Identifying Labor Market Sorting with Firm Dynamics*

Andreas Gulyas†

January 15, 2017

LINK TO MOST RECENT VERSION

Abstract

Studying wage inequality requires understanding how workers and firms match. I propose a novel strategy to identify the complementarities in production between unobserved worker and firm attributes, based on the idea that positive (negative) sorting implies that firms upgrade (downgrade) their workforce quality when they grow in size. I use German matched employer-employee data to estimate a search and matching model with worker-firm complementarities, job-to-job transitions, and firm dynamics. The relationship between changes in workforce quality and firm growth rates in the data informs the strength of complementarities in the model. Thus, this strategy bypasses the lack of identification inherent to environments with constant firm types. I find evidence of negative sorting and a significant dampening effect of worker-firm complementarities on wage inequality. Worker and firm heterogeneity, differential bargaining positions, and sorting contribute 71%, 20%, 32% and -23% to wage dispersion, respectively. Reallocating workers across firms to the first-best allocation without mismatch yields an output gain of less than one percent.

*I want to thank Lee Ohanian, Pablo Fajgelbaum, Till von Wachter and Pierre-Olivier Weill for invaluable support and advice. I also thank Andy Atkeson, Ariel Burstein, Kyle Herkenhoff, Andrea Di Miceli, and Ioannis Kospentaris and the seminar participants of UCLA, Vienna Institute for Advanced Studies and UC Louvain for useful comments. I gratefully acknowledge financial support from the Austrian Academy of Sciences. The usual disclaimer applies.

†University of California, Los Angeles; andreasgulyas@ucla.edu

1 Introduction

Why are observationally similar workers paid different wages? It has been established that observable worker and firm characteristics only account for some 30 percent of wage variation, see e.g. Abowd, Kramarz, and Margolis (1999). Therefore, understanding wage inequality requires identifying the distributions of unobserved worker and firm attributes, as well as how workers and firms sort. However, identifying the strength and direction of sorting has proven to be elusive. There is little consensus in the literature on the pattern of sorting and its importance for wage dispersion.

In this paper, I propose a novel strategy to identify 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 be determined by 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.

To leverage this idea, I develop a search and matching model with heterogeneous workers and firms, job-to-job transitions, sorting, and firm dynamics originating from idiosyncratic firm-level shocks. I estimate the structural model using German matched employer-employee data. I find that establishment growth is negatively related to changes in average workforce skills. This translates into negative sorting and an estimated correlation of -0.077 between worker and firm types. I then use the structural model to decompose the sources of wage variation. Worker heterogeneity explains the largest fraction with 71 percent. Firm heterogeneity, differential bargaining positions, and the complementarities in production contribute 20%, 32%, and -23% to wage variation, respectively. The estimated complementarities dampen wage dispersion because they induce negative sorting in equilibrium.

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. This will give rise to another source of wage dispersion through differential bargaining positions across workers.

I depart from most of the sorting literature by assuming that firms face idiosyncratic productivity shocks.[1] Given the large labor reallocation across firms observed every quarter

[1]To my best knowledge, Lise, Meghir, and Robin (2015) is the only other study considering firm-level shocks. Lise et al. (2015) do not use firm-level data and hence can only identify the strength of sorting, and

(Davis, Faberman, and Haltiwanger, 2006, 2012), it is highly implausible to assume fixed firm types over extended time periods. In my setup, firms adjust the size of their workforce in response to productivity shocks and also change the quality composition of their workforce. This reorganization happens in response to complementarities in production between firm and worker types. As in Becker (1973), if the two are complements, then positive assortative matching prevails in the labor market. In this case, 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 thus they reorganize their workforce towards higher skilled workers. With negative sorting, the exact opposite happens. Low type workers are more valued by high type firms and therefore firms downgrade their skill distribution after positive shocks and upgrade it after negative ones. I follow an extensive literature in economics explaining differences in firm size by productivity differences.[2] Because firms face convex job creation costs, firms with positive productivity shocks tend to expand 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. The key identification moment for the sign and degree of complementarities is the relationship between changes in average workforce quality and firm growth rates. 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.[3]

I estimate the model with German matched employer-employee data and find that establishments upgrade their worker skills when they downscale and downgrade them when they expand, after controlling for aggregate and industry-wide shocks. Establishments separate from low type workers when they shrink and hire less skilled workers as they grow. This result translates into weak negative sorting with a correlation coefficient of -0.077 between worker and firm types. My structural model implies four distinct sources of wage variation: worker and firm heterogeneity, differential bargaining positions, and sorting. First, variations in worker and firm types manifest themselves in wage dispersion. Second, due to firm shocks and job-to-job transitions, identical workers employed by the same firm type typically earn different wages. The same worker receives different wages if hired from unemployment or poached from a different firm. Additionally, the complementarities in production affect wage dispersion through sorting. I decompose wage variation into these four sources by

not the sign

[2]See e.g. Hopenhayn (1992), Melitz (2003), Luttmer (2007) and Lentz and Mortensen (2008)

[3]Lopes de Melo (2013) shows in a similar model that worker fixed effects capture the corresponding true worker types quite closely.

computing counterfactual economies adding one channel at a time. This reveals that worker heterogeneity alone explains 71 percent of wage dispersion. Firm heterogeneity and variations in bargaining positions add another 20 and 32 percent, respectively. The estimated negative sorting of workers across firm types dampens wage variation significantly by 23 percent compared to an economy without sorting.

To estimate the effects of search frictions in Germany on the extent of mismatch, I compute the output gains of reallocating workers across firms according to the frictionless allocation in Becker (1973). This reshuffling of workers yields an output gain of less than one percent.

My structural model builds on earlier papers studying wage inequality with search models without sorting. I borrow from 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.

My paper is joining a growing literature studying the sorting patterns in labor markets and its implication for wage inequality. Abowd, Kramarz, and Margolis (1999) pioneered the identificaiton of sorting by correlating worker and firm fixed effects from wage panel data. A large number of papers followed their approach and reached inconclusive results.[4] The fixed effect approach has recently been called into question by Eeckhout and Kircher (2011) and Lopes de Melo (2013). They point out that it relies on wages being monotonically increasing in firm types, which is violated in search models with sorting. Workers' wages typically peak at firm types providing the best match to their own type. In addition to this, firm types are not fixed my in framework, violating another identification assumption of the fixed effect approach.

My paper is closest to Bagger and Lentz (2015) and Hagedorn, Tzuo, and Manovskii (2014), who also provide identification strategies to identify both the sign and strength of sorting. I relax their assumption of fixed firm types. Furthermore, their identification strategies cannot be applied in my framework. I cannot rank workers within firms by their wage as in Hagedorn et al. (2014), because the same worker types might earn different wages if hired at different points in time in my framework. The poaching index proposed by Bagger and Lentz (2015) to rank firms relies on their assumption of no opportunity costs of matching

[4]Studies finding negative corrleation include amongst others Abowd, Kramarz, Lengermann, and Pérez-Duarte (2004), Andrews, Gill, Schank, and Upward (2008), Woodcock (2011). Iranzo, Schivardi, and Tosetti (2008) and Lopes de Melo (2013) find little sorting, and Card, Heining, and Kline (2013) and Song, Price, Guvenen, Bloom, and Von Wachter (2015) report positive sorting.

on the firm side. In my model, vacancies do not depreciate immediately and hence also lower type firms poach mismatched workers away from higher type firms, rendering the poaching index possibly non-monotonic in firm types.

Additionally, Lopes de Melo (2013), Lise and Robin (2014) and Lise, Meghir, and Robin (2015) study sorting in structural models of the labor market. Their approaches only allow them to identify the strength of sorting, whereas my procedure in addition identifies whether the labor market is characterized by positive or negative sorting. Bonhomme, Lamadon, and Manresa (2016) provide a semi-structural approach to study how firms and workers sort together. Abowd, Kramarz, Pérez-Duarte, and Schmutte (2014) study sorting between and within industries.

Bartolucci, Devicienti, and Monzon (2015) rank firms by profits, although the theoretical basis for this is not provided. Card, Heining, and Kline (2013), Kantenga and Law (2014) and Song, Price, Guvenen, Bloom, and Von Wachter (2015) study the effects of changes in sorting patterns over time on trends in wage inequality.

An additional contribution of my paper outside the sorting literature is to document how firms reorganize their workforce in response to shocks. There is surprisingly little evidence on this. 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 (2016) 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. Davis, Faberman, and Haltiwanger (2006, 2012) and Borovicková (2016) study job and worker flows, but not the skill composition of these flows.

This paper is structured as follows: the next section presents the key identification idea in a simplified search model. In section 3, I present the full model. Section 4 discusses the identification of all parameters, and section 5 provides the estimation results, which includes the estimated output loss due to mismatch and the decomposition of wage variation into worker and firm heterogeneity, sorting and the bargaining positions of workers. The last section concludes.

2 Simple Model

I begin with a simple search model to explain why considering firm dynamics will be useful for identifying the patterns of sorting. The simple model presented here is similar to Eeckhout and Kircher (2011). The full structural model shares the key building blocks that determine the sorting patterns and therefore the intuition presented here will carry through in my model.

Consider an economy populated by heterogenous workers and firms. Firms operate with constant returns to scale at the match level, therefore firms should be thought of a collection of individual matches. Worker types are denoted by x and firms' by y and both are uniformly distributed between 0 and 1. A match between types x and y produces output $f(x, y)$ with $f(0, 0) = 0$. The function $f(x, y)$ is twice continuously differentiable with $f_x(x, y) > 0$ and $f_y(x, y) > 0$. Thus, high types always have an absolute advantage over low types. As in Becker (1973), the cross partial derivative $f_{xy}(x, y)$ determines the pattern of assortative matching.

Becker's result in this frictionless environment is that if $f_{xy}(x, y)$ is positive, the equilibrium features positive assortative matching (PAM). Higher worker types have a relatively higher marginal productivity at high type firms, and thus they will end up working for those firms. Conversely, if $f_{xy}(x, y)$ is negative, high type workers are relatively more valued by low type firms, and negative assortative matching (NAM) will prevail in equilibrium.

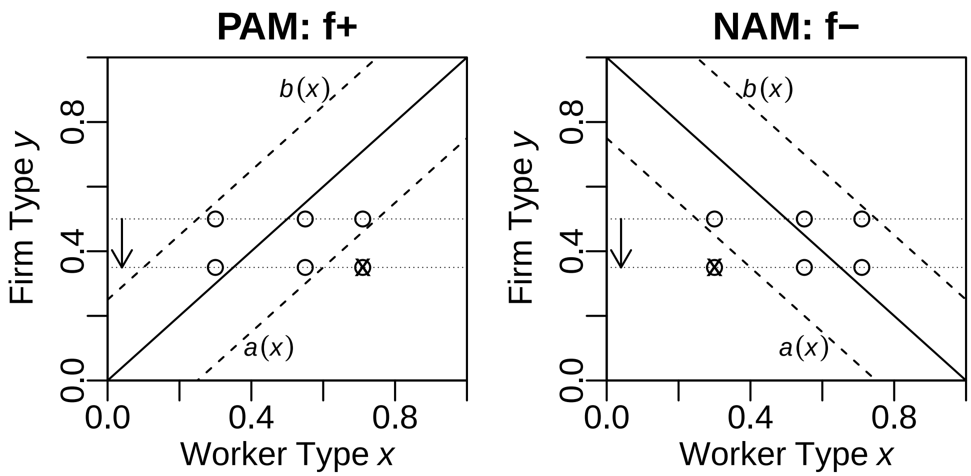
Now consider the following one period model: First, firms and workers start out being randomly matched together. If the parties decide to stay matched, they split the surplus evenly according to Nash bargaining. If they decide to separate, they both pay search cost c and receive the competitive payoffs corresponding to the frictionless assignment, which are given exogenously and denoted by $w^*(x)$ for workers and $\pi^*(y)$ for firms.

The surplus from matching is $f(x, y)$ minus the outside options. These are to pay the search cost c and get $\pi^*(y)$ and $w^*(x)$. The two parties will stay matched if the surplus is positive:

$$S(x, y) = f(x, y) - (w^*(x) + \pi^*(y) - 2c) \geq 0. \tag{1}$$

The matching sets are characterized by all combinations of worker and firm types that entail a positive surplus. Here, also a measure of mismatch arises. Clearly, the surplus will be the highest between firms and workers that would be partners in the frictionless case. Because agents have to pay a search cost c, everyone is willing to tolerate a match with a suboptimal partner, as long as the partner is not too different from their most preferred

Figure 1



Notes: Matching sets for production functions implying PAM and NAM.[5] With PAM, high type workers separate after negative shocks, whereas with NAM, it is the low type workers that move outside the matching bands and separate after negative productivity shocks.

one. Following Shimer and Smith (2000), PAM and NAM in a frictional environment can be defined by considering the slopes of matchings sets' bounds. There is PAM (NAM) if and only if the bound functions $a(x) \equiv \min\{y|S(x,y) \geq 0\}$ and $b(x) \equiv \max\{y|S(x,y) \geq 0\}$ are nondecreasing (nonincreasing). Intuitively, under PAM, types prefer to match with similar types whereas with NAM, agents prefer opposite types. Figure 1 plots an example of matching sets under PAM and NAM.

Let me now consider unexpected firm-level productivity shocks after the first stage. Figure 1 illustrates an example of a firm with a negative shock. The circles highlight some workers inside the original matching sets. Under positive sorting, lower productivity leads to a leftward shift in the matching set. This implies that the highest type worker move outside the matching set and separates from the firm (marked with a cross). This is in stark contrast to the NAM case as can be seen in the figure. In this case, a firm with a negative shock separates from its lower type employees. This simple intuition provides the basis for my novel identification strategy. Studying how firms reorganize the quality of their workforce in response to productivity shocks reveal the complementarities in production. Under PAM, firm growth after positive shocks will be associated with worker skill upgrading, whereas NAM will induce a negative relationship between growth rates and changes workforce quality. This allows me to sidestep the identification problem highlighted by Eeckhout and Kircher (2011):

In this framework, firm types cannot be identified using wage data alone. Workers' wages are not necessarily monotonically increasing in firm type, because they typically peak at the firm type that provides the best match for the worker's type.

The next section lays out the infinite horizon model with job-to-job transitions, firm dynamics, and endogenously emerging outside options. Nevertheless, the driving force behind sorting will be the same, and thus the basic intuition for the identification strategy explained in this section will still carry through.

3 Model

This section presents the full search model with multi-worker firms that is used for the estimation. 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).

The matching process of heterogeneous firms and workers is impeded by search frictions. Firms expand and shrink in response to productivity shocks, but also adjust the skill composition of their labor force, depending on productive complementarities at the match level.

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 and firms are heterogenous with respect to a one dimensional productivity type, denoted by x and y, respectively.[6] On the firm side, this comprises any characteristic that affects productivity such as managerial skills, capital intensity or quality of the capital stock. On the worker side any productive capacity of the worker not observed by the researcher. Worker types are fixed over time, whereas firm productivity is subject to idiosyncratic shocks. The stationary distribution of worker and firm types are give by the probability distribution functions $\phi_x(x)$ and $\phi_y(y)$ with support [0,1]. A match between a worker type x and firm type y produces $f(x, y)$, where $f(x, y)$ is twice continuously differentiable with $f_x(x, y) > 0$ and $f_y(x, y) > 0$. Thus, high types always have an absolute advantage over low types.

[5] $f^{PAM}(x, y) = \alpha xy + h(x) + g(y)$ and $f^{NAM}(x, y) = \alpha x(1-y) + h(x) + g(y)$. $h(x)$ and $g(x)$ are increasing functions such that $f_x^{PAM}(x, y) > 0, f_x^{NAM}(x, y) > 0, f_y^{PAM}(x, y) > 0, f_y^{NAM}(x, y) > 0$.

[6] Lindenlaub (2014) and Lise and Postel-Vinay (2014) provide notions of sorting based on multidimensional characteristics

Firms produce with a linear production technology.[7] Thus, the total production of a particular firm j is given by the integral over the distribution $\psi_j(x)$ of all of its individual matches, or

$$F_j(y) = \int f(x,y) d\psi_j(x). \tag{2}$$

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 it is filled or vacant. A (costless) vacancy is automatically posted for every unfilled jobs. Firm face idiosyncratic shocks to their 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 $c''(v^N) > 0$.[8] Firms will create new jobs until the marginal cost of establishing a new job is equal to the marginal value of a vacant job, i.e.

$$c'(v^N) = V(y), \tag{3}$$

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

$$v^N(y) = c'^{-1}(V(y)). \tag{4}$$

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

[7] 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.

[8] In my setup, it is better to think that firms create jobs and not just post vacancies. This is because jobs will stay around for a long time, they only depreciate slowly with rate d. This is in contrast to the setup in Bagger and Lentz (2015), 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.

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:

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

3.1 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 and negotiated wage rate w is $W(x, y, w)$. The value of a vacant job to a firm of type y is denoted as $V(y)$, whereas the value of a job occupied by a worker of type x with a wage rate w is $J(x, y, w)$. The value function are presented below in equations (16)-(19). The surplus of a match is consequently defined as

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

Wages are negotiated at the beginning of each employment spell and might be renegotiated after 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 $w^U(x, y)$ is set according to Nash bargaining with the worker's bargaining power α. Thus

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

As in Postel-Vinay and Robin (2002), Dey and Flinn (2005) and Cahuc et al. (2006), when a worker employed at a firm y meets another firm $\tilde{y}$ that would generate a higher surplus, the two companies engage in Bertrand competition. This drives up the wage to the point where the worker obtains the full surplus from his old job $S(x, y)$. Thus, after job-to-job transitions, the wage rate $w^E(x, y, \tilde{y})$ is set such that:

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

As in Hagedorn et al. (2014), I assume that workers cannot use outside offers from firms that would generate lower surpluses to negotiate their wages up. This assumption simplifies the exposition and has no effect on my identification. It can be rationalized with a small cost of writing an offer which prevents firms with no chance of poaching to engage in Bertrand competition.

After productivity shocks, I assume that wages are renegotiated if either the worker's value falls below her outside option ($W(x, y, w) - U(x) < 0$) or the firm's value falls below the value of a vacancy ($J(x, y, w) - V(x) < 0$). 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 shocks pushes the value of a worker below her participation threshold ($W(x, y, w) - U(x) < 0$), the firm extracts the full surplus and the wage $w^{NW}(x, y)$ is set such that

$$w^{NW}(x, y) \quad : \quad W(x, y, w) - U(x) = 0. \tag{10}$$

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

$$w^{NF}(x, y) \quad : \quad W(x, y, w) - U(x) = S(x, y). \tag{11}$$

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 (2014)). Second, it avoids the situation where inefficient separations happen 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 bad 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. 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.

For the presentation of the value functions in the next subsection it will be useful to define the events of wage renegotiation by indicator functions. I denote the wage renegotiation event as $A^{NW}(x, y, w)$ if triggered by the worker as $A^{NF}(x, y, w)$ if triggered by the firm:

$$A^{NW}(x, y, w) = \begin{cases} 1 & \text{if } W(x, y, w) - U(x) < 0 \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

$$A^{NF}(x, y, w) = \begin{cases} 1 & \text{if } W(x, y, w) - U(x) > S(x, y) \\ 0 & \text{otherwise .} \end{cases} \tag{13}$$

3.2 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 wage renegotiations or 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 matching sets are characterized by indicator functions. If an unemployed worker of type x meets a vacant job of productivity y, $A^U(x, y)$ takes on the value of 1 if the match is consummated and zero otherwise. Similarly, if a worker x employed at a type y firm is contacted by a poaching firm of type $\tilde{y}$, $A^E(x, y, \tilde{y})$ 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,

$$A^U(x,y) = \begin{cases} 1 & \text{if } S(x,y) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

$$A^E(x,y,\tilde{y}) = \begin{cases} 1 & \text{if } S(x,\tilde{y}) \geq S(x,y) \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

Equation (16) presents the value of a vacancy at the beginning of the period.

$$\begin{aligned} V(y) =& \beta(1-d) \int\limits_{y_{\min}}^{y_{\max}} \Bigg\{ V(y') + \lambda_f \Bigg(p^u \int\limits_{x_{\min}}^{x_{\max}} A^U(x,y')(1-\alpha) S(x,y') \frac{\mu_x(x)}{u} dx + \\ &+ (1-p^u) \int\limits_{y_{\min}}^{y_{\max}} \int\limits_{x_{\min}}^{x_{\max}} A^E(x,\tilde{y},y') \left(S(x,y') - S(x,\tilde{y}) \right) \frac{\psi^S(x,\tilde{y})}{e^s} dx d\tilde{y} \Bigg) \Bigg\} p(y'|y) dy'. \end{aligned} \tag{16}$$

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 ($A^U(x,y) = 1$) or the employed worker of type x is successfully poached away from her current employer of type $\tilde{y}$ ($A^E(x,\tilde{y},y') = 1$). 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 the measure of unemployed of type x $\mu_x(x)$ 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 $\tilde{y}$ is given by the probability mass of employed types at the search stage $\psi^S(x,\tilde{y})$ 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 generated with a previously unemployed worker. In case the employee has to be poached, Betrand competition implies that the firm is left over with the generated surplus $S(x, y')$ minus the surplus which the worker generated at her old job $S(x, \tilde{y})$.

Equation (17) represents the value of a filled job to a firm at the beginning of the production stage.

$$
\begin{aligned}
J(x, y, w) \quad = \quad & f(x, y) - w + \beta(1-d) \int_{y_{\min}}^{y_{\max}} \Bigg\{ V(y') + A^U(x, y') \int_{y_{\min}}^{y_{\max}} \Big\{ \left[1 - \lambda_e A^E(x, y', \tilde{y})\right] \\
& \times \left[\left(1 - A^{NW}(x, y', w) - A^{NF}(x, y', w)\right) \left(J(x, y', w) - V(y')\right) \right. \\
& \left. + A^{NW}(x, y', w) S(x, y') \right] \Big\} \frac{\mu_y(\tilde{y})}{v} d\tilde{y} \Bigg\} p(y'|y) dy'. \qquad (17)
\end{aligned}
$$

It consists of the flow output net of wages $f(x, y) - w$ plus the discounted continuation value. Since jobs are destroyed with probability d, the effective discount rate is $\beta(1-d)$. A number of other events affect the continuation value. The match only continues if the surplus is positive ($A^U(x, y') = 1$) and the worker is not poached. This is the case if the worker does not meet another job or the contacted employer $\tilde{y}$ has a lower surplus than the current one. The worker meets firm $\tilde{y}$ with probability $\mu_y(\tilde{y})/v$. Therefore, the probability that the worker does not experience a job-to-job transition to firm $\tilde{y}$ amounts to $1 - \lambda_e A^E(x, y', \tilde{y})$. If the match surplus becomes negative or the worker is poached the firm is left with a vacancy, which the firm values with $V(y')$. If the worker does not separate from the firm, the productivity shock might trigger a renegotiation of the wage rate. If it is demanded by the worker, the firm extracts the full surplus $S(x, y')$. 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 $J(x, y')$.

The worker's value functions are the mirror image of the firms' problems and are given in equation (18) and (19). Notice that the flow value of unemployment is normalized to zero.

$$
U(x) \quad = \quad \beta \left(U(x) + \lambda_w \int_{y_{\min}}^{y_{\max}} A^U(x, y) \alpha S(x, y) \frac{\mu_y(y)}{v} dy \right) \qquad (18)
$$

$$W(x,y,w) = w+\beta\left(U(x)+(1-d)\int\limits_{y_{\min}}^{y_{\max}}\left\{A^U(x,y')\int\limits_{y_{\min}}^{y_{\max}}\left\{\lambda_e A^E(x,y',\tilde{y})S(x,y')\right.\right.\right.$$
$$+(1-\lambda_e A^E(x,y',\tilde{y}))\left[\left(1-A^{NW}(x,y',w)-A^{NF}(x,y',w)\right)(W(x,y',w)-U(x))\right.$$
$$\left.\left.\left.+\ A^{NF}(x,y',w)S(x,y')\right]\frac{\mu_y(\tilde{y})}{v}\right\}d\tilde{y}\right\}p(y'|y)dy'\Bigg). \tag{19}$$

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

$$\begin{aligned}
S(x,y) =& f(x,y)+\beta(1-d)\int\limits_{y_{\min}}^{y_{\max}} A^U(x,y')S(x,y')p(y'|y)dy' \\
&- \beta\alpha\lambda_w \int\limits_{y_{\min}}^{y_{\max}} A^U(x,y)S(x,y)\frac{\mu_y(y)}{v}dy \\
&- \beta(1-d)\lambda_f \int\limits_{y_{\min}}^{y_{\max}} \left(p^u \int\limits_{x_{\min}}^{x_{\max}} \left(A^U(x,y')(1-\alpha)S(x,y')\right)\frac{\mu_x(x)}{u}dx+\right. \\
&+\ (1-p^u)\left.\int\limits_{y_{\min}}^{y_{\max}}\int\limits_{x_{\min}}^{x_{\max}} A^E(x,y',\tilde{y})\left(S(x,y')-S(x,\tilde{y})\right)\frac{\psi^S(x,\tilde{y})}{e^s}dxd\tilde{y}\right)p(y'|y)dy'.
\end{aligned} \tag{20}$$

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.

$$
\begin{aligned}
\mu_y(y) \;=\;& (1-d)\int\limits_{y_{\min}}^{y_{\max}} v^N(y)\phi(y) + \left((1-\lambda_f)+\lambda_f\left(p^u\int\limits_{x_{\min}}^{x_{\max}}(1-A^U(x,y'))\frac{\mu_x(x)}{u}dx\right.\right.\\
&+ (1-p^u)\int\limits_{y_{\min}}^{y_{\max}}\int\limits_{x_{\min}}^{x_{\max}}(1-A^E(\tilde{x},\tilde{y},y'))\frac{\psi^S(x,\tilde{y})}{e^S}dxd\tilde{y}\Bigg)\Bigg)\,p(y|y')\mu_y(y')dy'\\
&+\lambda_e\int_{y_{\min}}^{y_{\max}} A^E(x,y',y)\frac{\mu^y(y)}{v}\psi^S(x,y')dy'\\
&+\int\limits_{y_{\min}}^{y_{\max}}\int\limits_{x_{\min}}^{x_{\max}}\left(d+(1-d)\left(1-A^U(x,y)\right)\psi(x,y)\right)p(y|y')dxdy'. \qquad (21)
\end{aligned}
$$

Three distributions emerge endogenously in my model. In a stationary equilibrium the in- and outflows of the distributions of vacancies $\mu_y(y)$, unemployed $\mu_x(x)$ and employed workers across firm types $\psi(x,y)$ must balance each other. Equations (21) - (22) present the law of motions of these three distributions in steady state. Let me first consider the law of motion for the distribution of vacant jobs in equation (21). The first two lines comprise the unfilled jobs carried over from last period plus the newly created jobs that were not hit by a job destruction shock. The last two lines are the inflows from separations.

$$
\begin{aligned}
\psi(x,y) \;=\;& (1-\lambda_e)\psi^S(x,y)+\lambda_w A^U(x,y)\frac{\mu^y(y)}{v}\mu^x(x)\\
&+\lambda_e\int\left(1-A^E(x,y,\tilde{y})\right)\frac{\mu^y(\tilde{y})}{v}d\tilde{y}\psi^S(x,y)\\
&+\lambda_e\int_{y_{\min}}^{y_{\max}} A^E(x,y',y)\frac{\mu^y(y)}{v}\psi^S(x,y')dy'. \qquad (22)
\end{aligned}
$$

$$
\psi^s(x,y) \;=\; (1-d)A^U(x,y)\int\psi(x,y')p(y|y')d\tilde{y} \qquad (23)
$$

Equation (22) describes the law of motion for the joint distribution of matches at the production stage $\psi(x,y)$. The relevant measure for workers and firms is the distribution at the search stage $\psi^s(x,y)$, which is given in equation (23). 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 (23).

All the remaining workers engage in on-the-job search.

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

$$\mu^x(x) = \phi_x(x) - \int \psi(x,y)dy. \tag{24}$$

3.3 Identifying Sorting with Wages

Before I move on to the discussion of identification, I discuss why we cannot simply use fixed effects to identify firm types in wage regressions. For the fixed effects to correctly identify worker and firm types, it must be the case that wages are increasing in worker and firm type. Otherwise the ordering of firms by the estimated fixed effects would not recover the true ranking of firm types. To understand why wages may not satisfy this monotonicity condition, let us consider the following simplified case. I abstract from job to job transitions ($\lambda_e = 0$) and assume no firm shocks. In this case, wages are given by:

$$w(x,y) = \alpha \left[f(x,y) - (1-\beta(1-d))\, V(y) \right] + (1-\alpha)(1-\beta)U(x) \tag{25}$$

As will be shown later, $U'(x) > 0$ and $V'(y) > 0$ in this simplified model as well as in the richer model. This holds because by definition, higher types always produce more regardless of the match (i.e. $f_x(x,y) > 0$ and $f_y(x,y) > 0$). It follows that wages are monotonically increasing in worker type x, but it is not necessarily in firm productivity y. 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. This non-monotonicity in wages has been demonstrated before by Eeckhout and Kircher (2011) and Lopes de Melo (2013), amongst others.

Adding firm productivity shocks and job-to-job transitions complicates the identification further. The fixed-effects estimator identifies worker and firm effects off workers transitions across firms. Therefore, a connected set of firms and workers must be observed over a sufficiently long time span, typically around 10 years.[9] However, over such a time horizon, the assumption of constant firm types becomes implausible.

In my model, due to firm shocks and job-to-job transitions, identical workers employed

[9] Abowd et al. (1999), Song et al. (2015)

by the same firm type typically earn different wages. For example, the same worker receives different wages if hired from unemployment or poached from a different firm. In addition, wages also depend on the timing of the hire, since the bargaining setup implies that the bargaining positions are retained until one of the partners has a credible threat to terminate the employment spell.

As outlined in section 2, my identification strategy is to study how firms reorganize the quality of its employees in response to firm productivity shocks. This approach requires to identify both firm shocks and which types of workers join or separate from the firm. Let me first consider the identification of worker types. All is needed is a measure that is monotonically increasing in worker type x. Average lifetime earnings of workers provide such a measure and can be readily computed from typical matched employee-employer datasets. As I show in the next section, lifetime earnings, averaged over both employment and unemployment spells are monotonically increasing in worker type x.[10]

Firm shocks on the other hand can be identified by changes in firm size. More productive firms will grow larger in my setup. Thus, firms with positive productivity shocks tend to grow, and shrink after negative ones, regardless of the degree or strength of sorting. Here, I follow a large economic literature explaining variation in firm sizes by firm productivity differences.[11]

The details of the identification strategy are discussed in the following section.

4 Identification

The estimation follows an indirect interference approach. First, I choose a set of auxiliary statistics from the German Social Security data. 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.

[10]In contrast, averages wages might not be monotonically increasing in x. Since the matching sets are different for different sets of workers, it might be that some worker types have to be compensated for longer unemployment durations by higher average wage

[11]see e.g. Hopenhayn (1992), Melitz (2003), Luttmer (2007) and Lentz and Mortensen (2008)

4.1 Functional Form Assumptions

The model is estimated at a monthly frequency. The functional form assumptions are summarized in table 1. I use a CES production function of the form $f(x,y) = f_1 \left(x^{1/\rho} + y^{1/\rho}\right)^{\rho}$. It can generate a variety of different sorting patterns, depending on the complementary parameter ρ. In a frictionless economy as in Becker (1973), a value of $\rho < 1$ would generate negative sorting, whereas $\rho > 1$ would imply positive sorting in equilibrium. The production function also nests the no sorting case if $\rho = 1$.

Table 1: Functional forms

Worker distribution	Log-Normal(μ_x, σ_x)
Production function	$f_1 \left(x^{1/\rho} + y^{1/\rho}\right)^{\rho}$
Job creation cost function	$c_0 \left(\frac{v}{c_1}\right)^{c_1}$
Firm shocks	$f(y'\|y) = \begin{cases} y & \text{with prob. } 1-\phi \\ y' \sim \text{unif}(y-\bar{y}, y+\bar{y}) & \text{with prob. } \phi \end{cases}$

Notes: Log normal distribution is truncated to [0,1]. Since $y \in [0,1]$, the probability mass that falls outside this range is added to the stay probability. Thus, values of y close to the boundaries have a slightly higher probability of not changing.

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 choose a Markov process for the firm productivity shocks. Productivity shocks occur with Poisson frequency ϕ. In this case, the new productivity y' is drawn from a uniform distribution with support symmetrically around the old value, i.e. $y' \sim \text{unif}(y-\bar{y}, y+\bar{y})$.[12] A similar firm productivity process is assumed in Kaas and Kircher (2014). This Markov process implies a uniform steady state distribution of firms across types. The endogenous distribution of jobs across productivity types will be primarily governed by the job creation 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]

Three 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 1, and set the production function scale f_1 such that the maximum

[12] Since firm productivity y is bounded between [0,1], firm productivity might fall outside this range. To circumvent this, in cases where the support of y' would fall outside of [0,1], I add all the probability mass outside [0,1] to the probability of staying at the same level y. This implies that values close to the end points will have a slightly higher probability of not receiving a firm shock.

[13] Bagger and Lentz (2015), Coşar et al. (2010) and Merz and Yashiv (2007) amongst many others use this functional form

possible output is equal to 1. I set the discount rate to 0.995, which implies a yearly discount rate of about 6 per cent. Last, I fix the bargaining power of workers to 0.3, which is similar to the values used in Bagger and Lentz (2015) or Lise et al. (2015).

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 50 worker and 50 firm types. First, I obtain the acceptance sets by solving for a fixed point in the surplus function $S(x, y)$ and the endogenous distributions $\psi(x, y)$, $\mu^y(y)$ and $\mu^x(x)$. I then simulate 2500 firms over 18 years to construct a panel data set similar to the German social security data. Appendix B describes the numerical implementation in detail. I compute a set of auxiliary statistics on the model simulated data as on the German social security data, which selections I discuss in the following subsections.

Table 2 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.

4.2 Identifying the Complementarity Parameter ρ

The key parameter driving the sorting pattern is the complementarity parameter ρ. The basic idea is to study how firms adjust the quality of their employees in response to shocks. First, I show below that one can identify worker types by computing their average lifetime earnings. 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 firm productivities are complements in the production function, positive sorting prevails in equilibrium with similar types matching together (Becker, 1973). As a result, expanding firms reorganize their workforce towards higher quality, whereas shrinking firms reorganize towards lower quality workers. With negative sorting, low type firms employ high type workers and expanding firms 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.

Two results are useful for the identification and are stated in the following proposition:

Proposition. *U(x) and V(y) are increasing in their arguments.*

Both results are standard in search models and the proofs are given in the appendix. Intuitively, since higher type workers and firms always produce more independently of the match, higher types have a higher value of unemployment $U(x)$ and vacancies $V(y)$. First, $U(x)$ is tightly linked to average earnings of workers. Consider a worker at the beginning of her career. Her expected lifetime earnings are by definition the expected discounted sum of all per period payoffs $\pi_t(x)$, or simply $U(x)$. Then, we can write $U(x)$ as

$$U(x) = \mathbb{E}\sum_{t=0}^{\infty}\beta^t\pi_t(x) = \sum_{t=0}^{\infty}\beta^t\mathbb{E}\left[\pi_t(x)\right] = \frac{\mathbb{E}\left[\pi(x)\right]}{1-\beta} \tag{26}$$

The first equality is the definition of the value function. The second holds due to the linear utility assumption. The stationarity of the income process assures the last equality. The monotonicity of $U(x)$ implies that average per period earnings are monotonically increasing in x. Worker productivity represents any fixed non-observable productive characteristics of the worker in my model. Thus, for the mapping between data and the model, I follow Card et al. (2013) and Hagedorn et al. (2014) and filter out the explained portion of wages of age, education and their interaction term[14]. Average annual earnings after controlling for age and education identifies worker types.[15] Further details are described in appendix C.

Having a measure for worker skills, I can study how firms reorganize the skill composition of their workforce in response to productivity shocks. 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. This follows from $V'(y) > 0$ and the job creation equation (4), because more productive firms create more jobs and hence grow larger.[16]

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-

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

[15]In my model, I normalized the flow payoff from being unemployed to zero. When computing yearly earnings with the German social security data, I have to impute the flow value of unemployment. I calculated three different specifications: Imputing zero as in the model, the actual unemployment benefits the person is receiving and benefits plus a 20 percent premium representing non-monetary payoffs from unemployment such as home production and leisure. The correlation between the three different 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 use the first specification.

[16]It can happen that the job filling probability is lower for higher type firms as they might be "pickier". Since vacancies depreciate slowly with the same rate as filled jobs, it is nevertheless the case that more productive firms grow larger.

employer or model simulated data:

$$\Delta_{\%}\overline{Wquality}_{jt} = \alpha + \gamma growth_{jt} + \epsilon_{jt}. \tag{27}$$

Here, $\Delta_{\%}\overline{Wquality}_{jt}$ 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 beginning of each calendar year by averaging the employees' worker quality measure. Then $\Delta_{\%}\overline{Wquality}_{jt}$ 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 type is a complement to firm productivity ($\rho > 1$), 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 decrease the average level of their employees quality level. By the same argument, firms with negative shocks downgrade the average skills they employ. With negative sorting, the exact opposite is going to happen. Low 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 2, which shows the estimated relationship from the regression equation (27) with model simulated data. Under positive sorting ($\rho > 1$), expanding establishment upgrade the worker skills and shrinking ones downgrade them. If worker and firm types are substitutes ($\rho < 1$), a negative relationship between $\Delta_{\%}\overline{Wquality}_{jt}$ and $growth_{jt}$ is estimated. The regression yields a γ coefficient of virtually zero if there is no sorting ($\rho = 1$).

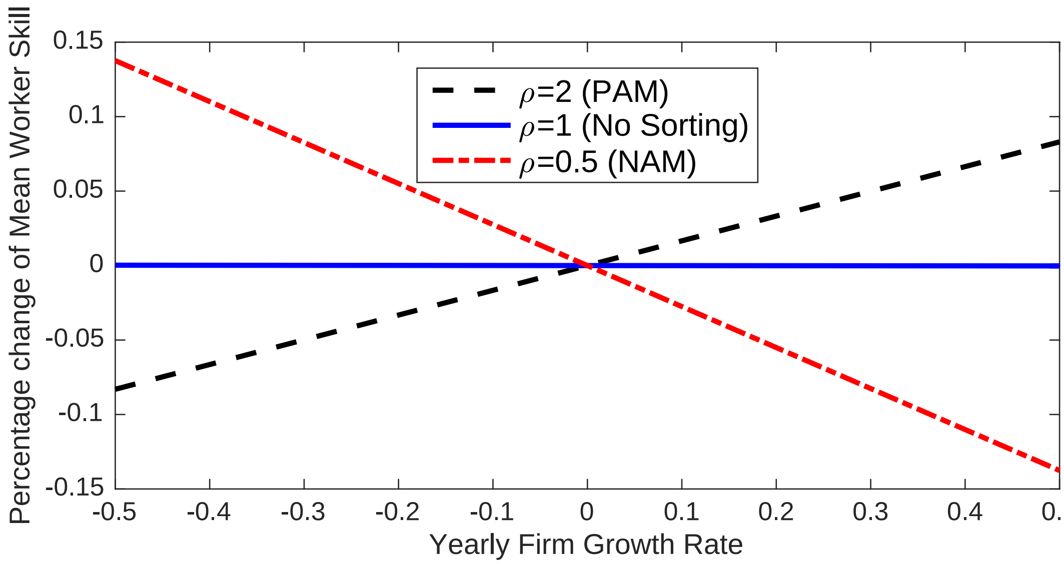
Figure 3 plots the results of regression equation (27) for German establishments using social security records. Instead of a a continuous measure of growth rates I use 5% establishment growth rate bins.[17] In my model, firms' adjustment of worker quality is driven by idiosyncratic productivity shocks. It is therefore important to filter out any business cycle or industry-wide effects from the empirical relationship. To address this, I include year dummies, 3-digits industry classifiers and the full interaction of the two as controls in regression equation (27). Furthermore, the regression is weighted by establishment employment.[18]

The results are suggestive of negative sorting. As figure 7 in the appendix shows, German

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

[18] The results are similar for unweighted regressions

Figure 2: Regression Slope for Different ρ



Notes: The figure shows the estimated relationship between firm growth rates and the percentage changes in average worker quality employed by firms using regression equation (27) for different values of ρ on model generated data. The rest of the parameters are held constant at the values reported in table 3.

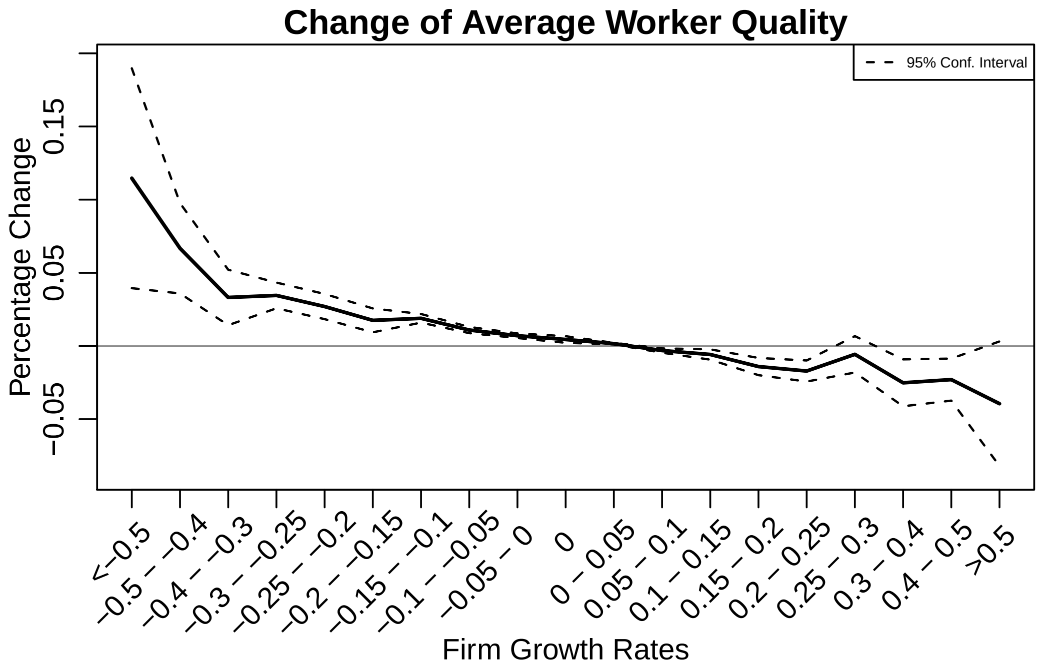
establishments shrink by separating from their lower type employees and expand by hiring low skilled workers. Thus, they upgrade the skill distribution of their workers when they scale back and downgrade it when they expand, as we would expect under negative sorting. Firms in general do not reorganize their workforce completely. Establishments that grow or shrink by less than 25% on average change the average worker quality by not more than 3 percent. Only firms with big shocks reorganize more aggressively. The coefficients are very precisely estimated, as the narrow 95% confidence intervals show.[19] This relationship says that, in shrinking firms, the workforce composition shifts towards workers with higher average lifetime 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 observables characteristics of workers (age and education).[20,21] This is another important advantage of identifying worker quality by their average lifetime earnings rather than ranking workers

[19] 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.

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

[21] 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 unobservable worker skills. It still holds that firms separate from their lowest earning workers and hire workers with wages below the current firm's median.

Figure 3: Reorganization of Worker Quality



Notes: The figure shows the percentage change of average employee fixed effect by establishment growth rates, controlling for year, 3-digit industry and interaction of year/industry effects. The sample consists of all establishments with size≥30. Estimates are weighted by employment and 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.

based on their current wage, which might be affected by factors outside the model such as match quality.

The relationship 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 (27) as one of the target moments in my indirect inference approach. Table 7 in the appendix presents the baseline estimate in column 1 that will be used as a target in the estimation. The estimated slope coefficient γ is -0.099, which mimics the slope of the relationship in figure 3.

In addition, the table reports a number of robustness exercises. The estimated coefficient is very robust across all specifications. One concern could be the that relationship only pertains to specific establishment sizes or ages. This is clearly rejected. First, including firm age and size as additional controls virtually leaves the slope coefficient unchanged. Second, if I only focus on the oldest or biggest establishments in my sample, I still find a negative and highly significant relationship, although it is slightly weaker. The relationship is also robust with respect to considering different time spans. I rerun the regression by considering three year windows instead of year-to-year changes and find very similar estimates. The relationship is also stable across different time periods.[22]

Table 8 in the appendix shows that the negative relationship between growth rates and worker quality adjustments is not driven by a few particular sectors. Although there is heterogeneity in the estimated relationship across 1-digit sectors, the results are indicative of negative sorting in all but one sector. Only in the sector comprising R&D, real estate, and software and hardware consulting are establishments upgrading their worker quality as they expand.

4.3 Identifying the Rest of the Parameters

The identification of the rest of the parameters is more standard. I target the total hire rate, together with the unemployment and job-to-job transition rate. The hire rate is defined as total number of hires normalized by employment.[23] I use the official German unemployment rates provided by the German Federal Employment Agency.[24] The unemployment rate averages to 8.24 percent between 1993 and 2010.[25] I count every transition from one firm to another with an intermitted spell of non-employment shorter than 31 days as a job-to-job transition.[26] Roughly speaking, these three parameters pin down the meeting rates for employed and unemployed workers λ_e, λ_w and the job destruction rate d.

The mean and the standard deviation of empirical fixed effect distribution will identify the scale and shape parameters of the worker type distribution μ_x and σ_x.[27]

The rest of the target moments mostly identify the parameters on the firm side. The parameters that affects the growth rate and establishment size distribution are the parameters of the job creation function c_0, c_1 and ϕ, $\bar{y}$ that govern the frequency and range of productivity shocks. To identify these parameters, I target the employment weighted stan-

[22] There was a significant labor market reform in Germany in 2004. It mostly affected the benefits for long term unemployed. I find no evidence of a break of the studied relationship.

[23] 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.

[24] In the social security data, I cannot distinguish between unemployment and non-participation. For this reason the official unemployment rate resembles the model implied unemployment rate more closely.

[25] Source: http://doku.iab.de/arbeitsmarktdaten/qualo_2015.xlsx

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

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

dard deviation of establishment growth rates, the employment weighted autocorrelation of establishment size, and the share of employment working in the 75 percent smallest establishments (labeled size distribution P75 in the table). The job filling rate will additionally help to identify the job creation cost function parameters. 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.[28][29]

The next section proceeds with the discussion of the estimation results.

5 Estimation Results

5.1 The Fit of the Moments

Table 2 presents the fit of all target moments and table 3 displays the estimated parameter values. The model closely matches all moments. The moments related to labor market transitions are fitted well. The rates of hiring, job-to-job transitions and job fillings are matched precisely. The estimated job destruction parameter d implies that jobs are on average exogenously destroyed every 5.5 years.

The standard deviation of employment-weighted growth rates is matched very closely. The estimated firm shock parameter implies that firms receive productivity shocks almost every two years, on average. This renders the assumption of fixed firm types even in very short time periods unrealistic.

There are small deviations from the targeted mean and standard deviation of the fixed effect distribution. The standard deviation of employment weighted growth rates, the employment share of the 75 percent smallest firms and the regression coefficient from equation (27) is exactly fitted. Although only one point in the firm size distribution is targeted, the model roughly replicates the shape of it, as presented in the right panel of figure 4. 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.

The fit of the coefficient from regression (27) is of particular interest, since it identifies the key parameter ρ which drives the sorting pattern. The linear regression on the model simulated data yields precisely the same coefficient as obtained from the German dataset.

[28] Source: `http://www.iab.de/stellenerhebung/download`

[29] Unfortunately, the data is only available from 2010 to 2015. This overlaps only in 2010 with the time period studied here. I nevertheless assume that the average of these 5 years is representative of the time period studied.

Table 2: Target Moments

Target Moment	Data	Model
Hire rate	0.024	0.025
Unemployment rate	0.082	0.082
Job-to-job transition rate	0.009	0.009
Job filling rate	0.388	0.388
Mean worker type distribution	0.460	0.478
Std of worker type distribution	0.228	0.260
Std of empl. weighted growth rates	0.123	0.123
Emp. weighted autocorr. of firm size	0.996	0.996
Size distribution P75	0.110	0.110
Regression Coeff Equation (27)	-0.099	-0.099

Notes: The standard deviation of yearly growth rates is employment-weighted. Size distribution P75 refers to the share of employment in the 75 percent smallest firms.

Even the more flexible representation of this relationship with firm growth rate categories instead of the continuous growth measure is captured remarkably well in the simulations. This relationship is shown in the left panel of figure 4. The coefficients on the growth rate dummies almost exactly match, except for firms declining the most. These establishment reorganize more aggressively than observed in the data. These outliers constitute only 0.33 percent of the sample and thus have not much weight in the linear regression.[30] The parameter ρ is estimated to be 0.644, which implies that worker and firm types are substitutes in the production function. This implies that negative sorting will prevail in equilibrium. The extent of sorting depends on the estimated importance of search frictions. To get a sense of how wide the confidence bands around the point estimate of ρ would be, I perform the following exercise: I re-estimate the model with targeting the lower and upper bound of the 95% confidence interval of the empirical slope coefficient from regression equation (27). This yields a confidence interval of the structural estimate ρ of [0.616,0.677].

[30]In addition, these are typically also smaller firms. Firms that shrink by more than 20 percent are on average half as big as the average establishment in the model simulations. Since I weigh by firm size, these observations are not only few but also have lower weights in the regression. Dropping the establishments with growth rates smaller than -0.2 from the sample has virtually no effect on the estimated regression coefficient from equation (27).

Table 3: Parameter Estimates

Parameters	Symbol	Value
Preassigned Parameters		
Discount Factor	β	0.995
Worker Bargaining Weight	α	0.300
Production function, scale	f_1	1.562
Calibrated Parameters		
Complementarity	ρ	0.644
Worker dist. location	μ_x	-0.252
Worker dist. scale	σ_x	0.709
Meeting rate workers	λ_w	0.166
Job-to-job meeting rate	λ_e	0.024
Job destruction rate	d	0.015
Job creation cost, scale	c_0	19.197
Job creation cost, convexity	c_1	1.101
Firm shocks, frequency	ϕ	0.035
Firm shocks, range	$\bar{y}$	0.140

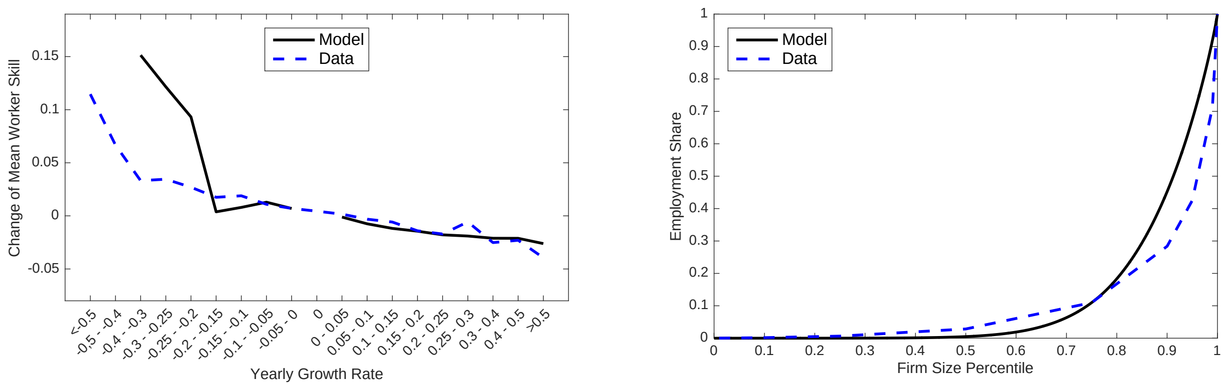
Notes: Confidence intervals on ρ are given by [0.616,0.677]. See text for explanation.

5.2 Sorting Patterns in the Labor Market

Figure 5 presents the estimated sorting patterns in Germany. The left panel plots a heatmap of $\psi(x, y)$, the equilibrium distribution of employed workers across firm types. Brighter colors represent higher densities. This distribution is driven by three forces. First, by the distribution of jobs across firm types, the distribution of workers across worker types and last the sorting pattern between the two types. The most evident pattern is that most employment is concentrated at the most productive establishments. Over 90 percent of workers are employed by firms above median productivity. The log-normal shape of the worker distribution is also recognizable, with its humped-shaped form. The sorting of worker types across firm types is not very pronounced. The overall correlation between firm and worker types is -0.077. Only for the most productive workers we see a pronounced impact of negative sorting, as these workers sort towards lower productivity firms compared to low quality workers. These findings are mirrored in the right panel of figure 5. It shows the cumulative distribution of workers conditional on firm type, which helps to better understand

Figure 4: Model Fit

(a) Reorganization of Worker Quality (b) Firm Size Distribution



Notes: The left panel compares the relationship between firm growth rates and percentage changes in average worker skills in the model and the Data. It is obtained by estimating equation (27) with firm growth rate bins. The relationship is captured very well, except for very fast shrinking firms. These firms constitute only a small fraction of the sample, amounting to less than 0.33 percent of all firms in the regression sample.
The right panel plots the share of total employment by firm size percentile.

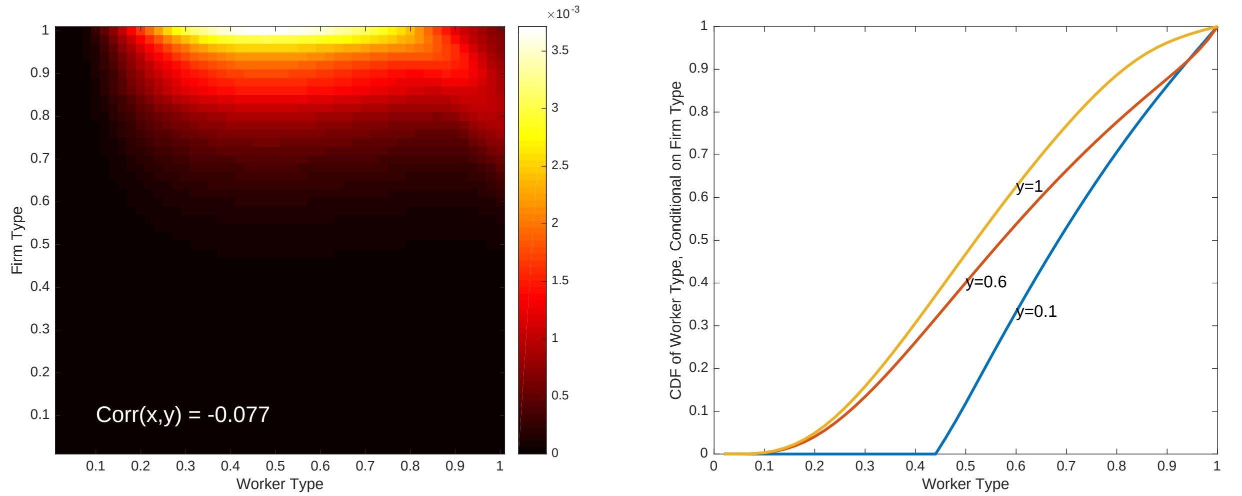
the matching patterns. The distribution of worker quality at lower type firms stochastically dominates the ones of at more productive firms. In high productivity firms a high proportion of their employees have low skills, whereas the labor force of low type firms consists mostly of high skilled workers. This is also reflected by the fact that median worker quality decreases monotonically with firm productivity.

How does this sorting pattern emerge in the labor market? Agents have two tools to optimize the quality of their matches. First, they decide how "picky" to be with respect to the quality of their partners. This is represented by the equilibrium matching sets characterized in figure 6a. Second, workers can also engage in on-the-job search to improve the quality of their matches. The probability of a worker quitting to another firm is displayed in the heatmap of job-to-job quit rates by worker and firm type in the right panel of figure 6. As before, lighter areas represent worker firm type combinations with high levels of quit rates, whereas the dark regions feature high retention rates of employees.

The sorting pattern in low productivity firms is mostly driven by choice of matching sets. They only match with workers above a certain skill level, which also can be seen in the conditional cdfs in figure 6b. Because of their low productivity, these firms are unattractive to prospective employees and unsuccessful at retaining current employees. Higher-type firms are willing to match with a broader set of agents and poach more often.

Figure 5: Sorting Pattern

(a) Employment Distribution $\psi(x, y)$ (b) CDF of Worker Distribution by Firm Type



Notes: The left figure plots the estimated equilibrium distribution of matched worker types across firm types $\psi(x, y)$. The right plot presents the cdf of worker type distributions conditional on firm types.

With the model solution at hand, I can turn to study the sources of wage variation, which I discuss in the following subsection.

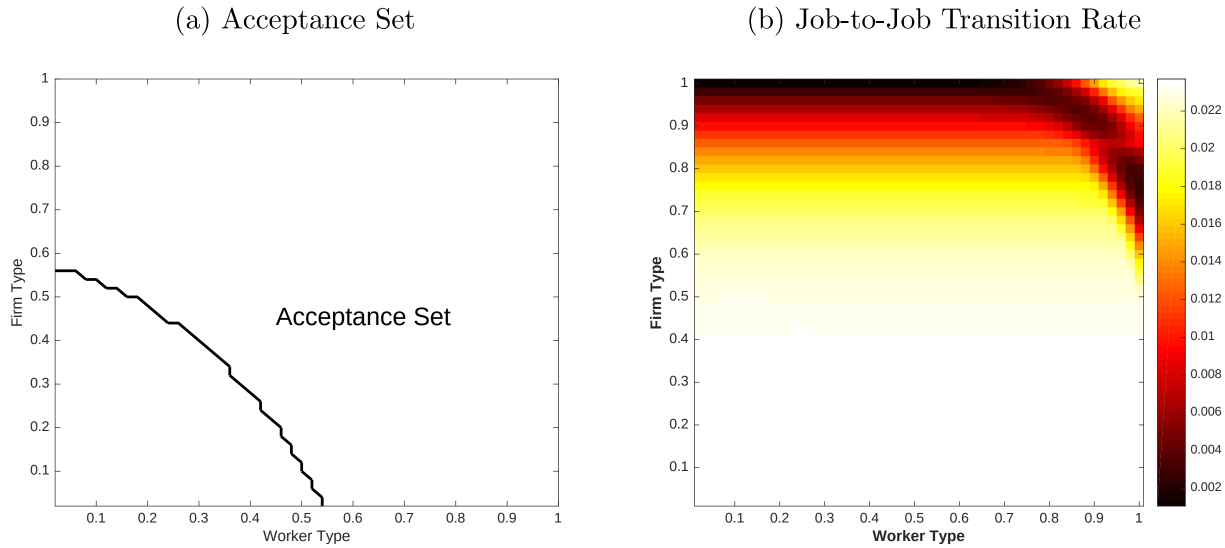
5.3 Sources of Wage Variation and Output Loss due to Mismatch

To understand the driving forces behind wage variation, we not only have to understand which factors determine wages, but also the underlying empirical distribution of those factors. In my framework, wages depend on worker skills, firm productivity, and the bargaining position of workers. These three determinants are not directly observable in the German data, but are readily observed in a simulated panel dataset obtained through the structural model. I simulate wages for 100,000 workers across 2,500 establishments over 50 years. As in the structural estimation, the first 32 years are burned in, and the wage variation is computed using the remaining 18 years.

$$\begin{aligned} V(w) &= E_x[V(w|x)] + V_x(E[w|x]) \\ &= \underbrace{E_x\left[E_y\left[V(w|x,y)\right]\right]}_{\text{Bargaining}} + \underbrace{E_x\left[V_y\left(E[w|x,y]\right)\right]}_{\text{Firms}} + \underbrace{V_x(E[w|x])}_{\text{Workers}}. \end{aligned} \tag{28}$$

Using the simulated wage data, I first consider a statistical within/between group wage

Figure 6: Matching Decisions



Notes: The left figure plots the acceptance policies of firms and workers. Viable matches lie to the north-east of the downward sloping frontier. The right plot presents a heatmap of the job-to-job transition rates by firm and worker type.

variance decomposition. Equation (28) shows the decomposition. In the first line, I decompose wage variation into within and between worker types, represented by the left and right terms, respectively. The within worker type wage variation can be further decomposed into variation that is originating from between and within firm types. The resulting decomposition has three terms. Identical workers employed by the same firm earn different wages because they hold differential bargaining positions. These originate from wage increases through job-to-job transitions and that past firm productivity levels manifest themselves in current wages through wage rigidity. This is captured by the first term of the second line labelled bargaining. The other two terms in equation (28) capture the wage variation between firms and between workers.

Table 4 shows this decomposition. Almost 74 percent of wage variation in the estimated model is between, and 26 percent within worker types. These 26 percent can be further decomposed into wage variation due to workers working at different firms and holding differential bargaining positions at wage negotiations. The differential bargaining positions explain 18 percent of total wage variation, whereas wages across different firm types contribute to wage inequality by 8 percent. The table also shows the breakdown between firm effects and bargaining positions conditioning on the four worker type quartiles. As we consider higher type workers, the variation in bargaining positions becomes increasingly important, whereas

Table 4: Variance Decomposition

	All Workers	Conditional on Worker Type Quartile			
		25%	50%	75%	100%
Worker Effect	73.5				
Firm Effect	8.2	33.1	32.4	31.4	27.2
Bargaining	18.2	66.9	67.6	68.6	72.8
Total	100.0	100.0	100.0	100.0	100.0

Notes: Between and within group decomposition of log-wages from model simulated panel dataset.

the contribution of firms decreases. This is due to the fact that higher type workers work on average for a smaller set of firms. Thus, differential bargaining positions play a bigger role in wage variation.

Although this statistical decomposition is suggestive of the underlying forces, it does not quantify the true contribution of heterogeneity in terms of model primitives. Consider the between worker wage variation as an example. This is not only driven by the underlying skill heterogeneity across workers, but also by their average bargaining positions and differential matching patterns across firm types. My findings of negative sorting suggest that high skill workers are employed by lower type firms, which might dampen wage variation between worker types. In addition, these numbers do not measure counter-factual outcomes without one of the forces at work. In a counterfactual economy where firms would not be heterogeneous for example, the labor market might price some skills differently than with firm heterogeneity. Therefore, to study the economic driving forces of wage dispersion, I use the estimated structural model as a laboratory.

To quantify the true contributions of worker and firm heterogeneity, differential bargaining positions, and the complementarities in production which induce sorting, I will recompute my model with the estimated parameters from section 5, shutting down particular channels at a time. The results are presented in table 5.[31] In the first column I consider a counterfactual economy with only firm heterogeneity. This model neither features worker heterogeneity, nor differential bargaining positions[32] and the complementarity parameter ρ is set to one. Firm heterogeneity alone explains about 20 percent of the wage variation found in the estimated

[31] I consider the standard deviation rather than the variance because average wages slightly adjust in the counterfactuals.

[32] I assume that all wages are bargained with the value of unemployment as outside option.

complete model. The next column adds differential bargaining positions and their inclusion more than doubles the standard deviation of wages. To compute their marginal contribution, I divide the marginal increase in standard deviation by the total wage variation of the baseline model. This yields a marginal contribution of about 32 percent. The third column represents the full model except that ρ is still one, which induces no sorting in equilibrium. Worker heterogeneity explains the largest fraction of wage variation, it alone contributes 71 percent of wage variation. Finally, the last column shows the results for the full model with all the features and parameter values from the structural estimation. The estimated complementarity parameter ρ in the baseline economy dampens wage variation significantly. In comparison to the economy with no sorting, wage dispersion decreases by 22.7 percent. The estimated $\rho < 1$ induces negative sorting, which dampens wage variation because low type workers are on average employed by higher type firms. These firms also pay a wage premium, since wages are bargained and higher type firms have a higher opportunity cost of waiting.

My estimated contributions of worker and firm heterogeneity largely echo findings in other studies that worker heterogeneity explains the largest part of wage dispersion and firm heterogeneity plays an important role as well.[33] The estimated effect of differential bargaining positions lies in the middle range of numbers previously reported.[34] The strong negative contribution of sorting is in contrast to previous findings. Although a number of studies have found negative sorting between estimated worker and firm types, they all report limited impacts on wage dispersion.[35] The reason behind this discrepancy lies in the misspecification of wages in the AKM approach, as I discuss in the next couple of paragraphs.

To better understand the contribution of my structural approach, I contrast my findings with results obtained from AKM on the same simulated wage data. The AKM approach assumes that log-wages can be decomposed additively into a worker and firm fixed effect, i.e. log wages for an individual i working for a firm j are given by

$$\log(w_{ij}) = \alpha_i + \psi_j + \epsilon_{ij}. \tag{29}$$

Table 6 presents the results of the AKM approach. The top panel shows the wage vari-

[33] Abowd et al. (1999); Bagger and Lentz (2015); Card et al. (2013)

[34] Postel-Vinay and Robin (2002) estimate its contributions to be higher in a model without sorting, and Bagger and Lentz (2015) find slightly lower contributions

[35] Abowd et al. (1999), Abowd et al. (2004), Andrews et al. (2008) and Woodcock (2011)

Table 5: Sources of Wage Dispersion

	Only firm heterogeneity	+bargaining positions	+ Worker heterogeneity	+Sorting $\rho = 0.644$
Standard deviation	0.071	0.186	0.444	0.362
Percentage Contribution	19.6	31.7	71.3	-22.7

Notes: The table presents the standard deviations in counterfactual economies. The first three economies feature no sorting, i.e. $\rho = 1$. The last column represents the full structural model. The last row measures the marginal contribution of each source. This row does not sum to 100 because of rounding.

Table 6: AKM Regression

Variance Decomposition				
Worker heterogeneity	Firm heterogeneity	Sorting	Residual	Total
82.6	4.3	-3.2	16.3	100.0

Type Correlations	
$Corr(x^{AKM}, x)$	$Corr(y^{AKM}, y)$
0.958	0.676

Notes: The table presents the variance decomposition based on AKM and the correlations between the AKM estimated worker and firm types and the true underlying types using model simulated wage data.

ance decomposition based on the estimated fixed effects.[36] This yields significantly different results than the decomposition based on counterfactuals. First, this is due to the different nature of the AKM decomposition: It is a statistical decomposition, whereas my structural decomposition takes into account general equilibrium effects when shutting down particular sources of wage dispersion.

Second, the additive specification of the wage equation (29) in AKM rules out complementarities between worker and firm types. As discussed earlier, in frameworks where sorting

[36]The log-wage variance decomposition is given by:

$$Var(log(w)) = \underbrace{Var(\hat{\alpha}_i)}_{\text{Worker}} + \underbrace{Var(\hat{\psi}_j)}_{\text{Firm}} + \underbrace{2Cov(\hat{\alpha}_i, \hat{\psi}_j)}_{\text{Sorting}} + \underbrace{Var(\hat{\epsilon})}_{\text{Residual}}, \tag{30}$$

where the hatted variables denote the estimated fixed effects and predicted residuals.

is driven by complementarities in production, this leads to a misidentification of firm types. This can be seen in the reported low correlation between the estimated firm fixed effects and the true firm type at the wage bargaining of 0.676. Worker types on the other hand are estimated relatively precisely by the AKM approach, confirming the findings in Lopes de Melo (2013).

To study the macroeconomic impact of mismatch, I conduct the following counter-factual experiment: I reshuffle workers across firms according to the frictionless allocation in Becker (1973). By comparing aggregate production in this counterfactual economy with total output in my model one can gauge the importance of mismatch. I find that aggregate output would only increase slightly by 0.6 percent if all mismatch would be eliminated. This implies that the labor market in Germany is flexible enough to eliminate the most severe mismatch.

6 Conclusion

Which factors explain wage inequality among observationally similar workers? To answer this question, I estimate a structural model of the labor market with German matched employer-employee data. In order to correctly understand the sources of wage dispersion, I have to identify the underlying complementarities in output between heterogeneous workers and firms. The introduction of firm dynamics into a labor search model with sorting allows me to tackle the identification of the sorting patterns from a new angle. I study how firms reorganize the quality of their workforce in response to shocks. German establishments reorganize the composition of their workforce towards higher skilled workers when they shrink and expand by lowering the average quality of their workers. This reveals that higher type workers are relatively more valued at lower productivity firms, and conversely low skilled workers are relatively higher valued at low type firms. Intuitively, but also in the structural estimation this implies that worker skills and firm productivity are substitutes in production. This induces negative sorting in the labor market, with an estimated correlation coefficient of -0.077 between worker and firm types.

I then perform a number of counterfactuals using my structural model to decompose the sources of wage variation stemming from worker and firm heterogeneity, sorting and workers' bargaining positions. Adding one channel at a time reveals that worker heterogeneity contributes with 71 percent the most to wage inequality. Differential bargaining positions and firm heterogeneity explain another 32 and 20 percent, respectively. The estimated complementarity in production which induces negative sorting dampens wage variation significantly.

The counterfactual economy without sorting features a 23 percent higher wage variation.

A comparison with the AKM approach on model simulated data reveals significant biases of the fixed effect approach. The misspecification of wages in AKM leads to an underprediction of the contribution of sorting to wage variation.

References

J Abowd, Francis Kramarz, Paul Lengermann, and Sébastien Pérez-Duarte. Are good workers employed by good firms? a test of a simple assortative matching model for france and the united states. 2004.

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

John M Abowd, Francis Kramarz, Sébastien Pérez-Duarte, and Ian M Schmutte. Sorting between and within industries: A testable model of assortative matching. Technical report, National Bureau of Economic Research, 2014.

Martyn J Andrews, Len Gill, Thorsten Schank, and Richard Upward. High wage workers and low wage firms: negative assortative matching or limited mobility bias? Journal of the Royal Statistical Society: Series A (Statistics in Society), 171(3):673–697, 2008.

Jesper Bagger and Rasmus Lentz. An empirical model of wage dispersion with sorting. Technical report, National Bureau of Economic Research, 2015.

Cristian Bartolucci, Francesco Devicienti, and Ignacio Monzon. Identifying sorting in practice. Technical report, 2015.

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

Stéphane Bonhomme, Thibaut Lamadon, and Elena Manresa. A distributional framework for matched employer employee data. Technical report, Working Paper, 2016.

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. 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. Technical report, National Bureau of Economic Research, 2010.

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.

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.

Shigeru Fujita and Giuseppe Moscarini. Recall and unemployment. Unpublished Manuscript, Yale University, 2012.

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

Marcus Hagedorn, Hann Law Tzuo, and Iourii Manovskii. Identifying equilibrium models of labor market sorting. 2014.

Nikolaus Hansen and Stefan Kern. Evaluating the cma evolution strategy on multimodal test functions. In International Conference on Parallel Problem Solving from Nature, pages 282–291. Springer, 2004.

Joerg Heining, Wolfram Klosterhuber, and Stefan Seth. Linked-employer-employee-daten des iab: Liab-lngsschnittmodell 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.

Susana Iranzo, Fabiano Schivardi, and Elisa Tosetti. Skill dispersion and firm productivity: An analysis with employer-employee matched data. Journal of Labor Economics, 26(2): 247–285, 2008.

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. 2014.

Kory Kantenga and Tzuo-Hann Law. Sorting and wage inequality. 2014.

Thibaut Lamadon. Productivity shocks, dynamic contracts and income uncertainty. 2014.

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. European University Institute WP, 2014.

Jeremy Lise and Fabien Postel-Vinay. Multidimensional skills, sorting, and human capital accumulation. Technical report, Mimeo, 2014.

Jeremy Lise and Jean-Marc Robin. The macro-dynamics of sorting between workers and firms. 2014.

Jeremy Lise, Costas Meghir, and Jean-Marc Robin. Mismatch, sorting and wage dynamics. 2015.

Rafael Lopes de Melo. Firm wage differentials and labor market sorting: Reconciling theory and evidence. Unpublished Manuscript, 2013.

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.

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. Technical report, National Bureau of Economic Research, 2015.

Sharon Traiberman. Occupations and import competition evidence from danish matched employee-employer data. 2016.

Simon D Woodcock. Match effects. unpublished manuscript, 2011.

A Appendix - Proofs

Here I proof the following proposition:

Proposition. *U(x) and V(y) are increasing in their arguments.*

This will hold because of the assumption that higher types always have an absolute advantage in production over lower types, i.e. $f_x(x, y) > 0$ and $f_y(x, y) > 0$. Consider two types of agents with $t_1 < t_2$. It must be the case that agent t_2 can achieve at least the utility level of t_1. This is because t_2 could just follow the acceptance and wage strategies of t_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 t_1 agents, they will also be willing to hire t_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 t_1 firms, they will also be willing to match with t_2 because wages and separation probabilities are the same by construction. Thus, t_2 agents will always have weakly higher payoffs as t_1 agents.

If I restrict my model, this can also be shown using the surplus functions. Consider $V(y)$ and assume when firms are hit by productivity shocks, they draw from $U[0, 1]$ instead of drawing from $U[y - \bar{y}, y + \bar{y}]$. In this case that the partial of $S(x, y)$ with respect to y is given by:

$$\frac{\partial S(x,y)}{\partial y} = f_y(x,y) + \beta(1-d)(1-\phi)\frac{\partial S(x,y)}{\partial y} - (V'(y)(1-\beta(1-d)(1-\phi)),$$

where ϕ is the probability of a productivity shock.

Deriving equation (16) with respect to y yields:

$$\begin{aligned} V'(y) &= \beta(1-d)(1-\phi)V'(y) + \lambda_f p^u(1-\alpha)(1-\phi)\int A^U(x,y)\frac{\partial S(x,y)}{\partial y}\frac{\mu_x(x)}{u}dx \\ &\quad +\lambda_f(1-\phi)\int\int A^E(x,y,\tilde{y})\frac{\partial}{\partial y}S(x,y)\frac{\psi(x,\tilde{y})}{e^S}dxd\tilde{y}. \end{aligned}$$

Substitution the expression for $\frac{\partial S(x,y)}{\partial y}$ yields:

$$\begin{aligned} V'(y) &= \beta(1-d)(1-\phi)V'(y) \\ &+\lambda_f p^u(1-\alpha)(1-\phi)\int A^U(x,y)\left(f_y(x,y)-V'(y)(1-\beta(1-d)(1-\phi))\right)\frac{\mu_x(x)}{u}dx \\ &+\lambda_f(1-\phi)\int\int A^E(x,y,\tilde{y})\left(f_y(x,y)-V'(y)(1-\beta(1-d)(1-\phi))\right)\frac{\psi(x,\tilde{y})}{e^S}dxd\tilde{y}. \end{aligned}$$

Collecting $V'(y)$ on the left hand side yields that $V'(y)>0$ since $f_y(x,y)>0$.

Let me now consider $U(x)$. In the special case of no firm shocks, taking the derivative of $S(x,y)$ with respect to x yields:

$$\frac{\partial}{\partial x}S(x,y) = \frac{\partial}{\partial x}f(x,y)+\beta(1-d)\frac{\partial}{\partial x}S(x,y)-(1-\beta)U'(x).$$

Rearraging yields:

$$\frac{\partial}{\partial x}S(x,y)\left(1-\beta(1-d)\right) = \frac{\partial}{\partial x}f(x,y)-(1-\beta)U'(x).$$

Instead of using the indicator function $A^U(x,y)$, I can rewrite equation (18) with the upper and lower matching bounds denoted by $a(x)$ and $b(x)$ as:

$$U(x)(1-\beta)=\alpha\beta\lambda_w\int_{a(x)}^{b(x)}S(x,y)\frac{\mu_y(y)}{V}dy.$$

Taking the derivative with respect to x yields

$$\begin{aligned} U'(x)\frac{(1-\beta)}{\alpha\beta\lambda_w} &= \int_{a(x)}^{b(x)}\frac{\partial S(x,y)}{\partial x}\frac{\mu_y(y)}{V}dy \\ &+b'(x)\left(S(x,b(x))\frac{\mu_y(b(x))}{V}\right)-a'(x)\left(S(x,a(x))\frac{\mu_y(a(x))}{V}\right) \end{aligned} \tag{31}$$

The second line of the equation (31) is equal to zero. At the interior boundaries of the matching sets we know that the surplus is zero, i.e. $S(x,b(x))=S(x,a(x))=0$. On the other hand, at the limits of the supports of the the agents types, the boundaries do not change, i.e. $a'(0)=b'(1)=0$.

Plugging in for $\frac{\partial S(x,y)}{\partial x}$ yields that $U'(x)>0$ since $\frac{f_x(x,y)}{\partial x}>0$.

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 50 worker and 50 firm types. The solution algorithm is the following iterative process:

1. Guess $S^0(x,y)$, $\psi^0(x,y)$, $\mu_x^0(x)$ and $\mu_y^0(y)$
2. Update $S^{i+1}(x,y)$ using equation (20)
3. Using the new value of $S(x,y)$, update acceptance policies $A^U(x,y)$ and $A^E(x,y,\tilde{y})$. It helps the convergence if one updates the indicator functions slowly.
4. Update the distributions $\psi(x,y)$, $\mu_x(x)$ and $\mu_y(y)$ using the updated acceptance policies. The distributions are updated by using the law of motion equations (22), (24) and (21).
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 $S(x,y)$ 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:

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

These randomizations only affect a tiny fraction of the state space. With the estimated parameters form section 5, only around 5 percent of all possible $A^E(x,y,\tilde{y})$ and no $A^U(x,y)$

are deviating from 0 or 1 by more than 10^{-6}. Similar smoothing strategies have been applied by Lopes de Melo (2013) and Hagedorn et al. (2014).

After obtaining the equilibrium solutions to value functions, acceptance rules and steady state distributions I simulate the evolution of 2500 firms over 600 months. I use the stationary distribution as initial conditions. 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 "covariance matrix adaptation evolution strategy" (CMA-ES) minimization procedure, which is well suited for highly non-linear and non-smooth minimization problems, for details see Hansen and Kern (2004). I use the Matlab code provided by the authors.

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[37] and below the upper earnings limit for statutory pension insurance are not reported. In

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

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.[38] 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.[39]

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:

$$e_{it} = \alpha_i + \beta X_{it} + \epsilon_{it}. \tag{32}$$

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 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 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 over all possible values of firm's

[38] 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.

[39] This is in line with recent evidence shown in Fujita and Moscarini (2012) and Nekoei and Weber (2015)

productivity next period and over all possible workers it might meet.

$$
\begin{aligned}
V(y) \quad &= \beta(1-d) \int_{y_{\min}}^{y_{\max}} ((1-\lambda_f)V(y') \\
&+\lambda_f \left(p^u \int_{x_{\min}}^{x_{\max}} \left(A^U(x,y')J(x,y',w^U(x,y')) + (1-A^U(x,y'))V(y')\right) \frac{\mu_x(x)}{u} dx + \right. \\
&\left. + \ (1-p^u) \int_{\tilde{y}} \int_{x} \left(A^E(x,y',\tilde{y})J(x,y',w^E(x,\tilde{y})) + (1-A^E(x,y',\tilde{y}))V(y')\right) \frac{\psi^S(x,\tilde{y})}{e^s} dx d\tilde{y} \right) p(y'|y) dy'
\end{aligned}
$$

A filled job produces a flow value of $f(x, y)$. If the match is not destroyed and the worker is not poached away, the firm receives a continuation value of $J^C(x, y', w)$. The continuation value will depend on whether the wage has to be renegotiated or not.

$$
\begin{aligned}
J(x,y,w) =& f(x,y) - w + \beta(1-d) \left(\int_{y_{\min}}^{y_{\max}} \left(1 - A^U(x,y')\right) V(y') \right. \\
&+ A^U(x,y') \left((1-\lambda_e) J^C(x,y',w) \right. \\
&\left.\left. + \ \lambda_e \int_{y_{\min}}^{y_{\max}} \left(\left(1 - A^E(x,y',\tilde{y})\right) J^C(x,y',w) + A^E(x,y',\tilde{y})V(y')\right) \right) p(y'|y) \frac{\mu_y(\tilde{y})}{V} d\tilde{y} dy' \right)
\end{aligned} \tag{33}
$$

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

$$
\begin{aligned}
U(x) \quad = \quad \beta & \left(\lambda_w \int_{y_{\min}}^{y_{\max}} \left(A^U(x,y)W(x,y,w^U(x,y)) + (1-A^U(x,y))U(x)\right) \frac{\mu_y(y)}{V} dy \right. \\
&\left. + (1-\lambda_w)U(x) \right)
\end{aligned} \tag{34}
$$

Workers receive the negotiated wage w this period. Next period, they either experience a separation, a job-to-job transition or continue to stay at the current job, which value is denoted by $W^C(x, y', w)$. Similar to firms, this continuation value depends on whether the wage will be renegotiated.

$$
\begin{aligned}
W(x,y,w) &= w + \beta \left(dU(x) + (1-d) \int_{y_{\min}}^{y_{\max}} A^U(x,y') \right. \\
&\left(\lambda_e \int_{y_{\min}}^{y_{\max}} \left(A^E(x,y',\tilde{y}) W(x,y',w) + (1 - A^E(x,y',\tilde{y})) W^C(x,y',w)\right) \frac{\mu_y(\tilde{y})}{V} d\tilde{y} \right. \\
&\left. \left. + \ (1-\lambda_e) W^C(x,y',w) \right) + (1 - A^U(x,y')) U(x) p(y'|y) dy' \right) \qquad (35)
\end{aligned}
$$

The continuation value for workers and firms $W^C(x,y,w)$, $J^C(x,y,w)$ in case no separation happens depends on whether a renegotiation of the wage contract is triggered. There are three possibilities. If none of the parties have a credible threat to end the relationship (i.e. neither $W(x,y',w) - U(x) < 0$, nor $J(x,y',w) - V(y') < 0$), the wage remains constant and the continuation value is $W(x,y',w)$ and $J(x,y',w)$. On the other hand, if either, the current wage w becomes unsustainably high for the firm ($A^F(x,y',w) = 1$) or to low to satisfy the worker's participation constrained, then the wage is renegotiated to either $w^{NF}(x,y')$ or $w^{NF}(x,y')$, depending on who triggers the renegotiation. Thus,

$$
\begin{aligned}
W^C(x,y',w) &= A^{NW}(x,y',w) W(x,y',w^{NW}(x,y')) + A^{NF}(x,y',w) W(x,y',w^{NF}(x,y')) \\
&\quad + \left(1 - A^{NW}(x,y',w) - A^{NF}(x,y',w)\right) W(x,y',w)
\end{aligned}
$$

$$
\begin{aligned}
J^C(x,y',w) &= A^{NW}(x,y',w) J(x,y',w^{NW}(x,y')) + A^{NF}(x,y',w) J(x,y',w^{NF}(x,y')) \\
&\quad + \left(1 - A^{NW}(x,y',w) - A^{NF}(x,y',w)\right) J(x,y',w)
\end{aligned}
$$

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:

$$
\begin{aligned}
S(x,y) \;=\;& f(x,y)+\beta(1-d)\left(\int\limits_{y_{\min}}^{y_{\max}} A^U(x,y')S(x,y')p(y'|y)dy'\right) \\
& -\beta\left(\alpha\lambda_w \int\limits_{y_{\min}}^{y_{\max}} A^U(x,y)S(x,y)\frac{\mu_y(y)}{V}dy\right) \\
& -\beta(1-d)\int\limits_{y_{\min}}^{y_{\max}} \lambda_f \left(p^u \int\limits_{x_{\min}}^{x_{\max}} \left(A^U(x,y')(1-\alpha)S(x,y')\right)\frac{\mu_x(x)}{u}dx+\right. \\
& \left. +\ (1-p^u)\int\limits_{y_{\min}}^{y_{\max}}\int\limits_{x_{\min}}^{x_{\max}} A^E(x,y',\tilde{y})\left(S(x,y')-S(x,\tilde{y})\right)\frac{\psi^S(x,\tilde{y})}{e^s}dxd\tilde{y}\right) p(y'|y)dy'
\end{aligned}
$$

E Firm Level Growth Regressions

Table 7: Regression Results

	$\Delta_{\%}\overline{Wquality}_{jt} = \alpha + \gamma growth_{jt} + \delta X_{jt} + \epsilon_{jt}$								
growth	-0.099	-0.100	-0.079	-0.061	-0.062	-0.100	-0.099	-0.084	-0.077
	0.016	0.016	0.021	0.015	0.013	0.019	0.017	0.009	0.01
Controls:									
Industry	x	x		x	x	x	x	x	x
Year	x	x	x	x	x	x	x	x	x
Industry x Year	x	x		x	x	x	x	x	x
Size		x							
Age		x							
Firm FE			x						
Sample	Baseline	Baseline	Baseline	Size>190	Age>15	Year<2004	Year≥2004	3 Yr. Chg.	5 Yr. Chg.
N	19981	19981	19981	6437	10060	9985	9996	15590	11756
Adj. R^2	0.380	0.383	0.076	0.573	0.650	0.222	0.526	0.347	0.271

Notes: Establishment level regressions of yearly firm growth rates on percentage change in average worker fixed effects. Regressions are weighted by establishment size. Standard errors are clustered at the 3 digit industry level. Regressions with establishment size and age (column x and y) include both a linear and a quadratic term. The baseline sample corresponds to establishments with more than 30 employees and growth rates between -0.75 and 0.75. See text for detailed explanation.

Table 8: Worker Quality Adjustments by Industry

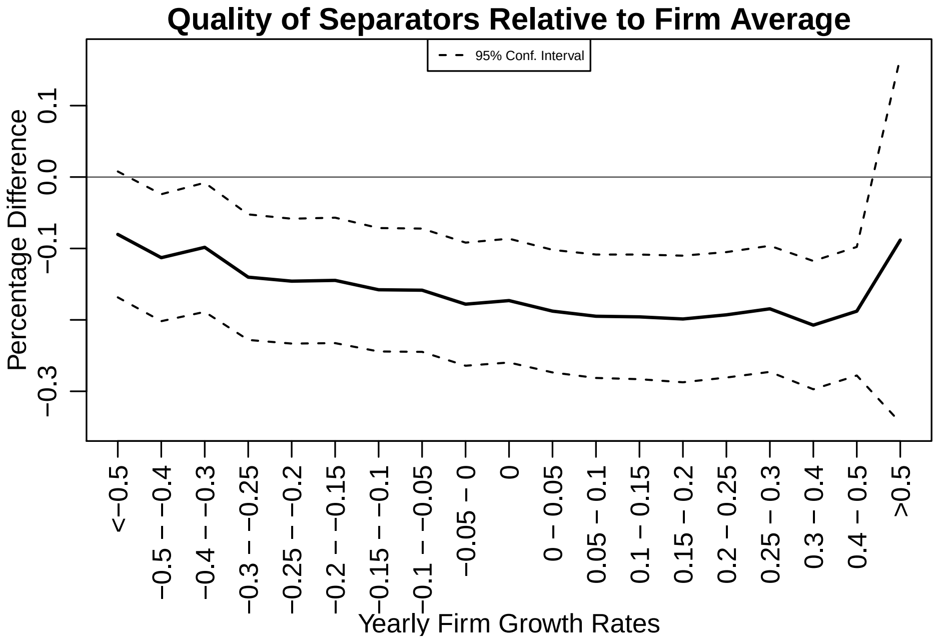
Industry	Point Estimate	Standard Error
Agriculture, hunting, forestry	-0.169	0.034
Mining, quarrying	-0.062	0.025
Manufacturing	-0.097	0.008
Construction, electricity, water and gas supply	-0.071	0.013
Wholesale & retail, hotels	-0.085	0.025
Transport, communications, financial services	-0.059	0.011
Real Estate, renting, business activities	0.076	0.121
Education	-0.063	0.013
Other community, social, personal service	-0.202	0.022

Notes: Slope coefficients and standard errors from regression equation (27) by 1-digit industry. Industry classifications follow WZ93.[40] Sample restrictions are the same as in table 7.

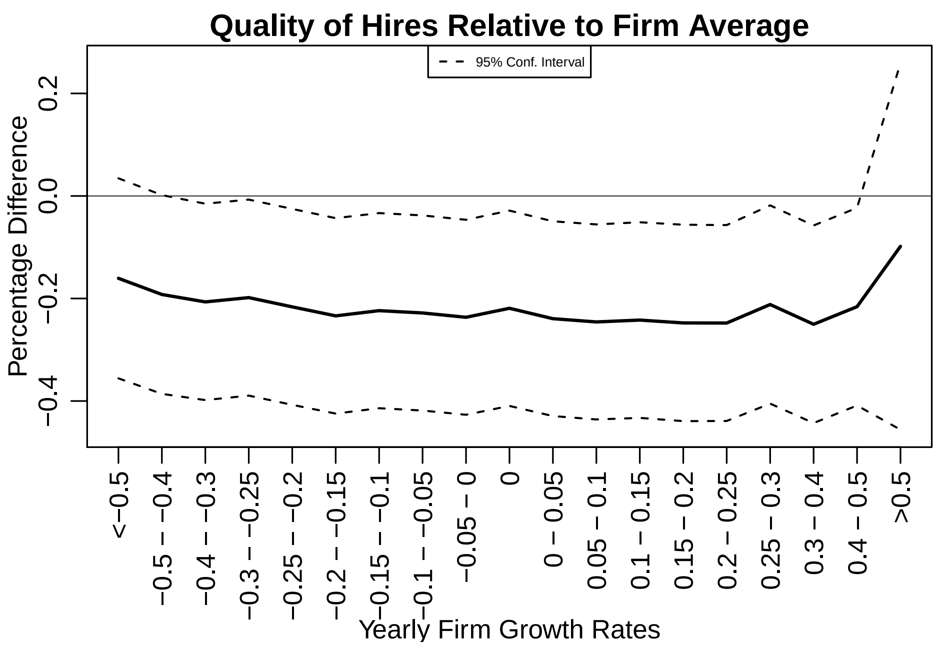
[40] https://www.destatis.de/DE/Methoden/Klassifikationen/GueterWirtschaftklassifikationen/klassifikationwz93englisch.pdf;jsessionid=BABDB27FF6747733D661FE86D0796687.cae2?__blob=publicationFile

Figure 7: Reorganization of Worker Quality

(a) Separations



(b) Hires



Notes: The figures show the percentage difference between the average worker type separating (top panel) and joining (bottom panel) the firm relative to the average at the beginning of the period by establishment growth rates, controlling for year, 3-digit industry and interaction of year/industry effects. The sample consists of all establishments with size≥30. Estimates are weighted by employment and standard errors are clustered at the 3-digit industry level. Broken lines indicate 95% confidence intervals. The time window is annual.