

Firm Productivity, Wages, and Sorting*

Benjamin Lochner¹ and Bastian Schulz²

¹ University of Erlangen-Nuremberg and Institute for Employment Research (IAB)

² Aarhus University, the Dale T. Mortensen Centre, and CESifo

March 15, 2021

Abstract

We study the link between firms' productivity and the wages firms pay. Guided by labor market sorting theory, we infer firm productivity from estimating firm-level production functions, taking into account that worker ability and firm productivity may interact at the match level. Using German data, we find that high wages are not necessarily a reflection of high firm productivity. Observed worker transitions towards higher wages are sometimes directed downwards on the firm-productivity ladder. Worker sorting into high-productivity firms is thus less pronounced than sorting into high-wage firms. Consequently, an implication of increasing wage sorting could be decreasing allocative efficiency.

Keywords: Assortative Matching, Labor Market Sorting, Wage Inequality, Job Mobility, Unobserved Heterogeneity, Firm Productivity, Production Function Estimation

JEL Classifications: J24, J31, J40, J62, J64, L25

*This paper is a substantially revised version of CESifo Working Paper No. 6066 (August 2016), which has been circulated under the title "Labor Market Sorting in Germany". First draft: February 2016. We thank John Abowd, Michele Battisti, Anja Bauer, Stéphane Bonhomme, Carlos Carillo-Tudela, Jason Faberman, Bernd Fitzenberger, Christian Hellwig, Timo Hener, Christian Holzner, Leo Kaas, Philipp Kircher, Patrick Kline, Thibaut Lamadon, Rasmus Lentz, Jonas Maibom, Christian Merkl, Annaïg Morin, Andreas Mueller, Helmut Rainer, Jean-Marc Robin, Johannes Schmieder, Heiko Stüber, Coen Teulings, Robert Ulbricht, Rune Vejlin, and six anonymous referees for feedback and many helpful suggestions. The input from seminar and conference audiences at Aarhus, Cambridge, CBS, the ECB, Glasgow, Nuremberg, LMU Munich, Regensburg, TSE, the ifo Institute, the IAB, SMYE 2016, Sam 2016, ES-NASM 2016, SED 2016, AEA/ASSA 2017, the DFG SPP 1764 conference 2017, SOLE 2017, and ES-WC 2020 has greatly improved the paper. We are indebted to Stefan Seth for helping us with data access. All errors are our own. Benjamin Lochner gratefully acknowledges financial support from the Graduate Program of the IAB and the School of Business and Economics of the University of Erlangen-Nuremberg (GradAB) as well as the Emerging Talents Initiative (ETI) of the University of Erlangen-Nuremberg. Bastian Schulz thanks the ifo Institute, the Dale T. Mortensen Centre, and the Aarhus University Research Foundation, particularly the Associate Professor Starting Grant awarded to Rune Vejlin, for financial support.

Corresponding author: Bastian Schulz, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark. Tel.: +4542836363. Email: bastian.schulz@econ.au.dk.

1 Introduction

The firms' contribution to wage inequality has been of interest for a long time.¹ Using the Abowd et al. (1999) (AKM) approach, Song et al. (2019) and Card et al. (2013) (CHK) show that increasing wage inequality is to a large extent driven by increasing sorting of high-wage workers into high-wage firms. However, we do not know much about what characterizes these high-wage firms. Are they also the most productive ones?

Presumably, the wages that firms pay are related to their productivity, although the mapping from productivity to wages is far from obvious. Highly productive firms may share high output with their workers in the form of high wages. But firms might also reduce wages by exploiting labor market imperfections.² Moreover, wages could be relatively low in high-productivity firms due to a high level of amenities. In turn, relatively unproductive (perhaps young) firms may have to pay high wages to retain workers or to expand their workforce if workers have high outside options. To link productivity and wages, we build upon the theory of labor market sorting. It highlights complementarities between worker ability and firm productivity as an important determinant of wages and the source of positive sorting in the data (Becker, 1973; Shimer and Smith, 2000; Eeckhout and Kircher, 2018).

The primary contribution of this paper is to shed light on the link between firms' productivity and the wages they pay. To this end, we estimate firm-level production functions to infer unobserved firm productivity from the data. Importantly, we rely on the wage determination mechanism of a sorting model with multi-worker firms, intra-firm bargaining, and matching frictions to ensure that our productivity estimation is consistent with sorting theory. In the model, the wages that firms pay include a match-specific component, which reflects that a given worker type may contribute a lot to output at some firm types but little at others in the presence of worker-firm complementarities. We measure this match-specific wage component using the AKM model and predict the model-consistent labor input for the production function to be estimated.

Our estimation of firm productivity proceeds in two steps. First, we estimate an AKM model on German matched employer-employee registers to decompose wages. We find that most of the wage variation left unexplained by worker effects, firm effects, and observables can be ascribed to match-specific effects. Second, we merge the estimated AKM effects with detailed German establishment-survey data.³ This allows us to estimate production functions with very detailed and model-consistent controls for

¹See Dickens and Katz (1987), Krueger and Summers (1988), Bell and Freeman (1991), and Gibbons and Katz (1992) for early work on industry wage differentials. Davis et al. (1991) and Groshen (1991) were among the first to study wages at the firm/establishment level.

²In the sequential auction framework of Postel-Vinay and Robin (2002), for example, firms can reduce workers' starting wages by offering steeper wage-tenure profiles.

³In the German data, we do not observe firms in the legal sense but establishments, i.e., single production units. We use the terms firm and establishment interchangeably throughout the paper.

heterogeneous worker ability. According to the model, the effect of workforce ability on output is inherently firm-specific because worker ability and firm productivity interact at the match-level. We show how the estimated AKM wage components capture this effect and use them to predict firm-level labor inputs. To estimate production functions, we use the refined control-function approach proposed by Ackerberg et al. (2015) (ACF). Specifically, ACF allow for dynamic implications of the firms' labor choice, a feature that is compatible with the frictional environment of our model.

Using these estimates, we study the mapping from firm productivity to wages, the degree of sorting between worker-ability and firm-productivity types, and implications for the interpretation of rising wage inequality. We find that the most productive firms do not pay the highest wages. In turn, many low-productivity firms pay relatively high wages. *Productivity sorting*, that is, the sorting of high-ability workers into high-productivity firms is less pronounced than sorting into high-wage firms. This implies that productivity sorting contributes less to rising wage inequality than wage sorting. This new insight complements earlier findings about the role of firms for increasing wage inequality.

The sorting between high-ability workers and high-productivity firms is in fact decreasing over time in our data. Matches between the most productive firms and the most able workers have become less common. We argue that this development is driven by the fact that almost all worker types earn lower wages in matches with the most productive firms as compared to somewhat less productive firms. For these workers, wage are not monotonically increasing in firm productivity. Studying worker transitions, we confirm that high-type workers move towards higher wages, even when this implies switching to a less-productive firm. Our findings suggest that widely-documented increasing wage sorting could be accompanied by decreasing allocative efficiency in the labor market and lower aggregate output.

Firms at the top of our estimated productivity distribution exhibit a number of interesting features. They have high revenues and high labor productivity (value added per worker), but they are not the largest firms in terms of headcount, capital stock, or the total wage bill. They do not pay the highest average wages, and their labor shares out of revenues and value added are low. These low labor shares could be related to the falling aggregate labor share, a trend observed in many developed economies.⁴

Lastly, we show that using productivity-based firm types as compared to wage-based measures is important for understanding the sources of increasing wage inequality. A decomposition of the variance of wages into shares explained within and between establishments reveals that the contribution of the between-firm component to overall wage dispersion has been rising by almost 10% in Germany between 1998 and 2008. This is in line with the findings of Song et al. (2019) for the U.S. Quantitatively, between-firm

⁴Our firm ranking appears to mirror the explanation emphasized in Autor et al. (2017) and Autor et al. (2020): the emergence of so-called “superstar” firms.

inequality is comparable in magnitude to the relatively stable within-firm component of wage dispersion. However, this picture changes when we decompose the variance of wages using our estimated firm-productivity and worker-ability types. We find that the share of wage variance explained between firm-productivity types is low in levels and increases only by around 4% over time. Its overall contribution to inequality is dwarfed by variance shares within firm-productivity types and between worker-ability types. Thus, we conclude that productivity sorting is quantitatively less important than wage sorting for rising wage inequality.

Contribution to the Literature

We show how the wage determination mechanism of a search-and-matching model with multi-worker firms, decreasing returns, intra-firm wage bargaining, worker-firm complementarities, and matching frictions can be used to facilitate the estimation of unobserved firm productivity. We build on Cahuc et al. (2008) who embed the Stole and Zwiebel (1996) intra-firm bargaining framework into the canonical search-and-matching model. The main difference to existing structural models in the sorting literature is that those models commonly adopt an assumption of one-worker-one-firm matches (e.g. Shimer and Smith, 2000; Atakan, 2006; Lise et al., 2016; Lise and Robin, 2017). They focus on worker and firm *quality*, while the *quantity* dimension of production is not taken into account. A notable exception is Bagger and Lentz (2019) (BL). Here, multi-worker firms are present, but production is linear, so firm size is limited by search frictions.

Eeckhout and Kircher (2018) also relax the one-worker-one-firm assumption and study both the quantity and the quality dimension of production. The firm must decide which worker type(s) to hire and, additionally, how many workers of each type. Firms optimally hire multiple workers of exactly one type. Eeckhout and Kircher (2018) show that this result holds for both frictionless matching and competitive search. In this paper, we are interested in the empirically-relevant case of firms that are simultaneously matched with multiple worker types. Thus, we focus on a model with random search, non-degenerate matching sets, and production structure that is geared to our empirical approach, whereas Eeckhout and Kircher (2018) study more general production functions.

Our findings contribute to the empirical literature on wage inequality in a number of ways. The aforementioned studies by CHK for Germany and Song et al. (2019) for the U.S. follow the AKM approach. They decompose wage dispersion into the contributions of unobserved worker ability, firm wage premia, and wage sorting in the labor market. Wage sorting measures the extent to which workers who receive high wages are matched with firms that pay high wages. We show that the way one measures firm heterogeneity, that is, by the wages firms pay or by their productivity, makes a difference for this kind of decomposition. In our data, firms with the highest estimated productivity do not pay

the highest wages. For this reason, sorting into high-productivity firms is quantitatively less important for rising inequality than wage sorting.

Only a small number of related papers in the empirical literature on wage dispersion uses non-wage-based measures of firm heterogeneity. Bartolucci et al. (2018) argue that firms' expected payoffs provide a summary statistic of firm heterogeneity. They use balance sheet data for a set of Italian firms and rank them by observed profits. Taber and Vejlin (2020) and BL rank firms by the share of workers they poach from other firms. Sorkin (2018) applies Google's page ranking algorithm to worker flows. Haltiwanger et al. (2018) and Bertheau et al. (2020) study the cyclical properties of worker flows by using, respectively, gross output and value added per worker as firm quality measures.

Another related literature studies the pass-through from time-varying firm productivity to wages (rent sharing). Card et al. (2018) survey this literature and link it to the AKM-inspired literature on wage dispersion. Bagger et al. (2014) estimate firm-level production functions with heterogeneous labor inputs to study wage dispersion using Danish data. Their wage equation includes occupation-specific worker effects and time and occupation-specific firm effects. It reduces to a log-linear form that allows estimation in the spirit of AKM. Chan et al. (2021) build on this approach to study heterogeneous pass-through from productivity shocks to wages. The distinguishing feature of our paper is the focus on sorting in the labor market. We show how firm productivity can be estimated in a way that is consistent with worker-firm specific marginal products of labor. Estimated productivity is highly persistent, so we use it to construct time-invariant firm types. We study how workers' wages vary across these firm types and study implications for worker transitions, the allocation of workers to firms, and wage inequality.

The main difference between the AKM-inspired literature and model-based analyses of labor market sorting is that models typically imply a non-linear wage equation. Due to the complementarities that the theory is based on, wages are not a monotonic function of firm productivity (e.g. Gautier and Teulings, 2006; Eeckhout and Kircher, 2011; Lise et al., 2016; Hagedorn et al., 2017; Lopes de Melo, 2018; Bagger and Lentz, 2019). By contrast, the log-linear AKM wage equation assumes that wages always increase in the firm effect. We document quantitatively important deviations from wage monotonicity at high and low-productivity firms for most worker types. Across firms in the middle of the productivity distribution, however, wages evolve largely monotonically.

Both model-based and AKM-inspired papers typically find the sign of sorting to be positive, reflecting positive assortative matching (PAM) in the labor market. However, there are large differences in the estimated strength of sorting. For Germany, the structural approach of Hagedorn et al. (2017) yields an estimated degree of sorting (correlation of worker and firm ranks) of 0.76, much higher than the correlation of estimated worker and firm effects in CHK, which is about 0.21.⁵ For Denmark, BL estimate a structural

⁵0.21 is the mean of two correlations measured by CHK for the time period in Hagedorn et al. (2017).

model in which sorting is driven by on-the-job search with endogenous search intensity. They find a correlation of 0.39. Lentz et al. (2018) use a variant of the Bonhomme et al. (2019) clustering technique and report a correlation coefficient of 0.28. The AKM model yields a correlation of 0.05 on Danish data according to Bagger et al. (2013). For comparison, our benchmark estimate of productivity sorting in Germany, that is, the rank correlation between wage-based worker types and productivity-based firm types, is 0.07.

Finally, important related papers in the empirical IO literature document the large extent of firm-productivity dispersion in the data (Syverson, 2011) and pioneer controlling for labor quality differences in production functions estimations. Fox and Smeets (2011) find that observable worker characteristics like education, gender, experience, and industry tenure explain about one fifth of the overall productivity dispersion across firms. Irarrazabal et al. (2013) find that 25% to 40% of the productivity premium of exporters is related to labor input quality differences, including unobserved worker heterogeneity.

The remainder of our paper is structured as follows: Section 2 introduces the model of large firms, sorting, and wages that guides our empirical approach. Section 3 describes our data. Section 4 explains our approaches to estimate worker and firm ranks and studies the properties of our rankings. Using the estimated ranks, Section 5 explores the extent of labor market sorting in Germany and documents changes over time. Section 6 relates our findings to wages and trends in wage inequality. Section 7 concludes.

2 Model

To fix ideas and motivate our empirical approach, this section develops a model of multi-worker firms, decreasing returns, intra-firm wage bargaining, worker-firm complementarities, and matching frictions. The model serves three purposes. First, it shows that the marginal product of different worker types is match-specific in the presence of worker-firm complementarities. Second, using the model, we take a stance on how these heterogeneous marginal products influence wages. We show how the model’s wage equation maps into the log-linear AKM wage equation. This step is key in constructing model-consistent firm-level labor inputs using estimated AKM wage components. Third, the model clarifies the assumptions under which our theory is compatible with the ACF production function estimation approach.

Consider an economy in which atomistic firms produce a numeraire good using multiple heterogeneous labor inputs. Worker heterogeneity is summarized by $n > 1$ ability types indexed by x . Worker types are time-invariant. Firm productivity, denoted Ω , may change due to idiosyncratic firm-level shocks that evolve according to a stochastic process characterized by the conditional CDF $G(\Omega'|\Omega)$.⁶ Workers and firms meet randomly.

⁶This feature is only added to ensure consistency with the ACF production function estimation. It is inconsequential for our derivation of the wage equation, which assumes a steady state.

Conditional on meeting, a match is not guaranteed because the surplus may be too low. Appendix A.1 spells out the details of the assumed matching mechanism.

To set up the problem of the firm, we build on Cahuc et al. (2008), who generalize the canonical search-and-matching model to allow for multi-worker firms with heterogeneous labor inputs and decreasing returns, strategic interactions, and intra-firm bargaining in the spirit of Stole and Zwiebel (1996). We add heterogeneous firm productivity and a production structure that is consistent with positive worker-firm sorting to the model. To facilitate a simple and transparent link to the empirical modes we use (AKM and ACF), we abstract from complementarities between worker types within the same firm (the focus of Cahuc et al., 2008) and span-of-control issues (studied by Eeckhout and Kircher, 2018).

We make a number of further simplifying assumptions. First, workers and firms are risk-neutral. Second, worker and firm heterogeneity are one-dimensional.⁷ Third, worker ability and firm productivity are known to all market participants and cardinally measurable. Fourth, infinitely-lived workers supply one infinitesimally small unit of labor (no extensive/intensive margin choice), so the labor input is a continuous variable. Fifth, we abstract from capital inputs for the discussion of the model.⁸ Sixth, we present the model in discrete time because the production function estimation following ACF also relies on a discrete-time model of dynamically optimizing firms. Time indices are omitted for brevity.

A general concave firm-level production function is

$$Y = F(L, \Omega), \quad (1)$$

where Y is value added. $L = \sum_x xL_x$ is a scalar composite labor input in units of worker ability. It combines all heterogeneous labor inputs L_x , the measure of type x workers employed by the firm. Ω is current productivity (or TFP) of the firm. Our focus lies on worker-firm sorting, so we assume that output is (log-)supermodular at the match level: a complementarity between firm productivity and worker ability determines the contribution of every single match to firm-level output. That is, the marginal product of an additional unit of worker ability is firm-specific due to the interaction with productivity Ω . We interpret the firm's productivity as a “non-rival” resource, that is, we do not consider the span-of-control problem of optimally allocating the firm's resources to heterogeneous workers.

⁷For recent explorations of sorting with multi-dimensional characteristics, see Lindenlaub (2017), Lindenlaub and Postel-Vinay (2020), and Lise and Postel-Vinay (2020).

⁸Both part-time labor and capital inputs are taken into account when we estimate production functions in Section 4.2.

A simple production structure in line with our assumptions is

$$F(L, \Omega) = \left(\sum_{x=1}^n (x \times \Omega) L_x \right)^{\beta_l}, \quad (2)$$

where $0 < \beta_l < 1$ is the output elasticity of the composite labor input. This production function is (weakly) log-supermodular at the match level, in line with the sufficient conditions for PAM derived by Shimer and Smith (2000).⁹ Under our assumptions, worker ability units are perfect substitutes at the firm level. That is, output depends on Ω and the efficiency units of labor employed. The resulting worker-firm-specific marginal product of labor (MPL) is

$$F_x(L, \Omega) = \frac{\partial F(L, \Omega)}{\partial L_x} = \beta_l \Omega^{\beta_l} L^{\beta_l - 1} x, \quad (3)$$

which depends positively on worker ability, firm productivity, and the output elasticity of labor. Due to decreasing returns, it decreases in the total composite labor input L . Note that the worker-firm-specific MPL does not depend on the composition of the workforce. Moreover, it is linear in worker ability x , a property that we will use below. As in Cahuc et al. (2008), employment is a state variable due to search frictions. The firm's problem amounts to optimally choosing how many vacancies to post given the expected profits from hiring. We assume that vacancies cannot be targeted to specific worker types and are subject to a productivity-dependent cost $c(\Omega)$. In Appendix A.2, we solve the firm's problem and derive the relevant optimality conditions in steady state.

The bargained wage solves the standard Nash sharing rule. The firm's surplus consists of the marginal profits from hiring an additional worker of type x (equation A.11). Its threat point is to renegotiate wages with all other employees (Stole and Zwiebel, 1996). The worker's surplus is the difference between the option values of employment and unemployment (equations A.14 and A.15, respectively). Wage (re)negotiations happen instantaneously, so firm-level employment remains fixed. In Appendix A.3, we show that the outcome of intra-firm wage bargaining in the model is described by the following differential equation

$$w(x, L, \Omega) = \alpha \left(F_x(L, \Omega) - L \frac{\partial w(k, L, \Omega)}{\partial L} \right) + (1 - \alpha)(1 - \beta)U(x). \quad (4)$$

This is a discrete-time version of the wage bargaining outcome derived by Cahuc et al. (2008) for their “single labor case”. We can relate to it due to our assumption of perfectly substitutable ability units. α is the workers' bargaining power parameter and

⁹Shimer and Smith (2000) establish the existence of an equilibrium in this environment. Positive assortative matching (PAM) arises with a log-supermodular match-level production function. Log-submodularity leads to negative assortative matching (NAM). In this empirical paper, we do not attempt to generalize the Eeckhout and Kircher (2018) large-firm sorting conditions for random search models. We leave this theoretical exercise for future work.

β is the common discount factor. The first term in brackets shows that the wage of a type x worker at a (L, Ω) firm is a function of the worker-firm-specific MPL , $F_x(L, \Omega)$. The second term captures the inframarginal effect that hiring the marginal worker has on all other workers' wages. It mirrors the finding that firms can reduce incumbent workers' wages by increasing employment in the presence of decreasing returns.¹⁰ Without complementarities between worker types, the inframarginal adjustment solely reflects decreasing returns and is unambiguously negative. Moreover, in the Stole and Zwiebel (1996) setting, firms instantaneously (re)negotiate with all workers as if they were the marginal worker, so the adjustment is the same for all worker types; it does not vary with x . Finally, the third term of equation (4) captures the worker-specific outside option, which we will return to below. The solution¹¹ to the differential equation (4) is

$$w(x, L, \Omega) = (1 - \alpha)(1 - \beta)U(x) + \int_0^1 z^{\frac{1-\alpha}{\alpha}} F_x(Lz, \Omega) dz, \quad (5)$$

so the wage depends on the worker's outside option and an integral expression combining the worker's marginal product and the inframarginal effect. The latter is weighted by the worker's bargaining power and decreasing in the distance to the margin L .

2.1 From Theory to Estimation

The model's wage equation makes clear that the wages firms pay are not directly informative about firm productivity. Equation (5) is non-linear in Ω , and the effect of productivity on wages is intertwined with firm size and the worker's bargaining power. Additionally, the wage reflects the worker's outside option. Suppose a high-ability worker bargains with a relatively unproductive firm. The worker's outside option is to wait until a matching opportunity with a more productive firm occurs. Upon matching, the low-productivity firm has to compensate the worker for this foregone option value, raising the worker's wage. Thus, according to our model, there are two distinct reasons for high (or low) wages according: the match-specific contribution to output (summarized by the integral term) and the worker's outside option.

To recover Ω empirically, we propose a two-step approach, combining techniques from the empirical wage dispersion and IO literatures. First, we estimate the log-linear AKM model on German matched employer-employee data, see Section 3.3. The estimation yields, for all individual workers and firms in the largest connected set, estimated worker-fixed effects, firm-fixed effects, and a residual. There is a direct correspondence between the worker-fixed effect, the AKM residual, and the wage equation of the model.

¹⁰Cahuc et al. (2008) allow for unrestricted substitutability/complementarity-patterns between worker types. In this case, effects on coworker wages can be either positive or negative. Firms may strategically over/under-employ specific worker types depending on their contribution to the total wage bill. Smith (1999) studies the efficiency of job creation in such a setting.

¹¹We follow Stole and Zwiebel (1996) and Cahuc et al. (2008). Details are relegated to Appendix A.3.

Consider the worker-fixed effect. Both the outside option and the integral expression in (5) are functions of worker ability. Their effect on wages is absorbed by the AKM worker-fixed effect under the condition that both wage components are linear in worker ability. Given our assumed production structure, the integral expression is indeed linear in worker ability x because the worker-firm-specific MPL (3) is also linear in x . Plugging in the MPL yields

$$w(x, L, \Omega) = (1 - \alpha)(1 - \beta)U(x) + x \int_0^1 z^{\frac{1-\alpha}{\alpha}} \beta_l \frac{F(Lz, \Omega)}{Lz} dz, \quad (6)$$

where x can be written in front of the integral. Thus, worker ability scales the worker's contribution to output and, thus, its effect on the wage. Having established that the integral term is linear in x , we show in Appendix A.4 that the worker's outside option is also linear in x under two additional assumption. First, the workers flow value of unemployment, denoted $b(x)$ in the model, is linear in x . This is a standard assumption. Second, matching sets cover the whole type space, that is, there are no unacceptable combinations of worker and firm types. This assumption is easily verified empirically.¹² Our result that estimated AKM worker-fixed effects fully capture the effect of worker ability on wages can also serve as a “micro-foundation” for ranking workers based on these wage effects, which we do in Section 4.1.

It should be noted that the model developed in Appendix A assumes that only unemployed workers search. That is, we consider a “reduced-form” of the worker's outside option. For the empirical application in our paper, it is inconsequential whether the worker's outside option reflects the value of unemployment or the value of employment at another firm with search on-the-job. The only condition that has to be fulfilled is that the outside option is linear in worker ability as discussed above. This ensures that the outside option is captured by the AKM worker-fixed effect.

Next, consider the AKM residual. The residual absorbs the non-linear part of the worker-firm-specific integral term. This expression comprises the wage effect of the worker's contribution to the firm's output through the production function, the output elasticity of labor, and the worker's bargaining power. According to our model, the information encoded in the AKM residuals is thus useful for capturing the effect that heterogeneous labor inputs have on output at the firm level.

Finally, AKM firm-fixed effects absorb time-invariant pay components that are not part of our model. These could be wage premia, e.g., wage effects of employee representation, or amenities. We assume that the pay components captured by the firm-fixed effect do not reflect the workers' labor input to the firm's production process.

To estimate Ω , we merge the estimated AKM wage components with rich survey-information on input and output choices for a representative sample of firms. We predict

¹²Appendix Figure C.3 shows that the joint density is positive for all worker-firm-type combinations.

the firms' labor input using worker effects, the effect of observable characteristics (not part of our model), and the AKM residual. Through the lens of our model, these estimated wage components summarize the effect that hiring a specific worker type has on the firm's output. Moreover, in contrast to using the headcount or full-time equivalents, the predicted labor inputs are comparable across firms and thus facilitate identification of Ω .¹³ Combining the predicted labor inputs with a measure of the capital stock and other controls enables us to estimate firm-level production functions in a way that is consistent with the theory of labor market sorting.

3 Data

Our analysis combines two data sets provided by the Institute for Employment Research (Institut für Arbeitsmarkt- und Berufsforschung, IAB) of the German Federal Employment Agency (Bundesagentur für Arbeit, BA).¹⁴ The first is the “IAB Employee History File” (BeH) which comprises the universe of employment spells recorded by the German social security system. The second data set is the “IAB Establishment Panel” (EP), which is a representative establishment-level survey that can be linked to the BeH.¹⁵

In this section, we describe the two data sets, explain how we prepare and combine them, and discuss the wage regressions we use to decompose observed wages. These wage components allow us to conduct the model-inspired adjustment of labor inputs at the establishment-level, which we then use to estimate productivity. We relegate additional details on sample selection and imputation procedures to Appendix B.

3.1 Data Sources and Preparation

The BeH employment spell data contain information on workers' gender, age, and education¹⁶, as well as start and end dates of the spells, total earnings, and occupation/industry codes. The data cover the vast majority of the German workforce. Only civil servants and self-employed workers, who do not pay social security contributions, are excluded. Each worker and each establishment have a unique identification number, which allows us to follow workers over time and from one establishment to another.

Regarding sample selection, we largely follow CHK and Lochner et al. (2020). We start from the universe of employment spells observed between 1998 and 2008. We observe daily wages. There is no exact information about hours worked in the BeH, so we restrict

¹³The idea to use observed input price variation to refine production function estimation techniques is well-known in the IO literature, see, e.g., Doraszelski and Jaumandreu (2013).

¹⁴The data sets are accessible through the research data center (FDZ) of the IAB.

¹⁵The EP is a random sample of all establishments in Germany, stratified according to size, industry, and federal state. See Kölling (2000) and Fischer et al. (2009) for a detailed description of the EP data.

¹⁶We use four education categories: 1 = “no degree”, 2 = “vocational training”, 3 = “some college degree”, 4 = “university degree”.

our sample to full-time employees (males and females) of ages 20 to 60. The age restriction avoids interference from apprenticeship training and early retirement programs. For each worker, we define the main job in a given year as the job with the highest total wage sum (including bonus payments).

The EP is a comprehensive annual survey of establishments, that is, single production units like factories or branches. It provides us with the necessary data to estimate production functions at the establishment level: we observe revenues, intermediate good purchases (reported as revenue shares), value added (calculated as revenues minus intermediate good purchases), and net investments in four different categories of capital goods (buildings, production machinery, IT, and transport equipment). We restrict our analysis to EP establishments with non-missing data on revenues and intermediate good purchases.¹⁷ Moreover, we drop establishments below the first and above the 99th percentile of the revenue distribution to ensure that our results are not driven by outliers.

We supplement the EP data with covariates from the “Establishment History Panel” (BHP).¹⁸ These include average wages, headcounts, shares of full-time/part-time workers, and worker shares by skill (education) group. Moreover, the BHP provides administrative information on firm age and a consistent industry classification.

A possible concern about working with establishment-level data is that firms (in the legal sense) may consist of multiple establishments that influence one another. The German economy, however, is well-known to be characterized by a broad basis of small and medium-sized enterprises. Accordingly, 80% of the establishments in our data (self-reportedly) belong to single-establishment firms.

3.2 Imputations

In the BeH, we observe nominal gross daily wages, which we deflate using the consumer price index from German national accounts. A limitation of the wage data is that earnings are tracked only up to a threshold, the social security contribution assessment ceiling (“Beitragsbemessungsgrenze”). We impute the upper tail of the wage distribution following Dustmann et al. (2009) by running a series of Tobit regressions, allowing for a maximum degree of heterogeneity by fitting the model separately for gender, time, education levels, and eight five-year age groups, see Appendix B.1. We impute missing and inconsistent education observations in the BeH using the methodology proposed in Fitzenberger et al. (2006), see Appendix B.2.

The EP data contain information on net investments. To estimate the level of the capital stock, which is an input to the production function, we use a perpetual inventory

¹⁷The main reason for missing information is that some firms choose not to report revenues as their measure of output. This applies mainly to financial institutions and public sector firms.

¹⁸The BHP covers all establishments with at least one employee liable to social security on a reference date (June 30th). See Spengler (2008) for a detailed description of the BHP data.

method following Müller (2008). This method approximates the establishment-level capital stock by combining net investments with average economic lives (depreciation rates, available from national accounts) for the different types of capital goods we observe.

3.3 Wage Regressions

We estimate an AKM-type wage regression on the largest connected set in the BeH data for our period of analysis, 1998–2008.¹⁹ We include both men and women. The connected set contains more than 233 million person-years, corresponding to 35 million individual workers at 3.3 million establishments. 55% of observations are movers between establishments. We use the CHK specification, that is, we estimate a log-linear wage equation for worker i who works at firm $j(i, t)$ in year t :

$$w_{it} = \alpha_i + \psi_{j(i,t)} + x'_{it}\beta + \varepsilon_{it}, \quad (7)$$

where w_{it} are log real daily wages, α_i is the worker-fixed effect, $\psi_{j(i,t)}$ is the establishment-fixed effect, and x'_{it} contains time-varying controls: an unrestricted set of year dummies and quadratic and cubic terms in age, fully interacted with educational attainment. ε_{it} is the residual. The regression model is identified for workers who move between establishments. Following CHK, we impute worker-fixed effects and residuals for “stayers”.²⁰

The adjusted R^2 of the regression is 0.92, broadly in line with CHK. The correlation of estimated worker and firm-fixed effects, often interpreted as a measure of wage sorting in the labor market, is 0.27 in our time interval. This is slightly higher than what CHK report, likely due to the longer time period and broader sample we consider.²¹ Estimated worker and firm-fixed effects as well as their covariance are likely to be biased due to limited mobility in the connected set (an incidental parameter problem).²² We tested the parametric correction suggested by Andrews et al. (2008) for two sub-periods of our data and find that the “limited mobility bias” is relatively small in our sample.²³

Table 1 shows variance decompositions for multiple groups of workers based on the estimated AKM wage components. Column (a) includes all person-years, (b) all women,

¹⁹The key assumptions underlying the AKM two-way fixed-effect model are additive separability of the wage equation and exogenous mobility, i.e., high-wage workers do not systematically sort into high-wage firms. See Abowd et al. (2002) for details.

²⁰Workers who work at the same establishment over the entire sample period.

²¹CHK report correlations for shorter time intervals: 0.17 (1996–2002) and 0.25 (2002–2009). Moreover, CHK include only men in former West Germany, whereas we include both men and women in reunified Germany from 1998–2008.

²²As shown by Andrews et al. (2008, 2012) and recently revisited by Borovičková and Shimer (2020) and Kline et al. (2020), estimated worker and firm-fixed effects are biased upwards and their covariance is biased downwards.

²³With the correction, the variance of worker effects increases by 5%/4% and the variance of firm effects increases by 4%/3% in the sub-periods 1998–2002/2003–2008, respectively. The covariance of worker and firm effects falls by 7%/5%.

(c) all men, and (d) all men in West Germany (the CHK sample). We replicate the well-known finding that the major share of wage variance is explained by unobserved worker heterogeneity. The worker-fixed effect explains almost half of the observed variation in wages, slightly more for women and slightly less for men. The second most important source of variation is firm-fixed effects. They explain roughly a quarter of the wage variance across the four groups. The third most important source is the covariance of worker and firm effects, which explains between 12 and 19% of wage variance.²⁴ With only 2%, the share of wage variance explained by time-varying observable characteristics is almost negligible. The same is true for the covariances of observable characteristics with worker and firm effects. Note that time-invariant covariates like education are absorbed by the worker effect.

In the log-linear wage equation (7), the residual absorbs the match-specific wage effects highlighted by our model (the integral expression in equation 6). This component explains wage variance shares between 7 and 9% across the four BeH samples. To separately assess the quantitative importance of match-specific effects, we follow CHK and decompose the residual into $\varepsilon_{it} = \eta_{i,j(i,t)} + r_{it}$, where η_{ij} is a match-specific wage effect for worker i at firm j and r_{it} is the remaining error term. The match effects alone explain 5.8% of wage variation. This accounts for almost 75% of the residual in specification (a).

Thus, most wage variation left unexplained by worker effects, firm effects, and observables can be ascribed to match-specific effects. 5.8% might appear to indicate that the quantitative importance of match-specific effects is small. Note, however, that this only measures the contribution of match-specific effects to wage dispersion. Recall equation (4) and suppose workers have low bargaining power. In this case, only a small fraction of workers' output contribution is reflected in wages. The quantitative importance of the match-level complementarity for output, highlighted by our model, could still be large.

3.4 Labor Inputs

We use the estimated AKM wage effects for men, women, and both movers and stayers from the full BeH sample (column (a) in Table 1) to construct labor inputs for the production function estimation. Formally, we define

$$W_{jt}^* = \sum_i \exp(\hat{\alpha}_i) + \exp(x'_{it}\hat{\gamma}) + \exp(\hat{\varepsilon}_{it}). \quad (8)$$

The wage bill W_{jt}^* includes the sum of (exponentiated) worker-fixed effects, effects of observable characteristics, and residuals for all individual workers i with estimated AKM effects at firm j in year t . According to our model, the residual absorbs worker-firm-

²⁴Interestingly, women are less positively sorted in terms of wages than men. This is in line with what Card et al. (2016) and Bruns (2019) find using Portuguese and German data, respectively.

Table 1: Wage Variance Decompositions

	(a)	(b)	(c)	(d)
	Regression (7) BeH, full	Regression (7) BeH, women	Regression (7) BeH, men	Regression (7) BeH, men, west
Var(w_{it})	0.276 (100%)	0.277 (100%)	0.245 (100%)	0.226 (100%)
Var($\hat{\alpha}_i$)	0.126 (46%)	0.138 (50%)	0.105 (43%)	0.106 (47%)
Var($\hat{\psi}_{j(i,t)}$)	0.068 (25%)	0.076 (27%)	0.061 (25%)	0.049 (22%)
Var($x'_{it}\hat{\beta}$)	0.005 (2%)	0.006 (2%)	0.005 (2%)	0.005 (2%)
$2 \times \text{Cov}(\hat{\alpha}_i, \hat{\psi}_{j(i,t)})$	0.049 (18%)	0.032 (12%)	0.047 (19%)	0.037 (16%)
$2 \times \text{Cov}(\hat{\alpha}_i, x'_{it}\hat{\beta})$	0.004 (0%)	0.001 (0%)	0.005 (2%)	0.006 (3%)
$2 \times \text{Cov}(\hat{\psi}_{j(i,t)}, x'_{it}\hat{\beta})$	0.003 (0%)	0.001 (0%)	0.004 (2%)	0.004 (2%)
Var($\hat{\varepsilon}_{it}$)	0.021 (8%)	0.025 (9%)	0.018 (7%)	0.019 (8%)
Sample mean wage	4.450	4.261	4.553	4.621
R^2	0.92	0.93	0.91	0.92
#Observations	233,117,492	82,267,794	150,849,698	123,087,610

Notes: Variance decompositions of log real daily wages according to regression model (7) for various BeH samples. Mean wages, variances, and covariances rounded to three decimal places. Source: BeH.

specific output effects. We argue that it is a useful proxy for match-specific interactions of worker and firm types.

The advantage of relying on the AKM residual is that it provides a measure of match-specific effects for both movers and stayers in the connected set, which are almost all individual workers.²⁵ A disadvantage is that the residual sums to zero across all workers and years for any firm.²⁶ Moreover, the residuals stem from a wage regression and, thus, reflect only the wage effect of match-level complementarities, not the output effects. According to the model, these wage effects are weighted by the workers' bargaining power. Thus, the residual underestimates the true worker-firm-specific output effect the more, the lower the workers' bargaining power. To the best of our knowledge, no better proxy for worker-firm-specific output effects is available, given that we do not observe output at the match-level.

²⁵Worker-fixed effects for stayers are imputed using the mean of observed wages, firm-fixed effects, and observables across all years. The residual for the stayers thus corresponds to the deviations from this mean in all single years.

²⁶One potential way to avoid this shortcoming would be the joint estimate of worker effects, firm effects, and match-specific effects following the approach suggested by Woodcock (2015) and Sørensen and Vejlin (2013). On the downside, match-specific effects would only be identified for movers, so we could measure them only for a subset of workers. This would lead to a downward bias in the predicted labor input.

3.5 Final Samples

We use two final samples in the subsequent analysis. We refer to the first sample as *All Matches*. It includes matches that started both before and after 1998. Here, we do not condition on the origin of the match, that is, we do not distinguish between job-to-job flows and matches formed out of non-employment. There are 4,695,108 employment spells of 1,344,382 workers employed at 10,004 EP establishments.

We refer to the second sample as *New Matches*. This sample includes matches formed from 1999 onward, and we distinguish between job-to-job flows and matches formed out of non-employment. Thus, we need one initial observation (1998) to see what the worker did before the start of the new match. There are 1,656,280 employment spells of 633,831 workers at 9,659 establishments. 64% (1,055,151) of new employment spells are job-to-job (J2J) moves from one employer to another. 36% (601,129) are spells formed out-of-non-employment (OON). Our definition of non-employment includes marginal employment, unemployment (benefit receipt), and inactivity.²⁷ Thus, OON spells also include young workers who enter the labor market.

In some parts of our empirical analysis, we study changes over time. To this end, we split our sample (1998–2008) into two halves, 1998–2002 and 2003–2008. A series of labor market reforms was implemented in Germany between 2003 and 2005.²⁸ Thus, one can loosely interpret 1998–2002 as the pre-reform period and 2003–2008 as the post-reform period. We do not attempt to interpret our empirical results in relation to these reforms.

4 Ranking Workers and Firms

4.1 The Worker Ranking

According to our model, the estimated AKM worker-fixed effects capture the effect of unobserved worker ability on output along with the worker's outside option. To study sorting, we rank individual workers based on this ability measure and create 50 bins of equal size.²⁹ Let $\bar{x}(i)$ denote the ability bin that worker i belongs to. In the following, individual workers in the same bin are thought of as workers of the same ability type.

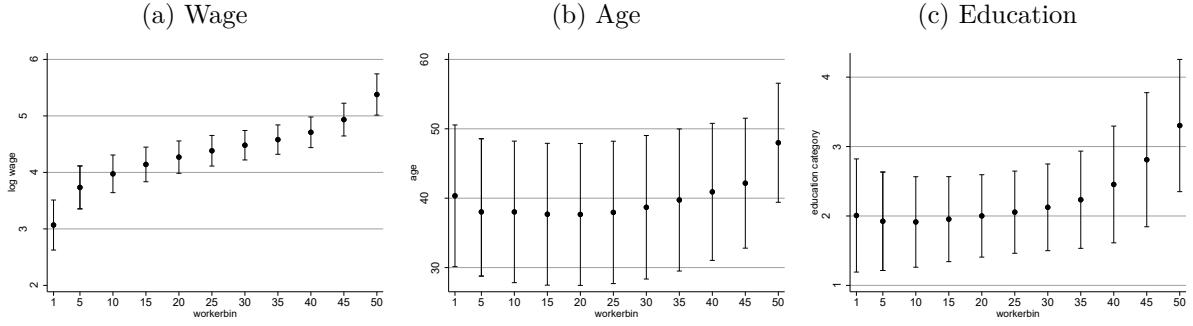
To show how worker heterogeneity is summarized by the bins, we decompose the respective variances of observed wages, age, and education into the shares explained within and between the bins. Little variance within the bins implies homogeneity of workers in the respective dimension. The ranking is based on worker fixed effects, which

²⁷We do not observe unemployment benefit payments in the BeH. Thus, we cannot distinguish between unemployment and inactivity.

²⁸The so-called Hartz reform package consisted of four reforms that were designed to increase labor demand (Hartz I and II), matching efficiency (Hartz III), and labor supply (Hartz II and IV).

²⁹The BeH sample includes nearly the full German labor force. There are 702,540 individual workers in each of the 50 bins.

Figure 1: Log Wage, Age, and Education Distributions across Worker Bins



Notes: Means \pm one standard deviations of log wages, age and education, plotted for every fifth worker bin. The age of individual workers in our sample ranges from 20 to 60. There are 4 education categories: 1 = “no degree”, 2 = “vocational training”, 3 = “some college degree”, 4 = “university degree”. Data source: BEH.

explain the major share of wage variation in the data. Thus, the share of wage variance between the bins is large (70%). In other words, workers in the same ability bin earn relatively similar wages. For age and education, the picture is quite different. 96% of the age variation and 74% of the education variation is found within the bins. We control for time-varying age and education effects in the AKM regression, so this finding is a direct reflection of the low correlation between estimated worker effects and time-varying observables (recall Table 1).

Figure 1 illustrates how log wages, age, and education vary across worker bins. Panel (a) shows that mean wages increase monotonically across worker bins as expected. In contrast, Panel (b) shows that the mean age across worker bins is essentially flat. Only the highest bin has a slightly higher mean. However, due to the large standard deviation, this difference is not significant. For education, Panel (c) shows that mean education increases above bin 35 but is flat below. This suggests that highly ranked workers are more likely to have tertiary degrees, while lower ranked workers tend to have vocational training only. But again, the differences are hardly significant. Note that the dispersion of education gets higher at the top of the worker ranking. It is more common to observe high-rank workers with little education than low-rank workers with tertiary degrees.

4.2 The Firm Ranking

We rank firms based on their unobserved productivity, which we infer by estimating production functions at the establishment level. We know from the empirical industrial organization (IO) literature that this approach is susceptible to two challenges. The first challenge is a classical endogeneity problem known as “transmission bias” (Marschak and Andrews, 1944). Input choices, e.g., the demand for labor or intermediate inputs, are likely to be correlated with firm productivity.³⁰ To address this challenge and estimate an

³⁰In every period, input demands are optimally chosen based on current firm-level productivity. The manager observes productivity when choosing those demands, but the econometrician does not.

unbiased measure of firm productivity, we rely on the ACF version of the “control function” approach.³¹ The key assumption is that intermediate input demand is a strictly increasing function of a (scalar) unobserved productivity shock. Under this assumption, the control function can be inverted. Effectively, one controls for unobserved firm productivity by substituting it out of the production function. ACF refine earlier approaches by allowing the intermediate input demand to depend on labor inputs, which makes their model a good fit for our focus on worker ability heterogeneity. Also, ACF allow for dynamic implications of the firms’ labor choice, which conforms to our model environment with matching frictions.

The second challenge is that, in the presence of heterogeneous worker ability, the quality of labor inputs varies across firms.³² Physical units, e.g., headcounts or hours worked, do not reflect worker ability differences and lead to measurement error. Moreover, if firm productivity and worker ability are complements, as our model presumes, even precise measures of worker ability are insufficient controls for the worker-firm-specific effects on output. This complicates the separate identification of firm productivity and the output elasticity of labor, leading to biased estimates. The model-based labor input measure we use addresses this challenge. It takes into account worker ability differences (absorbed by the worker-fixed effect) and worker-firm-specific output effects (absorbed by the residual), but leaves out time-invariant wage components unrelated to the firm’s output, e.g., amenities (absorbed by the firm-fixed effect).

Production Function Estimation

Our starting point is the production structure assumed in Section 2, equation 2. We rewrite it in terms of the composite labor input in ability units, L_{jt} , and add capital inputs K_{jt} as well as indices for individual firms j , workers i , and time t :

$$Y_{jt} = (\Omega_{jt} L_{jt})^{\beta_l} (\Omega_{jt} K_{jt})^{\beta_k}. \quad (9)$$

β_l and β_k are the output elasticities of labor and capital, respectively. Both inputs are scaled by the firm’s current productivity Ω_{jt} . Note that Y_{jt} is value added (revenue minus expenditures for intermediates).³³ Intermediate inputs are therefore not part of production function (9). Without assuming constant returns to scale, by the homogeneity

³¹This approach was originally developed by Olley and Pakes (1996), and refined by Levinsohn and Petrin (2003) and Wooldridge (2009).

³²Griliches (1957) was among the first to argue that labor inputs are not homogeneous within and across firms if workers are heterogeneous.

³³The ACF identification strategy is designed for value-added production functions. Gandhi et al. (2020) argue that the ACF method is not suitable to identify the parameters of the gross output production function without imposing further restrictions.

of the general Cobb-Douglas function (9) we get

$$Y_{jt} = \Omega_{jt}^{\beta_l + \beta_k} L_{jt}^{\beta_l} K_{jt}^{\beta_k}. \quad (10)$$

The sum of the output elasticities in the power of Ω_{jt} is irrelevant for the purpose of ranking firms. Thus, we define $\omega_{jt} = (\beta_l + \beta_k) \ln \Omega_{jt}$ when taking logs. We estimate the following value-added production function in logs:

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + \omega_{jt} + z'_{jt} \gamma + \epsilon_{jt}, \quad (11)$$

where lower case letters indicate logarithms. We add a constant β_0 , the residual ϵ_{jt} that absorbs transitory shocks, and a vector of additional controls z'_{jt} . The latter includes dummies for West German establishments, four firm age categories, the share of part-time workers, and a dummy for employee representation. We also include year and 32 sector dummies to account for cyclical fluctuations and differences in demand structures across sectors.

The ACF identification strategy assumes a discrete-time model of dynamically optimizing firms, compatible with our model in Section 2. The demand for labor and intermediate goods may change in response to realized firm productivity in the same period.³⁴ Capital is accumulated according to $k_{jt} = \kappa(k_{j,t-1}, i_{j,t-1})$, so investment in the previous period, $i_{j,t-1}$, predetermines the capital stock. The firm's information set when making dynamic input choices includes all past and present productivity shocks $\{\omega_{j\tau}\}_{\tau=0}^t$, but it does not include future productivity shocks. These are assumed to evolve according to a first-order Markov process:

$$\omega_{jt} = E(\omega_{jt} | \omega_{j,t-1}) + \xi_{jt} = \rho \omega_{j,t-1} + \xi_{jt}. \quad (12)$$

Thus, firm productivity in period t is a function of the conditional expectation for ω_{jt} based on last period's realization (the Markov property) and an innovation ξ_{jt} , which is assumed to be uncorrelated with ω_{jt} and the predetermined capital stock. In the following, we assume that ω_{jt} follows an AR(1) process with parameter ρ .

The control function we use is the demand for intermediate inputs

$$m_{jt} = f_t(l_{jt}, k_{jt}, \omega_{jt}), \quad (13)$$

which is a function of both the firm's labor input and the capital stock in addition to productivity.³⁵ Thus, conditional on both labor and capital, more productive firms use

³⁴The wage equation in Section 2 is derived for a steady state. However, it would be straightforward to allow for endogenous separations and changing labor demand in response to productivity shocks.

³⁵ACF use of the conditional (on labor) input demand function to improve the identification of the labor input parameter relative to Olley and Pakes (1996) and Levinsohn and Petrin (2003).

more intermediate goods in production. Due to the assumption of strict monotonicity in ω_{jt} , we can invert equation (13) and rewrite unobserved firm productivity as

$$\omega_{jt} = f_t^{-1}(l_{jt}, k_{jt}, m_{jt}), \quad (14)$$

which we then use to substitute ω_{jt} in (11) to get

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + f_t^{-1}(l_{jt}, k_{jt}, m_{jt}) + z'_{jt}\gamma + \epsilon_{jt} = \Phi_t(l_{jt}, k_{jt}, m_{jt}, z_{jt}) + \epsilon_{jt}. \quad (15)$$

This is the final production function we estimate. Following ACF, we adopt a two-stage procedure. First, value added is regressed on a polynomial approximation of $\Phi_t(l_{jt}, k_{jt}, m_{jt}, z_{jt})$. This does not identify any of the parameters but leads to an estimate $\hat{\Phi}_t(l_{jt}, k_{jt}, m_{jt}, z_{jt})$. In the second stage, estimated parameter values are calculated using GMM and a set of four moment conditions.

Estimation Results

Table 2 presents the results of the production function estimation. We show four different specifications in which we vary the way of controlling for worker ability. Column (a) presents our benchmark specification in which we use the wage bill W_{jt}^* , defined in equation (8), as the firm's labor input. It includes the AKM residual. Specification (b) follows the same approach but omits the AKM residual ($\tilde{W}_{jt} = \sum_i \exp(\hat{\alpha}_i) + \exp(x'_{it}\hat{\gamma})$). Specification (c) uses the total wage bill W_{jt} as labor input. In specification (d), we do not control for work force ability and use the headcount H_{jt} .

In the wage bill specifications (a)–(c), estimates of the output elasticity of labor and capital inputs are comparable in magnitude and imply decreasing returns to scale, in line with our model assumption. The headcount specification (d) is closer to the case of constant returns. Given that the headcount mismeasures the labor input in the presence of worker heterogeneity, the output elasticity of labor is likely biased upwards in specification (d). Consequently, estimated firm productivity is biased downwards in this case. We attempt to reduce the measurement error in the labor input by using (predicted) wage bills. Indeed, we observe that the variance of the estimated firm productivity $\hat{\omega}_{jt}$ increases as we move towards our benchmark specification, suggesting that we are successful in alleviating the measurement error problem. The variance of 0.074 in our benchmark specification is almost 50% higher as compared to column (d). All in all, our findings suggest that accounting for worker-firm-specific effects is important for separating the contributions of firm productivity and labor inputs to output.

We include the share of part-time workers as a control variable to capture variation in the prevalence of part-time work across firms. Recall that we do not estimate AKM wage components for part-time workers, so a high-part time share implies additional labor

Table 2: Production Function Estimation Results

Dependent Variable	<i>Value Added</i>			
	(a)	(b)	(c)	(d)
<i>Labor input</i>	0.713*** (0.001)	0.726*** (0.001)	0.710*** (0.001)	0.853*** (0.001)
<i>Capital input</i>	0.171*** (0.001)	0.185*** (0.001)	0.151*** (0.000)	0.132*** (0.001)
<i>Part-time worker share</i>	0.101*** (0.001)	0.116*** (0.002)	0.228*** (0.001)	-0.921*** (0.002)
<i>Firm age: 6–15 years</i>	0.121*** (0.001)	0.135*** (0.001)	0.122*** (0.001)	0.079*** (0.001)
<i>Firm age: 16–25 years</i>	0.075*** (0.001)	0.089*** (0.002)	0.079*** (0.001)	0.047*** (0.001)
<i>Firm age: >25 years</i>	0.047*** (0.002)	0.060*** (0.002)	0.049*** (0.001)	0.005*** (0.001)
<i>West German establishment</i>	0.310*** (0.001)	0.323*** (0.001)	0.141*** (0.000)	0.293*** (0.001)
<i>Employee representation</i>	0.356*** (0.001)	0.370*** (0.000)	0.279*** (0.001)	0.257*** (0.002)
<i>Labor input variable</i>	W_{jt}^*	\tilde{W}_{jt}	W_{jt}	H_{jt}
<i>Workforce quality control</i>	Predicted wage bill	Total wage bill	None	
<i>Year FEs</i>	Yes	Yes	Yes	Yes
<i>Sector FEs</i>	Yes	Yes	Yes	Yes
<i>Variance of $\hat{\omega}_{jt}$</i>	0.074	0.072	0.064	0.050
<i>Variance of residual</i>	0.628	0.626	0.597	0.583
<i>N</i>	38,598	38,598	38,598	39,808

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors (50 iterations) in parentheses. All estimated coefficients and standard errors are rounded to three decimal places. The reference category for the firm age dummies is a firm age of five years or less. Data sources: BHP, EP, BeH.

Table 3: Firm Ranking Correlations

	\bar{v}_j	\bar{v}_j/\bar{n}_j	$\bar{\pi}_j/\bar{n}_j$	\bar{n}_j	\bar{k}_j	\bar{k}_j/\bar{n}_j	workforce education
Correlation with $\hat{\omega}(j)$	0.43	0.25	0.16	0.40	0.24	-0.03	0.08

Notes: The table shows correlations of the time-invariant estimated firm ranks, $\hat{y}(j)$, with the means of the following firm statistics over time: log value added (\bar{v}_j), log value added per worker (\bar{v}_j/\bar{n}_j), profits per worker ($\bar{\pi}_j/\bar{n}_j$), the log size of the workforce (\bar{n}_j), the log capital stock (\bar{k}_j), the log capital stock per worker (\bar{k}_j/\bar{n}_j), and workforce education, as measured by the mean of the workers' education variable within the firm. Data sources: BHP, EP, BeH.

inputs beyond what is captured by the (predicted) wage bill.³⁶ One would thus expect a positive estimated coefficient, and this is indeed what we find. The headcount, H_{jt} , includes part-time workers. In this case, a higher share is associated with lower output.

Regarding the remaining controls, estimated effects of firm age are positive but decreasing compared to the reference category (firm age of five years or less). Dummies for West Germany and employee representation are always associated with higher value added, and more so in specifications (a) and (b). These coefficients likely capture persistent differences in wage structures, bargaining regimes, and, to a lesser extent, firm sizes between East and West Germany.

The variance of $\hat{\omega}_{jt}$ we find is an order of magnitude smaller than the residual variance. The residual absorbs transitory shocks and is thus more dispersed. The estimated autocorrelation of $\hat{\omega}_{jt}$ in the benchmark specification is 0.75. This is in line with an interpretation of ω as relatively persistent and slow-moving productivity process, reflecting enduring changes to the firms' production technologies. Such changes have repercussions for optimal choices, e.g., with respect to labor demand and wages.³⁷

Firm Ranking

Given the high degree of persistence of estimated productivity realizations over time, $\hat{\omega}_{jt}$, we take the mean at the firm level to rank firms. The empirical literature on wage inequality and labor market sorting, in relation to which we interpret our results, typically assumes permanency of worker and firm types. Our fixed-effects-based worker types are time-invariant, too. The estimated rank of firm j is denoted $\hat{\omega}(j)$.

Table 3 shows correlations of estimated firm ranks with several other firm-level statistics. The ranks are moderately positively correlated with firm size in heads (0.40) and

³⁶The share of part-time workers in the median firm in our data is only 5%, so excluding part-time workers implies little measurement error in the (predicted) wage bills. In an earlier version of this paper, Lochner and Schulz (2020), we constructed a ratio of wage bills to ability-adjust the full workforce. The results were only marginally different from what we find using (predicted) wage bills.

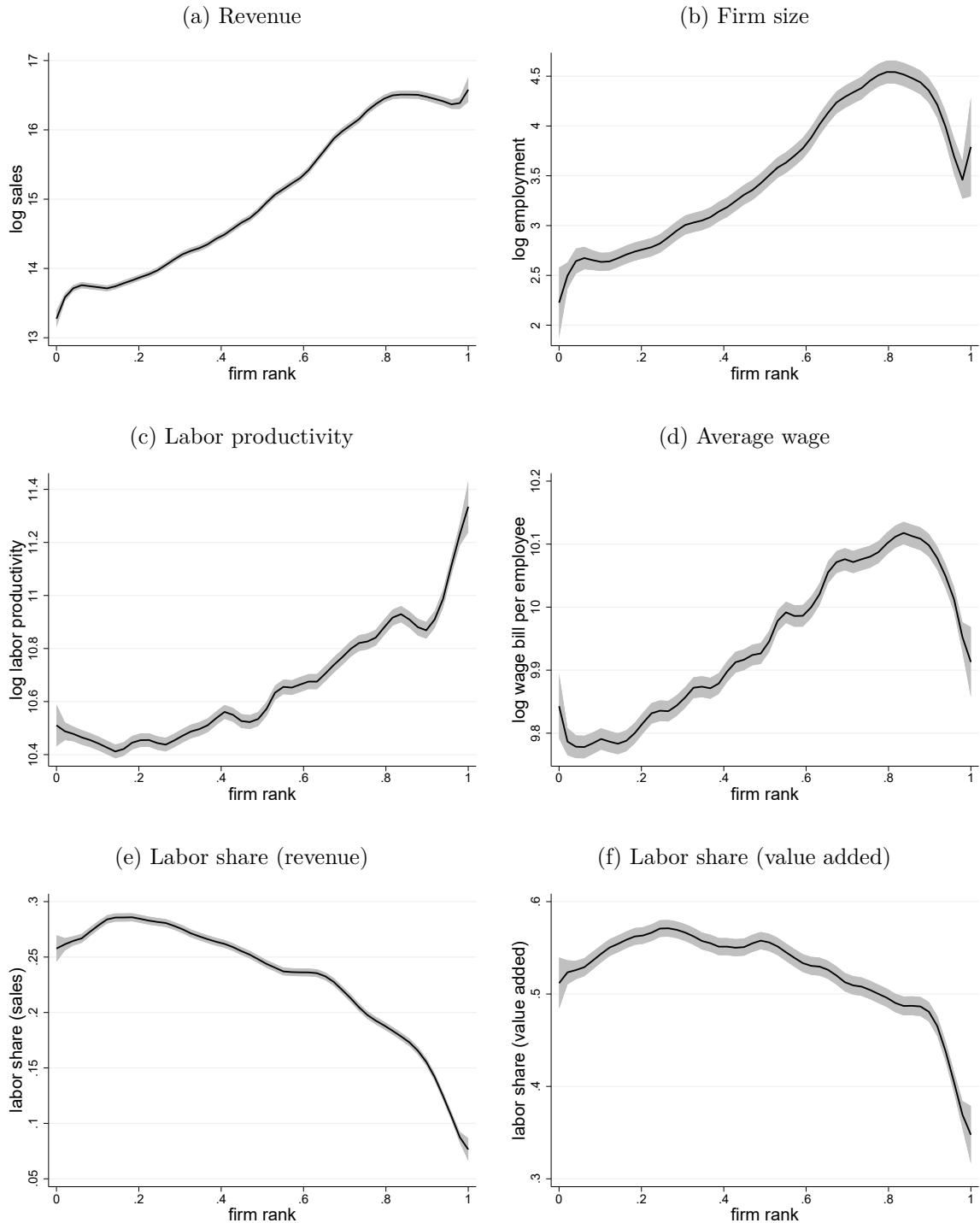
³⁷Guiso et al. (2005) show that firms fully *insure* their workers against transitory productivity shocks but only partially against the permanent productivity changes captured by $\hat{\omega}_{jt}$. This type of "pass-through" is fully in line with our model, where changes to ω do affect wages through the worker-firm-specific MPL . Due to our focus on labor market sorting, we do not study pass-through in this paper.

the capital stock (0.24). They are virtually uncorrelated with capital per worker (-0.03). These correlations suggest that the largest firms in our sample, both in terms of workforce and assets, are not the most productive ones. The correlation of firm ranks with the share of highly-educated workers (proxied by mean of workers' education within the firm) is only 0.08. Thus, highly productive firms are not just the ones with a highly-educated workforce. This low correlation also implies that workers' observable characteristics explain only a small share of productivity dispersion. Output measures (value added, both absolute and per worker) are also positively correlated with estimated firm ranks, but again only moderately so. The correlation with profits per worker is 0.16.

Figure 2 shows estimated kernel densities of six firm performance measures across firm ranks. Revenue (measured by log sales, Panel (a)) is mostly increasing in the rank, but becomes relatively flat at the top. The most productive firms do not have higher revenues than firms around the 80th percentile of the productivity distribution. In Panel (b), the relation between firm ranks and log employment (measured in heads) is shown. This size measure increases monotonically up until, roughly, the 80th percentile, but falls steeply thereafter. In line with the correlations reported above, the most productive firms in our sample are clearly not the biggest ones in terms of employment, despite the noise at the very top. Panel (c) shows that log labor productivity (value added per worker) is increasing almost everywhere in the firm rank. It is relatively flat in the bottom half of the distribution, but increases greatly for the most productive firms. Labor productivity of the firms at the top is twice as high as that of firms at the bottom of the productivity distribution. The log wage bill per employee, that is, the average wage that firms pay to their workers, is shown in Panel (d). For the most productive firms above the 80th percentile, average wages are falling. That is, the most productive firms do not pay the highest average wage. It is maximized around the 80th percentile, where also the largest firms in terms of employment are situated. Interestingly, the least productive firms pay higher average wages than other firms below the 25th percentile. We study in more detail how wages evolve across firm types in Section 6.

Finally, Panels (e) and (f) show labor shares, computed using both revenue and value added. The labor share is hump-shaped at the bottom but clearly falls in the estimated firm rank. For revenues, the labor share falls from just below 30% for the least productive firms to less than 10% for the most productive firms. For value added, the labor share falls from 50–60% to around 35% for the most productive firms. Thus, the average worker extracts a significantly lower share of output at high-productivity firms.

Figure 2: Firm Performance Measures by Estimated Firm Rank



Notes: Estimated univariate kernel densities of selected firm performance measures across estimated firm ranks, normalized between zero and one. The kernel is estimated using an Epanechnikov kernel function. The bandwidth is 0.01 for (a)–(d) and 0.02 for (e)–(f). The qualitative findings are robust to the bandwidth choice. 95% confidence bands in gray. Data sources: BHP, EP, BeH.

Table 4: Variance Decompositions with Firm Bins

	\bar{w}_j	\bar{v}_j	\bar{v}_j/\bar{n}_j	\bar{n}_j	\bar{k}_j
Overall	0.270	3.727	0.602	2.449	5.599
Between bins	0.016 (6%)	0.772 (21%)	0.044 (7%)	0.489 (20%)	0.483 (9%)
Within bins	0.253 (94%)	2.955 (79%)	0.558 (93%)	1.959 (80%)	5.116 (91%)

Notes: The table shows decompositions of the variance of the firm-level means of log wages (\bar{w}_{it}), log value added (\bar{v}_j), log value added per worker (\bar{v}_j/\bar{n}_j), log firm size (\bar{n}_j), and the capital stock (\bar{k}_j) into the respective shares explained within and between the firm bins. Data sources: BHP, EP, BeH.

Binning the Firms

In the next step, we group all individual firms into 15 bins of equal size.³⁸ Let $\bar{\omega}(j)$ denote the bin that firm j belongs to. In the following, individual firms in the same bin are thought of as firms of the same type. Similar to the worker bins before, we decompose the variance of some key variables into shares within and between the firm bins. Table 4 shows this decomposition.

The firm bins exhibit a high within-variance of wages. 94% of the observed variance of log wages is explained within the bins. This is a reflection of the fact that we control for heterogeneous worker ability when estimating firm productivity. It also implies that our firm types cannot be categorized into high-wage and low-wage firms—all firm types pay dispersed wages to their workers, depending on their ability and contribution to output.

The firms in the bins are also heterogeneous in terms of output (value added and value added per worker) and size (workforce size, assets) measures. The major share of the variance is always within the bins, suggesting that binning does not merely group firms of similar size and production capacity together. We also checked whether there is a relation between the firm bins and the industry a firm operates in or the prevalence of collective bargaining agreements and employee representation (not shown in Table 4). This is not the case: the dispersion of these attributes within the firm bins is also large.

Comparison with Alternative Rankings

Finally, we compare the productivity ranking with other ranking techniques used in the literature on wage dispersion and labor market sorting. We create two alternative firm rankings. The first is based on AKM firm-fixed effects (wage premia), the second on the poaching index used in BL and Taber and Vejlin (2020).³⁹ Overall, the correlations

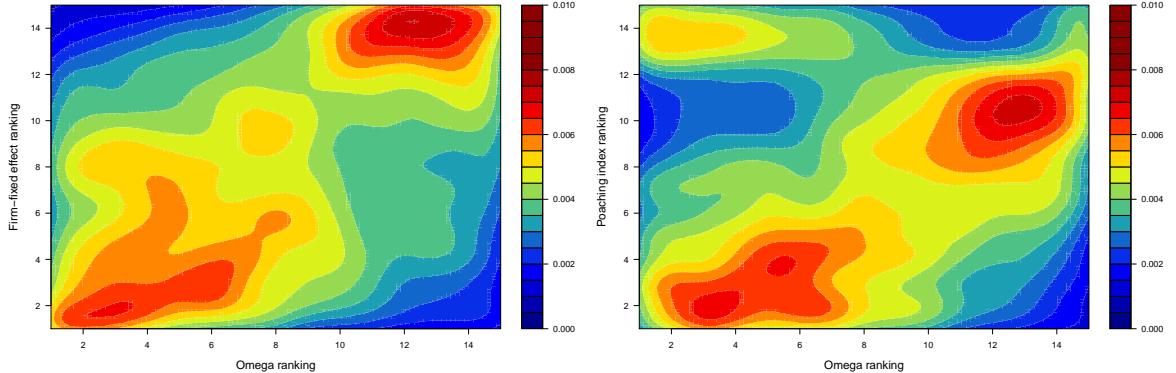
³⁸The EP sample used to estimate firm productivity contains 10,026 individual establishments, so there are 668 establishments in every bin.

³⁹The poaching index is based on the idea that high-paying firms poach workers from other firms rather than hiring unemployed workers. We compute it by comparing the yearly number of workers hired directly from other firms to all hires at the firm level and then rank firms based on the firm-level mean of the time-varying poaching index. We use the “Administrative Wage and Labor Market Flow

Figure 3: Comparison with Alternative Firm Rankings

(a) Firm-fixed effect (AKM)

(b) Poaching index



Notes: The two plots depict contours of the joint empirical distributions of firm years across combinations between the omega ranking (15 bins) and, respectively, AKM firm-effect ranks (15 bins, Panel (a)) and BL poaching index ranks (15 bins, Panel (b)). Data sources: BHP, EP, BeH.

of our omega ranking with rankings based on firms' estimated fixed-effect and poaching indices are positive and, coincidentally, of almost equal magnitude: 0.280 and 0.279, respectively. To analyze graphically how the different rankings are related, we create 15 firm bins based on both the firm-effect ranks and the poaching ranks. This allows us to compute the empirical distribution of firm years across the different firm bin combinations. We observe how firms with a given productivity rank are distributed across firm-fixed effect and poaching rank bins, respectively. Figure 3 plots contours of these empirical distributions.

In Panel (a), the AKM firm effect bins are depicted on the vertical axis and the omega ranking bins on the horizontal axis. We see that the mass of observations is concentrated around the diagonal, which is in line with the positive correlation reported above. In the upper-right quadrant of the plot, observations are highly concentrated somewhat above the diagonal, reaffirming our observation that the highest-paying firms (here in terms of AKM wage premia) are located just below the top of the productivity distribution. In the lower-left quadrant of the plot, observations are much more dispersed. It is not uncommon to observe firm years in which the estimated AKM wage premium is around the median but estimated productivity is very low and vice versa. Here, the disagreement between the two rankings is large. Figure C.1 in the Appendix shows that the observations in the lower-left quadrant stem mainly from young and small firms. The high-wage firms in the upper right quadrant are older and larger firms. Panels (a) and (b) in Figure C.1 show that low-wage (high-wage) firms are on average ranked low (high) with AKM, but there is a sizable overlap of both groups in terms of productivity.

In Panel (b), the mass of observations lies below the diagonal, that is, the poaching

Panel" (AWFP), see Stüber and Seth (2019). In this data set, the aggregated establishment-level worker flows needed to compute the poaching index are readily available.

rank tends to be lower than the productivity rank. Many high-productivity firms have high poaching ranks, but they are almost never at the top. At the same time, it is not uncommon even for medium-productivity firms to hire mainly out of non-employment, as evidenced by the high density of low-poaching rank firms that extends far to the right. Interestingly, the firm years with the highest poaching index are clustered in the upper-left quadrant of the plot. Apparently, some low-productivity firms are very active in poaching workers from other firms. Figure C.2 in the Appendix shows that many of the firms at the top of the poaching index distribution are small and young. One possible explanation is that these firms try to grow fast by poaching workers from other firms. The larger and older firms, which also pay the highest wages, are concentrated in the upper middle of the poaching rank distribution. Thus, they hire a non-negligible number of their employees out of non-employment.

In summary, the comparison of the different rankings shows that firm ranks based on firm wage premia and observed worker mobility are systematically different from our productivity-based firm ranking.

5 Labor Market Sorting

To analyze the allocation of workers to firms and measure the degree of labor market sorting in Germany, we merge the two data sets containing our estimated worker and firm rankings. In this step, we lose all employment spells at firms that are not part of the EP sample. For this reason, we re-bin workers and firms in the merged sample. The number of workers per bin decreases to 26,888. The number of firms per bin is almost unchanged at 667.⁴⁰ For all results presented in the following, we rely on the three samples defined in Section 3.5: (i) all existing matches and new matches subdivided into (ii) job-to-job flows and (iii) new matches formed out of non-employment.

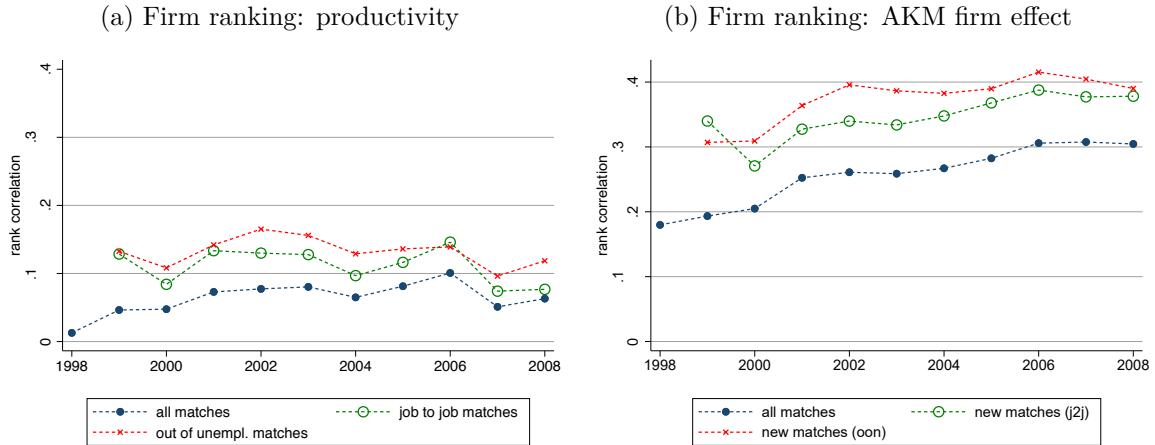
5.1 Rank Correlations

We use rank correlation coefficients (Spearman's ρ) to study the association between estimated worker and firm ranks in the data. The rank correlation is a simple yet informative summary statistic of how workers are allocated to firms. For all matches in all years, we find a significant positive but relatively low rank correlation of 0.07. For new matches, this correlation is 0.12. The correlation for job-to-job moves (0.11) is somewhat lower than the correlation for new matches out of non-employment (0.13).⁴¹

⁴⁰We are unable to merge the employment spell information for 22 out of 10,026 unique establishments.

⁴¹For clarity, the rank correlations presented here are calculated based on worker and firm bins. Correlations based on worker and firm ranks (not reported) are only marginally different. Table C.1 provides an overview of all estimated rank correlations for different samples and time periods.

Figure 4: Rank Correlation Coefficients over Time (1998–2008)



Notes: Panel (a) shows Spearman rank correlation coefficients based on our productivity-based (omega) firm ranking. Panel (b) ranks firms using AKM firm-fixed effects. Both plots use the worker-ability ranking introduced in Section 4.1. Data sources: BHP, EP, BeH.

The positive rank correlation we find is an indication of positive assortative matching (PAM) on the German labor market. The implied degree of PAM is low, both in an absolute sense and compared to earlier studies using German data. CHK and Hagedorn et al. (2017) find correlations of 0.17–0.25 and 0.76, respectively.⁴² To put our finding of low sorting into perspective, recall that we rank firms based on their estimated productivity, while the other two studies rely on wages and observed worker mobility across firms.

Next, we consider how the estimated rank correlations change over time. This gives us an indication of the time dynamics of productivity sorting. To this end, we present rank correlation coefficients for all samples and all years in Figure 4. First, consider Panel (a), in which the estimated rank correlations are calculated using the productivity-based firm ranking. The blue line depicts the correlations for all matches, while the red and green lines depict correlations for new matches out of non-employment and job-to-job switches, respectively. For all matches, the degree of sorting is first increasing over time and then relatively stable just below 0.1. We observe a distinct drop in 2007. For both types of new matches, the level of the correlation is always higher but falls toward the end. The degree of sorting for new matches out of non-employment is decreasing steadily after 2002. For job-to-job switches, the time dynamics closely mirror those for all matches. This is a reflection of the fact that about 65% of all new matches are job-to-job switches (see also Table C.1). In sum, new matches drive up the rank correlation for all matches in the beginning of the period. Once the correlation for new matches starts decreasing, the correlation for all matches levels off as well.

⁴²CHK use the log-linear AKM model and interpret the correlation of estimated worker and firm effects in the data as a measure of sorting. Hagedorn et al. (2017) study sorting through the lens of a structural model with worker and firm heterogeneity, search frictions, and on-the-job search. Their model allows identification of the sign and strength of sorting without assuming a log-linear wage equation.

We dig deeper into the sources of the changing dynamics of sorting for new and all matches in the next subsection. Panel (b) of Figure 4 facilitates comparison with popular wage-based approaches. Here, we compute the same rank correlation coefficients using the identical worker-ability ranks but firm ranks based on estimated AKM firm-fixed effects. Notably, this wage-based firm ranking implies higher rank correlations (more positive sorting) and a significant increase over time. The correlation for all matches reaches 0.3 towards the end of our period of analysis. These numbers are close in magnitude to the findings of CHK. For new matches, correlations are even higher and just below 0.4. New matches out of non-employment also exhibit the highest rank correlations with the AKM-based firm ranking.

5.2 Distributional Dynamics

Rank correlations are only a summary statistic for the allocation of workers to firms. To investigate more closely which worker-firm-type combinations contribute to changes of labor market sorting, we study which worker types the different firm types hire and how this has changed over time. This allows us to track precisely which worker types became “more sorted” and “less sorted” in their allocation to different firm types. Appendix Figure C.3 shows the full joint distribution of matches. Here, we find that the density is positive for all worker-firm-type combinations, in line with the technical assumption we make to linearize the worker’s outside option in Section 2.

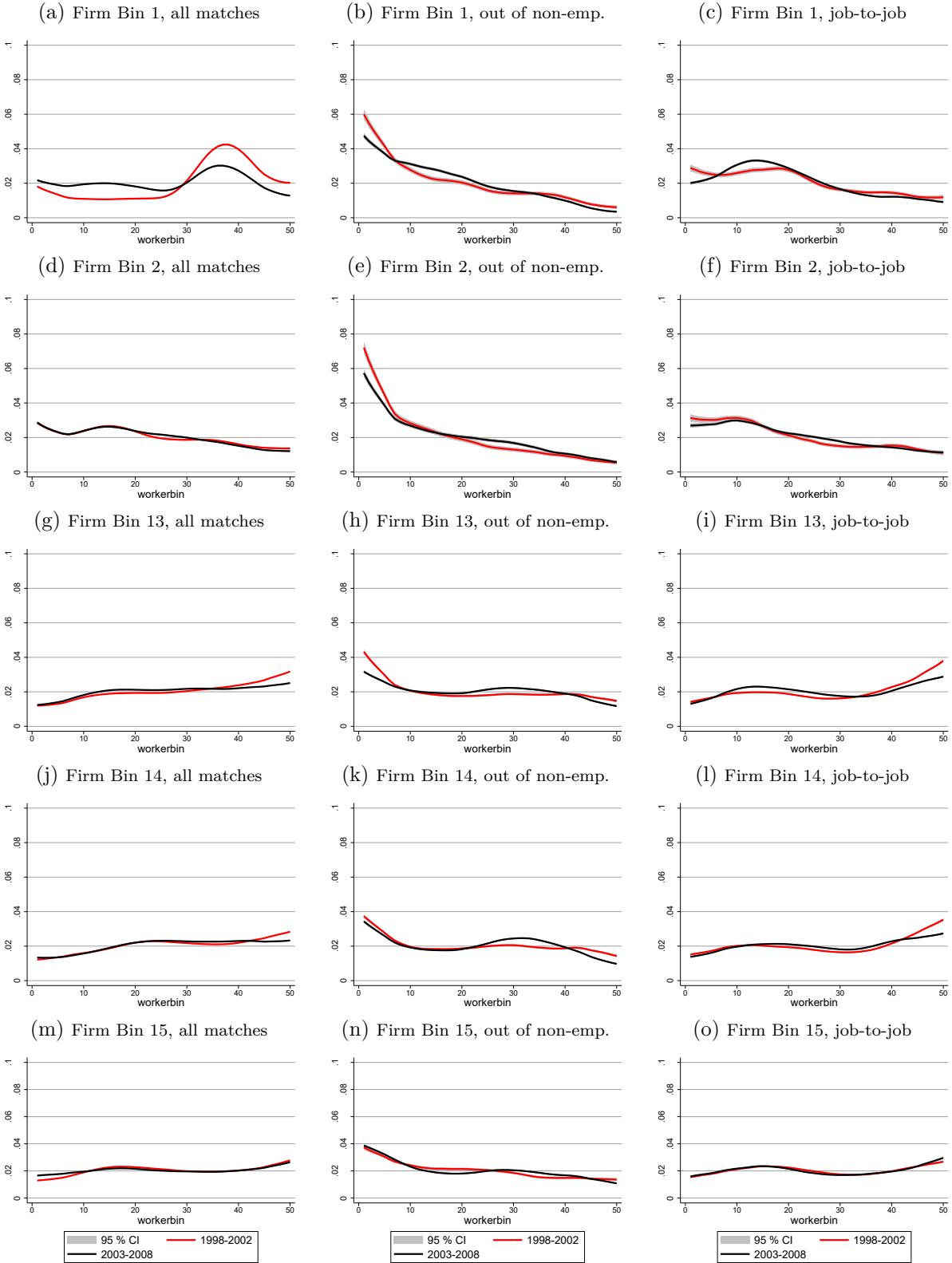
Figure 5 presents estimated univariate density functions of employed worker types for low-productivity firms (bins 1–2) and high-productivity firms (bins 13–15).⁴³ We compare the first half of our sample, 1998–2002 (red line), to the second half, 2003–2008 (black line), and show estimated densities for all matches, new matches out of non-employment, and job-to-job flows, along with 95% confidence intervals.⁴⁴

Low-type firms (bins 1–2) hire mainly low-ability workers out of non-employment. This is where the density of new matches is highest, and it is clearly falling in the worker type. This finding is consistent with PAM. However, we also find that low-productivity firms have increased the average worker quality of new hires over time. In Panels (b), (c), (e), and (f), we observe that the black densities for the second half of our sample display significantly higher values for medium-ability workers both out of non-employment and due to job-to-job switches as compared to the red densities. At the same time, the densities for low-ability workers decreased. Looking at the sample of all matches, the density of worker types hired by bin 1 firms (Panel (a)) has become more uniform over

⁴³Estimated densities for the remaining firm bins are available upon request.

⁴⁴In the plots, statistical significance can be determined based on the overlap of confidence intervals. This is a conservative approach: it is always true that with non-overlapping confidence intervals, two statistics are significantly different from each other. However, an overlap of the confidence intervals does not necessarily imply an insignificant difference.

Figure 5: Changes of Worker Type Distributions in Different Firm Bins



Notes: Estimated univariate kernel densities of all new matches conditional on worker bins, time, and match type. The kernel is estimated using an Epanechnikov kernel function. The bandwidth is calculated by Silverman's rule of thumb. Pointwise confidence intervals are calculated using a quantile of the standard normal distribution. Data sources: BHP, EP, BeH.

time. That is, in addition to a higher quality of new matches, these firms also have sizable outflows of high-ability workers. Taken together, the changes at low-productivity firms imply a trend towards less PAM in the labor market, consistent with the somewhat decreasing aggregate rank correlation we find.

For the most productive firms in bins 13–15, we observe increased hiring of medium-ability workers both through poaching and out of non-employment, see Panels (h)/(i), (k)/(l), and (n)/(o). At the same time, the density of new hires with high-ability workers has decreased. These changes also contribute to less PAM. Decreasing sorting of high-ability workers into high-productivity firms is also visible in the all matches samples for firm bins 13 and 14, see Panels (g) and (j). Overall, the extent of the described density changes is decreasing in firm productivity. At the most productive firms in bin 15, we see a small and significant increase of new matches though poaching of high-ability workers.

In summary, the distributional analysis reveals that worker sorting to the top of the firm-productivity distribution has decreased over time. High-ability workers have to some extent been replaced by medium-ability workers. At the same time, low-productivity firms reduced their hiring of low-ability worker types and increased the quality of their average worker. The relatively stable positive rank correlation we observe is therefore truly the result of two opposed trends: reduced sorting at the bottom and the top (fewer low-low and high-high matches) and increased sorting in the middle (more medium-medium matches). In the final section, we analyze how these changes in the allocation of workers to firms are related to the wages that workers earn at different firms.

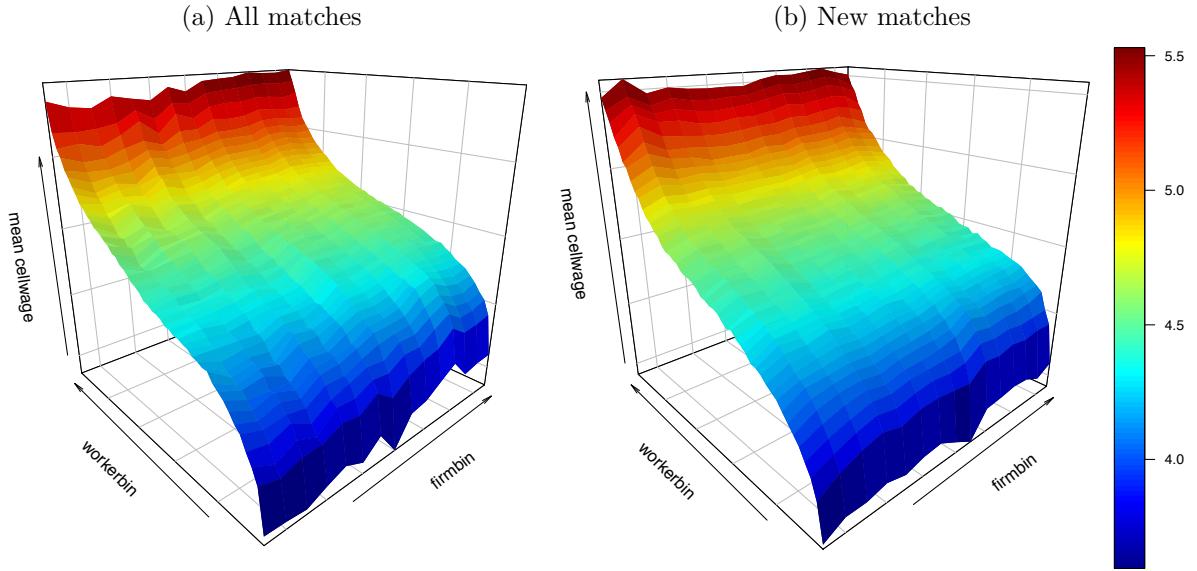
6 Wages and Inequality

First, we investigate how wages vary across worker and firm types. Second, we check whether observed worker transitions are directed towards higher wages and/or higher firm productivity. Third, we decompose the wage inequality trend using our productivity-based firm types and ability-based worker types to reassess the role of firms for wage inequality in Germany.

6.1 Wage Variation across Worker and Firm Bins

According to our theory, worker ability and firm productivity are complements in production. Their interaction determines both the output and the wage of a match. An implication is that, for a given worker type, moving to a more productive firm does not necessarily lead to a higher wage because the match-specific wage component might decrease. Many authors before us have argued that the presence of such complementarities

Figure 6: Mean Wages for all Worker-Firm Type Combinations (1998–2008)



Notes: The plots show the mean of the log real daily wage for all combinations of worker and firm types for both all matches and new matches on a grid with dimensions 50×15 (#worker types \times #firm types). Data sources: BHP, EP, BeH.

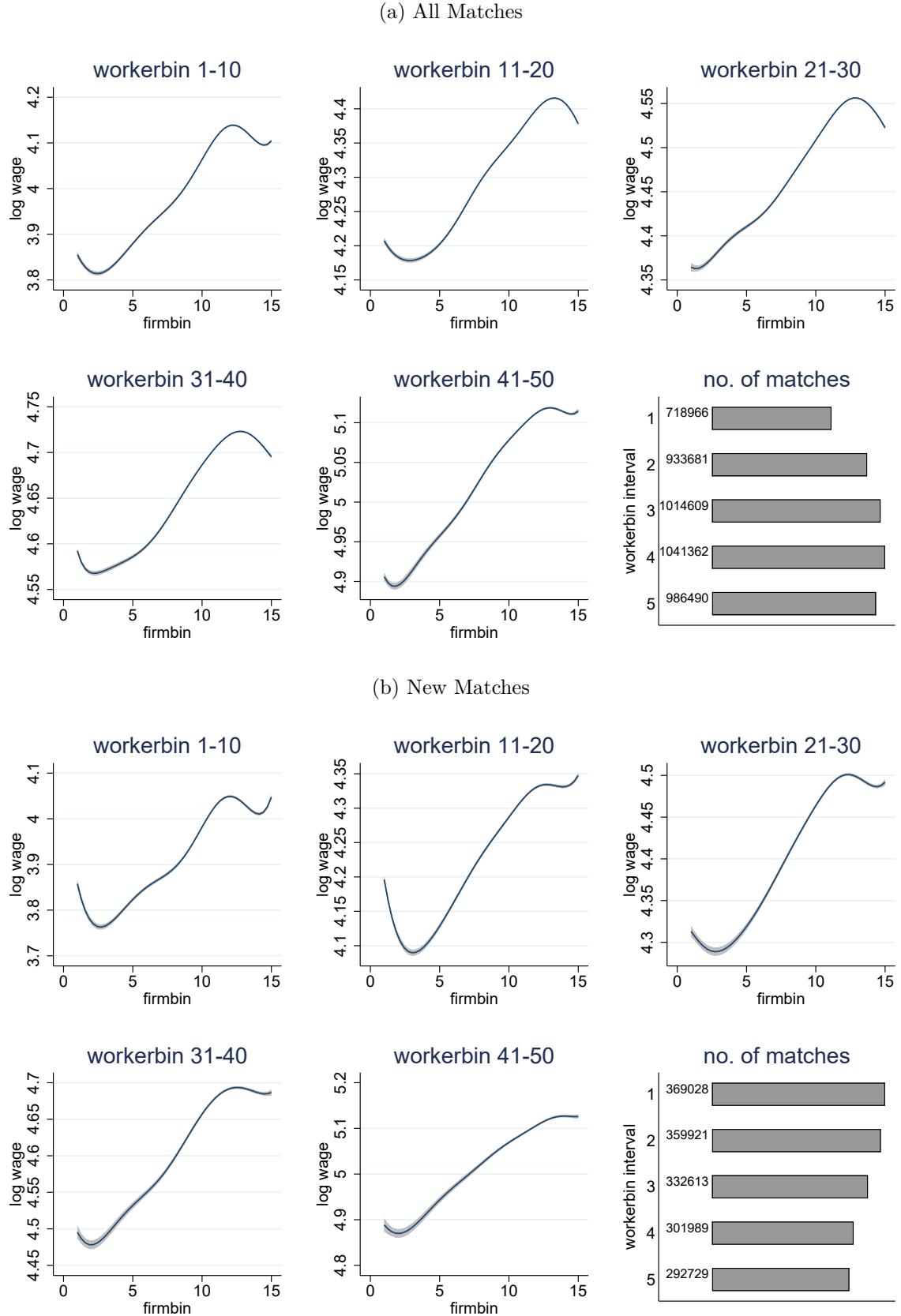
implies that the wage of a given worker type is non-monotonic in the firm type.⁴⁵ By contrast, the log-linear AKM wage equation assumes that wages always increase in the firm effect. Our approach of controlling for worker-firm-specific effects when estimating firm productivity allows for a simple test of wage monotonicity. We can simply plot wages across firm types for different worker types and examine the resulting wage profiles.

Figure 6 plots mean wages for all combinations of worker and firm types. Both for all matches and new matches, we find that log wages across cells appear to be well-approximable by a log-linear function. Wages increase strongly in the worker type but are rather flat in the firm dimension. This is consistent with the broader literature on wage dispersion: the steep increase in the worker dimension confirms that worker heterogeneity is the dominant source of wage dispersion. We also observe that mean wages fall in the firm dimension in some regions of the two plots, e.g., for high-type workers at low-type firms (upper-left corner) and high-type firms (top corner, new matches).

In the next step, we hold worker types fixed to analyze wage variation in firm productivity for specific worker types. In Figure 7, we consider groups of ten worker bins and plot the ordered set of mean wages that these workers receive at different firm types (wage profiles). All matches are shown in Panel (a) and new matches in Panel (b). We relegate plots for matches out of non-employment and job-to-job moves to Appendix Figure C.4. The broad patterns we discuss here do not depend on this distinction.

⁴⁵See, among others, Gautier and Teulings (2006), Eeckhout and Kircher (2011), Lise et al. (2016), Hagedorn et al. (2017), Lopes de Melo (2018), and Bagger and Lentz (2019).

Figure 7: Mean Wages across Worker and Firm Types



Notes: Plots show estimated wage profiles across firm bins for all matches (a) and new matches (b). The kernel is estimated using a Gaussian kernel function. The bandwidth is 2. 95% confidence bands in gray. Data sources: BHP, EP, BeH.

Figure 7 reveals quantitatively important deviations from monotonicity in both samples. Most wage profiles resemble a characteristic S-shape. Low-productivity firms pay relatively high wages. The lowest wages are typically paid around firm bins 3–4. Wages then increase monotonically up until firm bins 11–12 and decrease thereafter. This wage drop at the most productive firms is decreasing in worker ability. It is almost non-existent for new matches of high-ability workers in bins 41–50, although the wage still levels off. This is consistent with PAM. Finding lower wages at the most productive firms as compared to somewhat less productive firms is also consistent with the decrease of sorting between high-ability workers and high-productivity firms documented in Section 5.2.

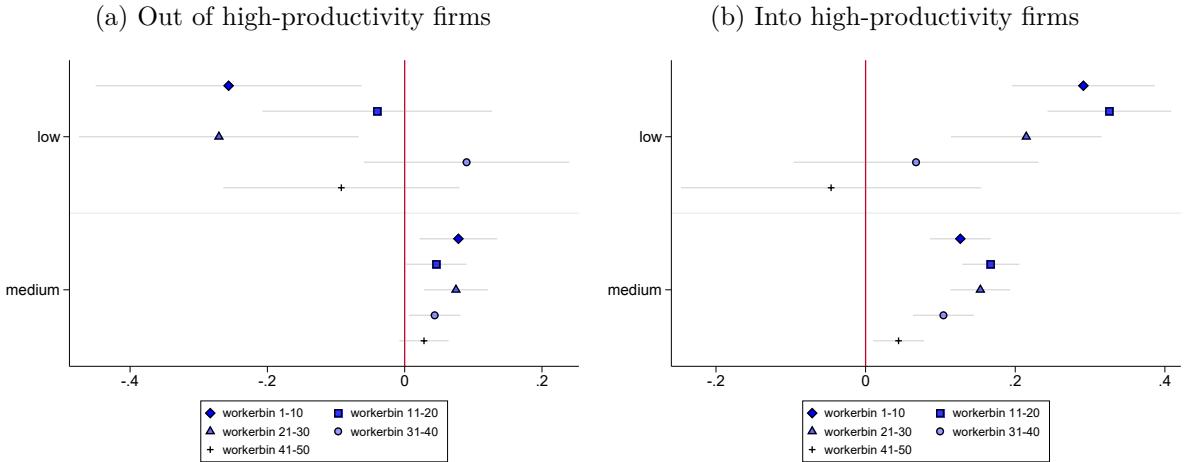
For the majority of worker types, wages are maximized around firm bin 12. We can relate this finding back to Figure 2, Section 4.2. The firms that pay the most are also the biggest in terms of headcount. Firms in the right tail of the productivity distribution are smaller, pay less on average, and also exhibit the lowest labor shares. The investigation of the wage setting mechanism at these firms is of particular interest. According to the model in Section 2, lower wages at these firms are either due to unexploited match-level complementarities or low worker outside options. Other explanations consistent with our results are monopsony power of high-productivity firms or positive amenities that make workers willing to accept lower wages. We confirm in Appendix Figure C.5 that the non-monotonicities we find are not driven by tenure effects. If anything, the wage drop at the top is more pronounced when considering the first match-year only. This implies that wages increase over time to alleviate lower starting wages at high productivity firms, especially for high-ability workers. Moreover, Figure C.6 shows that wage-based firm types constructed using AKM firm-fixed effects do not reveal any non-monotonicities.

Quantitatively, the wage drop is most pronounced for medium-ability workers. In the sample of all matches, the average wage loss from being employed in firm bin 15 instead of the best-paying firm for a bin 11-20 worker in terms of (deflated) log daily wages is about 4%. On a yearly basis, this translates into a wage loss of approximately 1,177 euros. We also find sizable non-monotonicities at low-productivity firms. Here, the least productive firms pay higher wages than firms ranked just above to workers of the same ability type. For new matches of workers in bins 11–20, the wage difference between working in a bin 1 firm and the lowest wage is roughly 10%. This translates into approximately 2,316 euros.

6.2 Worker Transitions

The negative wage effects (non-monotonicities) we depict in Figure 7 are simply mean wage differences across firm types for a given set of worker types. The next step is to check whether observed workers transitions between firm bins are consistent with the estimated wage profiles. The question is: do workers transition towards higher wages, even if this implies switching to a less-productive employer? Suppose this was the case.

Figure 8: Wage Changes for Observed Transitions



Notes: The plots show estimated coefficients and 95% confidence intervals (robust standard errors) from a linear regression of individual-level wage differences of transitioning workers on dummies for origin and destination firm bins. The sample consists of new matches (job-to-job switches, no intermittent non-employment spell) for five groups of worker types. The depicted coefficients are for transitions out of (Panel (a)) and into (Panel (b)) high-productivity firms (bins 13–15). The vertical axis captures destination/origin firm bin groups: low (bins 1–3) and center (bins 4–12). Data sources: BHP, EP, BeH.

Observed transitions out of the most productive firms should be accompanied by wage gains, at least for a subset of destination firms. Similarly, moving up the productivity ladder should yield wage gains, but one would expect them to be lower for higher origin firm bins of the transition.

Figure 8 shows that wage changes in observed transitions support our conjecture. Here, we regress log-wage differences for the same five groups of worker bins on a set of origin and destination firm bin dummies.⁴⁶ To simplify the graphical illustration, we also group firm bins into high (13–15), center (4–12), and low (1–3). Panel (a) shows that low and medium-type workers experience significant wage gains of 7.8% (bins 1–10), 4.6% (bins 11–20), 7.5% (bins 21–30), and 4.4% (bins 31–40). For the highest worker types, the wage differences are not significantly different from zero. Wage differences for transitions from high-productivity to low-productivity firms are quite noisy and in most cases not significantly different from zero. Overall, we find that transitions down the firm-productivity ladder lead to wage gains for many worker types and even transitions into low-productivity firms do not necessarily come with a negative wage change, in line with the idea that workers tend to transition towards higher wages. These findings support the idea that job mobility reflects endogenous, wage-based choices of workers (endogenous mobility).⁴⁷ Furthermore, the fact that transitions from high to medium

⁴⁶We measure the difference between the wage in the last spell in the pre-transition firm and the wage in the first spell in the post-transition firm.

⁴⁷CHK also test endogenous mobility, see their Figure V, p. 984. Here, origin and destination firms of transitioning workers are characterized by the quartile of the mean wage of coworkers at both firms. Using this wage-based classification of firms, CHK find that workers who move to firms with lower

productivity firms yield significant wage gains is in line with the non-monotonic wage profiles we document, and it can explain the trend of decreasing productivity sorting at the top (Section 5.2).

Panel (b) presents estimated wage changes for transitions into high-productivity firms in a similar manner. As conjectured, we find significant positive wage changes for most worker types, which tend to decrease in the firm types of origin, although there is some variation in the estimated wage coefficients. For workers in bins 31–50, we cannot reject the hypothesis that transitioning from a low to a high-productivity firm yields no wage gain. This observation is in line with S-shapes wage profiles. Low to medium-ability workers (1–30) can expect significant wage gains of more than 20% when moving from a low-productivity to a high-productivity firm. When the origin firm is of medium productivity (center), wage gains lie in the range between 4.4% and 16.7% and are decreasing in worker ability. Based on the non-monotonicities documented in Figure 7, one could have expected that some of these transitions yield wage losses. This is not the case, reaffirming that workers tend to move towards higher wages.

In summary, wage changes in observed transitions show that moves down the productivity ladder reflect the non-monotonic wage patterns in Figure 7. Workers can indeed increase their wages by moving to less-productive firms. Upward transitions tend to yield positive wage effects. We do not observe that workers transition to more productive firms at the cost of lower wages. Workers appear to select jobs to maximize their wages, even when the transition implies a move down the firm-productivity ladder.

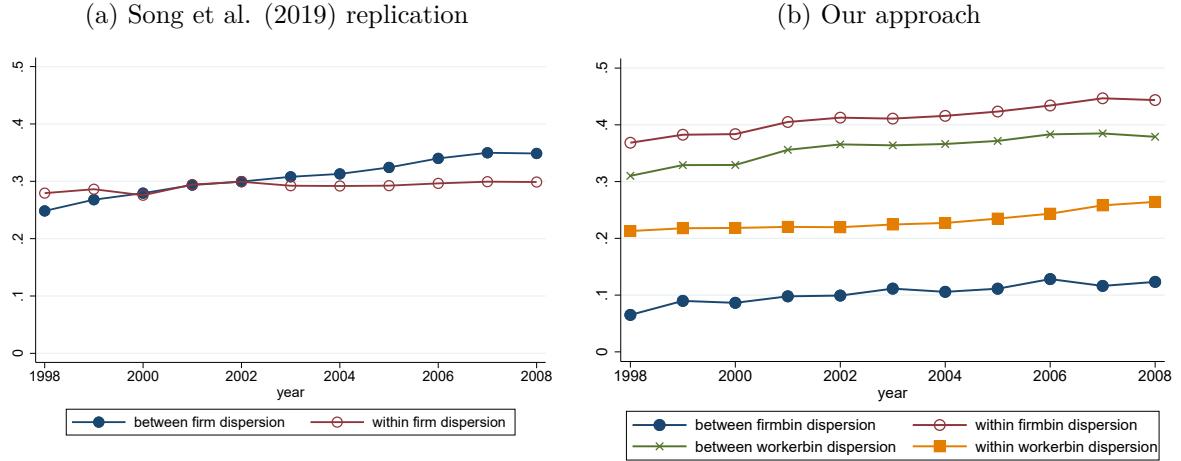
6.3 Wage Inequality

Song et al. (2019) argue that two thirds of the rise of wage inequality in the U.S. from 1978 to 2013 can be attributed to rising pay differences between firms. Using the AKM approach, they show that increasing wage sorting (high-wage workers into high-wage firms) and increasing segregation of workers contribute roughly equally to the firms' rising contribution to wage inequality. In the final step of our analysis, we check whether we can detect the same development in the German data between 1998 and 2008. Moreover, we decompose wage inequality based on our productivity-based firm and ability-based worker types.

Panel (a) of Figure 9 reveals that the Song et al. (2019) finding is also borne out by German data. Here, we simply use establishment identifiers to decompose the variance of wages into the respective shares explained within and between establishments. Wage inequality is increasing because the between-firm contribution (blue line) is growing by approximately 10% during our period of analysis. The within-establishment contribution (red line) to wage inequality is of similar magnitude but stable over time.

coworker mean wages indeed suffer wage losses.

Figure 9: Decomposition of Wage Dispersion over Time



Notes: Panel (a) shows yearly decompositions of wage dispersion using establishment identifiers. Panel (b) shows yearly decompositions of wage dispersion into the respective contributions within and between estimated worker bins and within and between estimated firm bins. Data sources: BHP, EP, BeH.

Panel (b) presents our decomposition. The wage variance between the firm bins (blue line) does increase but only by approximately 3%. It contributes little to overall inequality. The wage variance within firm bins (red line) is quantitatively more important (recall Table 4). Worker segregation, that is, wage differences between the worker bins (green line), is also a major contributing factor. Both components are increasing over time at a higher rate than between-firm-bin inequality. The contribution of within-worker-bin inequality (orange line) is moderately large and also increasing, but only in the second half of our sample. We conclude that rising pay differences between firm-productivity types are not the main driving force behind increasing wage inequality in Germany. Worker segregation and more dispersed wage distributions within firm-productivity types cause the rise.

Finally, to shed more light on the increasing between-workerbin and between-firmbin components of wage inequality in Germany, we show how the wage profiles depicted in Figure 7 have changed over time. Figure C.7 depicts them for the two sub-periods considered earlier, 1998–2002 and 2003–2008, for all matches as well as new matches out of non-employment and job-to-job switches. We observe large differences in wage growth across worker types. These differences correspond to increasing worker segregation, that is, the increasing between-workerbin variance component in Figure 9b. For low-type workers (bins 1–10), wages decreased in matches with all firm types above the very bottom and by more than 10% (all matches, firm bin 2–4). For the worker bins just above (11–20), wages were shrinking at low-productivity firms but relatively stable at more productive firms. For high-type workers (bins 41–50), wages increased in matches with almost all firm types and most notably at the top by more than 7% (all matches, firm bin

15). For medium-type workers, (bins 21–40), wages increased at more-productive firms and decreased at less-productive firms. Furthermore, the non-monotonic wage humps at the most productive firms became more pronounced over time, and the wage profiles have become steeper. More pronounced non-monotonicities at the top can explain the decreasing sorting of high-ability workers into high-productivity firms we documented above. The higher slope of wages across firm types corresponds to increasing between-firm bin wage inequality in Figure 9b.

7 Conclusions

Based on the wage equation of a sorting model with multi-worker firms, two-sided heterogeneity, and random search, we propose a novel strategy to estimate unobserved firm productivity. Our approach is inspired by the empirical IO literature and relies on estimated AKM wage components to measure the contribution of heterogeneous worker ability to output, which is worker-firm-specific in the presence of complementarities at the match level. Moreover, using predicted wage bills, the firms' labor input is made comparable across firms. Thereby, it facilitates the unbiased estimation of firm productivity. Based on our estimates, we rank firms and study productivity sorting, wages, and inequality trends in the German labor market.

Our analysis reveals a number of novel empirical facts. Productivity sorting is positive and relatively stable over time. At the most productive firms, however, sorting is decreasing as high-ability workers become more likely to be matched with slightly less productive firms that pay them higher wages. Observed transitions confirm that most workers' transitions are directed towards higher wages, also in cases where this implies moving down the firm-productivity ladder. Taken together, these findings imply that wages are not everywhere monotonically increasing in firm productivity. At the most productive firms, wages tend to decrease for almost all worker types. Moreover, low-productivity firms sometimes pay higher wages than somewhat more productive firms, perhaps to grow or to retain workers.

The finding that the most productive firms do not pay the highest wages enhances our understanding of increasing wage sorting. If workers move away from high-productivity firms to increase their wages, a side effect of increasing wage sorting could be decreasing allocative efficiency and lower aggregate output. Thus, we argue that our approach is a useful complement to the wage-based analysis of the allocation of workers to firms in the labor market. For future research, it seems promising to analyze the link between labor market sorting and increasing wage inequality, on the one hand, and much-debated trends like falling labor shares, rising market concentration, and increasing monopoly and monopsony power of certain firms, on the other hand.

References

- Abowd, John M., Robert H. Creecy, and Francis Kramarz (2002) “Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data,” *Longitudinal Employer-Household Dynamics Technical Papers 2002-06*, Center for Economic Studies, U.S. Census Bureau.
- Abowd, John M., Francis Kramarz, and David N. Margolis (1999) “High Wage Workers and High Wage Firms,” *Econometrica*, Vol. 67, pp. 251–333.
- Ackerberg, Daniel A., Kevin Caves, and Garth Frazer (2015) “Identification Properties of Recent Production Function Estimators,” *Econometrica*, Vol. 83, pp. 2411–2451.
- Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008) “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)*, Vol. 171, pp. 673–697.
- Andrews, M.J., L. Gill, T. Schank, and R. Upward (2012) “High Wage Workers Match with High Wage Firms: Clear Evidence of the Effects of Limited Mobility Bias,” *Economics Letters*, Vol. 117, pp. 824–827.
- Atakan, Alp E. (2006) “Assortative Matching with Explicit Search Costs,” *Econometrica*, Vol. 74, pp. 667–680.
- Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen (2017) “Concentrating on the Fall of the Labor Share,” *American Economic Review*, Vol. 107, pp. 180–85.
- Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen (2020) “The Fall of the Labor Share and the Rise of Superstar Firms,” *The Quarterly Journal of Economics*, Vol. 135, pp. 645–709.
- Bagger, Jesper, Bent Jesper Christensen, and Dale T. Mortensen (2014) “Productivity and Wage Dispersion: Heterogeneity or Misallocation?”, unpublished.
- Bagger, Jesper and Rasmus Lentz (2019) “An Empirical Model of Wage Dispersion with Sorting,” *Review of Economic Studies*, Vol. 86, pp. 153–190.
- Bagger, Jesper, Kenneth L. Sørensen, and Rune Vejlin (2013) “Wage sorting trends,” *Economics Letters*, Vol. 118, pp. 63–67.
- Bartolucci, Cristian, Francesco Devicienti, and Ignacio Monzón (2018) “Identifying Sorting in Practice,” *American Economic Journal: Applied Economics*, Vol. 10, pp. 408–38.

Becker, Gary S. (1973) “A Theory of Marriage: Part I,” *Journal of Political Economy*, Vol. 81, pp. 813–846.

Bell, Linda A. and Richard B. Freeman (1991) “The Causes of Increasing Interindustry Wage Dispersion in the United States,” *ILR Review*, Vol. 44, pp. 275–287.

Bertheau, Antoine, Henning Bunzel, and Rune Vejlin (2020) “Employment Reallocation over the Business Cycle: Evidence from Danish Data,” IZA Discussion Paper 13681, IZA Institute of Labor Economics.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa (2019) “A Distributional Framework for Matched Employer Employee Data,” *Econometrica*, Vol. 87, pp. 699–739.

Borovičková, Katarína and Robert Shimer (2020) “High Wage Workers Work for High Wage Firms,” unpublished.

Bruns, Benjamin (2019) “Changes in Workplace Heterogeneity and How They Widen the Gender Wage Gap,” *American Economic Journal: Applied Economics*, Vol. 11, pp. 74–113.

Cahuc, Pierre, Francois Marque, and Etienne Wasmer (2008) “A Theory of Wages and Labor Demand With Intra-Firm Bargaining and Matching Frictions,” *International Economic Review*, Vol. 49, pp. 943–972.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline (2018) “Firms and Labor Market Inequality: Evidence and Some Theory,” *Journal of Labor Economics*, Vol. 36, pp. S13–S70.

Card, David, Ana Rute Cardoso, and Patrick Kline (2016) “Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women,” *The Quarterly Journal of Economics*, Vol. 131, pp. 633–686.

Card, David, Jörg Heining, and Patrick Kline (2013) “Workplace Heterogeneity and the Rise of West German Wage Inequality,” *The Quarterly Journal of Economics*, Vol. 128, pp. 967–1015.

Chan, Mons, Sergio Salgado, and Ming Xu (2021) “Heterogeneous Passthrough from TFP to Wages,” January, unpublished.

Davis, Steve J., John Haltiwanger, Lawrence F. Katz, and Robert Topel (1991) “Wage Dispersion between and within U.S. Manufacturing Plants, 1963-86,” *Brookings Papers on Economic Activity*, Vol. 1991, pp. 115–200.

- Dickens, William T and Lawrence F Katz (1987) “Inter-Industry Wage Differences and Theories of Wage Determination,” Working Paper 2271, National Bureau of Economic Research.
- Doraszelski, Ulrich and Jordi Jaumandreu (2013) “R&D and Productivity: Estimating Endogenous Productivity,” *The Review of Economic Studies*, Vol. 80, pp. 1338–1383.
- Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg (2009) “Revisiting the German Wage Structure,” *The Quarterly Journal of Economics*, Vol. 124, pp. 843–881.
- EEckhout, Jan and Philipp Kircher (2011) “Identifying Sorting-In Theory,” *Review of Economic Studies*, Vol. 78, pp. 872–906.
- (2018) “Assortative Matching With Large Firms,” *Econometrica*, Vol. 86, pp. 85–132.
- Fischer, Gabriele, Florian Janik, Dana Müller, and Alexandra Schmucker (2009) “The IAB Establishment Panel - Things Users Should Know,” *Schmollers Jahrbuch: Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften*, Vol. 129, pp. 133–148.
- Fitzenberger, Bernd, Aderonke Osikominu, and Robert Völter (2006) “Imputation Rules to Improve the Education Variable in the IAB Employment Subsample,” *Schmollers Jahrbuch: Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften*, Vol. 126, pp. 405–436.
- Fox, Jeremy T. and Valérie Smeets (2011) “Does Input Quality Drive Measured Differences In Firm Productivity?,” *International Economic Review*, Vol. 52, pp. 961–989.
- Gandhi, Amit, Salvador Navarro, and David A. Rivers (2020) “On the Identification of Gross Output Production Functions,” *Journal of Political Economy*, Vol. 128, pp. 2973–3016.
- Gautier, Pieter A. and Coen N. Teulings (2006) “How large are Search Frictions?” *Journal of the European Economic Association*, Vol. 4, pp. 1193–1225.
- Gibbons, Robert and Lawrence Katz (1992) “Does Unmeasured Ability Explain Inter-Industry Wage Differentials?,” *Review of Economic Studies*, Vol. 59, pp. 515–535.
- Griliches, Zvi (1957) “Specification Bias in Estimates of Production Functions,” *Journal of Farm Economics*, Vol. 39, pp. 8–20.
- Groshen, Erica L. (1991) “Sources of Intra-Industry Wage Dispersion: How Much Do Employers Matter?” *The Quarterly Journal of Economics*, Vol. 106, pp. 869–884.

- Guiso, Luigi, Luigi Pistaferri, and Fabiano Schivardi (2005) “Insurance within the Firm,” *Journal of Political Economy*, Vol. 113, pp. 1054–1087.
- Hagedorn, Marcus, Tzuo Hann Law, and Iourii Manovskii (2017) “Identifying Equilibrium Models of Labor Market Sorting,” *Econometrica*, Vol. 85, pp. 29–65.
- Haltiwanger, John, Henry Hyatt, and Erika McEntarfer (2018) “Who Moves Up the Job Ladder?” *Journal of Labor Economics*, Vol. 36, pp. S301–S336.
- Irarrazabal, Alfonso, Andreas Moxnes, and Karen Helene Ulltveit-Moe (2013) “Heterogeneous Firms or Heterogeneous Workers? Implications for Exporter Premiums and the Gains from Trade,” *Review of Economics and Statistics*, Vol. 95, pp. 839–849.
- Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten (2020) “Leave-Out Estimation of Variance Components,” *Econometrica*, Vol. 88, pp. 1859–1898.
- Kölling, Arnd (2000) “The IAB-Establishment Panel,” *Schmollers Jahrbuch: Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften*, Vol. 120, pp. 291–300.
- Krueger, Alan B. and Lawrence H. Summers (1988) “Efficiency Wages and the Inter-Industry Wage Structure,” *Econometrica*, Vol. 56, pp. 259–293.
- Lentz, Rasmus, Suphanit Piyapromdee, and Jean-Marc Robin (2018) “On Worker and Firm Heterogeneity in Wages and Employment Mobility: Evidence from Danish Register Data,” unpublished.
- Levinsohn, James and Amil Petrin (2003) “Estimating Production Functions Using Inputs to Control for Unobservables,” *Review of Economic Studies*, Vol. 70, pp. 317–341.
- Lindenlaub, Ilse (2017) “Sorting Multidimensional Types: Theory and Application,” *Review of Economic Studies*, Vol. 84, pp. 718–789.
- Lindenlaub, Ilse and Fabien Postel-Vinay (2020) “Multidimensional Sorting Under Random Search,” unpublished.
- Lise, Jeremy, Costas Meghir, and Jean-Marc Robin (2016) “Matching, Sorting and Wages,” *Review of Economic Dynamics*, Vol. 19, pp. 63–87.
- Lise, Jeremy and Fabien Postel-Vinay (2020) “Multidimensional Skills, Sorting, and Human Capital Accumulation,” *American Economic Review*, Vol. 110, pp. 2328–76.
- Lise, Jeremy and Jean-Marc Robin (2017) “The Macrodynamics of Sorting between Workers and Firms,” *American Economic Review*, Vol. 107, pp. 1104–1135.

- Lochner, Benjamin and Bastian Schulz (2020) “Firm Productivity, Wages, and Sorting,” IAB Discussion Paper 04/2020, IAB Institute for Employment Research.
- Lochner, Benjamin, Stefan Seth, and Stefanie Wolter (2020) “Decomposing the large firm wage premium in Germany,” *Economics Letters*, Vol. 194, p. 109368.
- Marschak, Jacob and William H. Andrews (1944) “Random Simultaneous Equations and the Theory of Production,” *Econometrica*, Vol. 12, pp. 143–205.
- Lopes de Melo, Rafael (2018) “Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence,” *Journal of Political Economy*, Vol. 126, pp. 313–346.
- Müller, Steffen (2008) “Capital Stock Approximation Using Firm Level Panel Data,” *Jahrbücher für Nationalökonomie und Statistik*, Vol. 224, pp. 357–371.
- Olley, G. Steven and Ariel Pakes (1996) “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica*, Vol. 64, pp. 1263–1297.
- Petrongolo, Barbara and Christopher A. Pissarides (2001) “Looking into the Black Box: A Survey of the Matching Function,” *Journal of Economic Literature*, Vol. 39, pp. 390–431.
- Pissarides, Christopher A. (2000) *Equilibrium Unemployment Theory*: The MIT Press, 2nd edition.
- Postel-Vinay, Fabien and Jean-Marc Robin (2002) “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity,” *Econometrica*, Vol. 70, pp. 2295–2350.
- Shimer, Robert and Lones Smith (2000) “Assortative Matching and Search,” *Econometrica*, Vol. 68, pp. 343–369.
- Smith, Eric (1999) “Search, Concave Production, and Optimal Firm Size,” *Review of Economic Dynamics*, Vol. 2, pp. 456–471.
- Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter (2019) “Firming Up Inequality,” *The Quarterly Journal of Economics*, Vol. 134, pp. 1–50.
- Sorkin, Isaac (2018) “Ranking Firms Using Revealed Preference,” *The Quarterly Journal of Economics*, Vol. 133, pp. 1331–1393.
- Spengler, Anja (2008) “The Establishment History Panel,” *Schmollers Jahrbuch*, Vol. 128, pp. 501–509.
- Stole, Lars A. and Jeffrey Zwiebel (1996) “Intra-firm Bargaining under Non-binding Contracts,” *The Review of Economic Studies*, Vol. 63, pp. 375–410.

Stüber, Heiko and Stefan Seth (2019) “The FDZ sample of the Administrative Wage and Labor Market Flow Panel 1976 - 2014,” FDZ-Datenreport 01/2019, The Research Data Centre (FDZ) at the Institute for Employment Research (IAB), Nürnberg.

Syverson, Chad (2011) “What Determines Productivity?,” *Journal of Economic Literature*, Vol. 49, pp. 326–365.

Sørensen, Torben and Rune Vejlin (2013) “The importance of worker, firm and match effects in the formation of wages,” *Empirical Economics*, Vol. 45, pp. 435–464.

Taber, Christopher and Rune Vejlin (2020) “Estimation of a Roy/Search/Compensating Differential Model of the Labor Market,” *Econometrica*, Vol. 88, pp. 1031–1069.

Woodcock, Simon D. (2015) “Match effects,” *Research in Economics*, Vol. 69, pp. 100–121.

Wooldridge, Jeffrey M. (2009) “On estimating firm-level production functions using proxy variables to control for unobservables,” *Economics Letters*, Vol. 104, pp. 112–114.

Online Appendix

(not for publication)

A Model Details

A.1 Matching Technology

Due to random search, firms cannot target their vacancies to specific worker types. They post vacancies v subject to a productivity-dependent cost $c(\Omega)$. Meetings are generated by a Cobb-Douglas matching function with constant returns to scale (Pissarides, 2000; Petrongolo and Pissarides, 2001). Without loss of generality, let worker ability x and firm productivity Ω be distributed uniformly on $[0, 1]$. Meeting rates are functions of aggregate labor market tightness, $\theta = V/U$, where $V = \int g_v(\Omega) d\Omega$ and $U = \int g_u(x) dx$ are the aggregated numbers of vacancies and unemployed workers, respectively. $q_v(\theta)$ is the rate at which firms meet workers and $q_u(\theta)$ is the rate at which unemployed workers meet vacancies.

$g_v(\Omega)$ is the PDF of vacancies at type Ω firms and $g_u(x)$ is the PDF of unemployed workers of type x . In this environment, a match is not guaranteed conditional on meeting. Suppose a type x worker meets a productivity Ω firm. Both parties may prefer to continue searching in case the match surplus $S(x, L, \Omega)$, defined below, is negative. Note that the surplus depends on the worker type, the firm type, and the total composite labor input. In steady state, existing matches can only end at the exogenous rate δ . Endogenous separations may happen out of steady state in case a shock to firm productivity reduces the surplus with specific worker types below zero.

A.2 The Firm's Problem

In the outlined environment, the profit flow of a firm with productivity Ω solves the following Bellman equation. The firm's problem is to maximize output given the current composite labor input and productivity less the total wage bill and hiring costs. Current employment is a state variable. The firm controls future discounted profits by posting costly vacancies, given its expected evolution of productivity:

$$\Pi(L, \Omega) = \max_v \left\{ F(L, \Omega) - \sum_{x=1}^n w(x, L, \Omega) L_x - vc(\Omega) + \beta \int \Pi(L', \Omega') dG(\Omega' | \Omega) \right\}. \quad (\text{A.1})$$

This profit flow is maximized subject to n constraints that capture the evolution of employment for every worker type x at the firm:

$$L'_x = (1 - \delta)L_x + vq_v(\theta) \frac{g_u(x)}{U} \mu(x, L, \Omega) \quad \forall x. \quad (\text{A.2})$$

$g_u(x)/U$ is the probability that conditional on meeting the worker is of type x . The indicator function $\mu(x, L, \Omega)$ returns the value one if a match of a type x worker and productivity Ω firm with composite labor input L has a strictly positive surplus and zero otherwise. In case $\mu(x, L, \Omega) = 0$ no additional type x workers are hired.

The match surplus is defined as

$$S(x, L, \Omega) = J_x(L, \Omega) + E(x, \Omega) - U(x), \quad (\text{A.3})$$

and depends on three option value equations defined below. Thus, the indicator $\mu(x, L, \Omega)$ is defined as

$$\mu(x, \Omega) = \begin{cases} 1 & \text{if } S(x, L, \Omega) > 0 \\ 0 & \text{if } S(x, L, \Omega) \leq 0. \end{cases} \quad (\text{A.4})$$

Below, we indicate $\mu(x, L, \Omega) = 1$ ($\mu(x, L, \Omega_j) = 0$) by writing $\mu^+(x, L, \Omega)$ ($\mu^-(x, L, \Omega)$).

Optimality Conditions

We closely follow Cahuc et al. (2008) and define the marginal value of an additional worker of type x at a firm with productivity Ω and workforce L as

$$J_x(L, \Omega) = \frac{\partial \Pi(L, \Omega)}{\partial L_x}. \quad (\text{A.5})$$

The marginal product of type x labor (MPL) at a productivity- Ω firm is

$$F_x(L, \Omega) = \frac{\partial F(L, \Omega)}{\partial L_x}. \quad (\text{A.6})$$

The FOC of the maximization problem (A.1) with respect to v is

$$0 = -c(\Omega) + q_v(\theta) \frac{g_u(x)}{U} \mu(x, L, \Omega) J_x(L', \Omega'). \quad (\text{A.7})$$

The envelope theorem implies

$$J_x(L, \Omega) = \frac{\partial F(L, \Omega)}{\partial L_x} - \sum_{k=1}^n L_k \frac{\partial w_k(L, \Omega)}{\partial L_x} - w(x, L, \Omega) + \beta(1 - \delta) J_x(L', \Omega'). \quad (\text{A.8})$$

Assuming a steady state where $L' = L$ and $\Omega' = \Omega$, (A.7) can be rewritten as

$$J_x(L, \Omega) = \frac{c(\Omega)}{q_v(\theta) \frac{g_u(x)}{U} \mu^+(x, L, \Omega)}, \quad (\text{A.9})$$

so for every worker type within the firm's matching set ($\mu(x, L, \Omega) = 1$), the marginal profit is equal to the expected recruitment cost at the optimal level of employment. In

case a type x worker is not part of the firm's matching set, $\mu(x, L, \Omega) = 0$, marginal profits are undefined. Integrating the worker type out of (A.9) yields the firm's expected marginal profit of posting a vacancy.

$$J(L, \Omega) = \frac{c(\Omega)}{q_v(\theta) \int \frac{g_u(x)}{U} \mu^+(x, L, \Omega) dx}, \quad (\text{A.10})$$

Applying the steady state assumption to (A.8) yields

$$J_x(L, \Omega) = \frac{F_x(L, \Omega) - w(x, L, \Omega) - \sum_{k=1}^n L_k \frac{\partial w(k, L, \Omega)}{\partial L_x}}{1 - \beta(1 - \delta)}, \quad (\text{A.11})$$

so the marginal profit can also be expressed as the discounted marginal product, net of the individual wage and net of the effect of the marginal hire on the total wage bill.

Equating (A.9) and (A.11), one gets

$$F_x(L, \Omega) = w(x, L, \Omega) + \frac{c(\Omega)(1 - \beta(1 - \delta))}{q_v(\theta) \frac{g_u(x)}{U} \mu(x, L, \Omega)} + \sum_{k=1}^n L_k \frac{\partial w(k, L, \Omega)}{\partial L_x}, \quad (\text{A.12})$$

so the MPL of worker type x at a (L, Ω) firm equals the wage plus expected turnover costs and the marginal effect of the additional worker on the total wage bill.

A.3 Wage Determination

To derive the wage equation, we rely on the Nash sharing rule

$$\frac{\alpha}{1 - \alpha} J_x(L, \Omega) = E(x, L, \Omega) - U(x), \quad (\text{A.13})$$

where $\alpha \in (0, 1)$ is the workers' common bargaining parameter. The RHS captures the worker's surplus of working at a firm with productivity Ω and workforce L relative to the worker's outside option, the value of unemployment, $U(x)$. The firm's surplus consists of the marginal profits of hiring an additional worker of type x , $J_x(L, \Omega)$, as defined above. Its threat point is to fire the worker and renegotiate wages with all other employees (Stole and Zwiebel, 1996). Following Cahuc et al. (2008), we assume that wages are continuously and instantaneously (re)negotiated, so L is fixed during (re)negotiations.

In steady state, the worker's value of employment is

$$E(x, L, \Omega) = w(x, L, \Omega) + \underbrace{\beta \delta U(x)}_{\text{separation}} + \underbrace{\beta(1 - \delta) E(x, L, \Omega)}_{\text{continued employment}}. \quad (\text{A.14})$$

The value of unemployment is

$$U(x) = b(x) + \underbrace{\beta(1 - q_u(\theta))U(x)}_{\text{no meeting}} + \underbrace{\beta q_u(\theta) \int \frac{g_v(\Omega)}{V} \mu^+(x, L, \Omega) E(x, L, \Omega) d\Omega}_{\text{successful match}} \\ + \underbrace{\beta q_u(\theta) U(x) \int \frac{g_v(\Omega)}{V} \mu^-(x, L, \Omega) d\Omega}_{\text{meet unacceptable firm}}, \quad (\text{A.15})$$

where $b(x)$ is the flow value of unemployment, e.g. the value of increased leisure, home production or unemployment insurance benefits.

Next, we compute the difference $E(x, L, \Omega) - U(x)$ to be plugged into equation (A.13):

$$E(x, L, \Omega) - U(x) = w(x, L, \Omega) + \beta \delta U(x) + \beta(1 - \delta) E(x, L, \Omega) - U(x). \quad (\text{A.16})$$

After adding and subtracting $\beta U(x)$, this can be rearranged to

$$E(x, L, \Omega) - U(x) = \frac{w(x, L, \Omega) - (1 - \beta)U(x)}{1 - \beta(1 - \delta)}, \quad (\text{A.17})$$

which can be combined with (A.13) to get

$$\frac{\alpha}{1 - \alpha} J_x(L, \Omega) = \frac{w(x, L, \Omega) - (1 - \beta)U(x)}{1 - \beta(1 - \delta)}. \quad (\text{A.18})$$

Finally, substituting marginal profits according to equation (A.11) and rearranging yields the wage bargaining outcome:

$$w(x, L, \Omega) = \alpha \left(F_x(L, \Omega) - \sum_{k=1}^n L_k \frac{\partial w(k, L, \Omega)}{\partial L_x} \right) + (1 - \alpha)(1 - \beta)U(x). \quad (\text{A.19})$$

Due to our assumption of perfect substitutability of worker ability units at the firm level, the inframarginal adjustment term $\sum_{k=1}^n L_k \frac{\partial w(k, L, \Omega)}{\partial L_x}$ solely reflects decreasing returns and is unambiguously negative. Moreover, firms instantaneously (re)negotiate with all workers as if they were the marginal worker, so the adjustment does not vary with x . This yields the following simplified differential equation

$$w(x, L, \Omega) = \alpha \left(F_x(L, \Omega) - L \frac{\partial w(k, L, \Omega)}{\partial L} \right) + (1 - \alpha)(1 - \beta)U(x), \quad (\text{A.20})$$

which we can solve following the steps for the “single labor case” described in the Appendix of Cahuc et al. (2008). For the single steps and detailed technical assumptions, see their equations (B.1)–(B.6), p 961–962. The solution is

$$w(x, L, \Omega) = (1 - \alpha)(1 - \beta)U(x) + \int_0^1 z^{\frac{1-\alpha}{\alpha}} F_x(Lz, \Omega) dz. \quad (\text{A.21})$$

A.4 Linearity of the Wage Equation

Plugging in the worker-firm-specific MPL (3) into (A.21) yields

$$w(x, L, \Omega) = (1 - \alpha)(1 - \beta)U(x) + x \int_0^1 z^{\frac{1-\alpha}{\alpha}} \beta_l \frac{F(Lz, \Omega)}{Lz} dz, \quad (\text{A.22})$$

so the integral expression is only scaled by, and hence linear in, worker ability x .

To establish that the full wage equation is linear in x , the outside option $U(x)$ has to be linear in x , too. Consider the outside option according to equation (A.15). A straightforward assumption to ensure that $U(x)$ is indeed linear in x is that all worker types' matching sets cover the whole type space. In other words, conditional on meeting, there are no unacceptable firms. This claim is easily verified empirically, see Section 5. In the model, this implies that $\mu(x, L, \Omega) = 1$ holds for all potential matches and, thus, the last term of (A.15) is zero. Now rearrange equation (A.14) such that

$$E(x, L, \Omega) = \frac{w(x, L, \Omega) + \beta\delta U(x)}{1 - \beta(1 - \delta)}. \quad (\text{A.23})$$

Under our assumption, this can be plugged into equation (A.15) to yield an expression in the wage and the outside option only.

$$U(x) = b(x) + \beta(1 - q_u(\theta))U(x) + \beta q_u(\theta) \int \frac{g_v(\Omega)}{V} \frac{w(x, L, \Omega) + \beta\delta U(x)}{1 - \beta(1 - \delta)} d\Omega. \quad (\text{A.24})$$

Plugging in our solution for the wage, equation (A.22), into this expression and collecting the $U(x)$ terms in front of the integral yields

$$\begin{aligned} U(x) &= b(x) + \beta(1 - q_u(\theta))U(x) + \beta q_u(\theta) \frac{1 - \alpha + \beta(\alpha + \delta - 1)}{1 - \beta(1 - \delta)} U(x) \int \frac{g_v(\Omega)}{V} d\Omega \\ &\quad + \frac{\beta q_u(\theta)}{1 - \beta(1 - \delta)} \int \frac{g_v(\Omega)}{V} x \int_0^1 z^{\frac{1-\alpha}{\alpha}} \beta_l \frac{F(Lz, \Omega)}{Lz} dz d\Omega, \end{aligned} \quad (\text{A.25})$$

where $\int \frac{g_v(\Omega)}{V} d\Omega = 1$. After collecting all $U(x)$ terms on the LHS and dividing, we get the following expression for $U(x)$:

$$U(x) = \frac{(1 - \beta(1 - \delta))b(x) + x \frac{\beta q_u(\theta)}{1 - \beta} \int \frac{g_v(\Omega)}{V} \int_0^1 z^{\frac{1-\alpha}{\alpha}} \beta_l \frac{F(Lz, \Omega)}{Lz} dz d\Omega}{1 - \beta(1 - \delta - \alpha q_u(\theta))}, \quad (\text{A.26})$$

where worker ability x can be written in front of both integral sign. Thus, for $U(x)$ to be linear in x , we additionally have to assume that the workers flow value of unemployment, $b(x)$ is also linear in x , which is a standard assumption.

B Details of Data Preparation

B.1 Wage Imputation

In the BeH data, earnings are right censored at the contribution assessment ceiling ('Beitragsbemessungsgrenze'). This earning limit is given by the statutory pension fund and is adjusted annually due to changes in earnings. First we deflate daily wages by using the CPI. Then, in each year, we identify censored wage observations by comparing wages with the contribution assessment ceiling. We define a wage observation as censored whenever the reported wage is higher than 99% of the censoring threshold.

Following CHK and Dustmann et al. (2009), we fit a series of Tobit regression to impute the right tail of the wage distribution. We estimate Tobit regressions by year, sex, education and age group. In all these regressions we additionally control for the exact age, the mean log wage in other years, the fraction of censored wages in other years, the number of full time employees at the current establishment and its square, an indicator for large firms, the mean years of schooling and the fraction of university graduates at the current establishment, the mean log wage of co-workers and the fraction of co-workers with censored wages, an indicator for individuals observed only one year, an indicator for employees in one-worker establishments, and an indicator for regions. We assume that the error term is normally distributed but each education and age category can have a different variance. For each year, we impute censored wages as the sum of the predicted wage and a random component which is computed based on standard error of the forecast. This component is drawn from separate normal distributions with mean zero and the different variances for each education and age category.

B.2 Education Imputation

The employee education information is reported by employers after every year and whenever a job ends. Its quality may suffer because employers do not face consequences for non- and misreporting. However, the existence of a reporting rule allows for corrections. It prescribes that only the highest educational degree of an employee needs to be reported. Therefore the individual educational attainment should not decline over consecutive job spells. The imputation procedure (IP1) suggested by Fitzenberger et al. (2006) exploits this reporting rule by assuming that there is any over-reporting in the data.

The original education variable distinguishes the following four different educational degrees: high school, vocational training, technical college and university. By imputing following the IP1 procedure we extrapolate both back and forwards and do some additional adjustments using individual information on age and occupational status. As a result we get six education categories which can be ranked in increasing order. We drop the remaining 2% of observations for which education cannot be imputed.

C Additional Results

C.1 Rank Correlations

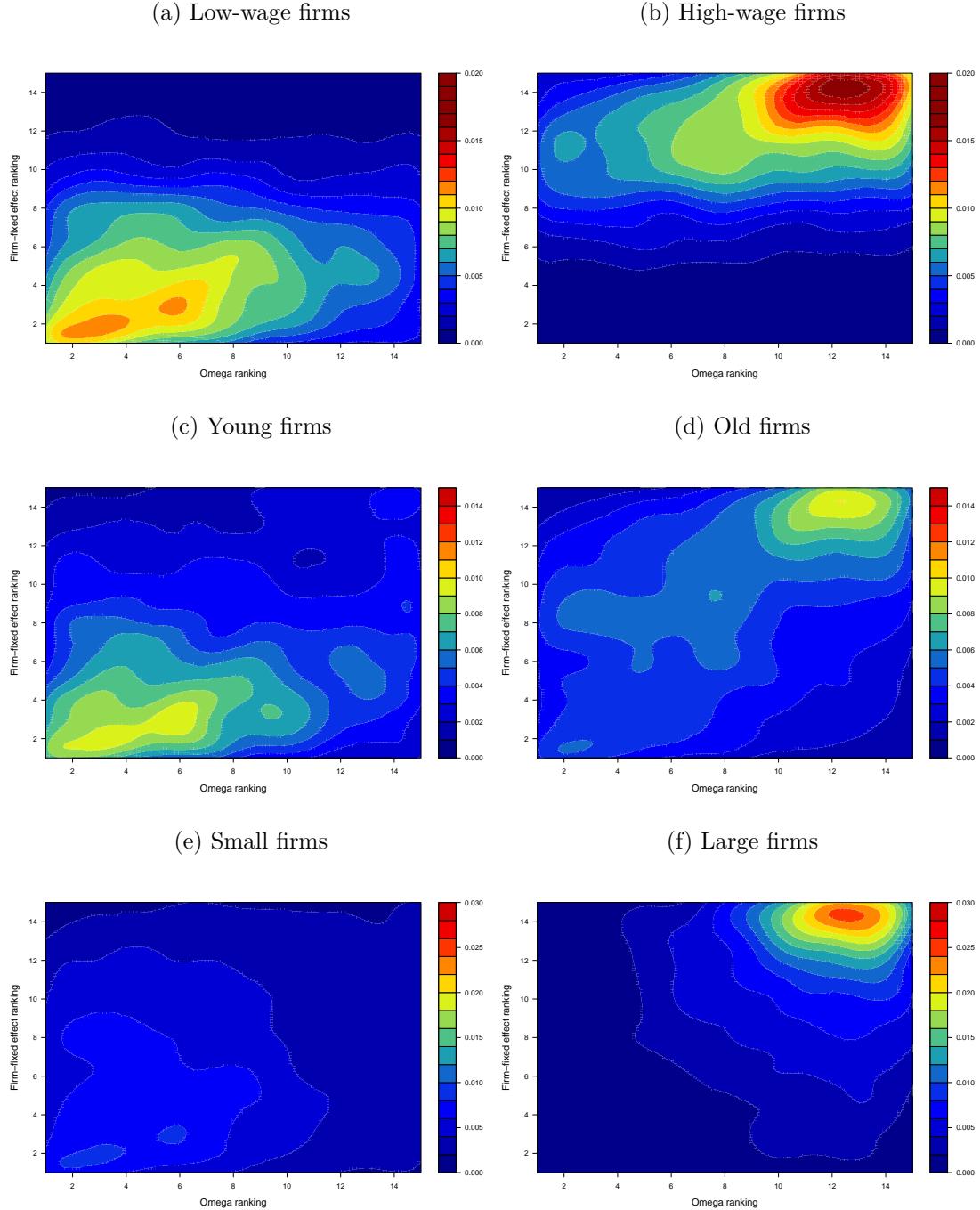
Table C.1: Spearman rank correlation coefficients and numbers of observations for different time intervals and samples

	All Matches	New Matches	Out of Nonemp.	Job-to-Job
1998-2008	0.065 (4,695,108)	0.124 (1,656,280)	0.132 (601,954)	0.110 (1,082,460)
1998-2002	0.055 (2,182,011)	0.139 (474,341)	0.141 (174,310)	0.120 (305,339)
2003-2008	0.074 (2,513,097)	0.118 (1,181,939)	0.129 (427,644)	0.107 (777,121)
1998	0.013 (311,861)	–	–	–
1999	0.046 (338,125)	0.140 (35,865)	0.133 (15,094)	0.129 (20,771)
2000	0.048 (493,323)	0.107 (107,740)	0.108 (41,731)	0.090 (66,009)
2001	0.073 (536,559)	0.148 (158,351)	0.142 (56,627)	0.137 (101,724)
2002	0.077 (502,143)	0.152 (172,385)	0.165 (60,777)	0.134 (111,608)
2003	0.080 (470,279)	0.146 (180,623)	0.156 (64,926)	0.131 (115,697)
2004	0.065 (458,467)	0.114 (191,207)	0.129 (67,932)	0.100 (123,275)
2005	0.081 (428,122)	0.129 (197,755)	0.136 (69,890)	0.118 (127,865)
2006	0.101 (415,153)	0.146 (206,984)	0.139 (74,340)	0.142 (132,644)
2007	0.051 (391,535)	0.084 (209,567)	0.096 (77,811)	0.074 (131,756)
2008	0.063 (349,541)	0.094 (195,803)	0.119 (72,001)	0.076 (123,802)

Notes: In all cells, we test the null hypothesis that worker and firm bins are statistically independent. All rank correlation coefficients are different from 0 at 1% level of significance. Rounded to 3 decimal places. Numbers of observations (matches according to the respective definition) are reported in brackets.

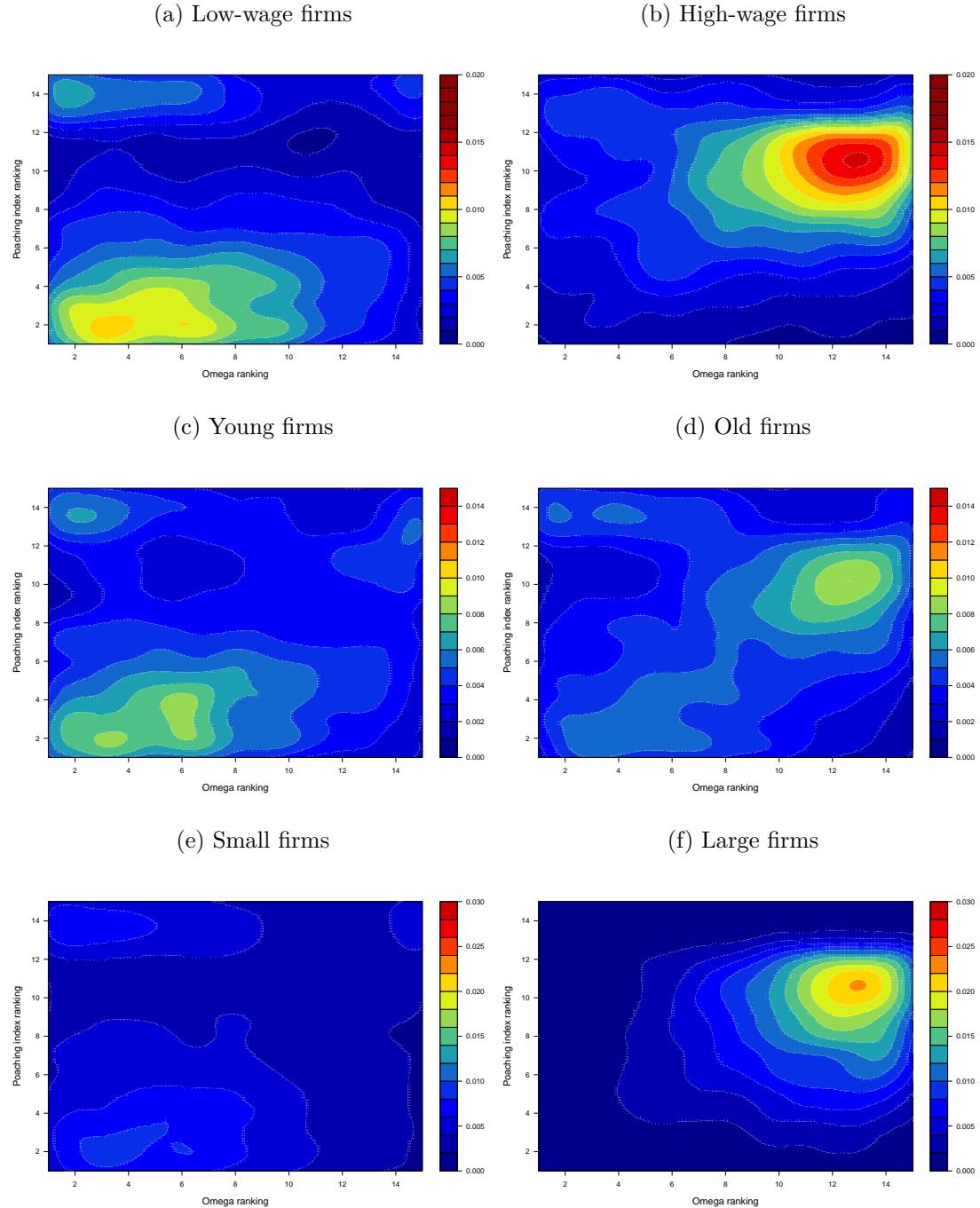
C.2 Ranking Comparisons

Figure C.1: Comparison of Productivity-based and fixed-effect-based Firm Ranking by Wages, Age, and Size



Notes: The six plots depict contours of the joint empirical distributions of firm years across combinations between the omega ranking (15 bins) and AKM firm-effect ranks (15 bins). In Panels (a) and (b), high-wage firms pay more than the grand mean of all firm-level mean wages and low-wage firms pay less. In Panels (c) and (d), the age of young firms is less than 15 years, old firms are 15 years and older. In Panels (e) and (f), small firms have less than 100 employees, large firms have more. Data sources: BHP, EP, BeH.

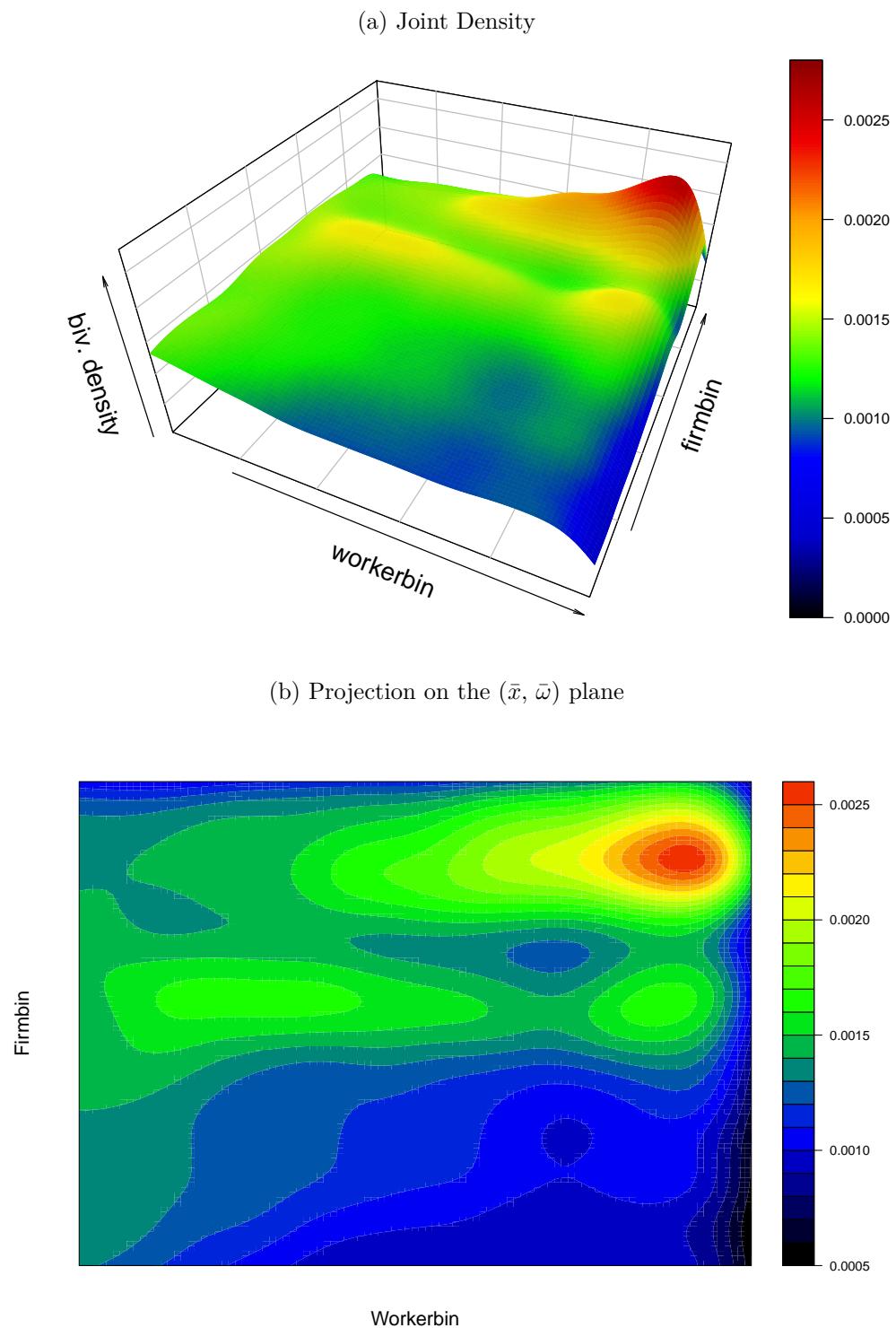
Figure C.2: Comparison of Productivity-based Firm Ranking and poaching-index-based Firm Ranking by Wages, Age, and Size



Notes: The six plots depict contours of the joint empirical distributions of firm years across combinations between the omega ranking (15 bins) and BL poaching index ranks (15 bins). In Panels (a) and (b), high-wage firms pay more than the grand mean of all firm-level mean wages and low-wage firms pay less. In Panels (c) and (d), the age of young firms is less than 15 years, old firms are 15 years and older. In Panels (e) and (f), small firms have less than 100 employees, large firms have more. Data sources: BHP, EP, BeH.

C.3 Joint Distribution of Matches

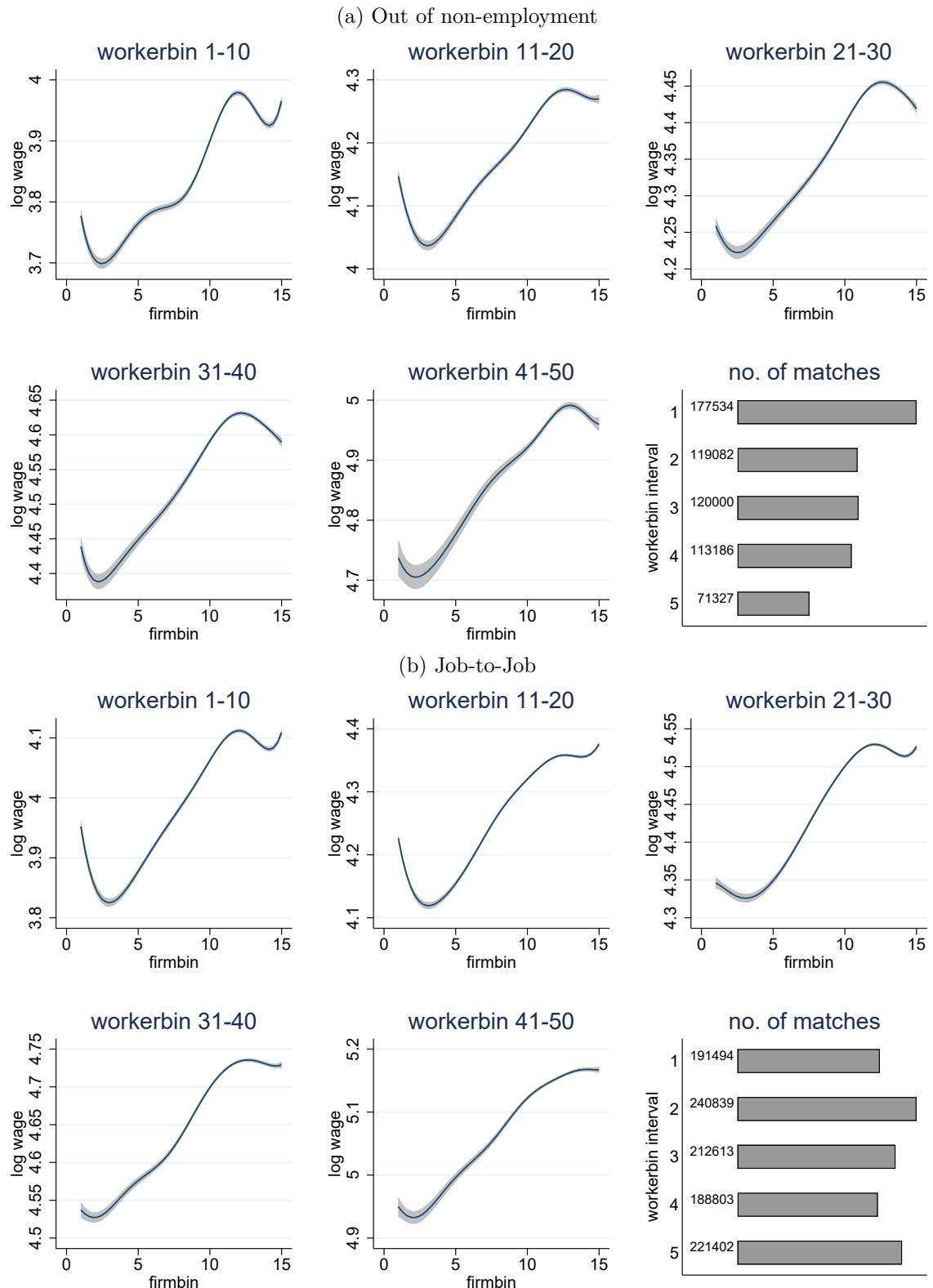
Figure C.3: Joint Distribution for all Worker-Firm Type Combinations (1998–2008)



Notes: The plots shows estimate joint kernel density of matches and its projection on the $(\bar{x}, \bar{\omega})$ plane for all combinations of worker and firm types in the sample of all matches on a grid with dimensions 50×15 (#worker types \times #firm types). Data sources: BHP, EP, BeH.

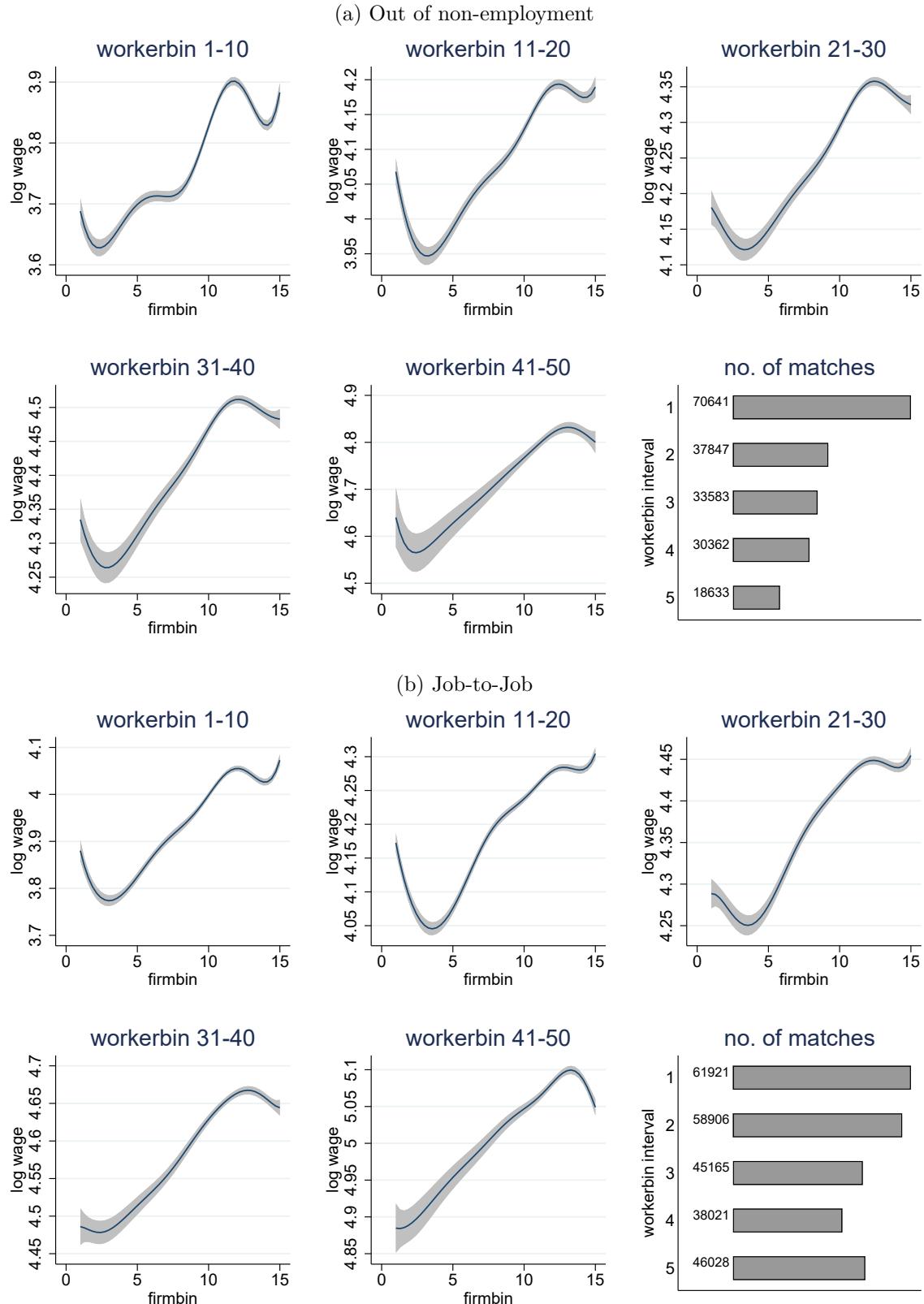
C.4 Wages

Figure C.4: Mean Wages across Worker Types, new Matches



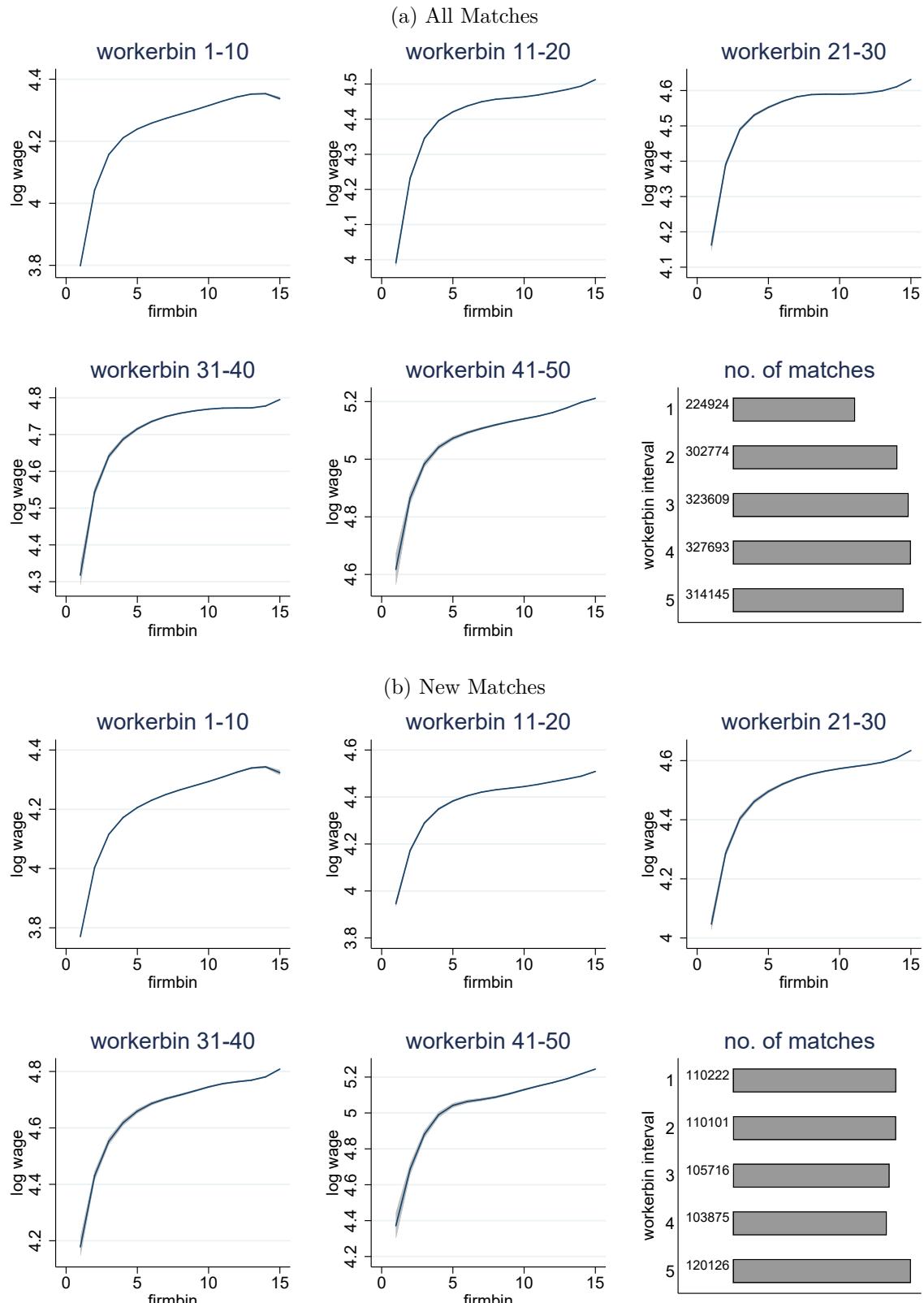
Notes: Plots show estimated wage profiles across firm bins for new matches out of non-employment and job-to-job. The kernel is estimated using a Gaussian kernel function. The bandwidth is 2. 95% confidence bands in gray. Data sources: BHP, EP, BeH.

Figure C.5: Mean Wages across Worker and Firm Types, new Matches, first match-year only to remove tenure effects



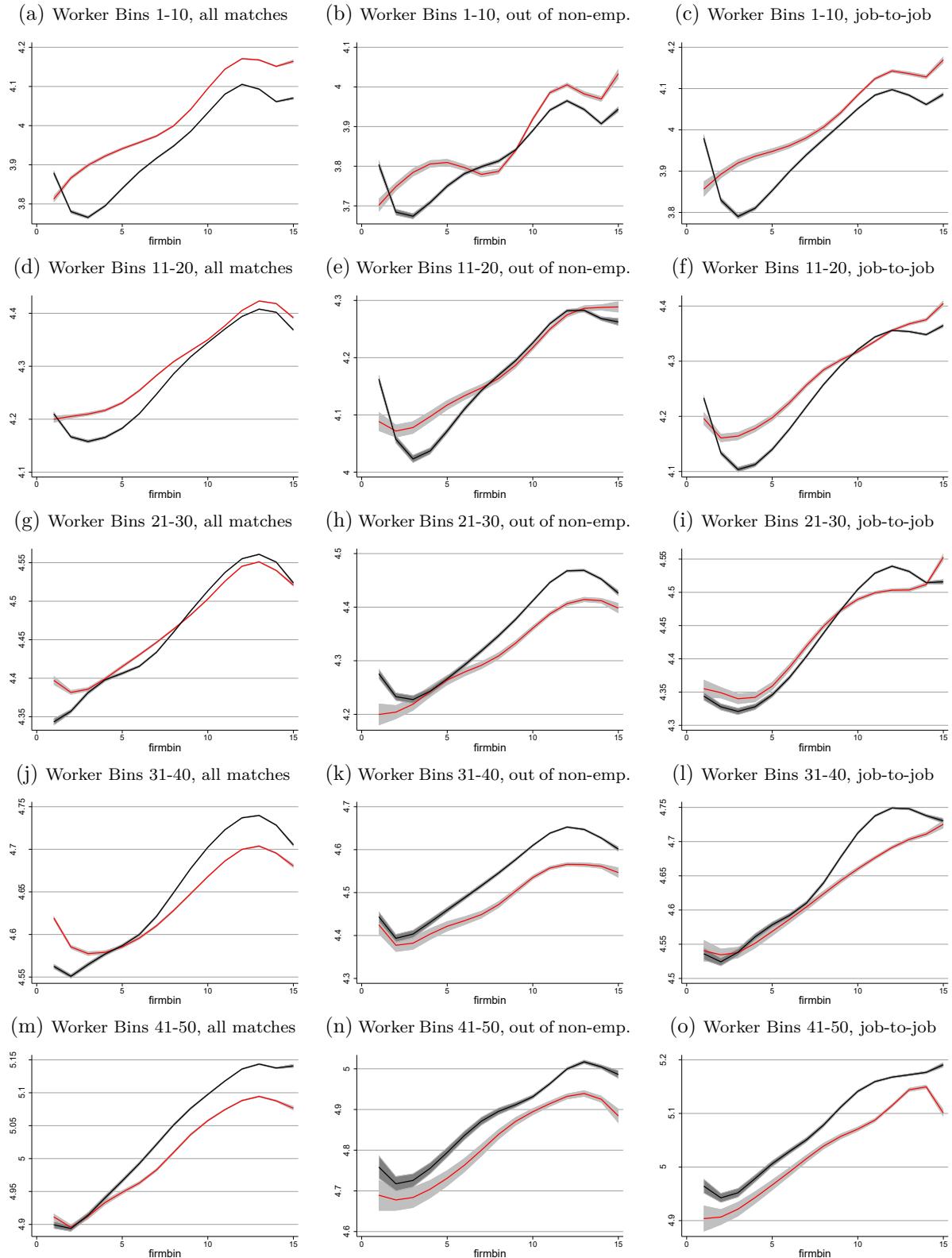
Notes: Plots show estimated wage profiles across firm bins constructed using only the first yearly wage observation for all matches and new matches. The kernel is estimated using an Gaussian kernel function. The bandwidth is 2. 95% confidence bands in gray. Data sources: BHP, EP, BeH.

Figure C.6: Mean Wages across Worker and Firm Types with AKM-based Firm Ranking



Notes: Plots show estimated wage profiles across firm bins constructed using AKM firm effects for all matches and new matches. The kernel is estimated using a Gaussian kernel function. The bandwidth is 2. 95% confidence bands in gray. Data sources: BHP, EP, BeH.

Figure C.7: Changes of Mean Wages across Worker and Firm Types: 1998-2002 (red) vs. 2003-2008 (black)



Notes: Plots show estimated wage profiles for two time periods for grouped worker bins across firm bins for all matches, new matches out of non-employment, and job-to-job moves. The kernel is estimated using an Gaussian kernel function. The bandwidth is 2. 95% confidence bands in gray. Data sources: BHP, EP, BeH.