in production. For example, it could be that worker skills are complements to the quality of

[6] see e.g. Eeckhout and Kircher (2011), Hagedorn et al. (2017), Bagger and Lentz (2018). Lindenlaub (2017) and Lise and Postel-Vinay (2020) provide notions of sorting based on multidimensional characteristics

[7] see Syverson (2011) for a review about the determinants of firm productivity.

the capital stock, but substitutes to managerial quality. To account for these possibilities, I depart from the literature by assuming that firms are heterogeneous in two components. On the one hand, firms are endowed with a fixed firm productivity component z, on the other hand they are also characterized by time-varying productivity y. In addition, the complementarities between worker types and the two different firm productivities is allowed to be different. A match between a worker type x and firm type y, z produces $f(x, y, z)$, where $f(x, y, z)$ is twice continuously differentiable with respect to all of its arguments, with $f_x(x, y, z) > 0$, $f_y(x, y, z) > 0$ and $f_z(x, y, z) > 0$. Thus, high types always have an absolute advantage over low types. Intuitively, which worker types firms will prefer will depend on whether high type workers will be have a higher marginal product at low type, or high type firms. This is determined by the two cross derivatives of the production function $f_{xy}(x, y, z)$ and $f_{xz}(x, y, z)$. The sign and strength of the cross derivatives $f_{xy}(x, y, z)$ and $f_{xz}(x, y, z)$ are left unrestricted and will be estimated. Worker and firm productivities are distriubted on the unit interval [0,1], with the stationary distributions of worker and firm types given by the probability distribution functions $\phi_x(x)$ and $\phi_f(y, z)$.

Firms produce with a linear production technology.[8] Thus, the total production of a particular firm j with productivity (y, z) is given by the integral over the distribution $\psi_j(x)$ of all of its individual matches, or

$$F_j(y, z) = \int f(x, y, z) d\psi_j(x). \tag{1}$$

Firms have a certain number of jobs available, which can be either filled or vacant. These jobs are costless to maintain, but depreciate at rate d each period, irrespective whether they are filled or vacant. A (costless) vacancy is automatically posted for every unfilled jobs. Firm face idiosyncratic shocks to their time-varying productivity, and the transition rate is given by $p(y'|y)$. Firms can costlessly downscale by separating from some of their workers. On the other hand, in order to expand, firms have to create new jobs v^N subject to a convex adjustment cost function $c(v^N)$, with $c'(v^N) > 0$ and $c''(v^N) > 0$.[9] Firms will create new jobs

[8] This assumption rules out any complementarities between workers within a firm. Studying such complementarities would render this model intractable as the surplus of each match would depend on the other matches within a firm.

[9] In my setup, it is better to think that firms create jobs and not just post vacancies. This is because jobs are presistent and only depreciate slowly with rate d. This is in contrast to the setup in Bagger and Lentz (2018), where firms also face a convex vacancy posting cost, but unfilled vacancies depreciate immediately. This implies that firms do not have an opportunity cost of matching, and will accept any worker they meet. In this sense my setup is in the tradition of Shimer and Smith (2000), where scarce jobs need to be allocated to the "right" type of workers.

until the marginal cost of establishing a new job is equal to the marginal value of a vacant job, i.e.

$$c'(v^N) = V(y, z), \tag{2}$$

where $V(y, z)$ is the value of a vacancy to a firm of type (y, z). Inverting this relationship yields the newly created jobs $v^N(y, z)$ for each firm type (y, z):

$$v^N(y, z) = c'^{-1}\left(V(y, z)\right). \tag{3}$$

The number of jobs at the firm level evolves according to the following law of motion

$$n_{t,j}(y, z) = n_{t-1,j}(1 - d) + v^N(y, z). \tag{4}$$

More productive firms will value vacancies higher, and thus will create more jobs and grow larger. Because of the linear job destruction rate and the convex job creation cost, sustaining larger firms becomes increasingly difficult. Thus over time, firms will converge to their steady state firm size. As jobs can either be filled or vacant, firm size will also depend on the matching process, to which I turn next.

Workers can search for jobs on and off the job, but contact potential jobs at different rates. Unemployed workers meet vacant jobs with rate λ_w, whereas employed workers contact them with rate λ_e. The search process in the labor market is undirected. This implies that agents sample from the distribution of searching firms and workers. Firms meet job applicants with rate λ_f, who can either be unemployed or employed at another firm. The mass of unemployed is denoted by u, whereas e^s represents the number of employed workers at the search and matching stage. The total mass of vacant jobs in the economy is v. Conditional on a meeting, the probability of a vacant job contacting an unemployed worker is given by the number of searching unemployed workers divided by all searching workers:

$$p^u = \frac{\lambda_w u}{\lambda_w u + \lambda_e e^s}. \tag{5}$$

Since it must be the case that the total number of meetings on the worker and firm side are the same, the following condition must hold:

$$\lambda_f v = \lambda_w u + \lambda_e e^s. \tag{6}$$

2.1 Wage Negotiation

When a job meets a suitable candidate, the two parties decide on a wage rate that is only renegotiated under certain circumstances. The assumed wage-setting mechanism together with linear utility will ensure bilateral efficiency. This implies that any match with a positive surplus will be formed and maintained. Therefore, the current wage rates will only affect the sharing of the surplus and not the surplus itself.

I denote the value of an unemployed worker of type x as $U(x)$. The value of an employed worker x matched together with a firm of type (y, z) and negotiated wage rate w is $W(x, y, z, w)$. The value of a vacant job to a firm of type y is denoted as $V(y, z)$, whereas the value of a job occupied by a worker of type x with a wage rate w is $J(x, y, z, w)$. The value functions are presented below in equations (14)-(17). The surplus of a match is consequently defined as

$$S(x, y, z) = W(x, y, z, w) - U(x) + J(x, y, z, w) - V(y, z). \tag{7}$$

Wages are negotiated at the beginning of each employment spell and might be renegotiated after outside offers or productivity shocks. At the beginning of the match, the share of the surplus appropriated by the worker depends on whether the worker is hired from unemployment or is poached from another firm. When the worker is hired from unemployment, the wage rate $w^U(x, y, z)$ is set according to Nash bargaining with the worker's bargaining power α. Thus

$$w^U(x, y, z) \quad : \quad W(x, y, z, w) - U(x) = \alpha S(x, y, z). \tag{8}$$

As in Postel-Vinay and Robin (2002), Dey and Flinn (2005) and Cahuc et al. (2006), when an employed worker meets another firm, the two firms engage in Bertrand competition. If the surplus of the poaching firm of type $(\tilde{y}, \tilde{z})$ is smaller than the surplus appropriated by the worker from its current employer, i.e. $S(x, \tilde{y}, \tilde{z}) < W(x, y, z, w) - U(x)$, then the meeting as no effect on the current wage and the worker stays at her current employer. If the surplus is at least higher than the worker's surplus, then the Bertrand competition drives up the wage to the point where the worker obtains the full surplus from the lower surplus firm $S(x, y, z)$. The worker moves to (or stays at) the higher surplus firm $(\tilde{y}, \tilde{z})$ and the wage rate $w^E(x, y, z, \tilde{y}, \tilde{z})$ is set such that:

$$w^E(x, y, z, \tilde{y}, \tilde{z}) \quad : \quad W(x, \tilde{y}, \tilde{z}, w)) - U(x) = S(x, y, z). \tag{9}$$

I denote the set of poaching firms that trigger a job-to-job transition as $\Upsilon^{EE}(x, y, z)$. The set of poaching firms that do not trigger a job-to-job transition, but yield a renegotiation of the wage is denoted by $\Upsilon^{BC}(x, y, z, w)$. Formally these are defined as

$$\Upsilon^{EE}(x, y, z) = \{\tilde{y}, \tilde{z} : S(x, \tilde{y}, \tilde{z}) \geq S(x, y, z)\} \tag{10}$$

$$\Upsilon^{BC}(x, y, z, w) = \{\tilde{y}, \tilde{z} : W(x, y, z, w) - U(x) < S(x, \tilde{y}, \tilde{z}) < S(x, y, z)\}. \tag{11}$$

After productivity shocks, I assume that wages are renegotiated if either the worker's value falls below her outside option ($W(x, y, z, w) - U(x) < 0$) or the firm's value falls below the value of a vacancy ($J(x, y, z, w) - V(y, z) < 0$). The idea behind this assumption is that wages are only renegotiated if one of the parties has a credible threat to leave the match.

The specific wage renegotiation process follows MacLeod and Malcomson (1993) and Postel-Vinay and Turon (2010). New wages are set such that the current wage moves the smallest amount necessary to bring them back into the bargaining set. This is achieved by assuming that the bargaining power of each party depends on which side demands the renegotiation. Intuitively, the side that requests the renegotiation has a weaker bargaining position than the side that prefers the current wage. If a productivity shock pushes the value of a worker below her participation threshold ($W(x, y, z, w) - U(x) < 0$), the firm extracts the full surplus and the wage $w^{NW}(x, y, z)$ is set such that

$$w^{NW}(x, y, z) \quad : \quad W(x, y, z, w) - U(x) = 0. \tag{12}$$

On the other hand, if the current wage becomes too high for the firm to sustain the match, i.e. $W(x, y, z, w) - U(x) > S(x, y, z)$, the worker has the better bargaining position and receives the full surplus. Thus,

$$w^{NF}(x, y, z) \quad : \quad W(x, y, z, w) - U(x) = S(x, y, z). \tag{13}$$

This wage setting mechanism has two appealing features. First, wages feature limited pass-through of productivity shocks, which is in line with recent evidence (See for example Haefke et al. (2013) and Lamadon (2016)). Second, it avoids the situation where inefficient separations occur despite the fact that both parties would have an incentive to renegotiate.

Wages may respond non-monotonically to productivity shocks. Positive productivity shocks might lead to wage cuts and/or separations. It all depends on the strength of sorting and thus on the degree of mismatch between worker and firm types. A match between a firm and a worker of a certain type might become more mismatched after a positive productivity shock because the firm now would prefer different types of workers. This causes the overall surplus to decrease, which might trigger either a separation or a wage cut. The same argument applies to negative productivity shocks. If the worker skill is now a better match to the productivity of the firm, the employee might receive a raise.

Let me consider, for example, the case of positive sorting (along the y dimension). Here, higher type workers are relatively more valued by higher type firms. The mismatch between high type workers and firms with negative productivity shocks will typically increase in this situation, whereas mismatch decreases for lower type workers within the firm. Thus, we might observe wage cuts for high type workers and wage increases for low type workers after negative productivity shocks. It all depends on the how mismatch changes in response to productivity shocks.

2.2 Timing, Matching Sets and Value Functions

Timing is as follows. First, production takes place. After production, a fraction d of jobs are exogenously destroyed and the idiosyncratic productivity shock is revealed. This can trigger endogenous separations. Workers who lost their job in a given period are not allowed to search in the same period again. After the separation stage, the search and matching stage takes place, which concludes the period.

The wage setting mechanism and the assumption of transferable utility assures that acceptance decisions jointly maximize the total surplus. Thus, agents are willing to match together if the match generates a positive surplus, and in case of job-to-job transitions, the prospective surplus is higher than the current one. This implies that the matching sets for unemployed workers are characterized by all productivity combinations that yield a positive surplus, i.e $S(x, y, z) \geq 0$. In the case of job-to-job transitions, a worker x employed at (y, z) is successfully poached by a firm $(\tilde{y}, \tilde{z})$ if the surplus of the poaching firm is higher than the surplus generated by the current employer, or formally if $S(x, \tilde{y}, \tilde{z}) \geq S(x, y, z)$. Equation (14) presents the value of a vacancy at the beginning of the period.

$$\begin{aligned} V(y,z) =& \beta(1-d) \int_{y'} \Bigg\{ V(y',z) + \lambda_f \Bigg(p^u \int_x (1-\alpha) S(x,y',z)^+ \frac{\mu_x(x)}{u} dx + \\ & + (1-p^u) \int_{\tilde{z}} \int_{\tilde{y}} \int_x (S(x,y',z) - S(x,\tilde{y},\tilde{z}))^+ \frac{\psi^S(x,\tilde{y},\tilde{z})}{e^S} dx d\tilde{y} d\tilde{z} \Bigg) \Bigg\} p(y'|y) dy', \end{aligned} \tag{14}$$

where $x^+ = \max\{x, 0\}$. The vacant job might be destroyed with probability d, thus the effective discount rate is given by $\beta(1-d)$. The job contacts an applicant with probability λ_f. The firm has to take into account which workers it might contact during search. First of all, the job seeker can be either employed or unemployed. The vacancy finds a suitable job applicant if either the unemployed worker's type x is inside the matching bands ($S(x,y,z) \geq 0$) or the employed worker of type x is successfully poached away from her current employer by a firm of type $(\tilde{y}, \tilde{z})$ ($S(x, \tilde{y}, \tilde{z}) \geq S(x, y, z)$). The probability of meeting an unemployed worker of type x is equal to the probability of meeting an unemployed p_u times the probability of the unemployed being of type x. The latter is given by $\mu_x(x)$, the measure of unemployed of type x, divided by the total number of unemployed u. Similarly, if the vacant job meets an employed worker, the probability of the job applicant being of type x working for a firm $(\tilde{y}, \tilde{z})$ is given by the probability mass of employed types at the search stage $\psi^S(x, \tilde{y}, \tilde{z})$ divided by the total mass of employed workers at the search stage e^S. As discussed in the previous section, the firm receives a fraction $(1-\alpha)$ of the surplus with a previously unemployed worker. In case the employee has to be poached, Bertrand competition implies that the firm obtains the surplus $S(x, y', z)$ minus the surplus which the worker generated at her old job $S(x, \tilde{y}, \tilde{z})$.

Equation (15) represents the value of a filled job to a firm at the beginning of the production stage.

$$
\begin{aligned}
J(x,y,z,w) = f(x,y,z) - w + \beta(1-d) \int_{y'} \Bigg\{ & V(y',z) + \mathbb{1}(S(x,y',z) \geq 0) \\
\times \int_{\tilde{z}} \int_{\tilde{y}} \Bigg\{ & \lambda_e \mathbb{1}\left((\tilde{y},\tilde{z}) \in \Upsilon^{BC}(x,y',z,w)\right) \left(S(x,y',z) - S(x,\tilde{y},\tilde{z})\right) \\
& + (1-\lambda_e \left(\mathbb{1}(\tilde{y},\tilde{z} \in \Upsilon^{BC}(x,y',z,w) \cup (\tilde{y},\tilde{z}) \in \Upsilon^{EE}(x,y',z)\right) \\
& \times \min\{\max\{J(x,y',z,w) - V(y',z), 0\}, S(x,y',z)\} \Bigg\} \times \frac{\mu_F(\tilde{y},\tilde{z})}{V} d\tilde{y} \Bigg\} p(y'|y) dy'.
\end{aligned}
\tag{15}
$$

It consists of the flow output net of wages $f(x,y,z) - w$ plus the discounted continuation value. Since jobs are destroyed with probability d, the effective discount rate is $\beta(1-d)$. The continuation value depends on the future value of the firm's transient productivity y', which conditional on today's productivity y, occurs with probability $p(y'|y)$. If the match surplus becomes negative or the worker is poached, the firm is left with a vacancy, which the firm values with $V(y',z)$. If the match surplus is positive after the productivity shock, a number of other events affect the continuation value. First, a worker might meet a job of type $(\tilde{y}, \tilde{z})$ from the distribution of vacant jobs $\mu_F(\tilde{y}, \tilde{z})$ with intensity λ_e. The second line in equation (15) represents the case where the poaching firm's surplus is not high enough to successfully poach the worker, but high enough to trigger a renegotiation of wages, i.e. $(\tilde{y}, \tilde{z}) \in \Upsilon^{BC}(x, y', z, w)$. In this case the current firm has to match the surplus of the poaching firm, and its continuation value is $S(x, y', z) - S(x, \tilde{y}, \tilde{z})$. If the worker does not contact a poaching firm, or the poaching firm neither triggers a wage renegotiation nor a job-to-job transition, the worker is retained. In this case, the expression in the last line captures the renegotiation triggered by a productivity shock that leads to a violated of the participation constrained on either side of the match. If the renegotiation is demanded by the worker, the firm extracts the full surplus $S(x, y', z)$. If the firm requires the negotiation, the worker receives the full surplus, thus this case does not feature in the formula above. If no party has a credible threat to leave the relationship, the wage rate remains unchanged and the firm receives $J(x, y', z, w)$.

The worker's value functions are the mirror image of the firms' problems. The value of unemployment is given in equation (16).

$$U(x) \quad = \quad b(x) + \beta \left(U(x) + \lambda_w \int_z \int_y \alpha S(x, y, z)^+ \frac{\mu_F(y, z)}{v} dydz \right) \tag{16}$$

Unemployed workers receive some flow value $b(x)$ during unemployment, which potentially depends on the worker type x. They contact jobs from the distribution of vacant jobs $\mu_F(y, z)$ with intensity λ_u. If the surplus is positive, the match materializes and the worker receives a fraction α of it. Equation (17) presents the value of an employed worker.

$$\begin{aligned} W(x, y, z, w) = w + \beta \Bigg(& U(x) + (1-d) \int_{y'} \Bigg\{ \mathbb{1}(S(x, y', z) \geq 0) \\ & \times \int_{\tilde{z}} \int_{\tilde{y}} \Bigg\{ \lambda_e \left[\mathbb{1}\left((\tilde{y}, \tilde{z}) \in \Upsilon^{EE}(x, y', z)\right) S(x, y', z) \right. \\ & \left. + \mathbb{1}\left((\tilde{y}, \tilde{z}) \in \Upsilon^{BC}(x, y', z, w)\right) S(x, \tilde{y}, \tilde{z}) \right] \\ & + \left(1 - \lambda_e \mathbb{1}((\tilde{y}, \tilde{z}) \in \Upsilon^{EE}(x, y', z) \cup \Upsilon^{BC}(x, y', z))\right) \\ & \times \min\{\max\{W(x, y', z, w) - U(x), 0\}, S(x, y', z)\} \Bigg\} \frac{\mu_F(\tilde{y}, \tilde{z})}{V} d\tilde{y} d\tilde{z} \Bigg\} p(y'|y) dy' \Bigg). \end{aligned} \tag{17}$$

I show in the appendix that the surplus can be expressed as

$$\begin{aligned}
S(x,y,z) =& f(x,y,z) - b(x) + \beta(1-d)\int_{y'} S(x,y',z)^{+} p(y'|y)dy' \\
&- \beta\alpha\lambda_w \int_y\int_z S(x,y,z)^{+}\frac{\mu_F(y,z)}{v}dzdy \\
&- \beta(1-d)\lambda_f \int_{y'}\Bigg(p^u\int_x (1-\alpha)S(x,y',z)^{+}\frac{\mu_x(x)}{u}dx \\
&+(1-p^u)\int_{\tilde{z}}\int_{\tilde{y}}\int_x (S(x,y',z)-S(x,\tilde{y},\tilde{z}))^{+}\,\frac{\psi^S(x,\tilde{y},\tilde{z})}{e^s}dxd\tilde{y}d\tilde{z}\Bigg)p(y'|y)dy'. \qquad (18)
\end{aligned}$$

The first line represents the flow output of the surplus, plus its continuation value, whereas the other terms in lines two - four originate from the outside options $V(y)$ and $U(x)$. The continuation value is independent of poaching events because in case of a job-to-job transition, the Bertand competition assumption implies that the worker will appropriate the current surplus at the new job. Therefore, the continuation value will be $S(x,y')$ independently of a poaching event. Notice how the surplus does not depend on current wages. This is due to the fact that wages only affect the surplus' distribution among the two parties. This simplifies the computational burden because I do not have to simultaneously solve for wage rates. In addition, it circumvents situations where feasible payoffs are non-convex, as studied by Shimer (2006). In that model, non-convex feasible payoffs arise because wages determine the expected duration of employments spells, since higher wages decrease the likelihood of successful poaching. Three distributions emerge endogenously in my model. In a stationary equilibrium the in- and outflows of the distributions of vacancies $\mu_F(y,z)$, unemployed workers $\mu_x(x)$ and employed workers across firm types $\psi(x,y,z)$ must balance each other. Equations (19) - (20) present the law of motions of these three distributions in steady state. Equation (19) first presents the law of motion for the distribution of vacant jobs:

$$\begin{aligned}
\mu_F(y,z) =& \int_{y'}\Bigg(\bigg((1-d)v^N(y',z)M\phi(y',z) + (1-\lambda_f) + \lambda_f\bigg(p^u\int_x \mathbb{1}(S(x,y',z)<0))\frac{\mu_x(x)}{u}dx \\
&+\ (1-p^u)\int_{\tilde{z}}\int_{\tilde{y}}\int_x \mathbb{1}(S(x,y',z)<S(x,\tilde{y},\tilde{z}))\frac{\psi^S(x,\tilde{y},\tilde{z})}{e^S}dxd\tilde{y}d\tilde{z}\bigg)\bigg)p(y|y')\mu_y(y')dy' \\
&+\lambda_e\int_{\tilde{z}}\int_{\tilde{y}}\int_x \mathbb{1}\left(S(x,\tilde{y},\tilde{z})\geq S(x,y,z)\right)\frac{\mu^F(\tilde{y},\tilde{z})}{v}\psi^S(x,y,z)dxd\tilde{y}d\tilde{z} \\
&+\int_{y'}\int_x (1-d)\mathbb{1}(S(x,y',z)<0))\psi(x,y',z)p(y|y')dxdy'. \qquad (19)
\end{aligned}$$

The first two lines comprise the newly created jobs and the unfilled jobs carried over from last period. The total mass of new jobs of type (y', z) created equals the number of jobs created by each firm $v^N(y', z)$ times the total mass of firms of the appropriate type $M\phi(y', z)$. Because jobs are created beginning of the period, they are subject to job destruction and firm productivity shocks. The remaining terms in line one and two describe the events of unsuccessful search. Line three and four represent the inflow from previously filled jobs through endogenous separations.

Equation (20) describes the law of motion for the distribution of worker types across firm types at the production stage $\psi(x, y, z)$.

$$
\begin{aligned}
\psi(x, y, z) =& \lambda_w \mathbb{1}\left(S(x, y, z) \geq 0\right) \frac{\mu_F(y, z)}{v} \mu_x(x) \\
&+ \lambda_e \int_{\tilde{z}} \int_{\tilde{y}} \mathbb{1}\left(y, z \in \Upsilon^{EE}(x, \tilde{y}, \tilde{z})\right) \psi^s(x, y, z) d\tilde{y} d\tilde{z} \frac{\mu_F(y, z)}{v} \\
&+ \left(1 - \lambda_e \int_{\tilde{z}} \int_{\tilde{y}} \mathbb{1}\left(\tilde{y}, \tilde{z} \notin \Upsilon^{EE}(x, y, z)\right) \frac{\mu_F(\tilde{y}, \tilde{z})}{v} d\tilde{y} d\tilde{z}\right) \psi^S(x, y, z), \qquad (20)
\end{aligned}
$$

where the $\psi^s(x, y)$ is the distribution of matches at the search stage. Line one and two describes the inflow through newly filled jobs from unemployment and from poaching, respectively. Line three presents the mass of retained jobs from last period. After production, firms receive productivity shocks and separate from the workers that now lie outside of the matching sets. In addition, a fraction d of jobs are destroyed. This process can be read off equation (21). All the remaining workers engage in on-the-job search.

$$
\psi^s(x, y, z) = (1 - d)\mathbb{1}(S(x, y, z) \geq 0) \int_{y'} \psi(x, y', z) p(y|y') d\tilde{y} \qquad (21)
$$

The distribution of unemployed workers can be readily computed as the residual between the distribution of workers $\phi_x(x)$ and distribution of employed workers $\psi(x, y)$. This yields:

$$
\mu^x(x) = \phi_x(x) - \int_z \int_y \psi(x, y, z) dy dz. \qquad (22)
$$

With the model description finished, the next section describes the identification and estimation strategy.

3 Identification

I use German social security records for the estimation of the model. The particular dataset is the longitudinal model of the Linked-Employer-Employee Data (LIAB LM 9310), provided by the German Federal Employment Agency.[10] The attractive feature of this dataset is that it covers the work histories of all employees of a representative sample of 5000-6000 establishments. These establishments are followed over 10 years, and it contains the social security records from 1993-2010 of each employee who worked in any of these establishment at any point between 1999 and 2009. Workers can therefore be followed before and after they are employed at one of the establishments in the sample. The panel structure of the data allows me to track changes in the worker skill composition in establishments over time. In total, the dataset contains 2,702 to 11,117 establishments per year, and 1,090,728 to 1,536,665 individuals per year. The exact working hours are not recorded, but the dataset contains information whether the worker is a part-time of full-time employee. Since hourly wages can only be constructed for full-time employees, I restrict the sample to full-time employees. The dataset contains a 3 digit occupation identifier, the beginning and end of all employment and unemployment spells precise to the day and the total daily wages and unemployment benefits received. All labor income is recorded up to the maximum social security contribution limit.[11] Further details of the dataset are described in appendix C.

The estimation follows an indirect inference approach. First, I choose a set of auxiliary statistics from the German Social Security data that identify the parameters of the model. Then, I search for a set of parameters that minimizes the distance between the computed auxiliary statistics from my model and the target values. This section describes the choice of functional forms and targeted moments and justifies their roles in the identification of the sorting pattern.

3.1 Functional Form Assumptions

The model is estimated at a monthly frequency. The functional form assumptions are summarized in table 1. I restrict the production function's cross derivative between y and z to be zero, ie. $f_{zy}(x, y, z) = 0$. This identification assumption will make it easier to empirically recover the complementarities between worker and firm types, which is the point of interest in this study. I use the sum of two CES production functions of the

[10] Heining et al. (2013) provide a detail data documentation.

[11] In the model estimation I also top-code income at the 90th percentile.

form $f(x,y,z) = \nu_y \left(x^{1/\rho_y} + y^{1/\rho_y}\right)^{\rho_y} + \nu_z \left(x^{1/\rho_z} + z^{1/\rho_z}\right)^{\rho_z}$, where ν_y and ν_z are normalizing constants, such that both parts are bounded between 0 and 1. ρ_y and ρ_z determine the cross products $f_{xy}(x,y,z)$ and $f_{xz}(x,y,z)$, and thus the sorting patterns in this economy. If $\rho_y > 1$ (or $\rho_z > 1$), then the production function is supermodular, which implies that high type worker are relatively more productive at firms with high transient (fixed) firm productivity. Thus conditional on z, following the logic from the frictionless case (see Becker (1973)), we would expect that high y firms value high type workers relatively more. The sorting patterns will not only depend on the complementarities in production, but on other factors such as the strength of search frictions, or the frequency of firm productivity shocks. Thus it is essential to estimate these together.

Table 1: Functional forms

Worker distribution	Log-Normal(μ_x, σ_x)
Firm type z distribution	Beta$(a_\beta$,$b_\beta)$
Production function	$\nu_y \left(x^{1/\rho_y} + y^{1/\rho_y}\right)^{\rho_y} + \nu_z \left(x^{1/\rho_z} + y^{1/\rho_z}\right)^{\rho_z}$
Flow utility unemployed	$b(x) = f(x,0,0)$
Job creation cost function	$c_0 \left(\frac{v}{c_1}\right)^{c_1}$
Firm shocks	$f(y'\mid y) = \begin{cases} y & \text{with prob. } 1-\phi \\ y' \sim N(y,\sigma_y) & \text{with prob. } \phi \end{cases}$

Notes: Log normal distribution is truncated to [0,1]. Since $y \in [0,1]$, the probability mass that falls outside the support of y is added to the respective extreme value.

The worker distribution is assumed to be log-normal, with location parameter μ_x and scale parameter σ_x truncated to the support [0,1]. I assume that higher type workers are more productive at home production. $b(x)$ set to the lowest possible market production level, i.e. $b(x) = f(x,0,0)$. The invariant firm type distribution is assumed to be Beta, with shape parameters a_β and b_β. Firm productivity shocks follow a Markov process. Shocks occur with Poisson frequency ϕ. In this case, the new productivity y' is drawn from a normal distribution with the old value as a mean and standard deviation σ_y, i.e. $y' \sim N(y, \sigma_y)$.[12] A similar firm productivity process is assumed in Kaas and Kircher (2015). This Markov process implies a normal steady state distribution of firms across y types. The endogenous distribution of jobs across productivity types will be primarily governed by the job creation

[12] Since firm productivity y is bounded between [0,1], firm productivity might fall outside this range. To circumvent this, all the probability mass that falls outside the support of y' is added to the extreme value.

cost function. Here I assume the standard form $c(v) = c_0 v^{c_1}/c_1$, where c_1 determines the convexity and c_0 the scale of the job creation costs.[13]

Six parameters are preassigned. First, since the job creation cost is measured in units of the final good, the model admits one normalization. I normalize the firm level output to lie between 0 and 2, and set the production function scale ν_y and ν_z such that both CES functions have maximum possible output levels equal to one. There is a unit mass of worker is the economy. The mass of firms M is set such that it equals the ratio of firms to workers in the German social security data. I set the discount rate to 0.995, which implies a yearly discount rate of about 6 per cent. The bargaining power of unemployed workers is set to 0.3, which is similar to the values used in Bagger and Lentz (2018) or Lise et al. (2016). I further restrict the shape parameter α_β to be one. The rest of the parameters are estimated to minimize the distance between the auxiliary statistics computed with the German social security data and model-generated data.

I discretize the model with 20 worker types and 160 firm types.[14] First, I obtain the acceptance sets by solving for a fixed point in the surplus function $S(x, y, z)$ and the endogenous distributions $\psi(x, y, z)$, $\mu_F(y, z)$ and $\mu_x(x)$. I then simulate 511239 workers and 6126 firms over 18 years to construct a panel data set that yields exactly the same number of person-year and firm-year observations as the German social security data. Appendix B describes the numerical implementation in detail. I compute a the set of auxiliary statistics on the model simulated data exactly the same way as on the German social security data, which selections I discuss in the following subsections.

Table 4 summarizes the target statistics and their values in the German social security data along with their values obtained from the model simulation. None of the parameters has a one-to-one relationship to the auxiliary statistics, but I provide a heuristic explanation of the underlying identification in the next subsections.

3.2 Identification of Parameters

The key identification challenge in search models with complementarities in production is to recover the complementarities in production and thus the sorting pattern. Before I discuss these, I turn to the identification of the other parameters, which are more standard. I target the total hire rate, together with the unemployment and job-to-job transition rate. The

[13] Bagger and Lentz (2018), Coşar et al. (2016) and Merz and Yashiv (2007) among many others use this functional form

[14] 10 grid points along the z dimension, and 16 grid points along the y dimension

hire rate is defined as total number of hires normalized by employment.[15] I compute the unemployment rate as the fraction of workers not employed each year on the reference day of December 31. I count every transition from one firm to another with an intermittent spell of non-employment shorter than 31 days as a job-to-job transition.[16] These three moments inform the meeting rates for employed and unemployed workers λ_e, λ_w and the job destruction rate d in the model.

To recover the worker productivity distribution, I leverage the insight that in any model where higher types have an absolute advantage in production, their utility level, and thus the discounted sum of all per-period earnings must be increasing in types.[17] Therefore, higher type workers will achieve higher average lifetime labor market earnings and a thus I use worker fixed effect as the worker quality measure.[18] In addition to ability, per period wages in my model also depend on firm productivity, as well as the bargaining position at the time of the hire. Over long time periods, these effects are washed out, but over shorter time periods, the worker fixed effect could still be influenced by these factors. To address this, I will use the estimated model to quantify how well the worker fixed effects recover worker types given the estimated labor mobility patterns and the 18 year horizon of the data. My model does not feature any life-cycle component and thus I follow Card et al. (2013) and Hagedorn et al. (2017) and compute wages after filtering out an education specific age profile and year effects.[19] Further details are described in appendix C. I recover the mean

[15]In computing labor market transitions, I exclude temporary layoffs where the non-employment spell is shorter than 31 days and the worker joins the same firm again.

[16]Although the sample and methodology differ slightly, Jung and Kuhn (2014) find similar hire and job-to-job transition rates in their study comparing worker and job flows in Germany and US.

[17]Consider two types of agents with $x_1 < x_2$. It must be the case that agent x_2 can achieve at least the utility level of x_1. This is because x_2 could just follow the acceptance and wage strategies of x_1. If all counter-parties will accept to match with her under these conditions, she will receive at least the value of the lower type. This must indeed be the case. If firms are willing to hire x_1 agents, they will also be willing to hire x_2 agents with the same conditions since these agents produce more and hence yield strictly higher profits. And if workers are willing to match with x_1 firms, they will also be willing to match with x_2 because wages and separation probabilities are the same by construction. Thus, x_2 agents will always have weakly higher payoffs as x_1 agents.

[18]Strictly speaking, higher lifetime labor market earnings do not necessarily translate to higher lifetime utilities $U(x)$ if high type workers experience longer unemployment duration. To address this, I calculate two additional specifications: First, I impute as the flow value of unemployment the actual unemployment benefits the person is receiving, second the benefits plus a 20 percent premium representing non-monetary payoffs from unemployment such as home production and leisure. The correlation between the baseline specification of only using the average labor market earnings and these two worker quality measures is between 0.9955 and 0.999. The reason behind this is simple: workers do not spend much time in unemployment. Concluding that the choice is inconsequential, I stick with the baseline specification.

[19]I compute the wage residual controlling for year effects and a cubic polynomial of age fully interacted with educational attainment.

and standard deviation of the worker productivity distribution by targeting the mean and standard deviation of the empirical worker fixed effect distribution. This will identify the scale and shape parameters of the worker type distribution μ_x and σ_x.[20]

The rest of the target moments mostly identify the parameters on the firm side. In order to map unobservable changes in productivity to observable changes in the dataset, I use the fact that firm employment expands after positive shocks whereas firms with negative shocks scale back their operations. Since more productive firms value vacant jobs higher,[21] i.e. $\frac{\partial V(z,y)}{\partial y} > 0$ and the job creation equation (3), more productive firms create more jobs and hence grow larger.[22]

The parameters that affect the growth rate and establishment size distribution are the parameters c_0, c_1 of the job creation function, ϕ, σ_y that govern the frequency and range of productivity shocks, as well as the shape parameter of the invariant firm productivity distribution b_β. To identify these parameters, I target the standard deviation of establishment growth rates, the autocorrelation of establishment size, five points over the firm size distribution, the standard deviation of the average worker fixed effect at the firm level, and the job filling rate. I compute the empirical job filling rate from the average time to fill a vacancy provided by the Institute for Employment Research, averaged over all time periods available.[23]

3.3 Identifying the Complementarity Parameters ρ_y and ρ_z

The only two parameters left to pin down are the complementarity parameters ρ_y and ρ_z. The identification of the complementarities in production in search models have proven to be challenging. (Eeckhout and Kircher, 2011). The underlying problem is that wages are potentially non-monotonic in firm types, with high type firms not paying necessarily higher wages. The intuition is simple: In a model with complementarities, a high type firm might only agree to hire a relatively unproductive worker type if the worker accepts a large enough wage cut to compensate the firm for option value of matching with a relatively more productive worker.

[20] I normalize both the empirical worker quality measure and the one obtained from the model to [0,1].

[21] The proof behind this follows the same logic as for workers. Lower type firms can always imitate the matching strategies of lower type firms, and thus obtain at least the same utility, as they have an absolute advantage in production

[22] It is theoretically possible that the job filling probability is lower for higher type firms as they might be "pickier". In practice, as high type firms have higher opportunity costs of waiting, the matching bands are wider and thus high types firms will have higher job filling rates.

[23] Unfortunately, the data is only available from 2010 onward. I use data from 2010 to 2015. Source: http://www.iab.de/stellenerhebung/download

This non-monotonicity in wages has been demonstrated before by Eeckhout and Kircher (2011) and Lopes de Melo (2018), amongst others. In my model the identification is further complicated by the fact that firms do not have a one dimensional productivity measure and that through productivity shocks, firms change types over time. But firm productivity shocks open up the possibility to study how firms adjust the quality of their workforce in response to the shocks. This will inform ρ_y, the complementarity between worker productivity and time-varying firm productivity. Firms that receive productivity shocks adjust the skill level of their workforce, but also their scale of operations. Firms with positive shocks expand and hire additional workers, whereas firms with negative shocks downscale. How employers change the quality of their workers depends on the complementarities in the production function. If worker and time-varying firm productivity are complements in the production function, positive sorting prevails in equilibrium along this dimension, with similar types matching together (Becker, 1973). As a result, expanding firms reorganize their workforce towards higher quality, whereas shrinking firms adjust towards lower quality workers. If worker and time-varying firm productivities are substitutes in production, firms with a positive productivity shock expand and reorganize towards lower type workers. The relation between firm growth rates and the change in the average quality of their workforce uncovers the underlying complementarities in production.

A compact way to summarize how firms reorganize their workforce composition in response to shocks is to run the following regression on either the German matched employee-employer or model simulated data:

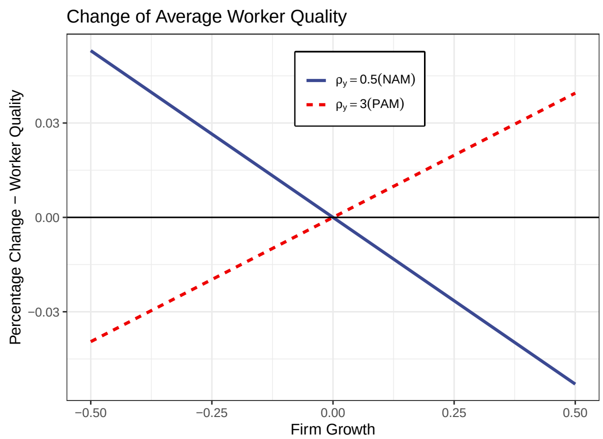
$$\Delta_{\%}\overline{WFE}_{jt} = \alpha + \gamma growth_{jt} + \epsilon_{jt}. \tag{23}$$

Here, $\Delta_{\%}\overline{WFE}_{jt}$ denotes the percentage change in average worker type at establishment j during year t, using the above described measure of worker types. I compute the average worker quality within establishments at the end of each calendar year by averaging the employees' worker quality measure. Then $\Delta_{\%}\overline{WFE}_{jt}$ is simply the yearly percentage change of this measure. $growth_{jt}$ is the percentage change of employment in establishment j during year t.

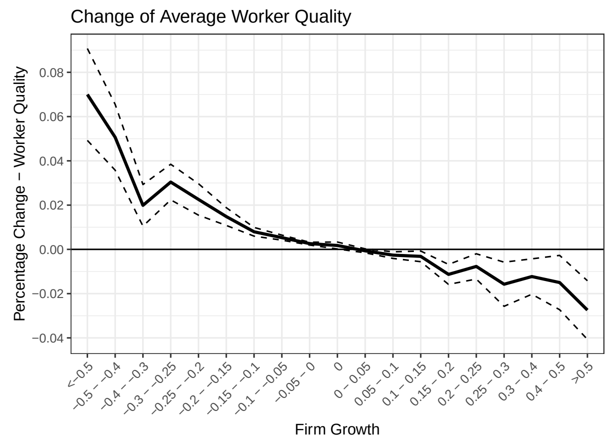
If worker productivity is a complement to y ($\rho_y > 1$), high type workers have a higher marginal productivity at high type firms. This implies that high type workers become more valuable to firms with positive productivity shocks and thus they increase the average level of their employees' skills. By the same argument, firms with negative shocks downgrade the average skills they employ. With negative sorting, the exact opposite occurs. Low

Figure 1: Regression Slope for Different ρ_y

(a) Worker Skill Adjustments - Theory



(b) Worker Skill Adjustments - Data



Notes: The left figure shows the estimated relationship between firm growth rates and the percentage changes in average worker quality employed by firms using regression equation (23) for different values of ρ_y on model generated data. The rest of the parameters are held constant at the values reported in table 5. The right panel shows the same relationship in the German social security data, using growth bins. The sample consists of all establishments with size≥20. Standard errors are clustered at the 3 digit industry level. Broken lines indicate 95% confidence intervals. Establishment growth rates and percentage changes in average worker quality are yearly.

type workers are more valued by high type firms. Therefore, firms downgrade their average worker skills after positive shocks and upgrade them after negative ones. This implies that under positive sorting, γ is be estimated to be positive, whereas it is negative under negative sorting. This logic can also be seen in figure 1a, which shows the estimated relationship from the regression equation (23) with model simulated data. Under positive sorting ($\rho_y > 1$), expanding establishment upgrade the worker skills and shrinking ones downgrade them. If worker and firm types are substitutes ($\rho_y < 1$), a negative relationship between $\Delta_{\%}\overline{WFE}_{jt}$ and $growth_{jt}$ is estimated.

Figure 1b plots the results of regression equation (23) for German establishments using social security records. Instead of a continuous measure of growth rates I use 5% establishment growth rate bins.[24] The results indicate that firms indeed adjust the skill composition of their workforce after shocks. Average worker quality decreases in expanding firms, whereas it increases in declining firms. Firms in general do not reorganize their workforce completely. Establishments that grow or shrink by less than 25% on average change the average worker quality by not more than three percent. Only firms with big shocks reorganize more aggres-

[24] Towards the extremes of the growth rate distribution where the sample size gets smaller, I use 10% bins.

sively. The coefficients are very precisely estimated, as the narrow 95% confidence intervals show.[25] This relationship shows that in shrinking firms, the workforce composition shifts towards workers with higher average lifetime residual earnings. Hence, this is not driven by firms separating from workers with low match qualities or with currently low wages, nor by selection based on the observable characteristics of workers (age and education).[26,27] This is another important advantage of identifying worker quality by their average lifetime earnings rather than ranking workers based on their current wage, which might be affected by the bargaining position or the economic conditions at the beginning of the employment relationship. That being said, as figure 9 in the appendix shows, the relationship is very similar if one considers worker wages instead of the worker fixed effects.

The relationship between firm growth and changes in the average worker fixed effect is almost perfectly linear over the entire range of the growth rate distribution, hence the regression with a continuous growth measure is a good representation. I will use the coefficient γ from regression (23) as one of the target moments in my indirect inference approach. Table 2 presents the baseline estimate in column 1 that will be used as a target in the estimation. The estimated slope coefficient γ is -0.068, which mimics the slope of the relationship in figure 1b.

In addition, table 2 and table 3 report a battery of robustness exercises to show that these adjustments are a robust and salient feature of firm dynamics. Columns two and three of table 2 show that the estimated coefficient barely changes if I use alternative specifications such as weighting by firm size, additionally controlling for firm age, firm size, as well as a 2 digit industry indicator fully interacted with a year effect, or using firm fixed effects as controls. Older firms further show very similar adjustments patterns. The relationship is also robust with respect to considering short versus long term fluctuations. The adjustments look identical if I use a three year window instead of the year-to-year changes (column 6). Column 7 shows that the negative relationship between firm growth and worker quality changes still hold if short term fluctuations are filtered out. For this, I regress the trends in firm growth onto trends in worker quality changes and still find the negative relationship, although the slope is somewhat lower.

[25] Only at the extremes of the growth rate distribution, the standard errors get larger because of the low number of establishments in those growth bins.

[26] The wage residuals are by constructions orthogonal to the observed characteristics.

[27] If I use changes in average wages instead of my worker quality measure, then selection based on observable characteristics are included in addition to selection based on permanent unobserved worker skills. It still holds that firms separate from their lowest earning workers and hire workers with wages below the current firm's median.

Table 2: Regression Results

Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Firmgrowth	-0.068	-0.068	-0.066	-0.073	-0.072	-0.068	-0.047
SE	0.005	0.009	0.007	0.005	0.006	0.005	0.01
Sample/Model	Baseline[1]	Weighted[2]	Additional Controls[3]	Firm FE[4]	Old Firms[5]	3 Yr Window[6]	Trends[7]
N	19912	19912	19912	19912	9327	14045	2066
Adj. R^2	0.058	0.059	0.097	0.087	0.076	0.053	0.01

Notes: Establishment level regressions of firm growth rates on percentage change in average worker fixed effects. Standard errors are clustered at the firm level. The samples are restricted to establishments with more than 20 employees and growth rates between -0.75 and 0.75. See text and notes below for detailed explanation of the different specifications.

[1] Baseline: Yearly changes

[2] Weighted by Firm Size, otherwise baseline specification

[3] With Controls: Additionally controlling for firm age, firm size, 2 digit industry, Year fixed effect and interaction of industry and year effect.

[4] Baseline specification, in addition controlling for firm fixed effect.

[5] Establishments older then 15 years

[6] Baseline regression with 3 year windows for changes

[7] Regression using 5 year trends in firm growth and worker FE changes.

Table 3: Regression Results

Specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Firmgrowth	-0.032	-0.049	-0.065	-0.059	-0.069	0.000	-0.065	-0.043
SE	0.004	0.005	0.007	0.008	0.010	0.000	0.008	0.005
Sample/Model	Core Occup.[1]	W/O Tenure[2]	High Urate[3]	Large Cities[4]	Large Firms[5]	Lagged[6]	No Labor Shortage[7]	Worker FE pre 1999[8]
N	14423	19912	9518	4970	4837	17165	5640	18375
Adj. R^2	0.016	0.03	0.051	0.055	0.067	0.000	0.053	0.025

Notes: Establishment level regressions of firm growth rates on percentage change in average worker fixed effects. Regressions are weighted by establishment size. Standard errors are clustered at the firm level. The samples are restricted to establishments with more than 20 employees and growth rates between -0.75 and 0.75. See text and notes below for detailed explanation of the different specifications.

[1] Baseline regression restricted to workers in mode 2-digit occupation within each establishment

[2] Worker fixed effect computed including a control for tenure, otherwise identical to baseline

[3] Baseline regression in years with above average unemployment rate

[4] Baseline regression restricted to cities with population above 500,000

[5] Baseline regression restricted to establishment with size ≥ 190

[6] Regression with worker quality adjustments regressed on previous years firm growth rate.

[7] Baseline regression restricted to establishment reporting neither "Difficulties in finding the required specialized personnel on the labor market" nor "Staff shortage" as expected hiring problems in the IAB establishment panel survey.

[8] Worker quality measure used is computed only with information pre 1999.

Before I turn to the estimation results in the next section, I discuss other potential explanations for the negative relationship that my model abstracts from. My model postulates that changes in worker quality are driven by the complementarities in production and thus after productivity shocks, firms prefer different worker types. Perhaps the negative relationship is purely driven by the fact that firms expand in lower paid occupation, and thus compositional changes in the occupational structure fully explain the changes in worker fixed effect. To address this, I reestimate the regression equation conditioning on the core 2 digit occupation in the establishment. Column 1 of table 3 shows that within its narrow core occupation, firms still downgrade worker skills when they expand. Column 2 also shows that the relationship is also not purely driven by a first-in first-out hiring and firing behavior. If I use the worker fixed effect after additionally controlling for tenure, I still find a significant negative relationship.

Another potential explanation why firms expand with lower skilled workers is that they could not find better workers. First, the relationship is linear and does not show a kink around zero, which implies that the motivation behind the adjustments do not change abruptly from shrinking to expanding firms. Second, in my model firms also take into account the availability of worker skills. The strength of adjustments depends not only on the production function, but also the distribution of worker skills, but also on the frequency of productivity shocks as well as the search frictions in the labor market. This highlights the importance to use a structural model to interpret the documented negative relationship between growth and changes in the worker skill composition, as the strength of the adjustments depend on many factors that my model captures. Both empirically and in my model firms poach away workers from other firms, so they are not only limited to the pool of unemployed workers. And indeed, as will be shown in the next section, in my model as well as empirically it is the case that expanding firms increasingly make use of poaching.

Nevertheless, my model abstracts from some details of the labor market, such as spatial geography, or monopsonistic firms. It is perceivable that the negative relationship is driven for example by monopsonist firms in remote locations that face a downward sloping labor supply curves in worker skills. If this would be the case, we would expect that the relationship would vanish or at least considerable weaken in times of high unemployment, or in large cities. In both situations, the pool of available workers is larger, and thus it is easier to find the right worker types. Column 3 and 4 show that this is not the case, the relationship is essentially unchanged during years with above average unemployment rates and when conditioning on the largest cities in Germany. If labor supply constraints are the driving force, we should see stronger adjustments in large firms, as these must run quicker into shortages. As column 5 shows this is not the case, large firms show quantitatively the same responses. In addition, if expanding firms are forced to match with lower type workers, we would expect that firms keep adjusting the quality in the years after. But this is not the case, the relationship between changes in worker skills and lagged growth is zero, as reported in column 6. Furthermore, there is direct survey evidence whether firms in the sample face difficulties hiring the required worker skills. As part of the IAB Establishment Panel Survey, a subsample of the establishments in the sample are surveyed about general hiring problems. Perhaps most convincingly, column 7 shows that the estimated relationship is virtually unchanged for firms that report neither "Difficulties in finding the required specialized personnel on the labor market" nor "Staff shortage" as expected hiring problems in the IAB establishment panel survey. A further robustness check concerns that firm growth might affect the estimation of worker fixed effects, which might introduce spurious correlations between firm growth and changes in firm-level worker fixed effects. To address this, I compute worker fixed effects only on observations before 1999, which predates the firm level analysis. Column 8 shows that I still find a significant negative relationship between firm growth and worker skill adjustments.

Figure 10 in the appendix shows that the negative relationship between growth rates and worker quality adjustments holds for all Nace-1 industries except hotels and restaurants, which is imprecisely estimated. Summarizing, the negative relationship between firm growth and changes in worker skills is a salient and robust fact of firm dynamics.

The identification of the last parameter, the complementarity in production between the time-invariant firm productivity and worker productivity ρ_z follows a similar logic. Conditional on all other parameters in the model, the strength of the relationship between firm size and the average worker quality within an establishment maps into ρ_z. Hereby I estimate the following cross-sectional regression both on the German social security data as well as model simulated data:

$$\log(\overline{WFE}_{jt}) = \alpha + \theta \log(fsize_{jt}) + \epsilon_{jt}, \tag{24}$$

where $\overline{WFE}_{jt}$ is the average worker quality and $fsize_{jt}$ is the number of employed workers in establishment j in period t. The higher the estimated coefficient θ, the stronger is the complementary between worker and time-invariant firm productivity, and hence the larger is $\rho_z > 1$. Large firms are typically more productive and pay more higher wages.[28] Since

[28]e.g. Idson and Oi (1999)

Table 4: Target Moments

Target Moment	Data	Model
Separation rate	0.021	0.021
Unemployment rate	0.129	0.121
Job-to-job transition rate	0.006	0.006
Job filling rate	0.388	0.391
Mean worker type distribution	0.480	0.480
Std. of worker type distribution	0.220	0.295
Std. of firm growth rates	0.314	0.313
Autocorr. of firm size	0.994	0.994
Std. of $\bar{W}_{j,t}$	0.160	0.080
Firm Size Distribution (5 Points)	see figure 2b	
Panel Regression Coeff Equation (23)	-0.068	-0.067
Cross-sectional Reg Coeff Equation (24)	0.119	0.108

Notes: Fit of identifying moments. See text for details.

the above findings show that expanding firms downgrade worker types, does this imply that larger firms employ on average lower types? Table 8 reveals that this is not the case, mirroring findings of postive sorting in the cross section reported by Card et al. (2013) and others. This highlights the importance of using a multidimensional sorting setup. With one-dimensional sorting, the worker skill adjustments would simply mirror the cross-sectional sorting direction. The next section proceeds with the discussion of the estimation results.

4 Estimation Results

4.1 The Fit of the Moments

Table 4 presents the fit of all target moments and table 5 displays the estimated parameter values. Overall, the model closely matches almost all moments. The rates of separation, job-to-job transitions and job fillings are matched closely. The estimated job destruction parameter d implies that jobs are on average exogenously destroyed every 5.5 years. There are small deviations from the targeted standard deviation of the worker fixed effect distribution, but its mean is perfectly matched.

The fit of the coefficient from regression (23) is of particular interest, since it identifies the key parameter ρ_y which drives the sorting pattern. The linear regression on model simulated data yields up to a rounding error the same coefficient as obtained from the

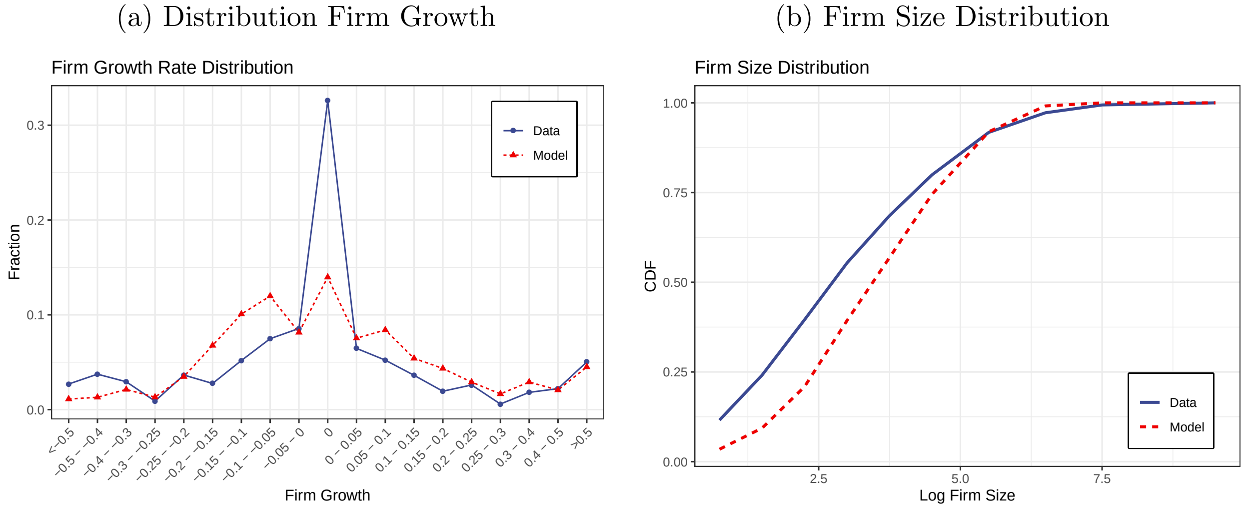
Table 5: Parameter Estimates

Parameters	Symbol	Value
Preassigned Parameters		
Discount Factor	β	0.995
Mass of firms	M	0.012
Worker Bargaining Weight	α	0.300
Time-invariant firm prod.	a_β	1.000
Calibrated Parameters		
Complementarity y	ρ_y	0.684
Complementarity z	ρ_z	9.501
Worker dist. location	μ_x	-0.642
Worker dist. scale	σ_x	1.499
Meeting rate workers	λ_w	0.114
Job-to-job meeting rate	λ_e	0.013
Job destruction rate	d	0.015
Job creation cost, scale	c_0	7.246
Job creation cost, convexity	c_1	1.206
Firm shocks, persistence	φ	0.851
Firm shocks, variance	σ_y	0.154
Time-invariant firm prod.	b_β	2.733

Notes: Estimated parameter values. See text for explanation.

German dataset. The parameter ρ_y is estimated to be 0.684, which implies that worker and time-varying firm types are substitutes in the production function. This implies that negative sorting prevails along the y dimension in equilibrium. In contrast, ρ_z is estimated to be 9.501, and hence worker and *time-invariant* firm productivity are complements in production, which leads to positive sorting in the cross section. The negative sorting in the y dimension is considerably weaker than the positive sorting along the z dimension. The implied correlation between worker types x and time-varying firm productivity y is -0.1, whereas the model estimates the correlation between x and time-invariant firm productivity z to be 0.221. This finding is not surprising as studies that abstract from the time-varying firm productivity component typically find positive sorting in Germany (see e.g. Card et al. (2013) or Hagedorn et al. (2017)). The apparent discrepancy between the positive cross-sectional relationship of firm size and worker skills and the negative adjustments along the

Figure 2: Model Fit



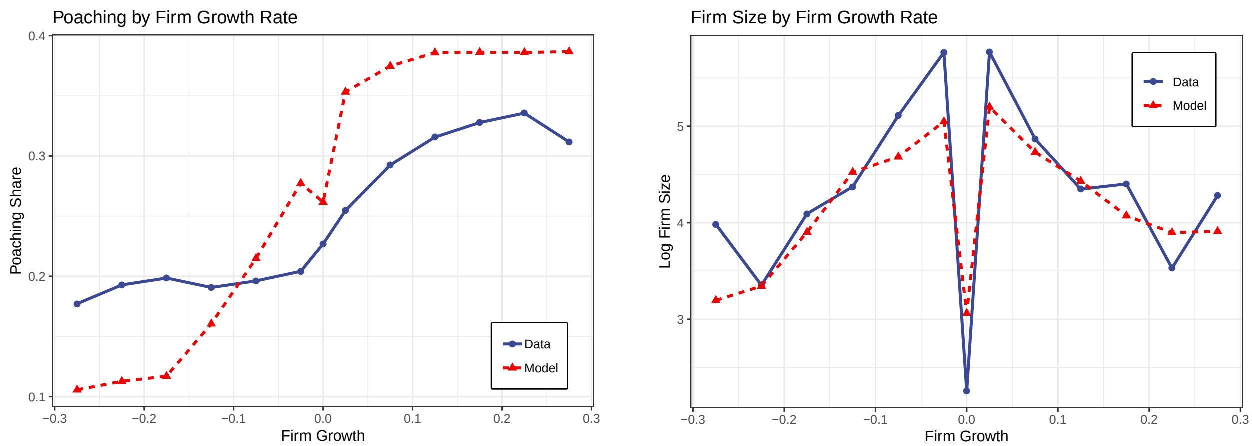
Notes: The left panel compares the firm growth rate distribution in the model versus the data. The model replicates the symmetric firm growth rate distribution very well. The right panel plots the firm size distribution in the model versus the data.

panel dimension highlights the importance of the multi-dimensional sorting setup considered in this paper. The estimated production function rationalizes this discrepancy and can generate the complex sorting patters observed in the data.

The standard deviation of firm growth rates is matched very closely. The estimated firm shock parameter implies that firms receive on average productivity shocks almost every half a year. This renders the assumption of fixed firm types even for very short time periods unrealistic. Figure 2 presents the model fit of the firm growth and size distribution. The overall standard deviation and autocorrelation of firm growth rates are well matched. As the left panel shows, the model implies a realistic growth rate distribution with its symmetric bell shape with mean zero. The empirical growth rate distribution has a mass point at zero, as many firms do not adjust the size of their operation in a given year. As my model does not feature fixed costs of adjusting the size of the operations, it underpredicts the fraction of firms with exactly zero growth. The shape of the firm distribution is overall well matched, although large firms are somewhat overrepresented in my model simulations. The shape of the size distribution is restricted by the particular choice of the job creation cost function $c(v)$, and thus it is not surprising that the model cannot replicate it precisely. Figure 3b further shows the untargeted relationship between firm size and growth. The model matches the observed pattern in Germany very closely. Firms with high growth or shrinking rates are smaller, whereas larger firms show smaller growth rates. The model in addition replicates

Figure 3: Untargeted Firm Dynamics Moments

(a) Poaching Share (b) Firm Size by Growth Rate



Notes: The left figure compares the share of total hiring made through poaching, by firm growth rate in the model and data. The right plot presents the empirical relationship of firms size and firm growth compared to the outcomes in the model.

the notable dip in firm size observed for firms that do not change size in a given year. The reason for the well matched relationship between firm size and growth rates is the identified production function. Constant and time-varying productivities enter additively, thus firm shocks matter less for highly productive and large firms. In turn, their growth path is more stable compared to smaller firms.

The estimated model also provides a laboratory to quantify how precisely the worker fixed effects recover the true worker types. The correlation between the worker fixed effects and the true worker type x is 0.987. Thus, the 18 year time span of the social security data is enough to recover the true worker types almost perfectly. The next section discusses the resulting sorting patterns in detail.

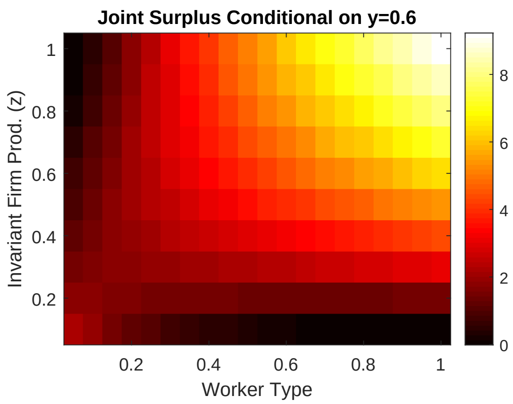
4.2 Sorting Patterns and Firm Dynamics

Worker flows, and thus the sorting pattern, are guided by the joint surplus $S(x, y, z)$. Figure 4 presents the surplus and sorting patterns along the time-invariant firm productivity z dimension. The left panel plots the surplus conditional on the time-varying firm productivity $y = 0.6$.[29] It shows that the surplus is the highest along the main diagonal. Hence workers along the main diagonal are well matched, whereas workers in the northwest and southeast

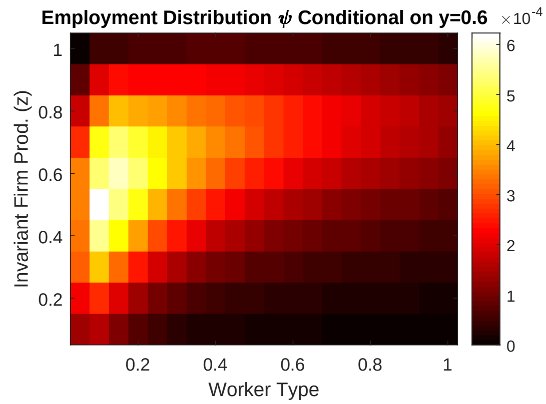
[29] For expositional purposes, I focus on presenting only one slice of the surplus function. The shape of the presented figures is very similar for different levels of y.

Figure 4: Sorting along z Dimension

(a) $S(x, y, z)$ with $y = 0.6$



(b) $\psi(x, y, z)$ with $y = 0.6$



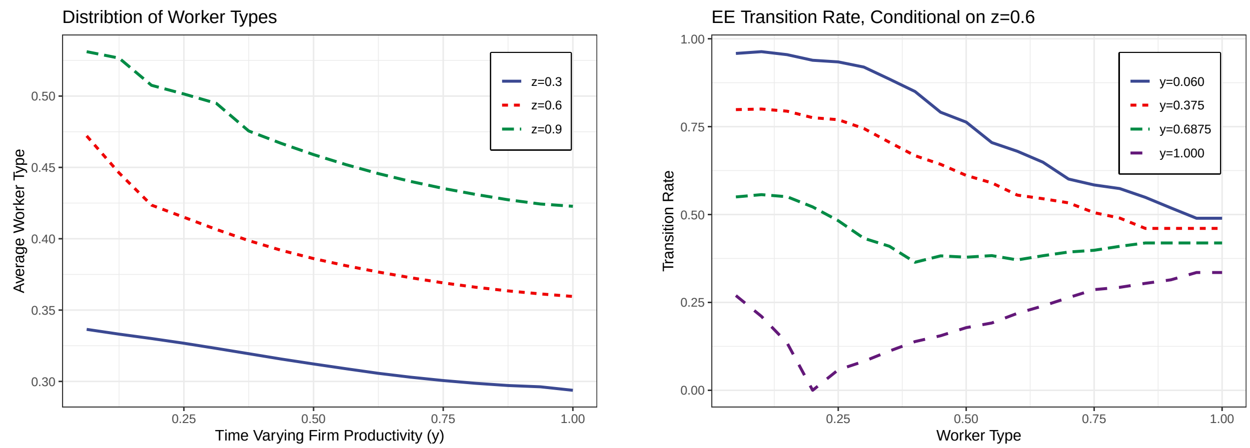
Notes: The left figure plots joint surplus conditional on y=0.6. It shows that the surplus is the highest along the main diagonal. The right figure shows the resulting positive sorting pattern along the z dimension.

corners are in worse matches. This result is not surprising given that ρ_z is estimated to be larger then 1 and thus worker productivity is complementary to time-invariant firm productivity z. Low type workers will therefore tend to flow towards low type firms, whereas high type workers move towards high z firms. The resulting positive sorting pattern is well visible in the right panel of figure 4, which shows the empirical employment distribution $\psi(x, y, z)$ conditional on $y = 0.6$. The employment distribution $\psi(x, y, z)$ is also affected by the worker and firm type distribution. The log-normal distribution of worker types is also visible in the employment distribution. More jobs are filled with low skilled workers as these are more abundant. The distribution is more balanced along the z dimension. This is because of two countervailing forces. Higher z firms grow endogenously larger, but there are fewer high type firms in the economy. The resulting positive sorting is nevertheless clearly visible. The higher the worker type, the more likely she is working for a high z firm. This in turn leads to the fact that on average, larger firms employ better workers.

More relevant for the understanding of the behavior of worker flows as firms expand and contract is the sorting pattern along the time-varying firm productivity dimension. Figure 5 presents the worker distribution and job-to-job transition rates across firm types conditional on z. The left panel shows the relationship between average worker quality and time-varying firm productivity y for different levels of z. The negative sorting along the y dimension is clearly visible. Independent of the level of z, firms that receive a positive productivity shock downgrade worker types. This generates the negative relationship between worker

Figure 5: Sorting along y Dimension

(a) Average Worker Type by Firm Type (b) Job-to-Job Transition Rate by Firm Type



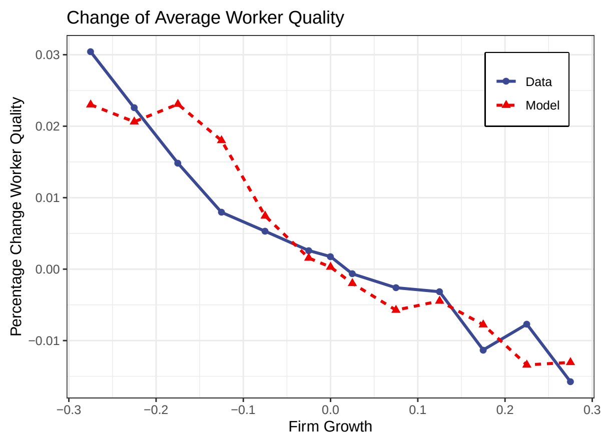
Notes: The left figure plots the estimated equilibrium surplus $S(x, y, z)$ conditional on $z = 0.6$. The right plot presents average worker type employed by different firm types. The left panel clearly shows the positive sorting of worker types along the z dimension, whereas the right panel shows the (weak) negative sorting along the y dimension.

skill changes and growth rates at the firm level in regression equation (23). The estimated complementary $\rho_y < 1$ implies that time-varying firm productivity and worker types are substitutes in production. Thus, after positive productivity shocks firms expand in size and capitalize on the increased productivity by substituting away from high type and expensive workers.

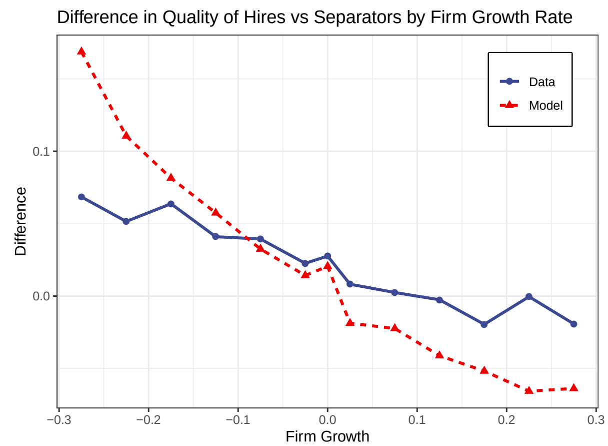
These adjustments are best seen by inspecting the job-to-job transition rates by worker and firm types, which are presented in figure 5b. It shows the job-to-job transition rates by worker type for different levels of y, conditional on $z = 0.6$. Job-to-job transition rates are directly related to the surplus of the match through equation (10). Higher y type firms have an absolute advantage in production, and thus yield a higher surplus which makes them more attractive employers in general. This is also reflected by the fact that faster growing firms increasingly make us of poaching and that each vacancy has a higher probability of being filled. Figure 3a shows that in my model as well as empirically, poaching makes up an increasing fraction of new hires in expanding firms. Although there is no good data equivalence in the German social security data, figure 7b shows that the vacancy yield is increasing by firm growth, that is, faster growing firms hire more workers per vacancy. The better the match between a specific worker and firm type, the higher is the surplus, and in turn the lower the job-to-job transition rate. The lowest rate by firm type indicates

Figure 6: Worker Quality Changes - Model Fit

(a) Change of Average Worker Quality



(b) Quality of Worker Flows

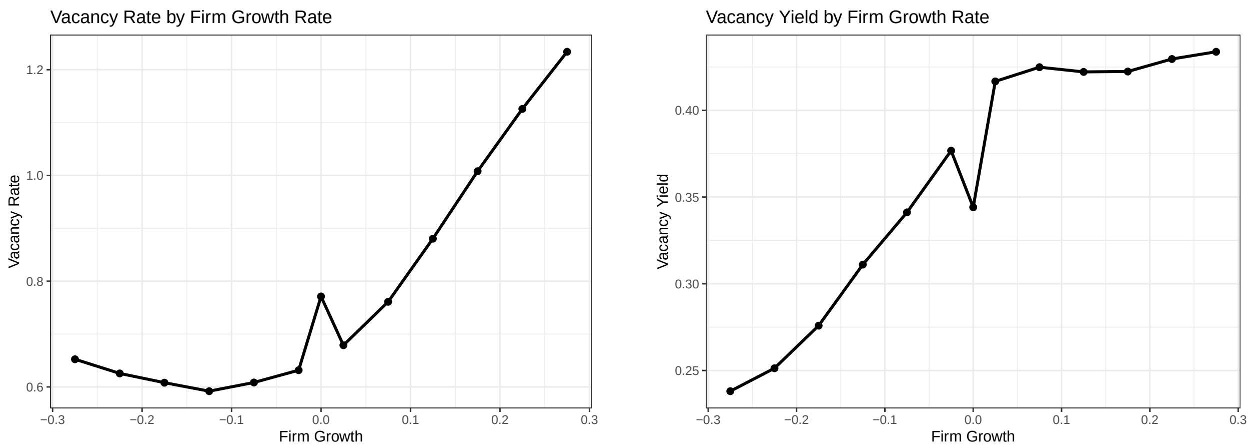


Notes: The left figure plots relationship between changes year-to-year percentage changes in the average worker quality and the firm growth rate. The right panel presents the average quality difference between the hires and separators by firm growth. Overall the two figures show that expanding firms upgrade worker quality, whereas expanding firms downgrade.

their most preferred worker type. In general, the figure shows that the job-to-job transition probability is declining in worker type in low y firms. This again is just a reflection of the estimated negative sorting along the y dimension. Low type y firms prefer to match with high type workers, hence these workers also have the highest retention rate. The most preferred worker type and thus the lowest job-to-job transition rate shifts towards lower and lower worker skills as we consider higher y firm types. Taken together, this implies that after positive productivity shocks, low type workers provide a better match, and hence the firm downgrades its worker skill. This then leads to the negative relationship between firm growth rate and worker quality changes. Figure 6a shows that the model replicates this relationship very precisely over the entire firm growth rate distribution. This not only occurs because growing firms expand with lower skilled workers. As we have seen in figure 5b, productivity shocks also affect the current workforce. Some workers are better matched after shocks, some become worse matched. This implies, that we should also see firms replacing existing workers with employees of a different skill level. Figure 6b reports the difference between the average skills of hires and separations by firm growth rate. It shows that both in the German social security data as well as in my model, contracting firms replace separating workers with higher type workers, whereas growing firms replace workers with lower type workers. In summary, the complementarities in production induce firms to

Figure 7: Vacancy Rate and Yield by Firm Growth Rate

(a) Vacancy Rate (b) Vacancy Yield

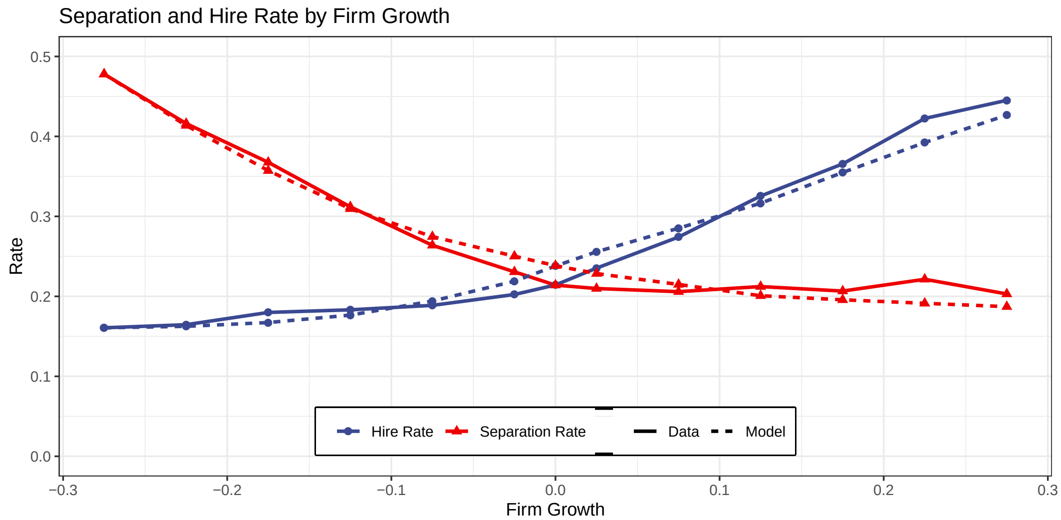


Notes: The left figure plots relationship between the vacancy rate (number of vacancies over employment) and the firm growth rate. The right panel presents the vacancy yield by firm growth. Both figures refer to model outcomes.

change the skill composition of their workforce after productivity shocks by hiring different workers than before in addition to replacing some of their existing workforce with different type of workers. This adjustment also generates a realistic relationship between the vacancy rate[30] and firm growth as depcited in figure 7a. There is no equivalent data in the social security data, but the shape replicates the relationship documented by Davis et al. (2013) for the US.[31] The vacancy rate increases sharply for expanding firms, but remains constant for contracting firms, as these firms replace separating workers with higher type workers. This also has implications for the worker turnover rate across the firm distribution. A salient fact of firm dynamics is that both expanding and contracting firms have excess job turnover rates. That is, in expanding firms, the separation rate does not decline but stays constant at some positive level, whereas in contracting firms the hire rate does not drop towards zero, but stay constant. This peculiar pattern gives rise to the ”hockey-stick” worker flow shapes at the firm level, first documented by Davis et al. (2006). But this result is puzzling from the point of view of standard models of firm dynamics. Why does an expanding firm not reduce worker churn to reduce its separation rate? At the same time, why do firms not reduce their hire rate as they contract faster? A very intuitive answer is that the workers separating from growing and contracting firms are different from the workers joining the firms. This study

[30]The vacancy rate is defined as the number of vacancies divided by firm size.
[31]Figure IV in Davis et al. (2013)

Figure 8: Worker Flows by Firm Growth



Notes: The figure shows the separation and hire rate over the firm growth distribution for model simulated data as well as the German social security data, controlling for firm fixed effects.

exactly provides evidence for this mechanism. Firms want to change the skill composition of their workforce when they expand and contract, which involves separating from workers that became worse matched after productivity shocks, and rehire worker types that provide a better match to the current firm productivity. This in turn leads to excess worker churn, where worker flows exceed the amount needed to achieve the change in employment. My model almost perfectly matches the empirically observed hire and separation flows, as is shown in figure 8. It compares the hire and separation rate by firm growth rates in the model to the empirically observed pattern estimated from the German social security data. The hockey-stick pattern is clearly visible in Germany as well. The separation rate for expanding firms remain roughly constant around 20 percent, whereas the hire rate remains close to 20 percent for all contracting firms. The model replicates this patterns extremely closely. In summary, labor market sorting is crucial to understand the behavior of firm dynamics.

4.3 Output Cost of Search Frictions

I next use the model to quantify the output costs of search frictions. As a counterfactual analysis, I decrease search frictions by increasing the job contact rate for unemployed and employed workers λ_w and λ_e by a factor of 1.5. All other parameter values are held constant

Table 6

Moments	Baseline	Low Frictions
Job-to-job transition rate	0.006	0.008
Unemployment rate	0.121	0.087
Job filling rate	0.392	0.498
Std. of firm growth rates	0.314	0.345
Emp. Share Largest 10% of Firms	0.507	0.525
Panel Regression Coeff Equation (23)	-0.067	-0.096
Cross-sectional Reg Coeff Equation (24)	0.108	0.146
Corr(x, y)	-0.100	-0.136
Corr(x, z)	0.221	0.311
Output	0.431	0.453

Notes: Effects of lower search frictions. λ_w and λ_e are increased by 50%, whereas all other parameters are left at the original value.

at their estimated values from table 5. Table 6 compares selected moments of the baseline model with the results from the low friction environment. Not surprisingly, as search frictions become less severe, job finding rates increase, which leads to lower unemployment, higher job-to-job transition and job-filling rates. Lower search frictions imply that firms can respond quicker to productivity shocks. First, firms can adjust their scale more intensively in response to shocks. This leads to an increase in the standard deviation of firm growth rates and to large firms gaining employment share. Second, they are able to better adjust the quality of their workforce. Because of the lower search frictions, firms can now be more "picky" during hiring. This leads to a strengthening of labor market sorting. On the one hand, after firm productivity shocks this leads to stronger adjustment in worker quality as measured by the relationship of firm growth and worker quality changes (regression coefficient of equation (23)). On the other hand, the worker composition in firms with high *time-invariant* productivity shift towards higher skilled workers. These two forces lead then to higher overall labor market sorting, as shown by the increased correlation in absolute terms between worker types and both firm types, $|corr(x, y)|$ and $|corr(x, z)|$.

Further, lower frictions overall lead to higher economic output. This is generated by three distinct forces. First, due to lower search frictions, more people are employed, which mechanically increases output. Second, as documented before, lower search frictions allow more productive firms to gain employment share. Third, they lead to an increase in sorting, which yields output gains because worker types are now better matched to firm types. To formally decompose the output gain into these three channels, I compute two intermediate employment distributions. First, I re-scale the original employment distribution $\psi(x, y, z)$, so that the total employment level equals employment in the counterfactual with low search frictions. Thus, $\psi^E(x, y, z) \equiv \psi(x, y, z)\frac{e}{e^N}$ captures the effect of increased employment, while holding the employment share of worker types across firms types constant. In a next step, I re-scale this new distribution such that it additionally reflects the reallocation of jobs towards more productive firm types. This intermediate distribution is computed $\psi^{FS}(x, y, z) = \psi^E(x, y, z)\frac{\int_x \psi^E(x,y,z)dx}{\int_x \psi(x,y,z)dx}$. Finally, the last step represents the effect coming from the better sorting of worker types across firm types. The intermediate distribution $\psi^{FS}(x, y, z)$ has the same employment level as under low search frictions as well as the same employment shares across firm types. The only difference to the employment distribution $\psi^N(x, y, z)$ under low frictions is the distribution of worker types across firm types, or sorting. Thus the last step measures the impact of increased sorting under lower search frictions, which is labeled the misallocation effect of workers across jobs. I then decompose the total output change $y^N - y$ between the low friction case (superscript N) and the baseline (no superscript) with the following equation:[32]

$$
\begin{aligned}
y^N - y &= \int_z \int_y \int_x f(x, y, z) \left[\psi^N(x, y, z) - \psi(x, y, z)\right] dxdydz \\
&= \int_z \int_y \int_x f(x, y, z) \\
&\times \left[\underbrace{\psi^N(x, y, z) - \psi^{FS}(x, y, z)}_{\text{Missall. Workers across Jobs}} + \underbrace{\psi^{FS}(x, y, z) - \psi^E(x, y, z)}_{\text{Missall. Jobs across Firms}} + \underbrace{\psi^E(x, y, z) - \psi(x, y, z)}_{\text{Employment Effect}} \right] dxdydz
\end{aligned}
\tag{25}
$$

Table 7 presents the results of this decomposition. It shows that the employment effect explains 79% of the output gain with low search frictions and is thus by far the most dominant force. Second comes the reallocation of jobs towards more productive firms, which accounts for 16%. The remaining 5% is explained by the increase in sorting and hence overall better match quality. The decomposition shows that misallocation of workers across firms is a significant source of the output cost of search frictions.

[32]The specific ordering of the decomposition is inconsequential.

Table 7: Decomposition of Output Cost of Search Frictions

	Total	Share
Total Difference	0.0215	1.000
Employment Effect	0.0169	0.785
Misallocation of Jobs Across Firms	0.0035	0.162
Misallocation of Workers across Jobs	0.0012	0.054

Notes: Decomposition of the output gain through lower search frictions.

5 Conclusion

Why do firms separate from some of their workforce while hiring other workers at the same time? This behavior is especially puzzling for expanding and contracting firms. In standard models of firm dynamics, growing or contracting firms reduce excess worker turnover to quickly achieve the desired scale of operation. In this paper I show that the behavior of worker flows at the establishment level is tightly linked to assortative matching in the labor market. With assortative sorting in the labor market, the ideal worker type changes for firms after productivity shocks, which leads firms to adjust the skill composition of their workforce. Therefore, firms separate from the some of existing workforce while hiring better suited workers, which generates excess worker turnover after firm productivity shocks. I develop a search and matching model with heterogeneous workers and firms, job-to-job transitions, firm dynamics originating from idiosyncratic firm-level shocks and multi-dimensional sorting. The typical identification strategy for assortative labor market sorting in the literature relies on fixed firm types. The introduction of firm dynamics therefore requires a novel identification strategy for the complementarities between unobserved worker and firm attributes that drive the pattern of sorting. The key idea is that firms' reorganization of their workforce in response to productivity shocks will identify the complementarities in production. Intuitively, in a world with positive (negative) assortative matching, firm growth is associated with worker quality upgrading, whereas shrinking firms will reorganize towards lower (higher) skilled workers. The strength of sorting is identified by the extent of these adjustments. The stronger the sorting motive, the more intensively firms are going to adjust the skill composition of their workforce in response to shocks. Using German social security data, I document that expanding firms hire lower worker types compared to before, and replace workers with lower skilled workers and contracting firms upgrade the skill composition of their workforce. I use this data to estimate the model and show that it replicates a number of salient empirical facts about firm dynamics: (1) a symmetric bell shaped firm growth distribution, (2) empirically consistent firm size distribution, (3) realistic relationship between firm size and growth and (4) that fast growing firms fill their vacancies disproportionately through poaching from other firms. In addition, it explains the excess worker turnover in response to firm productivity shocks. As an untargeted moment, it explains the establishment level behavior of separations and hires documented in Davis et al. (2006). I conclude that assortative matching is key to understand firm level dynamics of employment.

References

John M Abowd, Francis Kramarz, and David N Margolis. High wage workers and high wage firms. Econometrica, 67(2):251–333, 1999.

Rüdiger Bachmann, Christian Bayer, Christian Merkl, Stefan Seth, Heiko Stüber, and Felix Wellschmied. Worker churn in the cross section and over time: New evidence from germany. 2019.

Jesper Bagger and Rasmus Lentz. An empirical model of wage dispersion with sorting. The Review of Economic Studies, 86(1):153–190, 2018.

Gary S Becker. A theory of marriage: Part i. The Journal of Political Economy, pages 813–846, 1973.

Adrien G Bilal, Niklas Engbom, Simon Mongey, and Giovanni L Violante. Firm and worker dynamics in a frictional labor market. Technical report, National Bureau of Economic Research, 2019.

Stéphane Bonhomme, Thibaut Lamadon, and Elena Manresa. A distributional framework for matched employer employee data. Econometrica, 87(3):699–739, 2019.

Katarına Borovicková. Job flows, worker flows and labor market policies. Technical report, 2016.

Pierre Cahuc, Fabien Postel-Vinay, and Jean-Marc Robin. Wage bargaining with on-the-job search: Theory and evidence. Econometrica, 74(2):323–364, 2006.

Lorenzo Caliendo, Ferdinando Monte, and Esteban Rossi-Hansberg. The anatomy of french production hierarchies. Journal of Political Economy, 123(4):809–852, 2015.

David Card, Jörg Heining, and Patrick Kline. Workplace heterogeneity and the rise of west german wage inequality*. The Quarterly Journal of Economics, 128(3):967–1015, 2013.

A Kerem Coşar, Nezih Guner, and James Tybout. Firm dynamics, job turnover, and wage distributions in an open economy. American Economic Review, 106(3):625–63, 2016.

Steven J Davis, R Jason Faberman, and John Haltiwanger. The flow approach to labor markets: new data sources and micro–macro links. The Journal of Economic Perspectives, 20(3):3–26, 2006.

Steven J Davis, R Jason Faberman, and John Haltiwanger. Labor market flows in the cross section and over time. Journal of Monetary Economics, 59(1):1–18, 2012.

Steven J Davis, R Jason Faberman, and John C Haltiwanger. The establishment-level behavior of vacancies and hiring. The Quarterly Journal of Economics, 128(2):581–622, 2013.

Matthew S Dey and Christopher J Flinn. An equilibrium model of health insurance provision and wage determination. Econometrica, 73(2):571–627, 2005.

Jan Eeckhout and Philipp Kircher. Identifying sorting in theory. The Review of Economic Studies, page rdq034, 2011.

Niklas Engbom and Christian Moser. Earnings inequality and the minimum wage: Evidence from brazil. Working Paper, 2018.

Shigeru Fujita and Giuseppe Moscarini. Recall and unemployment. American Economic Review, 107(12):3875–3916, 2017.

Christian Haefke, Marcus Sonntag, and Thijs Van Rens. Wage rigidity and job creation. Journal of monetary economics, 60(8):887–899, 2013.

Marcus Hagedorn, Tzuo Hann Law, and Iourii Manovskii. Identifying equilibrium models of labor market sorting. Econometrica, 85(1):29–65, 2017. ISSN 1468-0262. doi: 10.3982/ECTA11301. URL http://dx.doi.org/10.3982/ECTA11301.

Joerg Heining, Wolfram Klosterhuber, and Stefan Seth. Linked-employer-employee-daten des iab: Liab-laengsschnittmodell 1993-2010 (liab lm 9310). FDZ-Datenreport, 08/2013, 2013.

Hugo A Hopenhayn. Entry, exit, and firm dynamics in long run equilibrium. Econometrica: Journal of the Econometric Society, pages 1127–1150, 1992.

Todd L Idson and Walter Y Oi. Workers are more productive in large firms. American Economic Review, 89(2):104–108, 1999.

Philip Jung and Moritz Kuhn. Labour market institutions and worker flows: comparing germany and the us. The Economic Journal, 124(581):1317–1342, 2014.

Leo Kaas and Philipp Kircher. Efficient firm dynamics in a frictional labor market. American Economic Review, 105(10):3030–60, 2015.

Thibaut Lamadon. Productivity shocks, long-term contracts and earnings dynamics. manuscript, University of Chicago, 2016.

Rasmus Lentz and Dale T Mortensen. An empirical model of growth through product innovation. Econometrica, 76(6):1317–1373, 2008.

Ilse Lindenlaub. Sorting multidimensional types: Theory and application. The Review of Economic Studies, 84(2):718–789, 2017.

Jeremy Lise and Fabien Postel-Vinay. Multidimensional skills, sorting, and human capital accumulation. forthcoming American Economic Review, 2020.

Jeremy Lise and Jean-Marc Robin. The macrodynamics of sorting between workers and firms. American Economic Review, 107(4):1104–35, 2017.

Jeremy Lise, Costas Meghir, and Jean-Marc Robin. Matching, sorting and wages. Review of Economic Dynamics, 19:63–87, 2016.

Rafael Lopes de Melo. Firm wage differentials and labor market sorting: Reconciling theory and evidence. Journal of Political Economy, 126(1):313–346, 2018. doi: 10.1086/695505. URL https://doi.org/10.1086/695505.

Erzo GJ Luttmer. Selection, growth, and the size distribution of firms. The Quarterly Journal of Economics, pages 1103–1144, 2007.

W Bentley MacLeod and James M Malcomson. Investments, holdup, and the form of market contracts. The American Economic Review, pages 811–837, 1993.

Marc J Melitz. The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica, 71(6):1695–1725, 2003.

Monika Merz and Eran Yashiv. Labor and the market value of the firm. The American Economic Review, 97(4):1419–1431, 2007.

Arash Nekoei and Andrea Weber. Recall expectations and duration dependence. The American Economic Review, 105(5):142–146, 2015.

F. Postel-Vinay and J.M. Robin. Equilibrium wage dispersion with worker and employer heterogeneity. Econometrica, 70(6):2295–2350, 2002.

F. Postel-Vinay and H. Turon. On-the-job search, productivity shocks, and the individual earnings process. International Economic Review, 51(3):599–629, 2010.

Fabien Postel-Vinay and Ilse Lindenlaub. Multidimensional sorting under random search. In 2017 Meeting Papers, number 501. Society for Economic Dynamics, 2017.

Edouard Schaal. Uncertainty and unemployment. Econometrica, 85(6):1675–1721, 2017.

Robert Shimer. On-the-job search and strategic bargaining. European Economic Review, 50(4):811–830, 2006.

Robert Shimer and Lones Smith. Assortative matching and search. Econometrica, 68(2): 343–369, 2000.

Jae Song, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter. Firming up inequality. The Quarterly Journal of Economics, 134(1):1–50, 2018.

Chad Syverson. What determines productivity? Journal of Economic Literature, 49(2): 326–365, 2011.

Sharon Traiberman. Occupations and import competition: Evidence from denmark. American Economic Review, 109(12):4260–4301, 2019.

A Value Functions and Derivation of Surplus

This appendix section presents the value functions and the derivation of the surplus function. To compute the value of a vacancy we have to integrate first integrate over all possible future time-varying firm productivity types. Second, we have to take into account whether the firm meets a suitable match with either an unemployed or employed worker. $A^U(x, y)$ is the indicator function that takes on the value of 1 if the match between an unemployed worker of type x and a firm of type y, z is consummated and zero otherwise. Similarly, if a worker x employed at a type y, z firm is contacted by a poaching firm of type $\tilde{y}, \tilde{z}$, $A^E(x, y, z, \tilde{y}, \tilde{z})$ is one if the job offer is accepted and zero otherwise. The wage setting mechanism and the assumption of transferable utility assures that acceptance decisions jointly maximize the total surplus. Thus, agents are willing to match together if the match generates a positive surplus, and in case of job-to-job transitions, the prospective surplus is higher than the current one. Formally,

$$A^U(x, y, z) \;=\; \begin{cases} 1 & \text{if } S(x, y, z) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{26}$$

$$A^E(x, y, \tilde{y}) \;=\; \begin{cases} 1 & \text{if } S(x, \tilde{y}, \tilde{z}) \geq S(x, y, z) \\ 0 & \text{otherwise.} \end{cases} \tag{27}$$

The value function of a vacant firm then reads

$$\begin{aligned} V(y, z) =& \beta(1-d) \int_{y'} \Bigg\{ (1-\lambda_f) V(y', z) \\ & + \lambda_f \Bigg(p^u \int_x A^U(x, y', z) J(x, y', z, w^U(x, y', z) + (1 - A^U(x, y', z)) V(y', z) \frac{\mu_x(x)}{u} dx + \\ & + (1-p^u) \int_{\tilde{y}} \int_{\tilde{z}} \int_x A^E(x, y', z, \tilde{y}, \tilde{z}) J(x, y', z, w^E(x, y', z, \tilde{y}, \tilde{z})) \\ & + (1 - A^E(x, y', z, \tilde{y}, \tilde{z})) V(y', z) \frac{\psi^S(x, \tilde{y}, \tilde{z})}{e^S} dx d\tilde{y} d\tilde{z} \Bigg) \Bigg\} p(y'|y) dy', \end{aligned}$$

An unemployed workers might either find a suitable match next period, or remains unemployed.

$$
\begin{aligned}
U(x) &= b(x) \\
&\quad + \beta\left(\lambda_w \int_z \int_y \left(A^U(x,y,z)W(x,y,z,w^U(x,y,z)) + (1-A^U(x,y,z))U(x)\right)\frac{\mu_F(y,z)}{V}dydz\right. \\
&\quad \left. +(1-\lambda_w)U(x)\right). \qquad (28)
\end{aligned}
$$

For ongoing matches, several different outcomes might occur. First, the firm productivity shock might be such that the match is not viable anymore ($A^U(x,y',z)=0$). Second, if the worker meets another firm of type $\tilde{y},\tilde{z}$, the poaching offer might either trigger a job-to-job transition ($\tilde{y},\tilde{z} \in \Upsilon^{EE}(x,y',z)$) or a wage renegotiation ($\tilde{y},\tilde{z} \in \Upsilon^{BC}(x,y',z)$). If the worker stays with its current employer, the firm productivity shock might lead to a violation of the participation constraint of one of the parties. This triggers a renegotiation as described in the main text. The set of y that violate the participation constraint for a worker of type x working for a firm of type z for a wage w is $\Upsilon^{NW}(x,z,w)$, whereas the corresponding set for the violation of the firms participation constraint is $\Upsilon^{NF}(x,z,w)$. Their formal definitions are

$$
\begin{aligned}
\Upsilon^{NW}(x,z,w) &= \{y : S(x,y,z) \geq 0 \geq W(x,y,z,w) - U(x)\} \\
\Upsilon^{NF}(x,z,w) &= \{y : S(x,y,z) \geq 0 \geq J(x,y,z,w) - V(y,z)\}.
\end{aligned}
$$

Equation (29) presents the value of a filled job to a firm.

$$
\begin{aligned}
J(x,y,z,w) = f(x,y,z) - w + \beta \int_{y'} \Big\{ & (1-(1-d)A^U(x,y',z))V(y',z) \\
& + (1-d)(A^U(x,y',z)) \int_{\tilde{z}} \int_{\tilde{y}} \big\{ \lambda_e \big[\mathbb{1}\left((\tilde{y},\tilde{z}) \in \Upsilon^{EE}(x,y',z)\right) V(y',z) \\
& + \mathbb{1}\left((\tilde{y},\tilde{z}) \in \Upsilon^{BC}(x,y',z)\right) J\left(x,y',z,w^E(x,\tilde{y},\tilde{z},y',z)\right) \\
& + (1 - \lambda_e \mathbb{1}\left((\tilde{y},\tilde{z}) \in \Upsilon^{BC}(x,y',z) \cup \Upsilon^{EE}(x,y',z)\right) \\
& \times \big[\mathbb{1}\left(y' \in \Upsilon^{NW}(x,z,w)\right) J\left(x,y',z,w^{NW}(x,y',z)\right) \\
& + \mathbb{1}\left(y' \in \Upsilon^{NF}(x,z,w)\right) J\left(x,y',z,w^{NF}(x,y',z)\right) \\
& + \mathbb{1}\left(y' \notin \Upsilon^{NF}(x,z,w) \cup \Upsilon^{NF}(x,z,w)\right) J(x,y',z,w) \big] \big\} \frac{\mu_F(\tilde{y},\tilde{z})}{v} d\tilde{y} \Big\} p(y'|y) dy'.
\end{aligned} \tag{29}
$$

Equation (30) presents the workers' valuation of a match, which is the mirror image of the firm's value of a filled job.

$$
W(x,y,z,w) = w + \beta \int_{y'} \Big\{ (1-(1-d)A^U(x,y',z))U(x) \tag{30}
$$

$$
+ (1-d)(A^U(x,y',z)) \int_{\tilde{z}} \int_{\tilde{y}} \big\{ \lambda_e \big[\mathbb{1}\left((\tilde{y},\tilde{z}) \in \Upsilon^{EE}(x,y',z)\right) W\left(x,\tilde{y},\tilde{z},w^E(x,y',z,\tilde{y},\tilde{z})\right) \tag{31}
$$

$$
+ \mathbb{1}\left((\tilde{y},\tilde{z}) \in \Upsilon^{BC}(x,y',z)\right) W\left(x,y',z,w^E(x,\tilde{y},\tilde{z},y',z)\right) \tag{32}
$$

$$
+ (1 - \lambda_e \mathbb{1}\left((\tilde{y},\tilde{z}) \in \Upsilon^{BC}(x,y',z) \cup \Upsilon^{EE}(x,y',z)\right) \tag{33}
$$

$$
\times \big[\mathbb{1}\left(y' \in \Upsilon^{NW}(x,z,w)\right) W\left(x,y',z,w^{NW}(x,y',z)\right) \tag{34}
$$

$$
+ \mathbb{1}\left(y' \in \Upsilon^{NF}(x,z,w)\right) W\left(x,y',z,w^{NF}(x,y',z)\right) \tag{35}
$$

$$
+ \mathbb{1}\left(y' \notin \Upsilon^{NF}(x,z,w) \cup \Upsilon^{NF}(x,z,w)\right) W(x,y',z,w) \big] \big\} \frac{\mu_F(\tilde{y},\tilde{z})}{v} d\tilde{y} \Big\} p(y'|y) dy'. \tag{36}
$$

The value function in the main text can be simply derived by using the specific bargaining rules defined in the wage setting mechanism. For deriving the surplus we first use the definition of the surplus $S(x,y) = J(x,y,w) - V(y) + W(x,y,w) - U(x)$. Then after some simplifications one can arrive at the surplus function:

$$\begin{aligned} S(x,y,z) =& f(x,y,z) - b(x) + \beta(1-d)\int_{y'} S(x,y',z)^+ p(y'|y)dy' \\ & - \beta\alpha\lambda_w \int_y\int_z S(x,y,z)^+ \frac{\mu_F(y,z)}{v} dzdy \\ & - \beta(1-d)\lambda_f \int_{y'} \Bigg(p^u \int_x (1-\alpha) S(x,y',z)^+ \frac{\mu_x(x)}{u} dx \\ & + (1-p^u) \int_{\tilde{z}}\int_{\tilde{y}}\int_x \left(S(x,y',z) - S(x,\tilde{y},\tilde{z})\right)^+ \frac{\psi^S(x,\tilde{y},\tilde{z})}{e^s} dx d\tilde{y} d\tilde{z} \Bigg) p(y'|y)dy'. \end{aligned} \tag{37}$$

B Appendix - Numerical Implementation

I apply the following numerical procedure to solve the model. First, I discretize the state space by using a equidistant grid of 20 worker types and 10 grid points over the support of z and 16 over the support of y. The solution algorithm is the following iterative process:

1. Guess $S^0(x,y,z)$, $\psi^0(x,y,z)$, $\mu_x^0(x)$ and $\mu_f^0(y,z)$
2. Update $S^{i+1}(x,y,z)$ using equation (18)
3. Using the new value of $S(x,y,z)$, update acceptance policies $A^U(x,y,z)$ and $A^E(x,y,z,\tilde{y},\tilde{z})$. The indicator functions are updated slowly.
4. Update the distributions $\psi(x,y,z)$, $\mu_x(x)$ and $\mu_f(y,z)$ using the updated acceptance policies. The distributions are updated by using the law of motion equations (20), (22) and (19).
5. Compute the sup norm of the absolute values of differences between the iteration outcomes and set set $i = i + 1$
6. Repeat steps 2-5 the until the surplus, acceptance strategies and the distributions converged. I use 10^{-6} as the convergence criteria for the surplus and acceptance strategies and 10^{-7} for the distributions.

Due to the discretization, infinitesimal changes in $S(x,y,z)$ lead to discontinuous changes in the distributions of agents. This could cause the algorithm to not converge at the desired convergence criteria. In order to smooth I assume that agents very close to the decision thresholds randomize between acceptance and rejection. I use the following randomization strategies:

$$
\begin{aligned}
A^U(x,y,z) &= \begin{cases} 1 & \text{if } S(x,y,z) \geq 10^{-2} \\ \frac{1-\left(10^{-2}-S(x,y,z)\right)}{10^{-2}} & \text{if } 0 \leq S(x,y,z) < 10^{-2} \\ 0 & \text{if } S(x,y,z) < 0 \end{cases} \\
A^E(x,y,\tilde{y},z,\tilde{z}) &= \frac{1}{1+\exp(-100(S(x,\tilde{y},\tilde{z})-S(x,y,z)))}
\end{aligned}
$$

These randomizations only affect a tiny fraction of the state space. With the estimated parameters from section 4, only around 5 percent of all possible $A^E(x,y,z,\tilde{y},\tilde{z})$ and no $A^U(x,y,z)$ are deviating from 0 or 1 by more than 10^{-6}. Similar smoothing strategies have been applied by Lopes de Melo (2018) and Hagedorn et al. (2017).

After obtaining the equilibrium solutions to value functions, acceptance rules and steady state distributions I simulate the evolution of 6126 firms and 511239 individuals. This exactly corresponds to the sample sizes in the German social security data. I use the stationary distribution as initial conditions and simulate labor market outcomes for 50 years. The first 32 years are burned in, thus the target moments are computed with the data of the remaining 18 years, which corresponds to the time frame of the German social security dataset. The calibration procedure minimizes the average percentage deviation from the target moments. I use the particle swarm optimization method. Swarm size is set to 96. In order to find a good initial swarm, I compute 15000 points using a Sobol sequence and use the best 96 as starting points.

C Appendix - Data Description:

The German social security data used in the empirical analysis is provided by the Research Data Centre of the German Federal Employment Agency. It is based on notifications of employers and several social insurance agencies for all workers and establishments covered by social security. This includes virtually every employees except of government employees. The particular dataset is the longitudinal model of the Linked-Employer-Employee Data (LIAB LM 9310). Heining et al. (2013) provide a detail data documentation.

This data set contains the complete work history of every worker that was employed at one of the selected establishments. The sample of establishments is based on the sample from IAB Establishment Survey. It is stratified according to industry, firm size, and federal state. In total, the dataset contains 2,702 to 11,117 establishments per year, and 1,090,728 to 1,536,665 individuals per year. It includes information on the foundation year of the establishment and a 3 digit industry identifier. For each worker employed at one of the establishments in the sample, the whole work history during 1993 and 2010 is recorded. This contains a 3 digit occupation identifier, part time and full time status, the beginning and end of all employment and unemployment spells precise to the day and the total daily wages and unemployment benefits received. All labor income is recorded that is subject to social security contribution. Only earnings that lie above the marginal part-time income threshold[33] and below the upper earnings limit for statutory pension insurance are not reported. In addition the dataset contains a number of socio demographic variables such as age, gender, nationality and education.

The exact working hours are not reported, only whether the employee is working part or full time. Since wages are recorded as daily wages, the hourly wage rate cannot be identified for part time employees.[34] Because of this, I focus on full time employees only in my analysis.

I use the following definitions for labor market transitions. I consider every worker transition from one employer to another firm as a job-to-job transition if the spell of non-employment between the two jobs was less than 30 days. In the computation of transition rates, I disregard any transition into unemployment and subsequent rehire if the person is rejoining the same firm within 30 days.[35]

I compute worker quality the following way. First, I deflate wages by the CPI index. Then, I compute annual earnings from full time jobs. I estimate a Mincer regression of the following form:

$$e_{it} = \alpha_i + \beta X_{it} + \epsilon_{it}. \tag{38}$$

Here e_{it} denotes the total anual earnings derived from employment and also potentially unemployment beneifts of individual i in year t. α_i represents the worker fixed effect and X_{it} a set of time-varying worker controls. I follow Card et al. (2013) and include a set of year dummies and quadratic and cubic terms in age fully interacted with educational attainment. The coding of the education variable follows exactly Card et al. (2013). The

[33]So called marginal part time jobs are not subject to social security contributions if the earnings do not exceed around 400 Euros a month

[34]The strict labor laws in Germany restrict the working week usually to around 40 hours. I therefore assume that the daily wages are a good measure for the true wage rate.

[35]This is in line with recent evidence shown in Fujita and Moscarini (2017) and Nekoei and Weber (2015)

social security data does not have information on the labor force status of workers. Thus, I assume that everyone with zero earnings from employment for a full calendar year (i.e. from 1st of January until 31st of December) is not part of the labor force. Years not spent in the labor force are excluded from the regression since my model does not feature a labor force participation margin. I trim the resulting fixed effects below the 0.5 and above the 99.5 percentile and normalize them to lie between 0 and 1.

D Additional Analysis

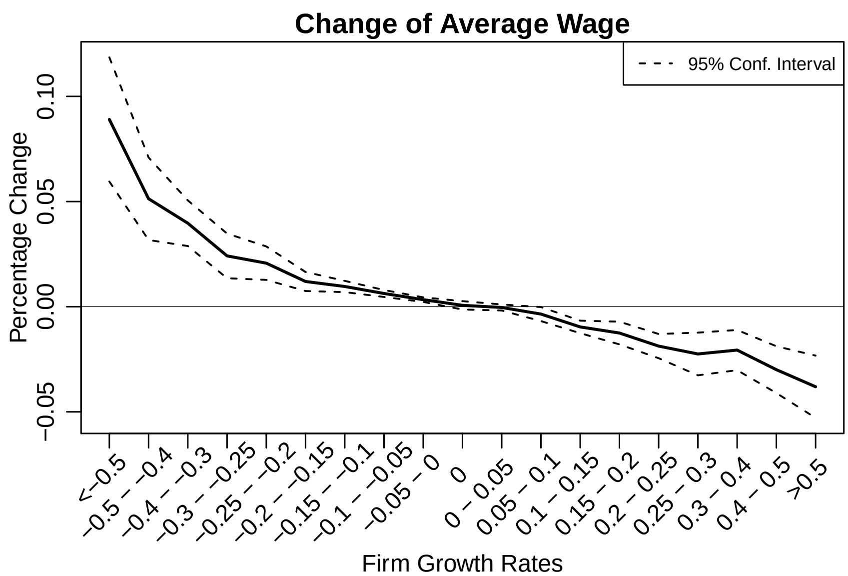


Figure 9: The figure shows the percentage change of average employee wage by establishment growth rates. The sample consists of all establishments with size≥20. Establishment growth rates and percentage changes in average wage are yearly.

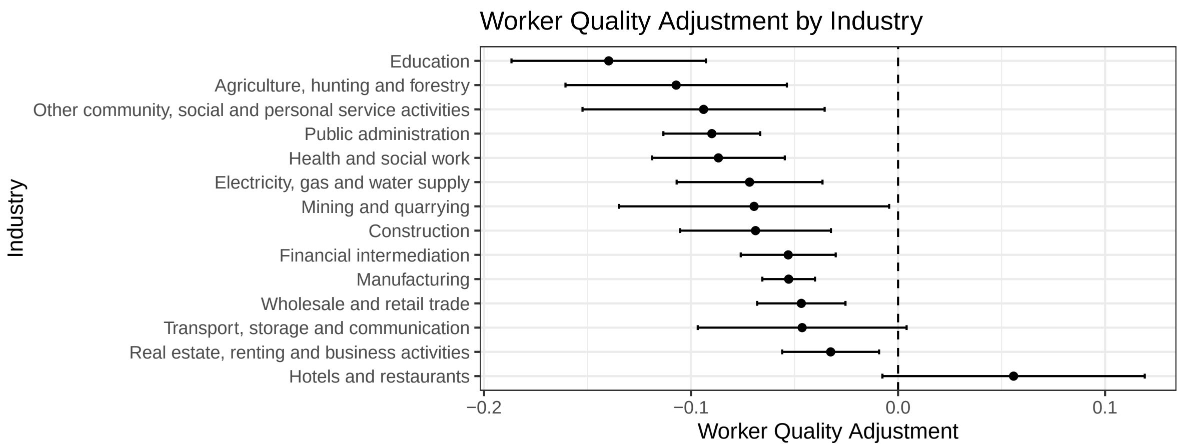


Figure 10: The figure shows the worker quality change (coefficient from regression equation 23) by NACE-1 industry classification. The sample consists of all establishments with size≥20 and growth rates $\in (-.75, 0.75)$. Establishment growth rates and percentage changes in worker quality are yearly.

Table 8: Regression Results

$\log(fsize)$	0.119
SE	0.003
N	50756
Adj. R^2	0.179

Notes: Cross section regression of log average worker fixed effects at the establishment level on establihsment size, as measured in workers employed. See text for details.