Firm Productivity, Wages, and Sorting*

Benjamin Lochner[1] and Bastian Schulz[2]

[1] *University of Erlangen-Nuremberg and Institute for Employment Research (IAB)*
[2] *Aarhus University, the Dale T. Mortensen Centre, and CESifo*

August 8, 2022

Abstract

We study the link between firm productivity and the wages that firms pay. Guided by a model of labor market sorting with large firms, we infer firm productivity by estimating firm-level production functions, recognizing that worker ability and firm productivity may interact at the match level. Using German data, we find that the most productive firms do not pay the highest wages. Worker transitions from high- to medium-productivity firms are on average associated with wage gains. Productivity sorting, that is, the sorting of high-ability workers into high-productivity firms, is less pronounced than sorting into high-wage firms.

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 was 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, Fabien Postel-Vinay, Helmut Rainer, Jean-Marc Robin, Johannes Schmieder, Heiko Stüber, Coen Teulings, Robert Ulbricht, Rune Vejlin, and eight anonymous referees for their feedback and many helpful suggestions. The input from seminar and conference participants 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 also 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

Using the seminal Abowd et al. (1999) (AKM) approach to decompose wages into worker- and firm-specific components, Card et al. (2013) (CHK) and Song et al. (2019) show that increases in wage inequality are driven by increases in the sorting of high-wage workers into high-wage firms. Presumably, the wages that these firms pay are related to their productivity. However, the mapping from productivity to wages is far from clear. Highly productive firms may share profits with their workers and pay high wages. However, the wages these firms pay could also be lower due to labor market imperfections or compensating differentials.

The primary contribution of this paper is to shed light on the link between firm productivity and the wages that firms pay. To this end, we estimate worker- and firm-specific wage components following AKM and firm productivity following Ackerberg et al. (2015) (ACF). We present a parsimonious search-matching model to elucidate the minimal set of assumptions necessary to combine the AKM and ACF approaches to study sorting. The model is compatible with both the discrete-time model of dynamically optimizing firms underlying ACF and the log-linear wage equation underlying AKM. It features matching frictions, decreasing returns, multiworker firms, intrafirm bargaining (following Cahuc et al., 2008), firm heterogeneity and worker-firm complementarities.[1]

Our empirical analysis proceeds in two steps. First, we estimate an AKM model for the full population of private sector employees in Germany and merge the estimated AKM wage components with detailed German establishment survey data.[2] Second, we estimate establishment-level production functions and calculate firm productivity using the ACF approach. We separate firm productivity from the effect of heterogeneous worker ability on output using the estimated AKM effects for all workers at the surveyed establishments.

We present empirical results concerning (i) productivity sorting, i.e., the sorting between worker ability (inferred from the AKM worker effects) and firm productivity (inferred using the ACF approach); (ii) the link between firm productivity and wages; and (iii) the role of productivity sorting in increasing wage inequality. (i) Productivity sorting is less pronounced than wage sorting; the average rank correlation between wage-based worker types and productivity-based firm types is 0.07, while the correlation between AKM worker and firm fixed effects is 0.27. Moreover, the AKM correlation increases over time, while our measure of productivity sorting remains relatively flat. Our study of the dynamics of productivity sorting reveals that this flat trend is the result of two opposing developments, namely, increasing sorting of medium-ability workers into medium-productivity firms and decreasing sorting of both high-ability workers into

[1] Theoretical sorting models highlight production complementarities as the primary reason for positive sorting (Becker, 1973; Shimer and Smith, 2000).

[2] We do not observe firms in the legal sense; rather, we observe establishments, i.e., individual production units. We use the terms “firm” and “establishment” interchangeably throughout the paper.

high-productivity firms and low-ability workers into low-productivity firms.

(ii) The empirical relationship between wages and firm productivity helps to explain these developments: it is hump-shaped. Specifically, the most productive firms pay lower wages than somewhat less productive firms, and this tendency has become more pronounced over time. Consistently, worker transitions from high- to medium-productivity firms are on average associated with wage gains. Moreover, the most productive firms are not the largest in terms of headcount or their total wage bill, despite very high labor productivity (value added per worker). Taken together, findings (i) and (ii) suggest that increases in wage sorting could be accompanied by a decrease in the allocative efficiency of the labor market.

The hump-shaped, nonmonotonic relation between wages and productivity that we document poses a challenge for widely used structural models of wage dispersion. In the parsimonious model that we use to outline the assumptions underlying our empirical strategy, wages are monotone in productivity. To make progress toward replicating the nonmonotonic relation with structural models, we discuss several possible model extensions. We evaluate their potential for breaking the tight link between wages and productivity and comment on their compatibility with the AKM approach (in the sense of retaining a log-linear wage equation). We consider on-the-job search, general production functions, and multidimensional firm and worker heterogeneity (firms: productivity and nonwage characteristics of jobs (amenities); workers: ability and preferences/networks).

(iii) To understand the role of productivity sorting in increasing wage inequality, we decompose the wage variance into its within- and between-firm components. The contribution of the between-firm component rose by almost 10% in Germany between 1998 and 2008 (in line with Song et al., 2019). It is comparable in magnitude to the relatively stable within-firm component. However, this picture changes when we decompose the wage variance using the estimated firm productivity and worker ability types. We find that the variance share explained by the between-firm productivity component is low and barely increases over time. Its contribution is dwarfed by the variance shares explained by the within-firm productivity types and between-worker ability types. We conclude that productivity sorting contributes less to rising wage inequality than does wage sorting.

We show how the wage determination mechanism of a search-matching model with multiworker firms, intrafirm wage bargaining, decreasing returns, and worker-firm complementarities 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) intrafirm bargaining framework into the canonical search-matching model. With multiworker firms, our approach differs from existing structural work in the sorting literature that assumes one-worker–one-firm matches. Under this assumption, the focus lies on worker *quality*, while the *quantity* dimension (i.e., the number of workers) is not considered.[3]

[3] An example is Lise et al. (2016). An exception is Bagger and Lentz (2019), where multiworker firms

Eeckhout and Kircher (2018) relax the one-worker–one-firm assumption and study both the quantity and the quality dimensions of production. The firm decides which worker type(s) to hire and, additionally, how many workers of each type to employ. The main result is that firms optimally hire multiple workers of exactly one type. 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, nondegenerate matching sets, and a production function that is geared toward our empirical approach.

Our findings contribute to the empirical literature on wage dispersion. The aforementioned study of Germany by CHK and that of the U.S. by Song et al. (2019) follow the AKM approach. These studies decompose wage dispersion into the contributions from unobserved worker ability, firm wage premia, and wage sorting, thus measuring the extent to which workers who receive high wages are matched with firms that pay high wages. We show that the way in which one measures firm heterogeneity, i.e., by the wages that firms pay or by firm productivity, affects these decompositions. In our data, the firms with the highest estimated productivity do not pay the highest wages.

The main difference between AKM-inspired and model-based analyses of labor market sorting is typically the implied (non)monotonicity of the wage equation. Due to complementarities, wages are not a monotonic function of firm type in many models (e.g., Gautier and Teulings, 2006; Eeckhout and Kircher, 2011; Hagedorn et al., 2017; Lopes de Melo, 2018). However, the log-linear AKM wage equation implies that wages are strictly increasing in the firm fixed effect. Abowd et al. (2018) find common ground by formally relating the equilibrium quantities derived from the directed-search sorting model of Shimer (2005) to the worker and firm heterogeneity components that can be identified with the AKM approach. We follow their example and discuss how our model's wage equation maps onto AKM and which assumptions are critical for this mapping. Moreover, we contribute to the discussion about wage (non)monotonicity. We find quantitatively important deviations from monotonicity among high- and low-productivity firms.

The literature has developed multiple ways to capture firm heterogeneity when analyzing labor markets, and the productivity-based firm ranking that we propose extends this toolbox. Bagger et al. (2014) estimate firm-level production functions with heterogeneous labor inputs to study wage dispersion. Hagedorn et al. (2017) rank firms based on the value of a vacancy, which can be identified from wage data in the context of their model. Bartolucci et al. (2018) use balance sheet data to rank firms by observed profits. Taber and Vejlin (2020) and Bagger and Lentz (2019) rank firms by the share of workers that they poach from other firms. Sorkin (2018) ranks firms by amenities based on a revealed-preference argument. Haltiwanger et al. (2018) study worker flows using gross output per worker as their firm productivity measure.

are present but production is linear and firm size is limited by search frictions, not decreasing returns.

Another related body of literature focuses on rent sharing and imperfect competition in the labor market. Card et al. (2018) show that rent-sharing elasticities are largely driven by worker quality (measured by AKM worker effects). Nevertheless, they find a sizable positive correlation between value added per worker and the AKM firm effect, which might be driven by compensating differentials. Lamadon et al. (2022) use U.S. data to estimate a model with compensating differentials, worker preference heterogeneity, and monopsony power. In line with our findings, they show that high-productivity firms pay relatively low wages; the reason is that these firms disproportionately compensate their workers with amenities.

2 Model

This section introduces a model of multiworker firms with decreasing returns, intrafirm wage bargaining, worker–firm complementarities, and matching frictions. This parsimonious model is designed to elucidate the minimal set of assumptions necessary to combine the AKM and ACF approaches. We show how the model's wage equation maps onto the log-linear AKM wage equation. Using this mapping and the estimated AKM wage components, we construct model-consistent firm-level labor inputs for production function estimation. We highlight the assumptions under which the model is compatible with the model of dynamically optimizing firms underlying the ACF approach. Model details are in the appendix, available online.

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 as Ω, may change due to idiosyncratic firm-level shocks that evolve according to a stochastic process characterized by the conditional CDF $G(\Omega'|\Omega)$. Workers and firms meet randomly. Conditional on meeting, a match is not guaranteed because the match surplus may be too low. Appendix A.1 provides details on the assumed matching mechanism.

To set up the firm's problem, we build on Cahuc et al. (2008), who generalize the canonical search-and-matching model to allow for multiworker firms with heterogeneous labor inputs and decreasing returns, strategic interactions, and intrafirm 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 facilitate a simple and transparent link to the empirical models that we use (i.e., the AKM and ACF models), we abstract away from complementarities between worker types within the same firm. Furthermore, we make six simplifying assumptions. First, workers and firms are risk neutral. Second, both worker and firm heterogeneity are one-dimensional.[4]

[4]We discuss how multidimensional firm/worker heterogeneity would affect our approach in Section 7.

Third, worker ability and firm productivity are known to all market participants and are cardinally measurable. Fourth, infinitely lived workers supply one infinitesimally small unit of labor (no extensive/intensive margin choice) so that labor input is a continuous variable.[5] Fifth, we abstract away from capital inputs when discussing the model.[6] Sixth, we present the model in discrete time because the ACF approach also relies on a discrete-time model of dynamically optimizing firms. Time indices are omitted.

The general concave, firm-level production function is as follows:

$$Y = F(L, \Omega), \tag{1}$$

where Y is value added and $L = \sum_x xL_x$ is a scalar composite labor input measured in units of worker ability. It combines all heterogeneous labor inputs L_x, which is the number of type x workers employed by the firm. Ω is the current productivity of the firm. Our focus is on worker–firm sorting. We assume that output is (log-)supermodular at the match level, i.e., the complementarity between firm productivity Ω and worker ability determines the contribution of every single match to firm-level output. The marginal product of an additional unit of worker ability is firm specific. We interpret firm productivity as a "nonrival" resource; i.e., we do not consider the span-of-control problem in the optimal allocation of resources to heterogeneous workers (Eeckhout and Kircher, 2018).

The following simple production structure is in line with our assumptions:

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

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, which is in line with the sufficient conditions for positive assortative matching (PAM) derived by Shimer and Smith (2000).[7] Under our assumptions, worker ability units are perfect substitutes at the firm level. That is, output depends on Ω and the number of efficiency units of labor employed. The marginal product of an additional unit of type x labor is as follows:

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

Worker ability x scales the marginal product at an (L, Ω) firm. We use this property below. The marginal product is increasing in firm productivity and the output elasticity

[5] We take part-time labor into account when we estimate the production functions in Section 4.2.

[6] Cahuc et al. (2008) explicitly consider a predetermined capital stock; thus, our model could be extended in this direction. ACF assume a similar capital accumulation mechanism; see Section 4.2.

[7] Shimer and Smith (2000) establish the existence of an equilibrium in this environment. PAM arises with a log-supermodular match-level production function, while 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.

of labor but decreasing in the composite labor input L due to decreasing returns.

As in Cahuc et al. (2008) and ACF, employment is a state variable. The firm's problem is to optimally choose how many vacancies to post given its expected profits from hiring. We assume that vacancies cannot be targeted to specific worker types and are subject to a flow cost c, which could depend on the firm type. In Appendix A.2, we solve the firm's problem and derive the relevant optimality conditions in the steady state.

The bargained wage satisfies the Nash sharing rule. The firm's surplus consists of its marginal profit 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). Wage (re)negotiations happen instantaneously so that firm-level employment remains fixed. In Appendix A.3, we show that the outcome of intrafirm wage bargaining in the model is described by the following differential equation:

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

This is a discrete-time version of the wage bargaining outcome derived by Cahuc et al. (2008) for the "single labor case". We can utilize their result due to our assumption of perfectly substitutable ability units. α is the workers' bargaining power parameter, and β is the common discount factor. The first term in parentheses shows that the wage of a type x worker at an (L, Ω) firm is a function of the worker–firm-specific marginal product, $F_x(L, \Omega)$. The second term captures the inframarginal effect that hiring the marginal worker has on all other workers' wages. This mirrors the finding that firms can reduce incumbent workers' wages by increasing employment in the presence of decreasing returns.[8] Absent complementarities between worker types, the inframarginal adjustment reflects decreasing returns only and is unambiguously negative. Finally, the third term in equation (4) captures the worker-specific outside option, which we return to below. The solution[9] to the differential equation (4) is as follows:

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

which indicates that the wage depends on the worker's outside option and an integral expression that combines the worker's marginal product and the inframarginal effect. The latter is weighted by the worker's bargaining power and decreases in the distance from the margin. With the assumed production function, equation (5) also implies that wages are monotone in firm productivity, which we show formally in Appendix A.4.

[8] Cahuc et al. (2008) allows for unrestricted substitutability/complementarity patterns between worker types. In this case, the effects on coworker wages can be either positive or negative. Firms may strategically over/underemploy specific worker types depending on their contribution to the total wage bill.

[9] We follow Stole and Zwiebel (1996) and Cahuc et al. (2008); see Appendix A.3.

2.1 From Theory to Estimation

To empirically recover Ω, we use a two-step approach that combines techniques from the empirical wage dispersion and industrial organization (IO) literature. First, we estimate the AKM model with German data (Section 3, Appendix D.1). The estimation yields, for all individual workers and firms in the largest connected set, worker fixed effects, firm fixed effects, and a residual. There is a direct correspondence between the AKM wage components and the wage equation in our model. Second, we estimate the production function (Section 4.2, Appendix C). We merge rich survey data on inputs and output for a representative sample of firms with the estimated AKM wage components for their full workforce to predict the firm-level labor input (wage bill). According to the model's separability property, the worker fixed effects capture the full effect of worker ability on wages. Thus, this wage component allows us to separate firm productivity from the effect of worker ability on output. To predict labor inputs, we include the effect of observable worker characteristics (not part of our model). Moreover, we include either the firm fixed effects or the AKM residuals, or both. This allows us to investigate how firm-specific and time-(in)variant wage components matter for the inferred productivity measure.

The model's wage equation maps onto the AKM model as follows. Both the outside option and the integral expression in (5) are functions of worker ability. Their effects on wages are absorbed by the AKM worker fixed effects when both wage components are linear with respect to worker ability x. Given our production structure, the integral expression is indeed linear in x because worker ability scales the marginal product of labor. This feature is crucial for the applicability of the AKM approach.[10] Plugging the marginal product of labor into the wage equation (5) yields:

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

where x can be written in front of the integral sign. We show in Appendix A.5 that the worker's outside option in the first term is also linear in x under two additional assumptions. First, the flow value of unemployment for the worker, $b(x)$, must be proportional to x. This is a standard assumption. Second, matching sets must cover the whole type space; that is, there are no unacceptable combinations of worker and firm types. An implication of this assumption combined with our matching technology is that all worker types are unemployed for the same (expected) duration. In Section 4.1, we verify that this holds in our data (conditional on observables). Another implication is that there is no endogenous sorting despite the complementarity in the production function.

Next, consider firm fixed effects. How the integral expression in (6) is captured in the AKM model depends on the underlying data-generating process (DGP). Suppose that

[10] Linearity in x implies that the wage equation is additive in logs (see Appendix A.5).

our wage equation describes reality without error. The economy is in a steady state, and firm productivity does not change. In this case, (a log-affine transformation of) the firm-specific integral, which represents the effects of firm productivity and decreasing returns on the wage, is absorbed by the time-invariant firm fixed effect, and the residuals are zero. With a more complex DGP, the value of the integral changes over time. Suppose productivity Ω follows an AR(1) process. In this case, the AKM firm fixed effect captures a time average of the effect of changing productivity on the wage. Thus, the more persistent the productivity process, the more informative is the time-invariant firm effect. The AKM wage component that remains is the residual, which absorbs fluctuations due to transitory shocks around the time average of the effect of productivity on the wage.[11]

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). 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.[12] The BeH data contain information on worker gender, age, and education (5 categories); start and end dates for the employment 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 make social security contributions are excluded. Each worker and each establishment have a unique identification number, which allows us to follow workers over time and across establishments. Regarding sample selection, we follow CHK and Lochner et al. (2020). We start with the universe of employment spells observed between 1998 and 2008. There is no exact information on hours worked, so we restrict our sample to full-time employees (males and females) aged 20 to 60. We define the job with the highest total wage as each worker's main job.

The EP is a comprehensive annual survey of establishments.[13] It provides us with the data needed to estimate production functions; we observe revenue, intermediate goods purchases (reported as revenue shares), value added (calculated as revenue minus intermediate goods purchases), and net investments in four categories of capital goods (buildings, production machinery, IT, and transport equipment). We restrict our analysis to establishments with nonmissing data for revenue and intermediate goods and trim the data by

[11] The mapping that we describe could be obscured when using real data, e.g., due to measurement/sampling error in short panels. We assess the robustness of our results in Section 6.

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

[13] These are single production units. A total of 80% report that they are single-establishment firms.

dropping establishments below the 1st and above the 99th percentile of the revenue distribution. We supplement the EP data with covariates from the "Establishment History Panel" (BHP; see Spengler, 2008). These include average wages, headcounts, shares of full-time/part-time workers, and worker shares by education group. Moreover, the BHP provides administrative information on firm age and a consistent industry classification.

We observe nominal gross daily wages, which we deflate using the consumer price index. Earnings are tracked up to a threshold, so we follow Dustmann et al. (2009) and impute the upper tail of the wage distribution by running a series of tobit regressions; see Appendix B.1. We impute missing and inconsistent education data using the methodology proposed in Fitzenberger et al. (2006); see Appendix B.2. The EP data contain information on net investments. To estimate the capital stock, we use a perpetual inventory method (Müller, 2008). This method approximates the establishment-level capital stock by combining net investments with the average economic lives (depreciation rates, available from the national accounts) of the different types of capital goods.

Following CHK, we estimate an AKM-type wage regression on the largest connected set in the BeH for our period of analysis, namely, 1998–2008. 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 the observations are classified as movers between establishments, and the regression model is identified for such workers.[14] The estimation results are in Appendix D.1. We use the estimated AKM wage effects for men, women, and both movers and stayers from the full BeH sample (column (a) in Table D.1) to construct the labor input for the production function estimation at the firm level. Based on the discussion in Section 2.1, we include (exponentiated) worker fixed effects, the effects of observable characteristics, and, in some specifications, firm fixed effects and residuals to distinguish among the different channels through which wages and firm productivity are related according to our theory.

We construct two final samples for analysis. The first sample includes *all matches*, i.e., matches that formed both before and after 1998. We do not distinguish between job-to-job (J2J) flows and matches formed out of nonemployment (OON). There are 4,695,108 employment spells for 1,344,382 workers employed at 10,004 EP establishments. The second sample includes *new matches*, i.e., matches formed from 1999 onward. We distinguish between J2J flows and matches that formed OON and use the initial observation (1998) to determine employment status before the new match. There are 1,656,280 employment spells for 633,831 workers at 9,659 establishments. A total of 64% (1,055,151) of new employment spells are J2J moves from one employer to another, while 36% (601,129) are spells formed OON, which also include young workers who enter the labor market.[15]

[14] Following CHK, we impute worker fixed effects and residuals for "stayers" (workers who do not change their employer over the entire sample period).

[15] Our definition of nonemployment includes marginal employment, unemployment (benefit receipt), and inactivity.

4 Ranking Workers and Firms

4.1 Worker Ranking

According to our model, the estimated AKM worker fixed effects fully capture the effect of unobserved worker ability on wages. This result can serve as a "microfoundation" for using the AKM effects to rank workers. We create 50 ability bins of equal size.[16]

To show how the bins summarize worker heterogeneity, we decompose the variances of observed wages, age, and education into the shares explained within and between bins. Little within-bin variance implies homogeneity among workers along the corresponding dimension. The rankings are based on the worker fixed effects, which explain the majority of the wage variation in the data. Thus, the between-bin share of the wage variance is large (70%). For age and education, the picture is quite different. We find that 96% of the age variation and 74% of the education variation are within bins. Because we control for time-varying age and education effects in the AKM regression, this finding is a direct reflection of the low correlation between the estimated worker effects and the time-varying observables (see Table D.1).

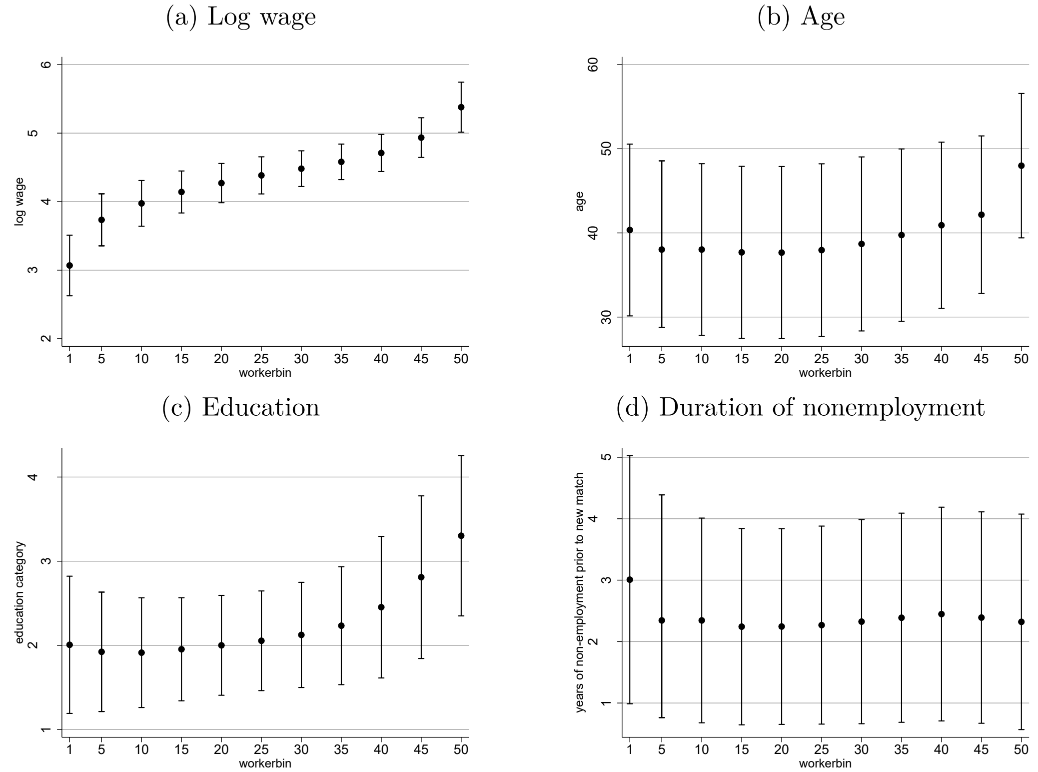
Figure 1 illustrates how log wages, age, and education vary across worker bins. Panel (a) shows that mean wages increase monotonically across worker bins, as expected. In contrast, Panel (b) shows that the mean age across worker bins is relatively flat. Only the highest bin has a somewhat higher mean age. However, due to the large standard deviation, this difference is not significant. Panel (c) shows that mean education increases above bin 35 but is flat below it. This suggests that highly ranked workers are more likely to have tertiary degrees, while lower-ranked workers tend to have vocational training only, but these differences are hardly significant. Note that the dispersion of education increases among the highest-ranked workers. It is more common to observe high-ranked workers with little education than low-ranked workers with tertiary degrees. Finally, Panel (d) shows the duration of nonemployment across worker types. The mean observed nonemployment duration is virtually flat. Recall that we control for the effect of observable characteristics. The finding that, conditional on observables, nonemployment durations do not change with worker type supports the assumption that the matching sets cover the whole type space (see Section 2).

4.2 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 IO literature that this approach is susceptible to two challenges. The first is that input choices are likely correlated with firm productivity. To address this challenge, we rely on the ACF

[16] There are 702,540 individual workers in each of the 50 bins.

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



Notes: The figures plot the means ± one standard deviation for log wages, age, education, and nonemployment duration for every fifth worker bin. The ages of the individual workers in our sample range from 20 to 60. There are 5 education categories, which are defined in Appendix B.2. Data source: BeH.

version of the control function approach (Olley and Pakes, 1996; Levinsohn and Petrin, 2003). The key assumption behind this approach is that intermediate input demand (the control function) is strictly increasing in (scalar and unobserved) firm productivity. This implies that the control function can be inverted, thus effectively controlling for unobserved firm productivity by substituting it out of the production function.

The second challenge is that the quality of the labor inputs varies across firms. Physical units, e.g., headcounts or hours worked, do not reflect differences in worker ability and imply measurement error. This complicates the separate identification of firm productivity and the output elasticity of labor. The empirical IO literature addresses this challenge by controlling for labor quality differences (e.g., Fox and Smeets, 2011; Irarrazabal et al., 2013). We construct model-based labor input measures that isolate the effect of worker ability on output (see Section 2.1).

To estimate the production function, we rewrite equation (2) using the composite labor input, L_{jt}. We add capital, K_{jt}, and indices for firms (j) and time (t):

$$Y_{jt} = (\Omega_{jt} L_{jt})^{\beta_l} (\Omega_{jt} K_{jt})^{\beta_k}. \tag{7}$$

β_l and β_k are the output elasticities of labor and capital. Both inputs are scaled by the firm's current productivity Ω_{jt}. Note that Y_{jt} is value added (revenue minus expenditures on intermediate goods). Without assuming constant returns to scale, we obtain

$$Y_{jt} = \Omega_{jt}^{\beta_l+\beta_k} L_{jt}^{\beta_l} K_{jt}^{\beta_k}. \tag{8}$$

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

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + \omega_{jt} + z'_{jt}\gamma + \epsilon_{jt}. \tag{9}$$

We add a constant β_0, the residual ϵ_{jt}, a vector of controls z'_{jt} (indicators for West German establishments and employee representation in management, four firm age categories, and the share of part-time workers), and time and sector fixed effects.

ACF assume that the capital stock is predetermined by investment $i_{j,t-1}$:

$$k_{jt} = \kappa(k_{j,t-1}, i_{j,t-1}). \tag{10}$$

We abstract away from capital in Section 2, but our model builds on that of Cahuc et al. (2008), who explicitly consider a capital accumulation process similar to equation (10), so the models are compatible in the capital dimension. The demand for other inputs, labor and intermediate goods, may change in response to firm productivity shocks.[17] The conditional expectation of firm productivity is modeled using a first-order Markov process:

$$\omega_{jt} = E(\omega_{jt}|\omega_{j,t-1}) + \xi_{jt} = \rho\omega_{j,t-1} + \xi_{jt}, \tag{11}$$

where ξ_{jt} is an innovation that is assumed to be uncorrelated with ω_{jt} or k_{jt}. For the estimation, we follow the ACF model and assume that ω_{jt} follows an AR(1) process with parameter ρ. In our model, both the firm's vacancy-posting decision and its matching set depend on current productivity; see Appendix A.2. Moreover, the firm's dynamic problem involves the expected future evolution of productivity. Recall that we abstract away from productivity shocks when deriving the wage equation (4) and establishing its linearity in worker type x. This simplifies the derivation but is not critical for the result. As long as the evolution of firm productivity is independent of the employed worker types, the wage equation will still be (log-)linear in x under our assumed production structure.[18]

Finally, the ACF approach allows for (exogenous) across-firm differences in labor adjustment costs and wage setting. This flexibility makes it a good fit for search-matching

[17] The firm's information set when making input choices includes all past and present productivity shocks $\{\omega_{j\tau}\}_{\tau=0}^{t}$, but it does not include future productivity shocks.

[18] Consider an environment with productivity shocks and, accordingly, changing labor demand and endogenous separations. In this case, the wage equation contains an additional forward-looking component that captures the expected evolution in firm productivity. Worker type x can still be written in front of the integral sign under the mild assumption that x and the firm productivity process are independent.

Table 1: Production Function Estimation Results

Dependent Variable	*Value Added*				
	(a)	(b)	(c)	(d)	(e)
Labor input	0.7129	0.7131	0.7125	0.7099	0.8573
	(0.0141)	(0.0141)	(0.0141)	(0.0135)	(0.0133)
Capital input	0.1706	0.1703	0.1707	0.1512	0.1358
	(0.0117)	(0.0117)	(0.0117)	(0.0116)	(0.0110)
Labor input variable	AKM Predicted Wage Bill				Headcount
Worker fixed effects	Yes	Yes	Yes	Yes	–
Worker observables	Yes	Yes	Yes	Yes	–
Firm fixed effects	No	Yes	No	Yes	–
Residual	Yes	No	No	Yes	–
Variance of $\hat{\omega}_{jt}$	0.074	0.073	0.072	0.064	0.050
Variance of ϵ_{jt}	0.628	0.627	0.626	0.597	0.583
N	38,598	38,598	38,598	38,598	39,808

Notes: Bootstrapped standard errors (500 iterations) are given in parentheses. All estimated coefficients and standard errors are rounded to four decimal places. All specifications include firm-level controls as well as year and sector fixed effects (32 categories). Data sources: BHP, EP, BeH.

models. ACF use the intermediate input demand conditional on labor as the control function. This conditioning eliminates the link between intermediate input demand and the labor input, which, in turn, depends on wage conditions and adjustment costs.[19] In the context of our model, differences in adjustment costs and wage setting across firms arise for two reasons. First, vacancy posting costs can be heterogeneous, and firms have limited control over the hiring process.[20] Second, due to decreasing returns to labor and intrafirm bargaining, wages have a firm-size component that varies across firms.[21]

We describe the technical details of the ACF production function estimation in Appendix C. Table 1 presents the results. We show the results from five specifications in which we vary labor input, i.e., the way we control for worker ability. All specifications include the aforementioned firm-level controls, as well as year and sector fixed effects.

Specifications (a)–(d) use predicted wage bills based on different combinations of the estimated AKM wage components. We always include worker fixed effects, which,

[19] Earlier approaches that use unconditional intermediate input demand are more restrictive. As explained in ACF (p. 2431), their approach can handle serially correlated, exogenous, and unobserved shocks to the prices of capital and labor. Levinsohn and Petrin (2003) can only allow for such shocks to the price of capital inputs. Neither approach allows for shocks to the price of intermediates.

[20] With random search, the time that passes until a suitable worker arrives is stochastic.

[21] Our model implies a nondegenerate firm-size distribution due to heterogeneous productivity.

according to our model, capture the effect of worker ability on output, along with the effects of worker observables. In addition, we include the AKM residual in (a), the firm fixed effect in (b), neither the residual nor the firm fixed effect in (c), and both in (d). The labor input in (d) is thus equal to the total observed wage bill for the firm. For the sake of comparison, we also show the results of a specification that uses the worker headcount (e).

The first interesting finding is the large difference in the estimation results for (a)–(d) and those for (e). The headcount mismeasures the labor input in the presence of worker heterogeneity; thus, the output elasticity of labor is biased upward in specification (e), while the capital elasticity and the estimated productivity $\hat{\omega}_{jt}$ are biased downward. Across specifications (a)–(d), the results are very consistent, with an estimated output elasticity of labor of approximately 0.71. Together with the capital elasticity, these estimates imply decreasing returns to scale, in line with our model's assumption.

The second interesting finding is that the AKM wage components beyond the worker fixed effects and observable effects have little effect on the estimation results. In light of our model, this is unsurprising. The worker fixed effects capture the effect of worker ability on output. The firm effect, however, does not reflect contributions to output. It captures the time average of the effect of firm productivity on wages (see Section 2.1). The AKM residual absorbs the effect of firm productivity shocks on wages according to our model. Such shocks have been shown to affect wages ("pass-through").[22] A correlation between the contemporaneous productivity realization and the labor input does not threaten our estimation strategy.[23] Notably, adding the residual has little impact on the estimated output elasticities. However, capturing the reaction of wages to productivity shocks could still be important. If worker ability and firm productivity are complements, the contribution of worker ability to firm output could shrink after a negative shock. For this reason, we select specification (a), which includes the AKM residuals, as our benchmark specification.

Finally, we examine the dispersion and persistence of estimated firm productivity $\hat{\omega}_{jt}$. The estimated variance of 0.074 in our benchmark specification is almost 50% higher than that in the headcount specification (e). It is, however, an order of magnitude smaller than the residual variance. The residual absorbs transitory shocks, whereas productivity ω_{jt} is the realization of a persistent stochastic process, and the estimated autocorrelation in $\hat{\omega}_{jt}$ in the benchmark specification is indeed high (0.75). Given the link between firm productivity and the AKM firm effects suggested by our model, this high degree of persistence implies that firm fixed effects should be quite stable over time.[24] Thus, we rank firms by the firm-level mean of productivity.

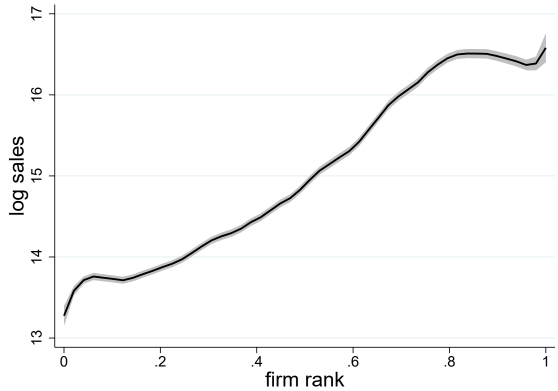
[22]Guiso et al. (2005) show that firms fully insure their workers against transitory productivity shocks but only partially against the persistent productivity changes captured by $\hat{\omega}_{jt}$.

[23]See the ACF moment conditions, which we discuss in Appendix C, equation (C.5).

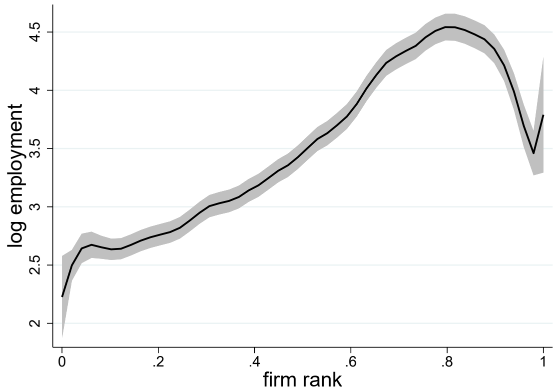
[24]Recent evidence by Lachowska et al. (2022) supports this conjecture.

Figure 2: Firm Performance Measures by Estimated Firm Rank

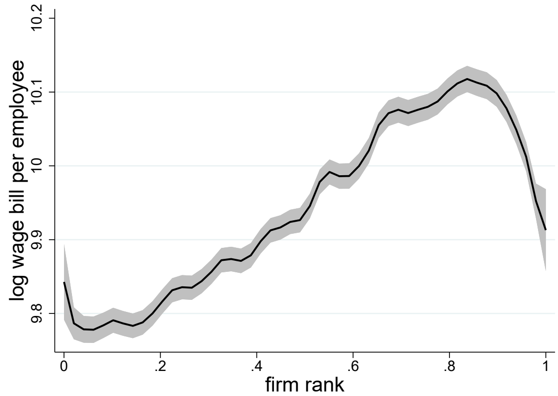
(a) Revenue



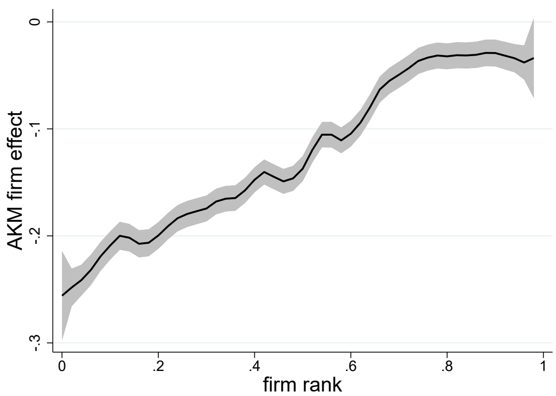
(b) Firm size



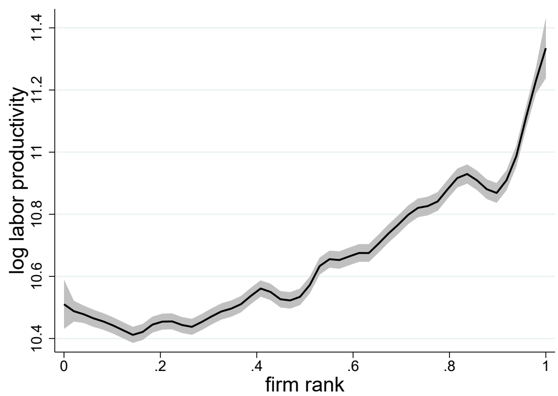
(c) Average wage



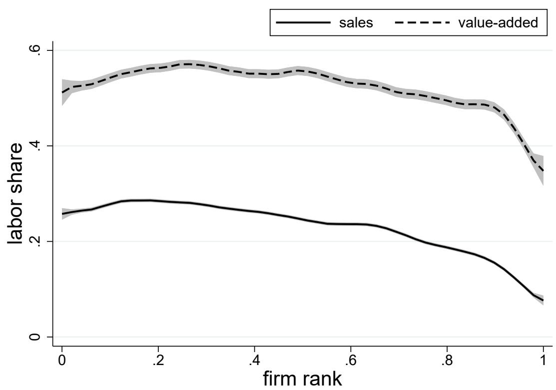
(d) AKM firm effects



(e) Labor productivity



(f) Labor shares



Notes: Estimated univariate kernel densities for selected firm performance measures across estimated firm ranks, normalized to be between zero and one. Kernel: Epanechnikov. The bandwidth is 0.05 for (a)–(c) and (e) and 0.04 for (d) and (f). The qualitative findings are robust to the bandwidth choice. The 95% confidence bands are shown in gray. Data sources: BHP, EP, BeH.

Productivity ranks are moderately positively correlated with headcount (0.40), value added (0.43), value added per worker (0.25), the capital stock (0.24), and profit (0.16). They are not correlated with capital per worker (-0.03). These correlations suggest that the largest firms in terms of output, workforce, and assets are not the most productive firms. The correlation with the share of workers with tertiary degrees is rather low (0.08), so productive firms do not necessarily have a highly educated workforce. In Appendix D.2, we compare the productivity ranking with other rankings used in the literature.

Figure 2 shows estimated kernel densities for six firm performance measures across ranks. Revenue (measured by log sales, Panel (a)) mostly increases in rank but becomes relatively flat at the top. The most productive firms do not have greater revenue than firms around the 80th percentile of the productivity distribution. In Panel (b), the relation between firm rank and log employment (measured by the headcount) is shown. This size measure increases until approximately the 80th percentile and falls steeply thereafter. The most productive firms are comparable to median-productivity firms in terms of size.

The log wage bill per employee (the average wage) is shown in Panel (c). Above the 80th percentile, average wages fall. Notably, the largest firms are also situated near the 80th percentile; therefore, average wages are closely related to size but less related to productivity. Interestingly, the least productive firms pay higher average wages than other firms below the 25th percentile. The AKM firm effect in Panel (d) is essentially flat above the 80th percentile. Thus, firms within the top productivity quintile are indistinguishable in terms of AKM firm effects.

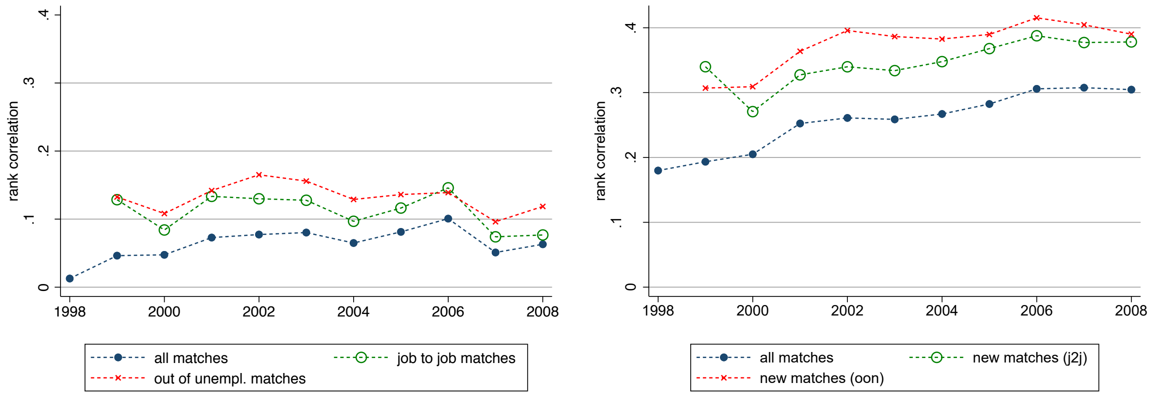
Panel (e) shows that log labor productivity (value added per worker) is relatively flat in the bottom half of the distribution but increases substantially for the most productive firms. At the most productive firms in the distribution, it is approximately twice as high as that of the firms at the bottom. Panel (f) shows the labor shares for revenue and value added. Both labor shares exhibit a hump shape but clearly fall in the estimated firm rank. The labor share of revenue falls from just below 30% among the least productive firms to less than 10% among the most productive firms. For the value added measure, the labor share falls from 50–60% to approximately 35%. Thus, the average worker obtains a significantly lower share of the output at high-productivity firms.

To proceed, we group all individual firms into 15 productivity bins of equal size.[25] The firm bins exhibit a high within-bin wage variance (94% of the total variance). This is due to controlling for worker ability when estimating firm productivity, and it implies that our firms cannot be categorized into high-wage and low-wage firms. Within-bin heterogeneity in output (value added, value added per worker) and size (headcount, assets) is sizable, which implies that firms of similar sizes or production capacities are not necessarily in the same bin.

[25] The EP sample used to estimate firm productivity contains 10,026 individual establishments, so there are 668 establishments in each bin.

Figure 3: Rank Correlation Coefficients over Time (1998–2008)

(a) Firm ranking: productivity (b) Firm ranking: AKM firm effects



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

5 Productivity Sorting

To measure the degree of productivity sorting, 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 rebin workers and firms in the merged sample.[26] We rely on the final samples defined in Section 3.

5.1 Rank Correlations

We compute Spearman's ρ to study the associations between the estimated worker and firm ranks in the data. For all matches in all years, we find a significantly positive but relatively low rank correlation of 0.07. For new matches, this correlation is 0.12. The correlation for J2J moves (0.11) is somewhat lower than the correlation for new matches OON (0.13).[27]

The positive rank correlation that we find reflects PAM based on worker ability and firm productivity. However, the degree of productivity sorting is low, both in an absolute sense and compared to that found in earlier studies.[28] The key difference is that we rank firms based on their estimated productivity, while, e.g., CHK and Hagedorn et al. (2017) rely on wages and observed worker mobility to rank firms.

Next, we consider how the estimated rank correlations change over time. Figure 3

[26]The number of workers per bin decreases to 26,888. The number of firms per bin is stable (667).

[27]For clarity, the rank correlations presented here are calculated based on the worker and firm bins. Correlations based on worker and firm ranks (not reported) differ only marginally. Table D.2 provides an overview of all estimated rank correlations for different samples and time periods.

[28]Using German data, CHK and Hagedorn et al. (2017) find correlations of 0.17–0.25 and 0.71, respectively. CHK use the AKM model and compute their measure of wage sorting by correlating estimated worker and firm effects. Hagedorn et al. (2017) study sorting through the lens of a structural model with worker and firm heterogeneity.

presents rank correlations for all samples and all years. In Panel (a), the correlations are calculated using the productivity-based firm ranking. The solid circles depict the correlations for all matches. It first increases and then stabilizes just below 0.1 before dropping in 2007. For both types of new matches, the correlations are higher but also fall toward the end. The degree of sorting OON (crosses) decreases steadily after 2002. For J2J switches (hollow circles), the time dynamics closely mirror those for all matches because approximately 65% of all new matches are J2J switches (see also Table D.2). Thus, new matches drive up the overall rank correlation in the beginning; however, once their correlation decreases, the correlation for all matches also levels off.

Panel (b) of Figure 3 facilitates the comparison with wage-based approaches. We compute rank correlations using the same worker ranking but use the AKM firm fixed effects to rank firms. Notably, this implies higher rank correlations (more positive sorting) and a significant increase over time. The correlation for all matches reaches 0.3 near the end of our sample period, which is close in magnitude to that obtained by CHK. For new matches, the correlations are even higher, just below 0.4. New matches OON exhibit the highest rank correlations with both the AKM and the productivity-based firm rankings.

5.2 Distributional Dynamics

Rank correlations are only a summary statistic for the allocation of workers to firms. To investigate which worker–firm-type combinations contribute to changes in productivity sorting, we study how the matching of worker and firm types has changed over time. To this end, we estimate univariate density functions for the employed worker types. Figure D.4 presents these densities for low-productivity firms (bins 1–2), medium-productivity firms (bin 8), and high-productivity firms (bins 14–15). We compare the first half of our sample period (1998–2002) to the second half (2003–2008) and show estimated densities for all matches, new matches OON, and J2J flows.

Low-type firms hire mainly low-ability workers OON. This is where the density of new matches is highest, and it clearly falls in the worker type. This is consistent with positive productivity sorting. However, the average quality of new hires in low-productivity firms has increased over time. Low-type firms now hire more medium-ability workers and fewer low-ability workers, both OON and through J2J flows. Considering all matches, the densities of the worker types hired by bin 1 firms have become more uniform over time. That is, in addition to making new matches of higher quality, these firms also have sizable outflows of high-ability workers. Taken together, these changes in low-productivity firms confirm the aggregate trend toward less productivity sorting.

For medium-type firms, we observe a strong increase in the hiring of medium-ability workers both through J2J flows and OON. For very productive firms, an increase in the hiring of medium-ability workers is also present but less pronounced. This increase is paired with a notable decrease in the hiring of high-ability workers for new matches made both through J2J flows and OON. Considering all matches, we see a significant reduction in the hiring of high-ability workers and an increase for medium-ability workers, although that increase among all firms is smaller than for medium-productivity firms alone. Only for the most productive firms can we discern a small but significant increase in new matches with high-ability workers through J2J flows, although the density for all matches indicates that the allocation of high-ability workers at the very top is quite stable.

This distributional analysis reveals why productivity sorting is relatively stable over time. High-ability workers have, to some extent, been replaced by medium-ability workers. Moreover, low-productivity firms have reduced their hiring of low-ability worker types and increased the quality of their average worker. The stable rank correlation is thus the result of two opposing trends: reduced productivity 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 allocations are related to the wages that workers earn at different firms.

6 Wages and Inequality

6.1 Wage Variation across Worker and Firm Bins

We investigate how the wages of different worker types vary across firm types. Figure 4 depicts the resulting wage–productivity profiles for five groups that include ten worker bins each. Panel (a) includes all matches, while new matches are shown in Panel (b).[29] The wage–productivity profiles exhibit a characteristic S-shape, i.e., low-productivity firms pay relatively high wages. The lowest wages are paid by firms in bins 3–4. Wages then increase monotonically up to firm bins 11–12 and decrease thereafter.[30] Moreover, the wage drop among the most productive firms decreases in worker ability. It is most pronounced for low- and medium-ability workers.[31] For high-ability workers, the wage levels off (bins 41–50, new matches).

This nonmonotonic, hump-shaped relation between wages and firm productivity poses a challenge for empirical models of wage dispersion. From a theoretical perspective, complementarities may imply that wages are a hump-shaped function of firm type (e.g., Gautier and Teulings, 2006; Eeckhout and Kircher, 2011). However, many structural models cannot generate a hump-shaped relation between wages and firm productivity.

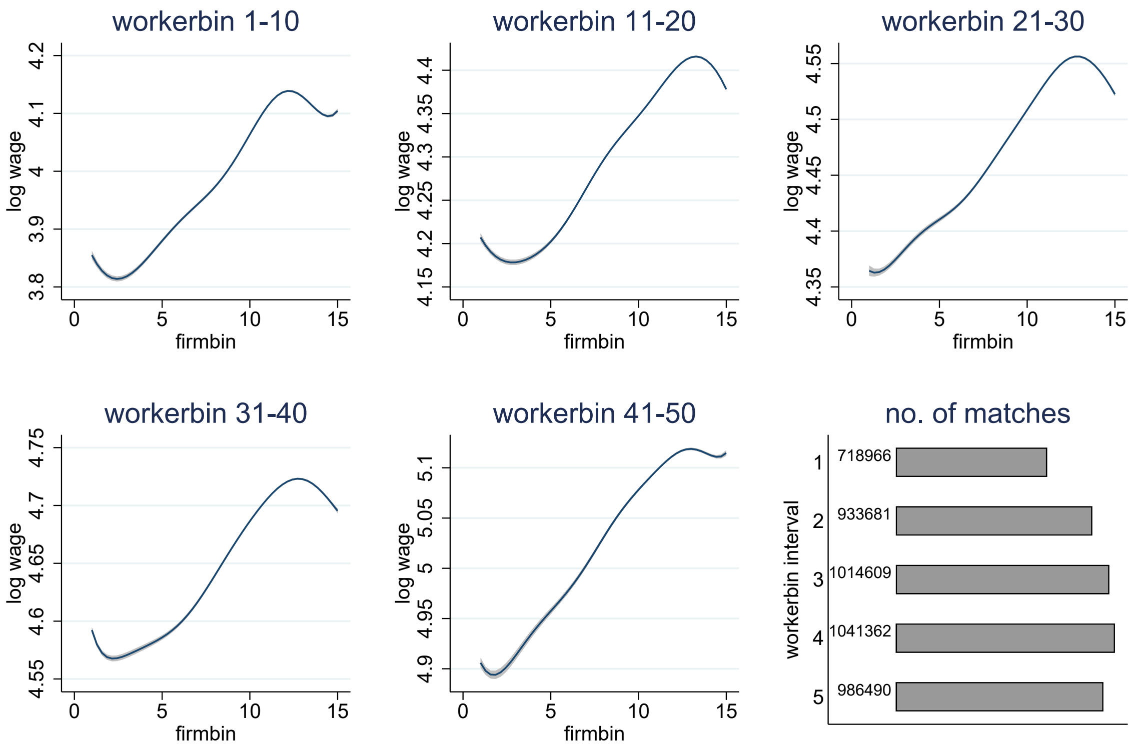
[29] The respective plots for matches OON and J2J flows are relegated to Appendix Figure D.6.

[30] The maximum around firm bin 12 is consistent with the firm performance measures shown in Figure 2, Section 4.2. These firms (around the 80th percentile of the productivity distribution) pay the highest average wages and are the largest firms in terms of headcount. The firms above are smaller, pay less on average, and exhibit greater labor productivity as well as lower labor shares.

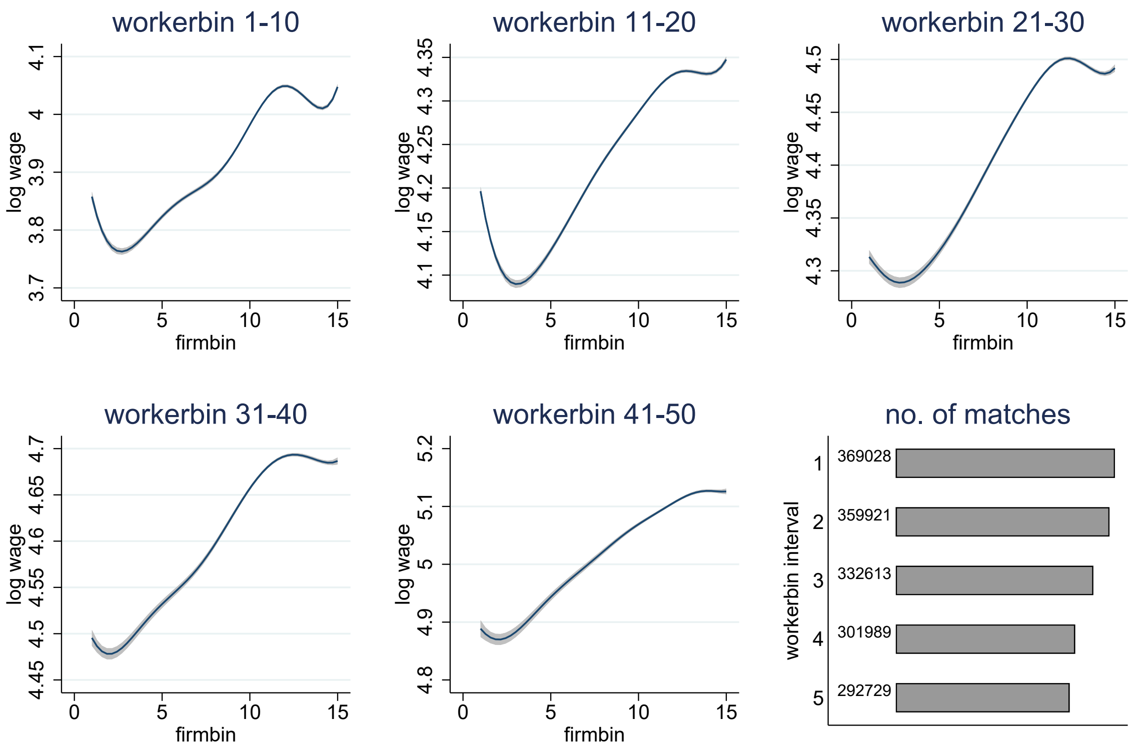
[31] The average wage difference between bin 15 and bin 12 firms for a bin 11–20 worker in terms of (deflated) log daily wages is approximately 4% (all matches), or approximately 1,177 euros annually.

Figure 4: Wage–Productivity Profiles

(a) All Matches



(b) New Matches



Notes: The plots show the estimated wage–productivity profiles across firm bins for all matches (a) and new matches (b). Plots are based on kernel-weighted local polynomial regressions. Kernel: Gaussian. The bandwidth is 2. Shaded areas are 95% confidence bands. Data sources: BHP, EP, BeH.

Consider two widely used models of on-the-job search (OJS). The most productive firm pays the highest wage in the wage-posting model of Burdett and Mortensen (1998). In the sequential auction model of Postel-Vinay and Robin (2002), wages decline in the poaching firm's productivity. This can explain the U-shaped relation at the bottom (firm bins 1–5) but not the hump-shaped relation at the top. We further discuss the potential for OJS models to rationalize our findings in Section 7.

In the model that we use to motivate our empirical strategy, wages are monotonic in firm productivity. They increase in firm productivity at a decreasing rate (see Appendix A.4). To generate a hump shape, the model needs a mechanism that allows the wage to be locally decreasing in firm productivity. To explore how such a mechanism could arise in an extended model, we take a closer look at the two key features of the empirical wage–productivity profiles: the observed wage drop among the most productive firms and the fact that this drop decreases in worker ability. To this end, we rely on the estimated AKM wage components. Appendix Figures D.7 and D.8 reproduce Figure 4 using the firm fixed effects and AKM residuals on the vertical axis instead of wages.

Looking at the firm fixed effects (Figure D.7), we observe a hump shape at the top of the distribution across all worker types, including high-ability workers. This suggests that the wage decrease in the most productive firms is driven by firm-specific factors. According to our model, the firm fixed effect captures a time average of the effect of firm productivity on wages. In an extended model with multidimensional firm heterogeneity, the firm effect would also absorb the wage effect of other firm characteristics, e.g., the nonwage characteristics of jobs (amenities), which have the potential to reduce wages. We discuss how adding multidimensional heterogeneity would change our model and empirical approach in Section 7.

Looking at the AKM residuals (Figure D.8), we observe large positive values for the matches between the most productive firms and high-ability workers. For these worker types, the positive wage residuals compensate for the relatively small firm-specific pay components among high-productivity firms, so the hump shape is less pronounced. According to our model, the residual absorbs the effect of transitory productivity shocks on wages. The quantitative importance of this channel might be higher among the most productive firms, and it might also covary with worker ability. Moreover, the positive residuals could indicate that production complementarities are more complex than those in our simple model. For example, high-ability workers might increase the marginal product of other worker types, which could, in turn, add to their wage. Cahuc et al. (2008) study such mechanisms in detail, but they are absent in our simple model due to the assumptions needed to nest the AKM approach. In Section 7, we explore how more complex production functions relax the link between wages and productivity and how they would affect our empirical approach.

Finally, we check the robustness of the empirical wage–productivity profiles. Concerns about their validity might be related to measurement error in the data used to rank workers and firms. On the worker side, we use the universe of employment spells available through the German social security registers. The estimated AKM worker effects could be sensitive to the number of years included. On the firm side, we rely on survey data, which might be prone to sampling error. Moreover, the firm-year productivities $\hat{\omega}_{jt}$ could be measured with error. We first assess the sensitivity of the wage–productivity profiles to the inclusion of fewer years in the sample used to rank workers (2003–2008). The S-shape is robust (see Figure D.9). Second, we re-estimate the wage-productivity profiles using 20 random firm subsamples (drawn with replacement and clustered at the firm level). The S-shape is robust (see Figure D.10). Third, to alleviate concerns about measurement error in ω_{jt}, we bootstrap the ACF estimation and assess the dispersion of estimated productivities at the firm-year level (see Figure D.11a). The mean deviation is less than 0.2% and more than 90% of errors lie within $\pm$ 9% of the bootstrapped firm-year-level mean. For every firm-year, we draw a new omega from the respective 95% confidence interval using uniform random numbers, rerank and rebin the firms, and reestimate the wage-productivity profiles. We repeat this 50 times. Again, the S-shape is robust (see Figure D.11b). Fourth, Figure D.12 confirms that the nonmonotonicities are not related to tenure effects. We re-estimate the wage–productivity profiles using the first match-year only. If anything, the wage drop at the top is more pronounced in this case. Additionally, Figure D.13 shows that firm types based on AKM firm fixed effects do not reveal any nonmonotonicities, as one would expect.

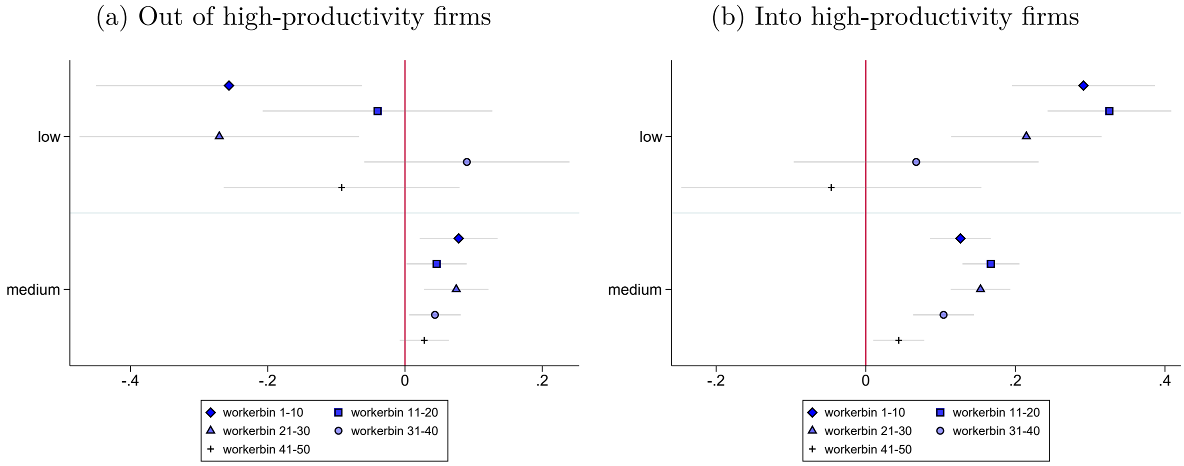
6.2 Worker Transitions

The nonmonotonicities documented in Figure 4 reflect the mean wage differences across firm types for a given set of worker types. The next step is to check whether observed worker transitions between firm bins are consistent with these wage–productivity profiles; i.e., do workers transition toward higher wages, even if this implies switching to a less-productive employer? In this case, transitions out of the most productive firms should lead to wage gains, at least for a subset of destination firm bins. Similarly, when moving up the productivity ladder, wage gains should be decreasing in the origin firm bin and might even be negative for transitions to the most productive firms.

Figure 5 shows that the wage changes for observed J2J transitions largely support these conjectures. We regress log wage differences for the same five groups of worker bins on a set of origin and destination firm bin dummies.[32] To simplify the graphical illustration, we also group the firm bins into high (13–15), medium (4–12), and low (1–3) bins. Panel (a) shows the estimated wage changes for transitions out of high-

[32]We measure the difference between the wage during the last spell in the pretransition firm and the wage during the first spell in the posttransition firm.

Figure 5: 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 among transitioning workers on dummies for origin and destination firm bins. The sample consists of new matches (J2J switches, no intermittent nonemployment 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 the destination/origin firm bin groups: low (bins 1–3) and medium (bins 4–12). Data sources: BHP, EP, BeH.

productivity firms. 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) when they transition from high-productivity to medium-productivity firms. For the highest-ability worker types, the wage differences are not significantly different from zero, which is consistent with the muted wage drop for these worker types. 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 that even transitions into low-productivity firms do not always come with a negative wage change. These findings suggest that job mobility and sorting are guided by wages and not by firm productivity. This also explains the trend of decreasing productivity sorting at the top.

Panel (b) presents the 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, and they decrease in the origin firm type. 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, which is in line with the S-shaped wage-productivity 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, wage gains are between 4.4% and 16.7% and decrease in worker ability. Based on the wage drops at the top in Figure 4, one might expect that some of these transitions lead to wage cuts. This is not the case, which reaffirms that workers tend to move toward higher wages.

In summary, workers select jobs to maximize their wages, even when this implies moving to a less productive firm. Upward transitions typically lead to wage increases as well, so workers do not accept wage cuts to move to the most productive firms. We again test the robustness of our findings by re-estimating the wage changes using random subsamples (drawn with replacement and clustered at the firm level). All findings are robust (see Figure D.14).

6.3 Wage Inequality

Song et al. (2019) show that two-thirds of the increase in wage inequality in the U.S. from 1978 to 2013 can be attributed to increasing pay differences between firms. Increasing wage sorting (high-wage workers into high-wage firms) and increasing segregation of workers contribute roughly equally to the rising contribution of firms to wage inequality. Panel (a) of Figure 6 shows that this trend is also present in the German data. We use establishment identifiers to decompose the variance in wages into the respective shares explained within and between establishments. Between 1998 and 2008, wage inequality grew because the between-firm component (filled circles) grew by approximately 10%. The within-establishment contribution (hollow circles) is of similar magnitude but stable.

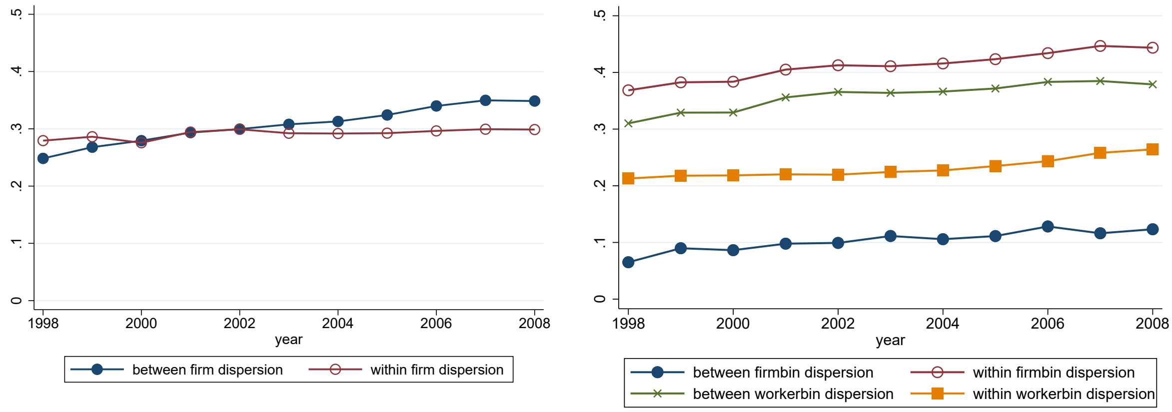
In Panel (b), we decompose wage inequality based on firm productivity and worker ability. The variance between the firm bins (filled circles) does increase but only by approximately 3%. It contributes little to overall inequality. The variance within firm bins (hollow circles) is most important. Worker segregation, that is, wage differences between worker bins (crosses), is also a major contributing factor. The increase in these two components dominates the increase in between-firm inequality. The contribution of within-worker bin inequality (squares) is moderately large, but it increases only in the second half of the sample period. We conclude that increasing pay differences between firm productivity types are not the major reason for increasing wage inequality in Germany.

We link the increasing variance of the different pay components to the wage–productivity profiles in Figure 4 by analyzing how the components change over time (1998–2002 and 2003–2008); see Figure D.15. There are large differences in wage growth across worker types, and they correspond to increasing worker segregation (the between-worker component in Figure 6b). For low-ability workers, wages decrease by more than 10% for all firm types above the very bottom. For medium-ability workers, wages shrink among low-productivity firms but are relatively stable or somewhat increasing among more productive firms. For high-ability workers, wages increase among almost all firm types but most notably at the top. However, the hump shape has also become more pronounced over time, which explains the decreasing sorting of high-ability workers into high-productivity firms. Overall, the wage–productivity profiles become steeper, which implies an increase in between-firm wage inequality consistent with Figure 6b.

Figure 6: Decomposition of Wage Dispersion over Time

(a) Song et al. (2019) replication

(b) Our approach



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

7 Extensions

In this section, we discuss model extensions from two perspectives: (i) compatibility with the AKM approach and (ii) compatibility with nonmonotonic wage–productivity profiles. We consider OJS, general production functions, and multidimensional heterogeneity.

The model in Section 2 assumes that only unemployed workers search for jobs. Can the AKM and ACF approaches still be combined if the DGP features OJS? Workers' wages could depend on their employment history, and thus, worker types would be identifiable only through wages out of unemployment. To circumvent this issue, it is common to use functional form assumptions that separate worker ability and labor market history. For example, Postel-Vinay and Robin (2002) show that the wage is log-linear in the worker type in the sequential auction model if the utility function features CRRA. Our model relies on a comparable assumption to obtain a log-linear wage equation. We assume that worker ability units are perfectly substitutable at the firm level, so the marginal product of labor scales with the worker type, which implies additive separability in logs. This similarity leads us to suspect that an OJS version of our model can, for specific production functions, retain a wage equation that is log-linear in the worker type. In this case, the AKM approach can be used to rank workers in the presence of OJS.[33]

The second question is whether our model could generate wage productivity that is hump-shaped through OJS. In the sequential auction model, the productivities of both the poaching firm and the previous employer affect the wage. On the one hand, the wage

[33] DiAddario et al. (2021) show that an augmented AKM model with origin- and destination-firm effects resembles the wage equation of sequential auction models for specific utility functions. They find a small covariance between the worker fixed effect and the origin-firm effect in Italian data, which supports the separability of worker ability and labor market history.

increases in the productivity of the previous employer. On the other hand, it decreases in the productivity of the poaching firm due to the option value of future wage gains. Overall, the effect of firm productivity on wages is nonlinear. However, the magnitude of the second effect decreases in firm productivity because it becomes increasingly unlikely for workers to meet an even more productive firm as they move up the productivity ladder. This implies a U-shaped relation between firm productivity and wages, not a hump-shaped relation. In summary, while canonical OJS models can yield a log-linear (in the worker type) wage equation, the implied relation between wages and firm productivity is at odds with the hump-shaped wage–productivity profiles.

Next, we consider more general production functions. Although the wage–productivity relation flattens among high-productivity firms, the simple production function motivating our empirical strategy is inconsistent with a hump-shaped relation between productivity and wages. This reflects our assumption that firm productivity augments the effect of worker ability at the match level so that decreasing returns to labor dampen higher productivity's effect on output. This implies (see Appendix A.4) that wages increase with firm productivity at a decreasing rate. Cahuc et al. (2008) show that firms exploit decreasing returns by strategically overemploying workers to reduce their marginal product and pay lower wages. With complementarities among worker types, employing a worker of one type can increase the other types' marginal products and, thus, their wages. This effect could outweigh the effect of decreasing returns and interact with firm productivity, heterogeneous worker bargaining power, or the firm's capital stock.

Such mechanisms seem flexible enough to generate a nonmonotonic relation between productivity and wages and generate a less pronounced hump shape for high-ability workers. A more complex production function would probably make the wage equation nonlinear in the worker type, and the AKM model would not be nested therein. However, the literature provides more flexible methods for ranking workers (and firms) based on unobserved heterogeneity. For example, Bonhomme et al. (2019) allow for nonlinear wage equations. Future research could explore combining their method with ACF-style production function estimates.

Finally, consider the addition of multidimensional heterogeneity. First, we discuss multidimensional *firm* heterogeneity. Specifically, we consider heterogeneous amenities in addition to productivity differences. Amenities, i.e., the positively valued nonwage characteristics of jobs, can give rise to wage differences between equally skilled workers (Rosen, 1986; Mortensen, 2003). To explain the relatively low wages at the top, the most productive firms would need to offer a higher level of amenities than somewhat less productive firms. To test this hypothesis, we need a firm-level measure of amenities. The EP survey contains a number of questions whose answers are indicative of the prevalence of amenities. We use them to construct a firm-level amenity index.[34] We regress this index

[34]The questions relate to the flexibility of working hours, working time accounts, parental benefits,

on the productivity rankings, controlling for (2-digit) industry and firm size (headcount). The most productive firms indeed offer the most amenities (see Figure D.16). Another amenity to consider is job security. High-productivity firms might be relatively unlikely to lay off workers, and this perceived job security could entice workers to accept lower wages (as in Jarosch, 2021). We compute the rate of employer-initiated separations. The most productive firms have the lowest firing rates (see Figure D.17).

The finding that high-productivity firms provide a high level of amenities can explain the hump shape at the top. However, less productive firms pay low wages despite offering few amenities, so wages depend on both firm productivity and amenities (and worker ability). Important issues in setting up a model with two-dimensional firm heterogeneity include how the amenity level is determined, how it influences wages, and how these extensions impact the applicability of the AKM approach. Firms might invest in amenities after favorable productivity shocks. This would be consistent with the positive correlation between amenities and productivity. Whether AKM can be applied in this case depends on the production function. With homogeneous worker preferences for amenities and assuming that amenities do not directly affect the marginal product of labor, nothing changes. The wage effect of amenities would be absorbed by the firm fixed effect.

Second, we consider multidimensional *worker* heterogeneity. Specifically, we consider heterogeneous preferences for amenities or outside options in addition to ability differences. With heterogeneous preferences for amenities, the composition of the workforce may depend on amenities (as in Lamadon et al., 2022).[35] If the composition of the workforce does not matter for output, e.g., due to the perfect substitutability of ability units, nothing changes. If heterogeneous amenity preferences are combined with richer complementarities in the production function, the marginal product of a given worker type could depend on amenities because those amenities might attract other types that affect the marginal product. This would again call for a more flexible approach to ranking workers. With similar implications, one could allow for heterogeneity in the workers' outside options, e.g., due to coworker networks or referrals (e.g., Dustmann et al., 2015).[36]

Another issue is whether amenities change the estimated productivity rankings. We take into account ability heterogeneity when constructing labor inputs, so the composition

equal opportunity agreements, equity participation, and the prevalence of on-the-job training. We define an indicator as 1 if an amenity exists and as 0 otherwise. For each establishment year, we sum up the indicator values and divide by the total number of indicators for that year.

[35]Lamadon et al. (2022) combine heterogeneous preferences for amenities with monopsonistic wage-setting in a model of imperfect competition. Consistent with our findings, they find that high-productivity firms offer a high level of amenities and pay relatively low wages.

[36]This type of heterogeneity could also explain wage differences conditional on firm productivity and thus influence how pronounced the hump shape is. For our empirical strategy, considering differences in workers' outside options would be comparable to allowing for heterogeneous amenity preferences. As long as one is willing to assume that such differences do not affect the workers' contribution to output, nothing changes. However, with richer complementarities in the production function and the influence of networks on firms' hiring decisions, a more flexible approach is needed to rank workers.

of the workforce may depend on amenities. Beyond this composition, amenities do not affect output because they are not a production input. Thus, adding amenities as a second dimension of firm heterogeneity to our model should not affect the estimated firm rankings. Reassuringly, adding the amenity index as a control variable into the production function estimation leaves the rankings virtually unchanged.[37]

8 Conclusions

We present a parsimonious search-matching model that elucidates the assumptions necessary to combine the AKM and ACF approaches. The model is compatible with both the discrete-time model of dynamically optimizing firms underlying the ACF approach and the log-linear wage equation underlying the AKM approach. It features worker and firm heterogeneity, matching frictions, decreasing returns, multiworker firms, intrafirm bargaining, and worker–firm complementarities.

Based on worker ability (inferred from the AKM worker effects) and firm productivity (inferred using the ACF approach), we study productivity sorting, wages across firm productivity types, and inequality trends in the German labor market. Productivity sorting is positive, low, and relatively stable over time. Among the most productive firms, sorting decreases as high-ability workers become more likely to work at somewhat less productive firms that pay higher wages. This implies that sorting into high-productivity firms is quantitatively less important for rising wage inequality than wage sorting. Moreover, if workers move away from high-productivity firms to increase their wages, a side effect of this increasing wage sorting could be decreasing allocative efficiency in the labor market. Thus, studying productivity sorting is a useful complement to the well-established wage-based analysis of sorting.

We discuss several model extensions in regard to their potential to match the facts that we document. Canonical OJS models are compatible with log-linear (in the worker type) wage equations but incompatible with the hump-shaped relation between wages and productivity. With more general production functions, our framework could generate the hump shape, but the AKM model would no longer be nested. Models with multidimensional firm and/or worker heterogeneity can generate a hump shape with relatively simple production functions and thus remain compatible with the AKM approach.

References

Abowd, John M., Francis Kramarz, and David N. Margolis (1999) “High Wage Workers and High Wage Firms,” *Econometrica*, Vol. 67, pp. 251–333.

[37]These results are available upon request.

Abowd, John M., Francis Kramarz, Sébastien Pérez-Duarte, and Ian M. Schmutte (2018) "Sorting Between and Within Industries: A Testable Model of Assortative Matching," *Annals of Economics and Statistics*, Vol. 129, pp. 1–32.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer (2015) "Identification Properties of Recent Production Function Estimators," *Econometrica*, Vol. 83, pp. 2411–2451.

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.

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.

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

Burdett, Kenneth and Dale T. Mortensen (1998) "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, Vol. 39, pp. 257–273.

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, 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.

DiAddario, Sabrina L., Patrick M Kline, Raffaele Saggio, and Mikkel Sølvsten (2021) "It Ain't Where You're From, It's Where You're At: Hiring Origins, Firm Heterogeneity, and Wages.," Working Paper 28917, National Bureau of Economic Research.

Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker (2015) "Referral-based Job Search Networks," *The Review of Economic Studies*, Vol. 83, pp. 514–546.

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*, 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*, 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.

Gautier, Pieter A. and Coen N. Teulings (2006) "How large are Search Frictions?" *Journal of the European Economic Association*, Vol. 4, pp. 1193–1225.

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.

Jarosch, Gregor (2021) "Searching for Job Security and the Consequences of Job Loss," Working Paper 28481, National Bureau of Economic Research.

Kölling, Arnd (2000) "The IAB-Establishment Panel," *Schmollers Jahrbuch: Journal of Applied Social Science Studies*, Vol. 120, pp. 291–300.

Lachowska, Marta, Alexandre Mas, Raffaele Saggio, and Stephen A. Woodbury (2022) "Do firm effects drift? Evidence from Washington administrative data," *Journal of Econometrics*, forthcoming.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler (2022) "Imperfect Competition, Compensating Differentials, and Rent Sharing in the US Labor Market," *American Economic Review*, Vol. 112, pp. 169–212.

Levinsohn, James and Amil Petrin (2003) "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, Vol. 70, pp. 317–341.

Lise, Jeremy, Costas Meghir, and Jean-Marc Robin (2016) "Matching, Sorting and Wages," *Review of Economic Dynamics*, Vol. 19, pp. 63–87.

Lochner, Benjamin, Stefan Seth, and Stefanie Wolter (2020) "Decomposing the large firm wage premium in Germany," *Economics Letters*, Vol. 194, p. 109368.

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.

Mortensen, Dale (2003) *Wage dispersion: why are similar workers paid differently?*: MIT press.

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.

Postel-Vinay, Fabien and Jean-Marc Robin (2002) "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, Vol. 70, pp. 2295–2350.

Rosen, Sherwin (1986) "The Theory of Equalizing Differences," *Handbook of Labor Economics*, Vol. 1, pp. 641–692.

Shimer, Robert (2005) "The Assignment of Workers to Jobs in an Economy with Coordination Frictions," *Journal of Political Economy*, Vol. 113, pp. 996–1025.

Shimer, Robert and Lones Smith (2000) "Assortative Matching and Search," *Econometrica*, Vol. 68, pp. 343–369.

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: Journal of Applied Social Science Studies*, 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.

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

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 over $[0, 1]$. Meeting rates are functions of aggregate labor market tightness, $\theta = {}^{V}\!/_{U}$, where $V = \int g_v(\Omega)\mathrm{d}\Omega$ and $U = \int g_u(x)\mathrm{d}x$ are the aggregate 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 a meeting. Suppose that a type x worker meets a productivity Ω firm. Both parties may prefer to continue searching if 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 the steady state, existing matches can end only at the exogenous rate δ. Endogenous separations may happen outside of the steady state in the case that 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 is the solution to 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 its future discounted profits by posting costly vacancies given the expected evolution of its productivity:

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

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

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

${g_u(x)}/{U}$ 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 the case that $\mu(x, L, \Omega) = 0$, no additional type x workers are hired.

The match surplus is defined as

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

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

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

Below, we denote $\mu(x, L, \Omega) = 1$ ($\mu(x, L, \Omega_j) = 0$) 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

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

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

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

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

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

The envelope theorem implies that

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

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

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

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

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

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

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

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), we obtain

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

Therefore, the MPL for worker type x at an (L,Ω) firm equals the wage plus the 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

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

where $\alpha\in(0,1)$ is the workers' common bargaining parameter. The RHS captures the worker's surplus from 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 from 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 the steady state, the value of employment for the worker is

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

The value of unemployment is

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

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):

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

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

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

which can be combined with equation (A.13) to obtain

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

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

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

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

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

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

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

A.4 Wages, Labor Demand, and Productivity

It is worth considering how wages vary with firm productivity. Partially differentiating the wage equation (A.21) twice with respect to Ω yields

$$\frac{\partial w(x,L,\Omega)}{\partial \Omega} = \int_0^1 z^{\frac{1-\alpha}{\alpha}} \beta_l \frac{F_x(Lz,\Omega)}{\Omega} \mathrm{d}z > 0, \tag{A.22}$$

$$\frac{\partial^2 w(x,L,\Omega)}{\partial \Omega^2} = (\beta_l - 1) \int_0^1 z^{\frac{1-\alpha}{\alpha}} \beta_l \frac{F_x(Lz,\Omega)}{\Omega^2} \mathrm{d}z < 0, \tag{A.23}$$

which implies that wages increase in firm productivity but at a decreasing rate. This model property is a direct consequence of the assumed production function, which exhibits complementarities at the match level; see equation (2). Worker ability and firm productivity jointly determine how much one unit of labor of a given type contributes to output, but the total composite labor input is subject to decreasing returns. In other words, the positive effect of higher productivity on output is dampened by decreasing returns to labor, and this effect carries over into wages. For this reason, the second derivative of the wage with respect to productivity is negative.

Based on the solution to the wage equation in equation (A.21), we can also compute the derivative of the last term of (A.12) for the single labor case:

$$F_x(L,\Omega) = w(x,L,\Omega) + \frac{c(\Omega)(1-\beta(1-\delta))}{q_v(\theta)\frac{g_u(x)}{U}\mu(x,L,\Omega)} + (\beta_l - 1) \int_0^1 z^{\frac{1-\alpha}{\alpha}+\beta_l-1} F_x(Lz,\Omega) \mathrm{d}z. \tag{A.24}$$

This expression pins down labor demand because it equalizes the marginal product and the labor cost, which consists of the wage, turnover costs, and the effect of employment on the wage. The factor that we write in front of the integral is negative when $\beta_l < 1$. This implies that firms can reduce their labor costs by expanding employment in the presence of decreasing returns (overemployment, as in Stole and Zwiebel, 1996; Smith, 1999; Cahuc et al., 2008).

A.5 Linearity of the Wage Equation

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

$$w(x,L,\Omega) = (1-\alpha)(1-\beta)U(x) + x \int_0^1 z^{\frac{1-\alpha}{\alpha}} \beta_l \frac{F(Lz,\Omega)}{Lz} \mathrm{d}z. \tag{A.25}$$

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

To establish that the full wage equation is linear in x, the outside option $U(x)$ must also be linear in x. Consider the outside option according to equation (A.15). A straightforward assumption that ensures 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. In the model, this implies that $\mu(x, L, \Omega) = 1$ holds for all potential matches, and thus, the last term of equation (A.15) is zero. Now rearrange equation (A.14) such that

$$E(x, L, \Omega) = \frac{w(x, L, \Omega) + \beta\delta U(x)}{1 - \beta(1 - \delta)}. \tag{A.26}$$

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

$$U(x) = b(x) + \beta(1 - q_u(\theta))U(x) + \beta q_u(\theta) \int \frac{g_v(\Omega)}{V} \frac{w(x, L, \Omega) + \beta\delta U(x)}{1 - \beta(1 - \delta)} \mathrm{d}\Omega. \tag{A.27}$$

Plugging in our solution for the wage, equation (A.25), into this expression and collecting the $U(x)$ terms in front of the integral yields

$$\begin{aligned} U(x) = b(x) + \beta(1 - q_u(\theta))U(x) + \beta q_u(\theta) \frac{1 - \alpha + \beta(\alpha + \delta - 1)}{1 - \beta(1 - \delta)} U(x) \int \frac{g_v(\Omega)}{V} \mathrm{d}\Omega \\ + \frac{\beta q_u(\theta)}{1 - \beta(1 - \delta)} \int \frac{g_v(\Omega)}{V} x \int_0^1 z^{\frac{1-\alpha}{\alpha}} \beta_l \frac{F(Lz, \Omega)}{Lz} \mathrm{d}z \mathrm{d}\Omega, \end{aligned} \tag{A.28}$$

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

$$U(x) = \frac{(1 - \beta(1 - \delta))b(x) + x \frac{\beta q_u(\theta)}{1 - \beta} \int \frac{g_v(\Omega)}{V} \int_0^1 z^{\frac{1-\alpha}{\alpha}} \beta_l \frac{F(Lz, \Omega)}{Lz} \mathrm{d}z \mathrm{d}\Omega}{1 - \beta(1 - \delta - \alpha q_u(\theta))}, \tag{A.29}$$

where worker ability x can be written in front of both integral signs. Thus, for $U(x)$ to be linear in x, we must additionally assume that the workers' flow value of unemployment, $b(x)$, is proportional to x (i.e., $b(x) \propto x$), which is a standard assumption used in, e.g., Postel-Vinay and Robin (2002). Note that $U(x) = x\bar{U}$ is log-additive in x by the product rule: $\ln U(x) = \ln x + \ln \bar{U}$. This also applies to the integral in the wage equation (A.25).

B Details of Data Preparation

B.1 Wage Imputation

In the BeH data, earnings are right-censored at the contribution assessment ceiling ("Beitragsbemessungsgrenze"). This earnings limit is given by the statutory pension fund and is adjusted annually due to changes in earnings. In each year, we identify censored wage observations by comparing wages to the contribution assessment ceiling. We define a wage observation as censored whenever the reported wage is higher than 99% of the censoring thresholds. Following CHK and Dustmann et al. (2009), we fit a series of tobit regressions to impute the right tail of the wage distribution. We estimate the tobit regressions by year, sex, education and age group. In all these regressions, we also control for 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, mean years of schooling and the fraction of university graduates at the current establishment, the mean log wage of coworkers and the fraction of coworkers with censored wages, an indicator for individuals observed in only one year, an indicator for employees in one-worker establishments, and an indicator for region. 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 the censored wages as the sum of the predicted wage and a random component that is computed based on the 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

Employee education information is reported by employers each year and whenever a job ends. Its quality may suffer because employers do not face consequences for nonreporting or 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, individual educational attainment should not decline over consecutive job periods. We follow the procedure suggested by Fitzenberger et al. (2006), which exploits this reporting rule by assuming that there is some overreporting in the data.

The original education variable in the BeH defines six categories, plus an additional category for missing information. Following CHK, we convert these into the following categories: (0) missing; (1) primary/lower secondary or intermediate school leaving certificate, or equivalent, with no vocational training; (2) primary/lower secondary or intermediate school leaving certificate, or equivalent, with vocational training; (3) upper secondary school certificate with or without vocational training; and (4) some university degree. Within each job, we assign the modal education category observed for an individual during the years that he/she is at the same job.

C Production Function Estimation Details

We start with the production function to be estimated, which corresponds to equation (9). Lower-case letters indicate logs, and here, we ignore firm-level controls.

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + \omega_{jt} + \epsilon_{jt}, \tag{C.1}$$

Following ACF, the control function that we use is the demand for intermediate inputs *conditional on labor*[38] and capital:

$$m_{jt} = f_t(l_{jt}, k_{jt}, \omega_{jt}). \tag{C.2}$$

A natural interpretation of this demand function is that conditional on both labor and capital, more productive firms use more intermediate goods in production. Intermediate inputs m_{jt} are chosen either simultaneously with the labor input l_{jt} or afterwards, reflecting the fact that labor is a dynamic input.

Under the assumption of strict monotonicity in ω_{jt}, we can invert equation (C.2) to obtain $\omega_{jt} = f_t^{-1}(l_{jt}, k_{jt}, m_{jt})$, and substitute ω_{jt} out of (C.1):

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + f_t^{-1}(l_{jt}, k_{jt}, m_{jt}) + \epsilon_{jt} = \Phi_t(l_{jt}, k_{jt}, m_{jt}) + \epsilon_{jt}. \tag{C.3}$$

We adopt the two-stage procedure proposed by ACF. First, value added is regressed on a polynomial approximation of $\Phi_t(l_{jt}, k_{jt}, m_{jt})$. This does not identify any of the parameters but delivers the estimate $\hat{\Phi}_t(l_{jt}, k_{jt}, m_{jt})$ and separates productivity ω_{jt} from the transitory shocks absorbed by ϵ_{jt}.

To derive the second-stage estimation equation, we use the assumption that firm productivity follows an AR(1) process (equation 11). Thus, the conditional expectation of ω_{jt} formed at $t-1$ is $E\left[\omega_{jt}|\omega_{j,t-1}\right] = \rho\omega_{j,t-1}$, and the productivity innovation is denoted ξ_{jt}. Additionally, the following hold: $\omega_{j,t-1} = y_{j,t-1} - \beta_0 - \beta_l l_{j,t-1} - \beta_k k_{j,t-1} - \epsilon_{j,t-1}$ (from equation C.1) and $\epsilon_{j,t-1} = y_{j,t-1} - \hat{\Phi}_{t-1}(l_{j,t-1}, k_{j,t-1}, m_{j,t-1})$ (from equation C.3 with the first-stage estimate plugged in). Using these expressions along with the conditional expectation of firm productivity in equation (C.1) yields:

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + \rho(\hat{\Phi}_{t-1}(.) - \beta_0 - \beta_l l_{jt-1} - \beta_k k_{jt-1}) + \xi_{jt} + \epsilon_{jt}, \tag{C.4}$$

where $\xi_{jt} + \epsilon_{jt}$ is a composite error term. Following ACF, we use GMM and four (unconditional) second-stage moment conditions to identify the four parameters $(\beta_0, \beta_l, \beta_k, \rho)$:

$$E\left[\left(\underbrace{y_{jt} - \beta_0 - \beta_l l_{jt} - \beta_k k_{jt} - \rho(\hat{\Phi}_{t-1}(.) - \beta_0 - \beta_l l_{jt-1} - \beta_k k_{jt-1})}_{\xi_{jt}+\epsilon_{jt}}\right) \otimes \begin{pmatrix} 1 \\ k_{jt} \\ \hat{\Phi}_{t-1}(.) \\ l_{j,t-1} \end{pmatrix}\right] = 0 \tag{C.5}$$

[38] ACF use the conditional (on labor) input demand function to address the problem of functional dependence and improve identification of the labor input parameter relative to Olley and Pakes (1996) and Levinsohn and Petrin (2003).

The first three moment conditions are standard in the literature and identical to those in Levinsohn and Petrin (2003): (i) The composite error term $\xi_{jt} + \epsilon_{jt}$ is zero in expectation. (ii) The capital stock is predetermined—recall equation (10)—and therefore uncorrelated with the error term. (iii) The polynomial representation of the control function, estimated in the first stage and evaluated at $t-1$, $\hat{\Phi}_{t-1}$, is also uncorrelated with the composite error term because the intermediate input choice during the last period should be uncorrelated with the current productivity innovation and the transitory shock.

The fourth moment condition is specific to the ACF approach and reflects the fact that labor is a dynamic input. Only $l_{j,t-1}$ needs to be uncorrelated with the error term, while l_{jt} is allowed to covary freely with ξ_{jt}. In other words, labor is chosen in response to the current firm productivity realization, subject to across-firm differences in adjustment costs and wage setting (firm-specific, serially correlated, unobserved shocks to the price of labor). The conditional intermediate input demand (that is, conditional on l_{jt}) does not depend on these costs/shocks. The ACF approach does not allow for shocks to the price of intermediate inputs (due to the unobservable scalar assumption, which is required for the inversion of $f_t(l_{jt}, k_{jt}, \omega_{jt})$).[39] Olley and Pakes (1996) cannot allow for any input price shocks, Levinsohn and Petrin (2003) can allow for shocks related to k_{jt} but not to l_{jt} and m_{jt}, and ACF can allow for shocks related to k_{jt} and l_{jt} but not to m_{jt}.

Finally, note that the exclusion of m_{jt} from the production function implies that we estimate a value added production function (value added is simply defined as revenue minus the cost of intermediate inputs). One structural interpretation of this fact is that the production function is Leontief in the intermediate input, and this intermediate input needs to be proportional to output. As explained by ACF, their approach is designed for value added production functions and is not suitable for identifying the parameters of gross output production functions without further assumptions (see also Bond and Söderbom (2005) and Gandhi et al. (2020)).

D Additional Results

D.1 Wage Regressions

We estimate a log-linear wage equation for worker i who works at firm $j(i,t)$ in year t:

$$w_{it} = \alpha_i + \psi_{j(i,t)} + x'_{it}\beta + \varepsilon_{it}. \tag{C.6}$$

w_{it} represents log real daily wages, α_i represents the worker fixed effects, $\psi_{j(i,t)}$ represents the firm fixed effects, and x'_{it} contains time-varying controls (an unrestricted set of year dummies, quadratic/cubic terms for age interacted with education). ε_{it} is the residual.

[39] Arguably, intermediate inputs are often commodities, implying only little price variation across firms.

Table D.1: Wage Variance Decompositions

	(a) Regression (C.6) BeH, full		(b) Regression (C.6) BeH, women		(c) Regression (C.6) BeH, men		(d) Regression (C.6) BeH, men, West	
$\text{Var}(w_{it})$	0.276	(100%)	0.277	(100%)	0.245	(100%)	0.226	(100%)
$\text{Var}(\hat{\alpha}_i)$	0.126	(46%)	0.138	(50%)	0.105	(43%)	0.106	(47%)
$\text{Var}(\hat{\psi}_{j(i,t)})$	0.068	(25%)	0.076	(27%)	0.061	(25%)	0.049	(22%)
$\text{Var}(x'_{it}\hat{\beta})$	0.005	(2%)	0.006	(2%)	0.005	(2%)	0.005	(2%)
$2 \times \text{Cov}(\hat{\alpha}_i, \hat{\psi}_{(i,t)})$	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%)
$\text{Var}(\hat{\varepsilon}_{it})$	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 for log real daily wages based on regression model (C.6) for various BeH samples. Mean wages, variances, and covariances are rounded to three decimal places. Source: BeH.

Table D.1 shows the variance decompositions for multiple groups of workers based on the estimated AKM wage components. Column (a) includes all person-years, (b) all women, (c) all men, and (d) all men in West Germany (the CHK sample). We replicate the well-known finding that the majority of the wage variance is explained by unobserved worker heterogeneity; i.e., 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 the firm fixed effects, which explain roughly one-quarter of the wage variation across the four groups. The third largest determinant is the covariance between the worker and firm effects, which explains between 12% and 19% of the wage variance.[40] At only 2%, the share of the wage variance explained by the time-varying observable characteristics is almost negligible. The same is true for the covariances of the observable characteristics with the worker and firm effects. Note that time-invariant covariates such as education are absorbed by the worker effect. The residual explains between 7% and 9% of the wage variance across the four BeH samples. Estimated worker and firm fixed effects, as well as their covariance, could be biased due to limited mobility in the connected set.[41] We apply the parametric correction suggested by Andrews et al. (2008) for two subperiods in our data and find that the limited mobility bias is small.[42]

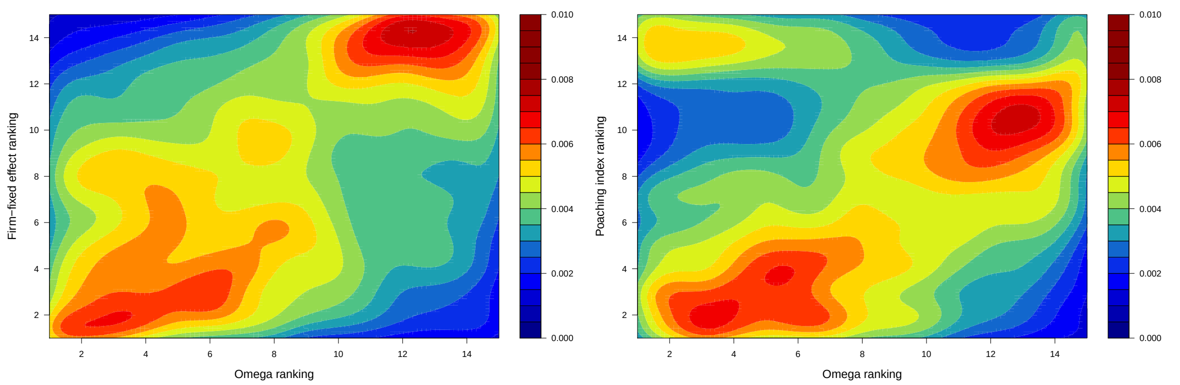
[40]Interestingly, women are less positively sorted in terms of wages than men. This is in line with what Card et al. (2016) and Bruns (2019) find using Portuguese and German data, respectively.

[41]The estimated variances of worker and firm fixed effects could be biased upward, and their covariance biased downward (Andrews et al., 2008, 2012; Borovičková and Shimer, 2020; Kline et al., 2020).

[42]The fact that we estimate the AKM model on the universe of German register data over a ten-year

Figure D.1: Comparison with Alternative Firm Rankings

(a) Firm fixed effects (b) Poaching index



Notes: The two plots depict the contours of the joint empirical distribution of firm-years across combinations of the ω ranking (15 bins) with the AKM firm-effects ranking (15 bins, Panel (a)) and the poaching index ranking (15 bins, Panel (b)). Data sources: BHP, EP, BeH.

D.2 Comparison with Alternative Rankings

We compare our productivity ranking with the results of 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 the AKM firm fixed effects and the second on the poaching index used in Bagger and Lentz (2019) and Taber and Vejlin (2020).[43] Correlations with the productivity ranking are positive, i.e., 0.280 and 0.119. To graphically analyze how the rankings are related, we create two sets of 15-firm bins based on both the firm fixed effects ranking and the poaching ranking. This approach 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 the firm fixed effect and poaching rank bins. Figure D.1 plots the contours of these empirical distributions.

In Panel (a), the AKM firm fixed effects bins are depicted on the vertical axis, and the ω ranking bins are depicted on the horizontal axis. The mass of observations is concentrated along the diagonal, which is in line with the positive correlation reported above. In the upper-right quadrant of the plot, the observations are highly concentrated somewhat

timeframe and include both men and women mitigates concerns regarding severe limited mobility bias due to the large number of observed transitions. Lochner et al. (2020) and Lachowska et al. (2022) confirm that this bias is limited in data sets with a large number of transitions. The Andrews et al. (2008) correction implies that in our setting, the variance in worker effects decreases by 5% (4%) and the variance in firm effects decreases by 4% (3%) in the 1998–2002 (2003–2008) subperiod. The covariance between the worker and firm effects increases by 7% (5%).

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

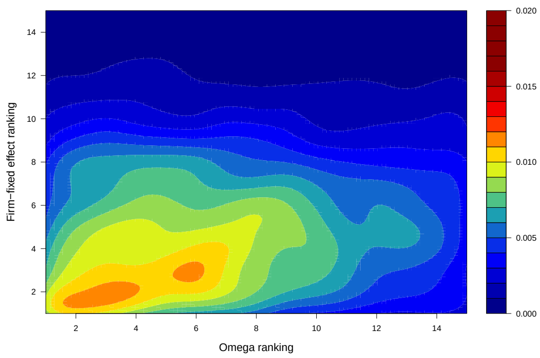
above the diagonal, thereby reaffirming our observation that the highest-paying firms (here, in terms of the AKM wage premia) are located below the top of the productivity distribution. In the lower-left quadrant, the observations are more dispersed. It is not uncommon to observe firm-years in which 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 D.2 shows that the observations in the lower-left quadrant are mainly young and small firms. The high-wage firms in the upper-right quadrant are older and larger firms. Panels (a) and (b) in Figure D.2 show that low-wage (high-wage) firms have, on average, a low (high) AKM rank; however, there is a sizable overlap between the two groups in terms of productivity.

In Panel (b), the mass of observations lies below the diagonal; that is, a firm's poaching rank tends to be lower than its productivity rank. Many high-productivity firms have high poaching ranks, but they are almost never located at the top. Additionally, it is not uncommon even for medium-productivity firms to hire mainly OON, 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. Apparently, some low-productivity firms very actively poach workers from other firms. Figure D.3 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 attempt to grow quickly by poaching workers from other firms. The larger and older firms, which also pay the highest wages, are concentrated in the upper-middle portion of the poaching rank distribution. They hire a nonnegligible number of employees OON.

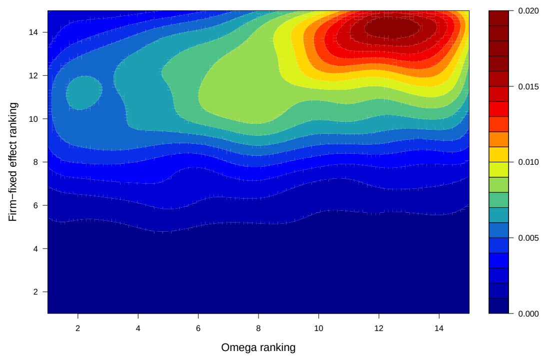
In summary, the comparison of the different rankings shows that firm ranks based on firm wage premia and observed worker mobility are systematically different from our productivity-based firm ranking.

Figure D.2: Comparison of Productivity-Based and Fixed Effect-Based Firm Rankings by Wages, Age, and Size

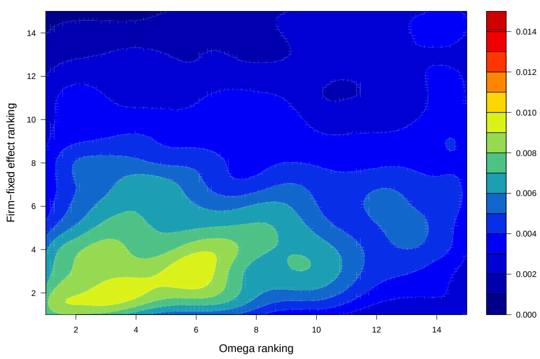
(a) Low-wage firms



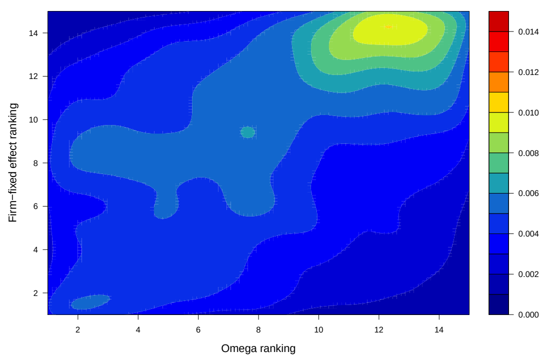
(b) High-wage firms



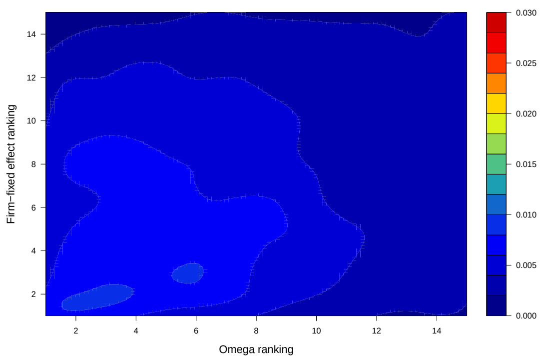
(c) Young firms



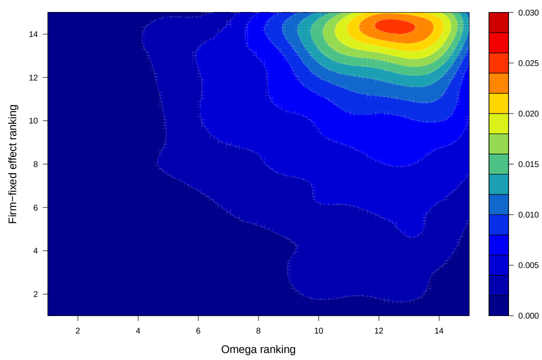
(d) Old firms



(e) Small firms



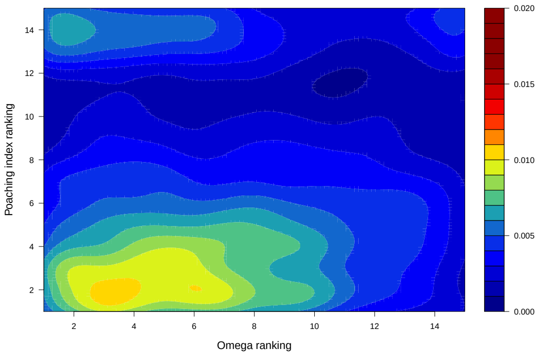
(f) Large firms



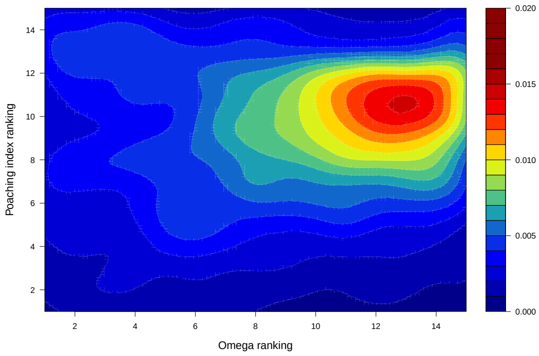
Notes: These six plots depict the contours of the joint empirical distributions of firm-years across combinations of the Ω ranking (15 bins) and AKM firm fixed effect ranks (15 bins). In Panels (a) and (b), high-wage firms are those that pay more than the grand mean of all firm-level mean wages, and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

Figure D.3: Comparison of Productivity-Based Firm Ranking and Poaching Index-Based Firm Ranking by Wages, Age, and Size

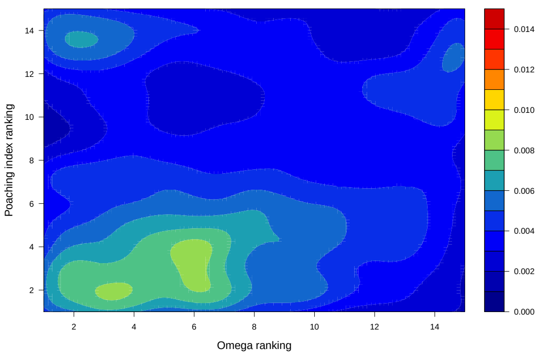
(a) Low-wage firms



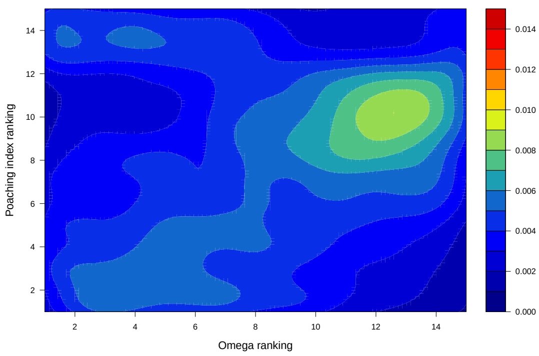
(b) High-wage firms



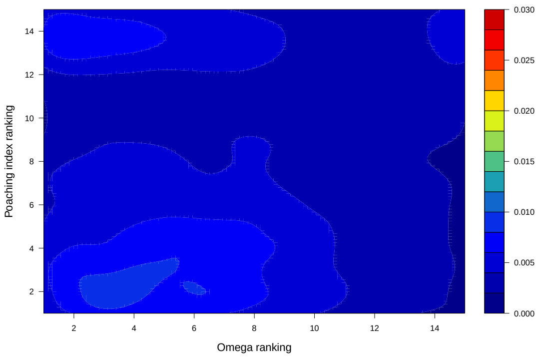
(c) Young firms



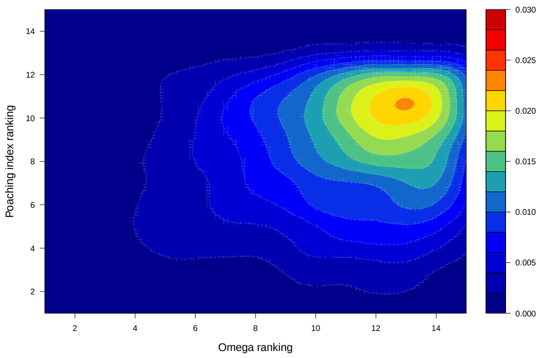
(d) Old firms



(e) Small firms



(f) Large firms



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

D.3 Rank Correlations

Table D.2: Spearman Rank Correlation Coefficients and Numbers of Observations for Different Time Intervals and Samples

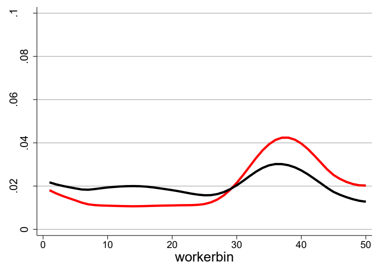
	All Matches	New Matches	Out of Nonemp.	Job-to-Job
1998-2008	0.065 (4,695,108)	0.124 (1,656,280)	0.132 (601,954)	0.110 (1,082,460)
1998-2002	0.055 (2,182,011)	0.139 (474,341)	0.141 (174,310)	0.120 (305,339)
2003-2008	0.074 (2,513,097)	0.118 (1,181,939)	0.129 (427,644)	0.107 (777,121)
1998	0.013 (311,861)	–	–	–
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 the worker and firm bins are statistically independent. All rank correlation coefficients are different from 0 at the 1% level of significance. Results are rounded to 3 decimal places. Numbers of observations (matches according to the respective definition) are reported in brackets. Data sources: BHP, EP, BeH.

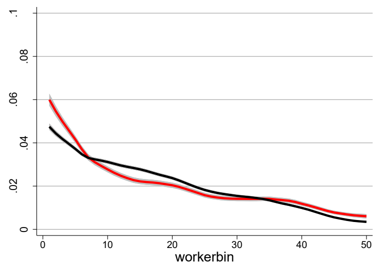
D.4 Distributional Dynamics

Figure D.4: Changes in the Worker Type Distributions within Different Firm Bins

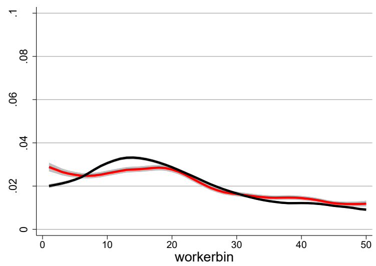
(a) Firm Bin 1, all matches



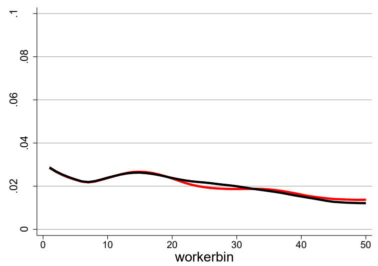
(b) Firm Bin 1, out of nonemp.



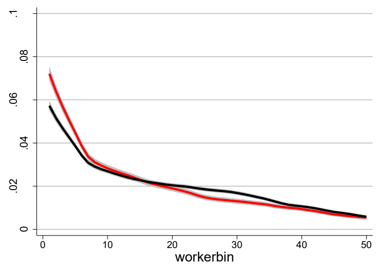
(c) Firm Bin 1, job-to-job



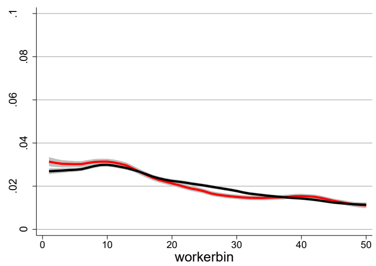
(d) Firm Bin 2, all matches



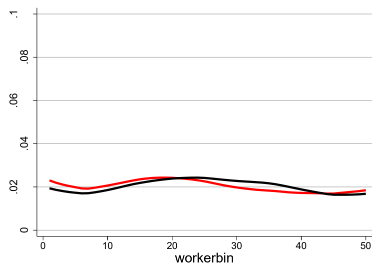
(e) Firm Bin 2, out of nonemp.



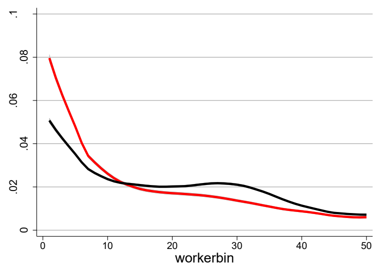
(f) Firm Bin 2, job-to-job



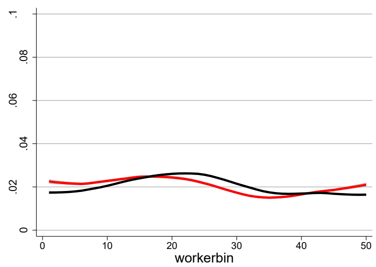
(g) Firm Bin 8, all matches



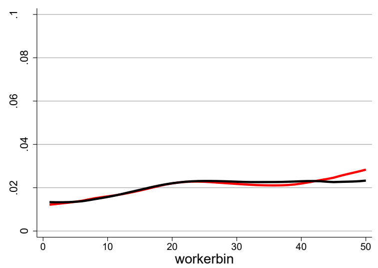
(h) Firm Bin 8, out of nonemp.



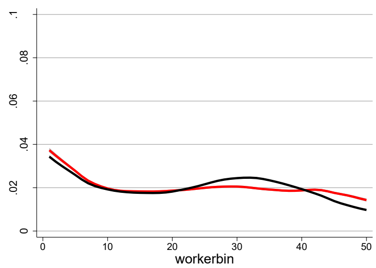
(i) Firm Bin 8, job-to-job



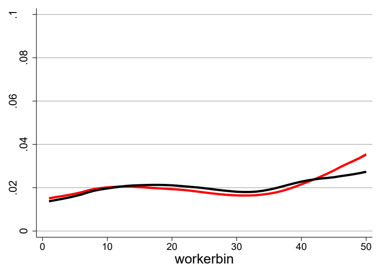
(j) Firm Bin 14, all matches



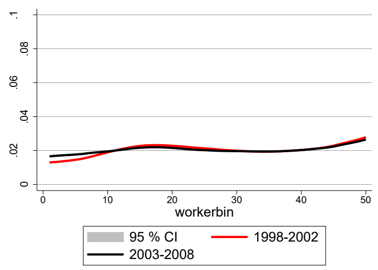
(k) Firm Bin 14, out of nonemp.



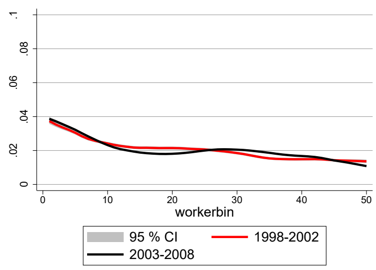
(l) Firm Bin 14, job-to-job



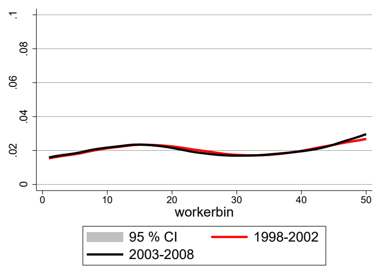
(m) Firm Bin 15, all matches



(n) Firm Bin 15, out of nonemp.



(o) Firm Bin 15, job-to-job

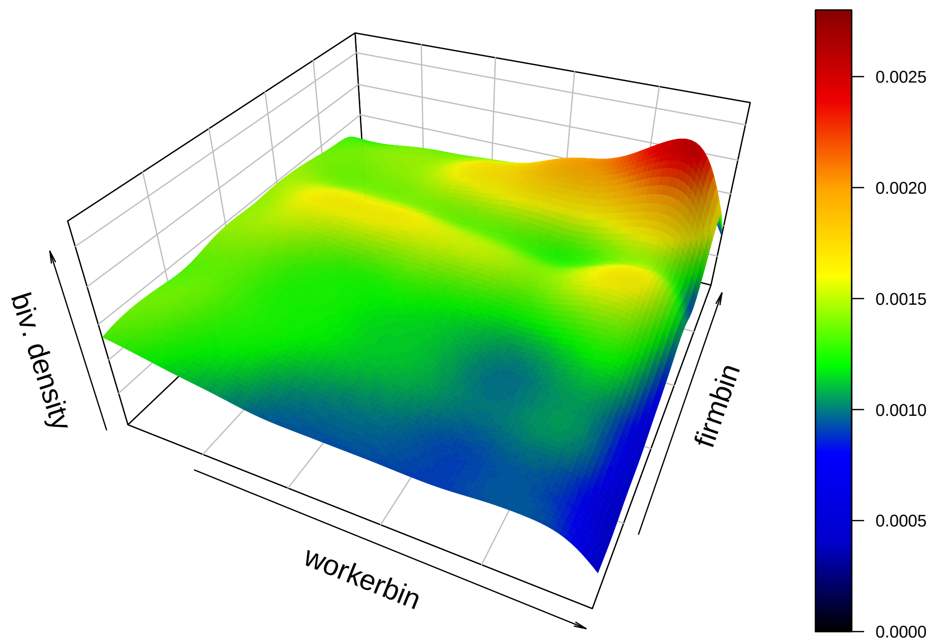


Notes: Estimated univariate kernel densities for all new matches conditional on worker bin, time, and match type. Kernel: Epanechnikov. The bandwidth is calculated with Silverman's rule of thumb. Pointwise confidence intervals are calculated using quantiles of the standard normal distribution. Data sources: BHP, EP, BeH.

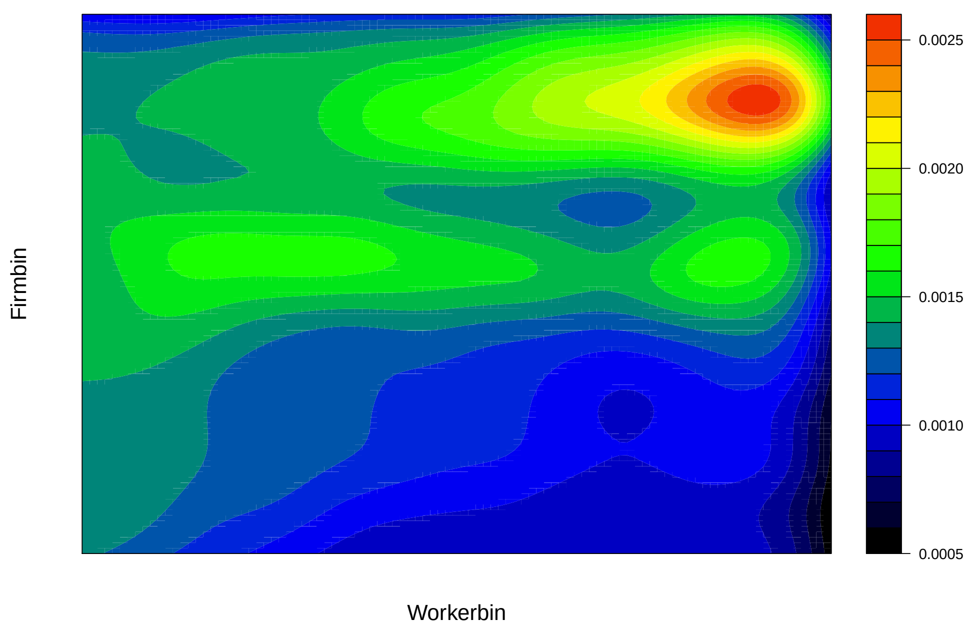
D.5 Joint Distribution of Matches

Figure D.5: Joint Distribution of All Worker-Firm Type Combinations (1998–2008)

(a) Joint Density



(b) Projection on the $(\bar{x}, \bar{\omega})$ plane

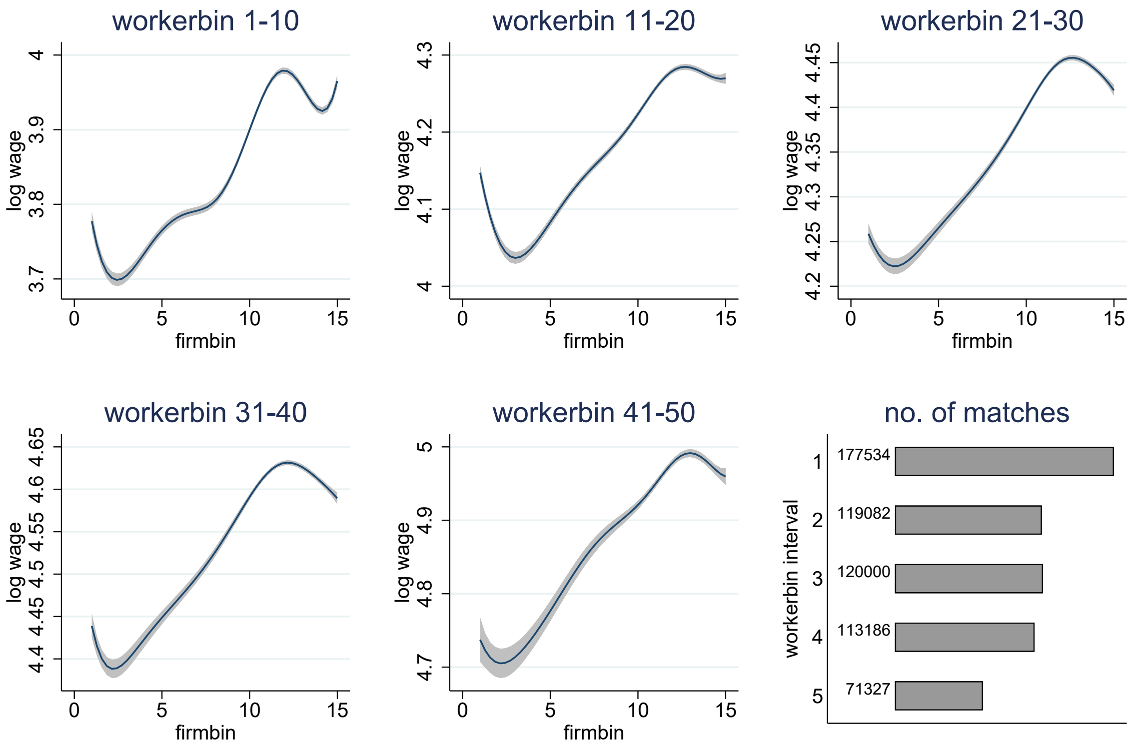


Notes: The plots show the estimated joint kernel density of matches and its projection onto the $(\bar{x}, \bar{\omega}$ plane for all combinations of worker and firm types in the sample of all matches on a grid with dimensions 50×15 (#worker types $\times$ #firm types). Data sources: BHP, EP, BeH.

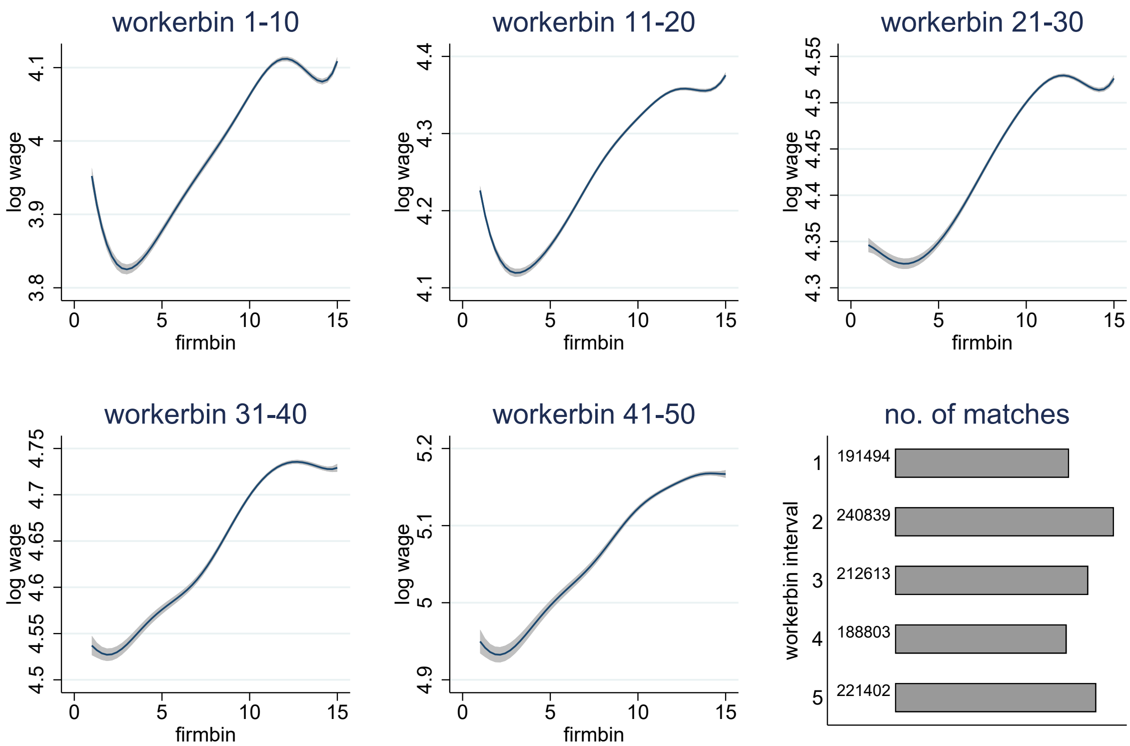
D.6 Wages and Transitions

Figure D.6: Wage–Productivity Profiles, New Matches

(a) Out of Nonemployment



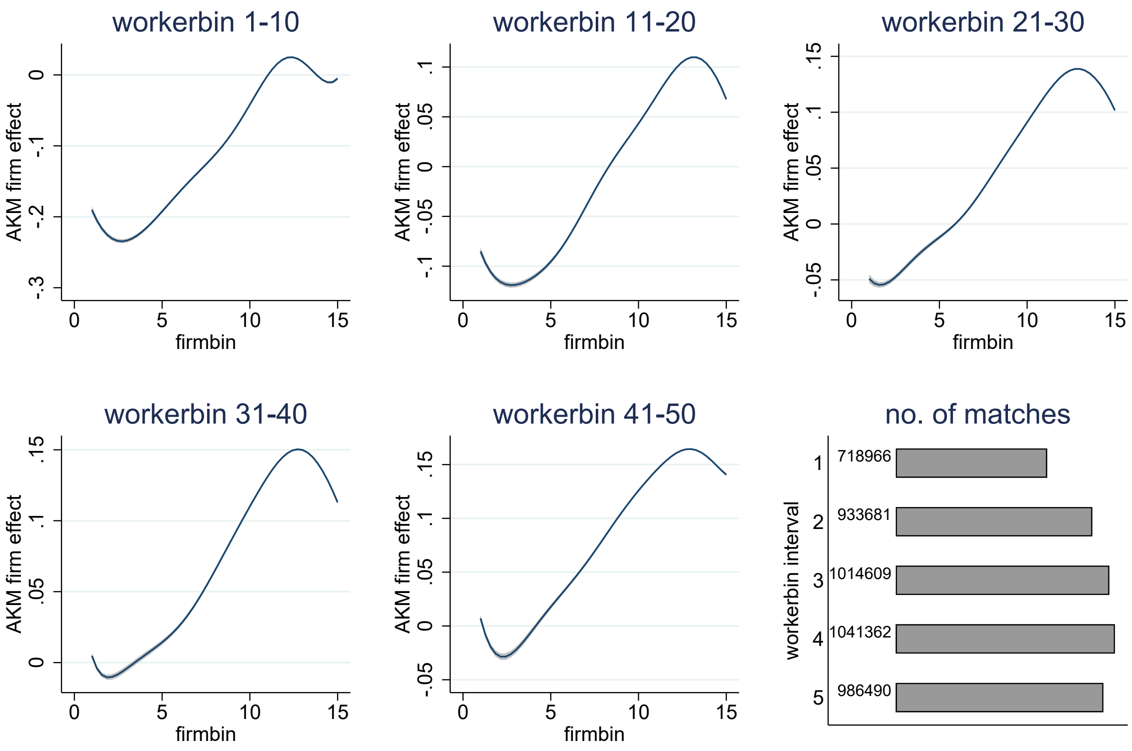
(b) Job-to-Job



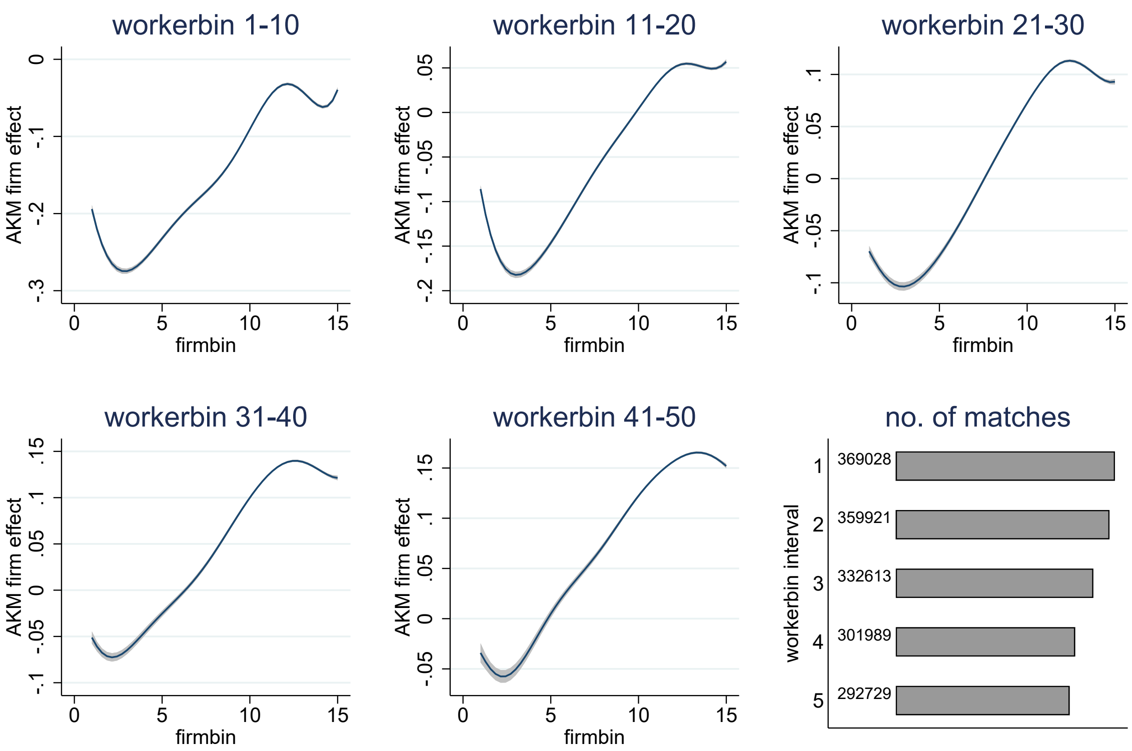
Notes: Plots show the estimated wage–productivity profiles across firm bins for new matches out of nonemployment and new job-to-job matches. Plots are based on kernel-weighted local polynomial regressions. Kernel: Gaussian. The bandwidth is 2. 95% confidence bands in gray. Data sources: BHP, EP, BeH.

Figure D.7: Wage–Productivity Profiles, Firm Fixed Effects Only

(a) All Matches



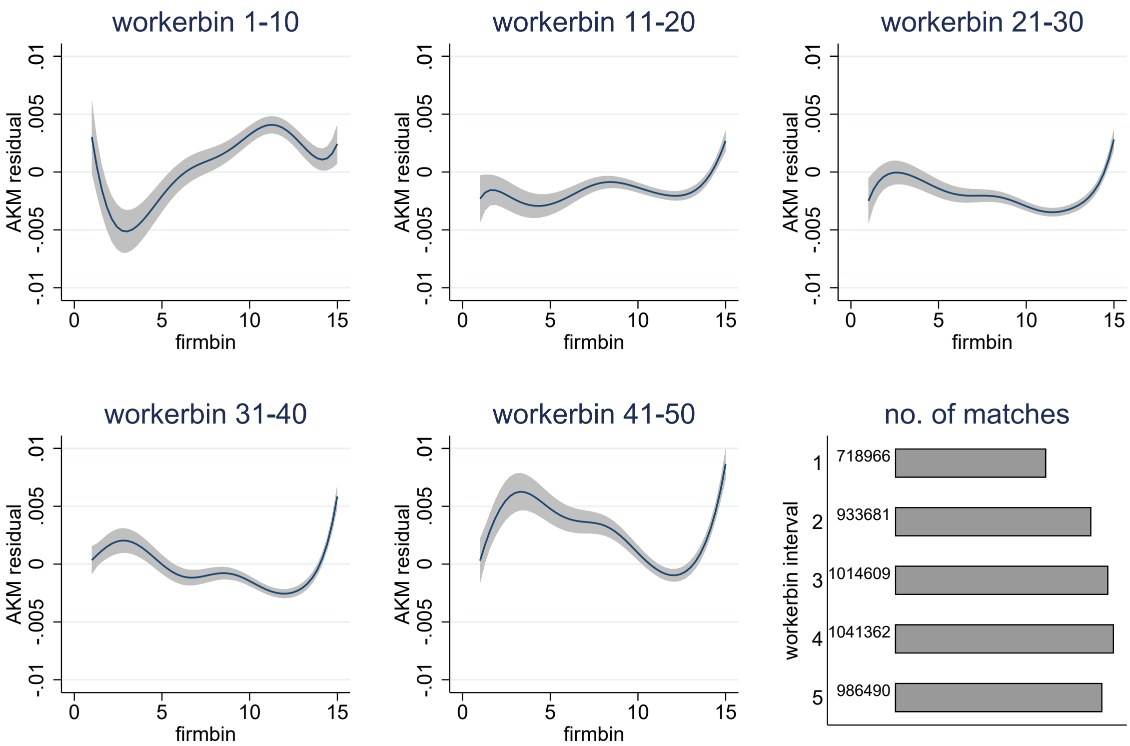
(b) New Matches



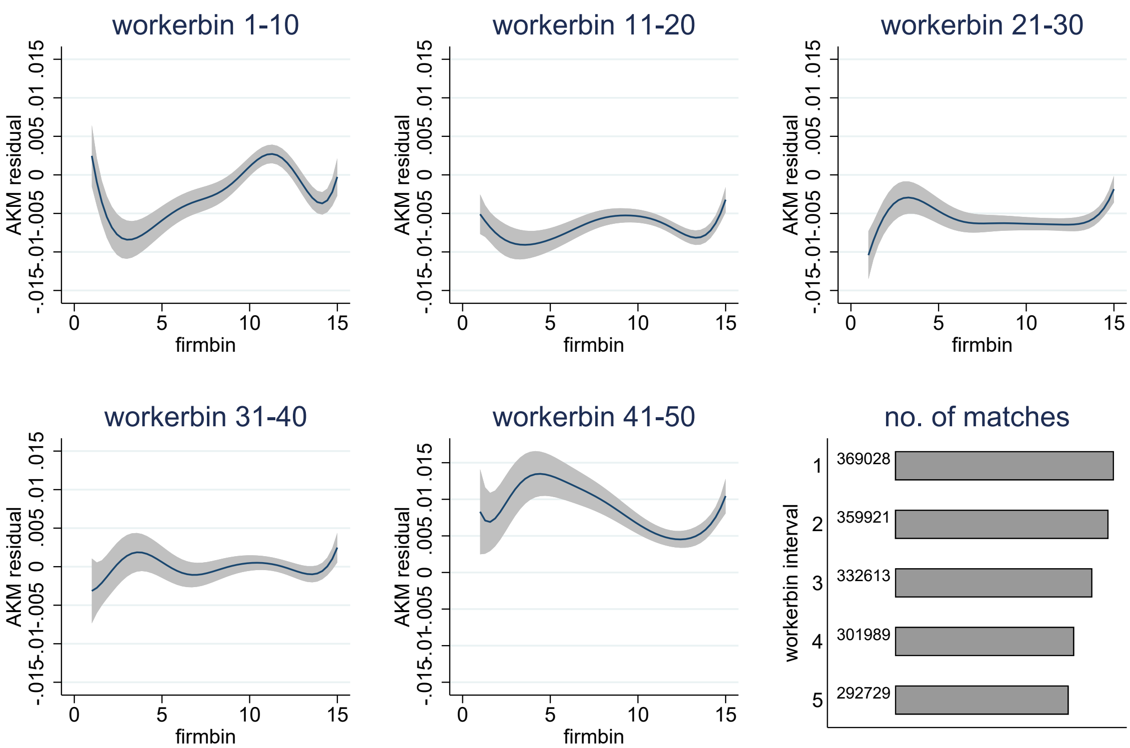
Notes: Plots show the estimated wage–productivity profiles across firm bins for all matches and new matches when using the AKM firm fixed effects as the wage variable. Plots are based on kernel-weighted local polynomial regressions. Kernel: Gaussian. The bandwidth is 2. 95% confidence bands in gray. Data sources: BHP, EP, BeH.

Figure D.8: Wage–Productivity Profiles, Residuals Only

(a) All Matches



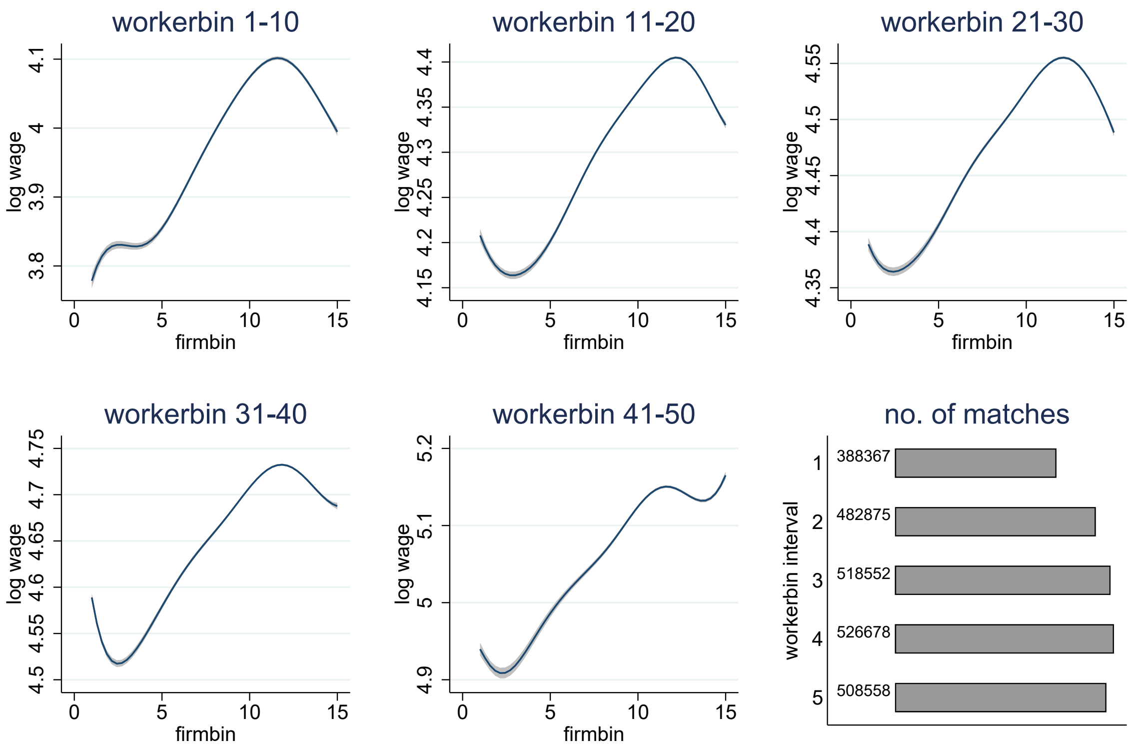
(b) New Matches



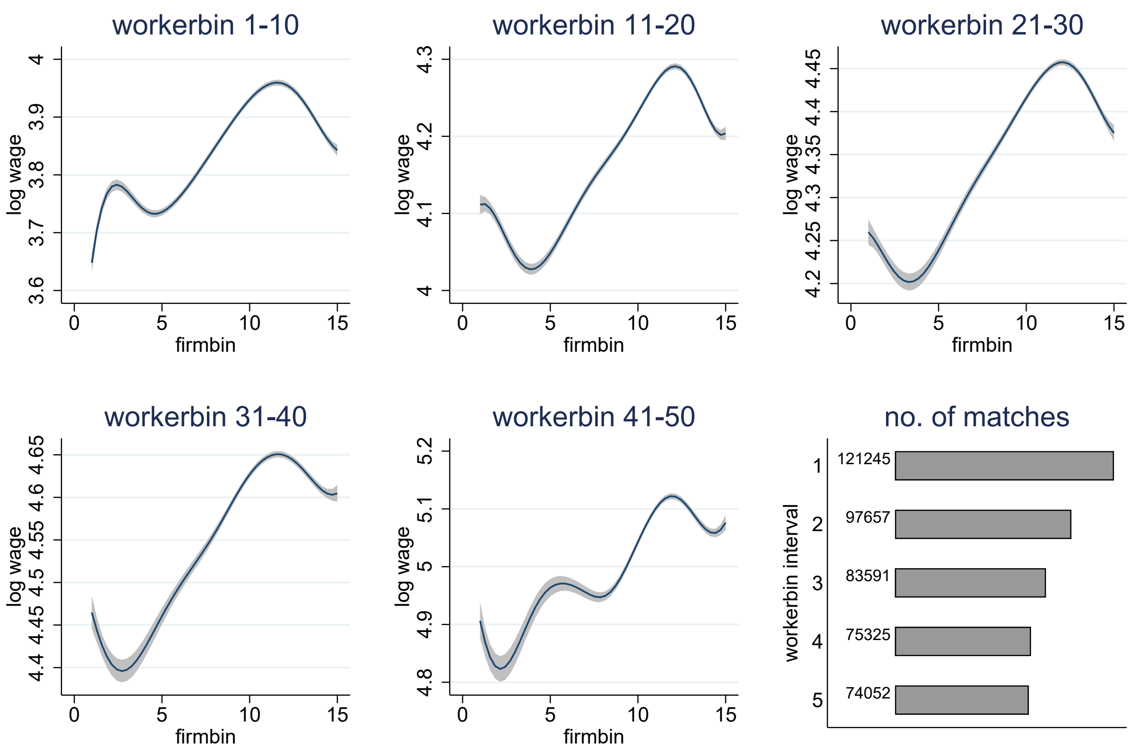
Notes: Plots show the estimated wage–productivity profiles across firm bins for all matches and new matches when using the AKM residuals as the wage variable. Plots are based on kernel-weighted local polynomial regressions. Kernel: Gaussian. The bandwidth is 2. 95% confidence bands in gray. Data sources: BHP, EP, BeH.

Figure D.9: Wage–Productivity Profiles, Worker Subsamples

(a) All Matches



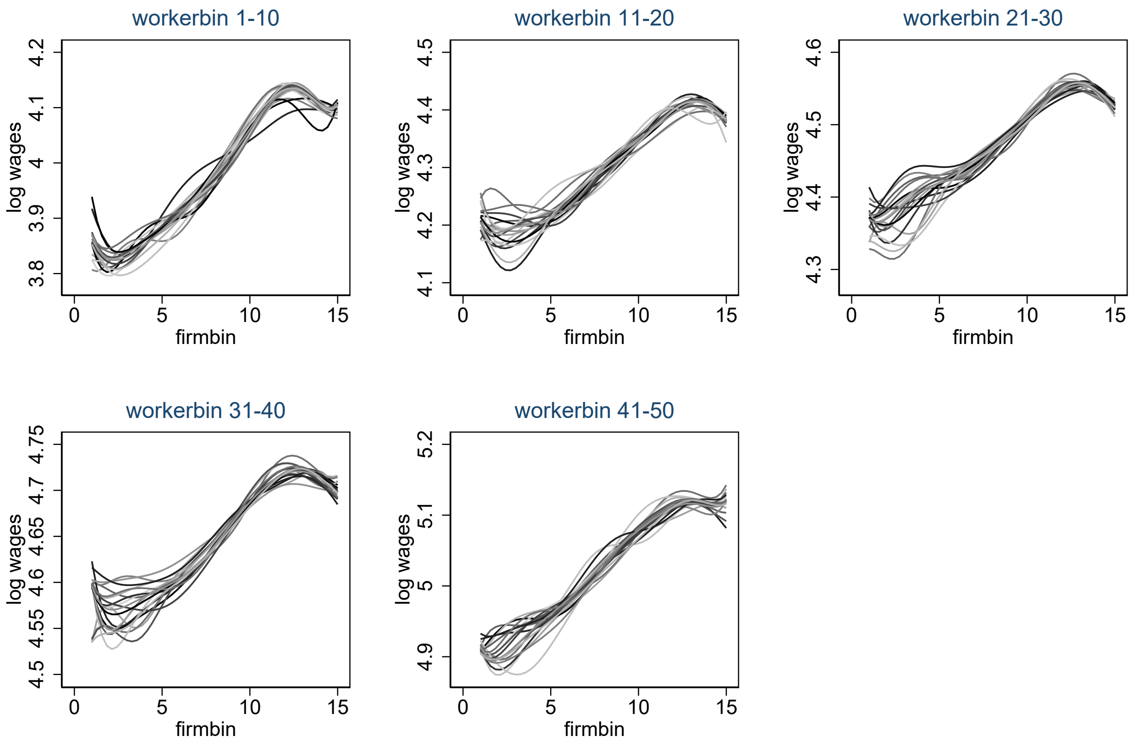
(b) New Matches



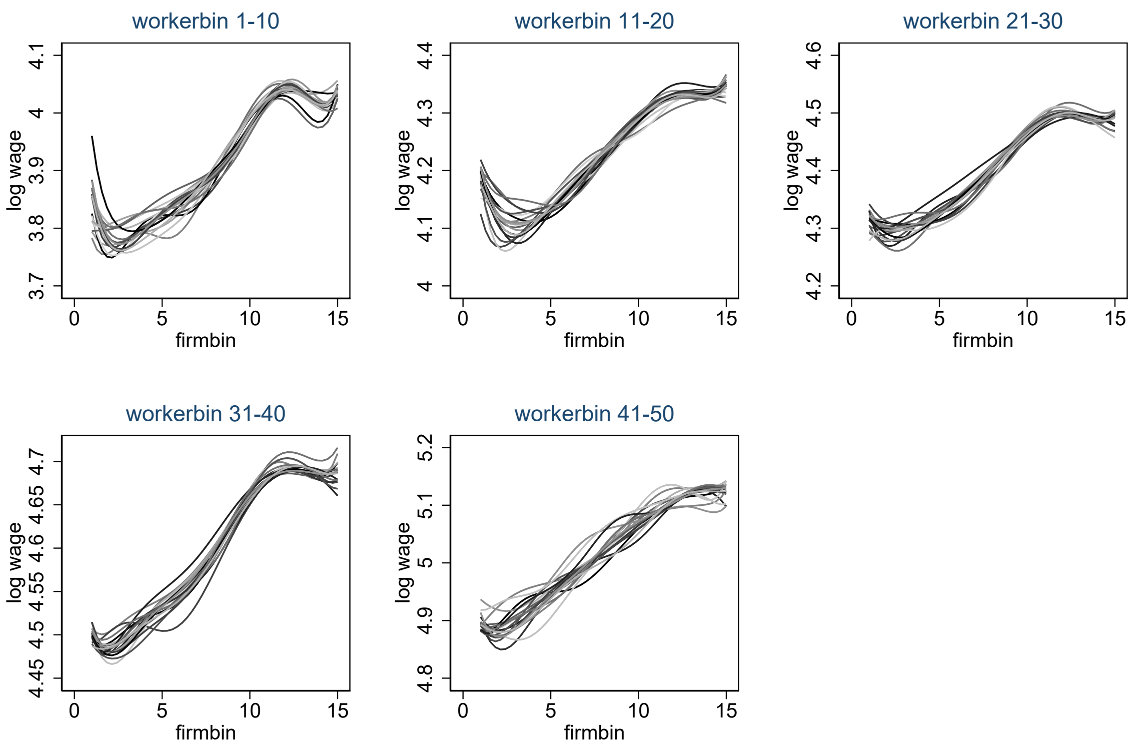
Notes: The plots show the estimated wage–productivity profiles across firm bins for all matches (a) and new matches (b). We use an alternative worker ranking based on the AKM worker effects from a shorter panel (2003–2008). Plots are based on kernel-weighted local polynomial regressions. Kernel: Gaussian. The bandwidth is 2. 95% confidence bands in gray. Data sources: BHP, EP, BeH.

Figure D.10: Wage–Productivity Profiles, Firm Subsamples

(a) All Matches



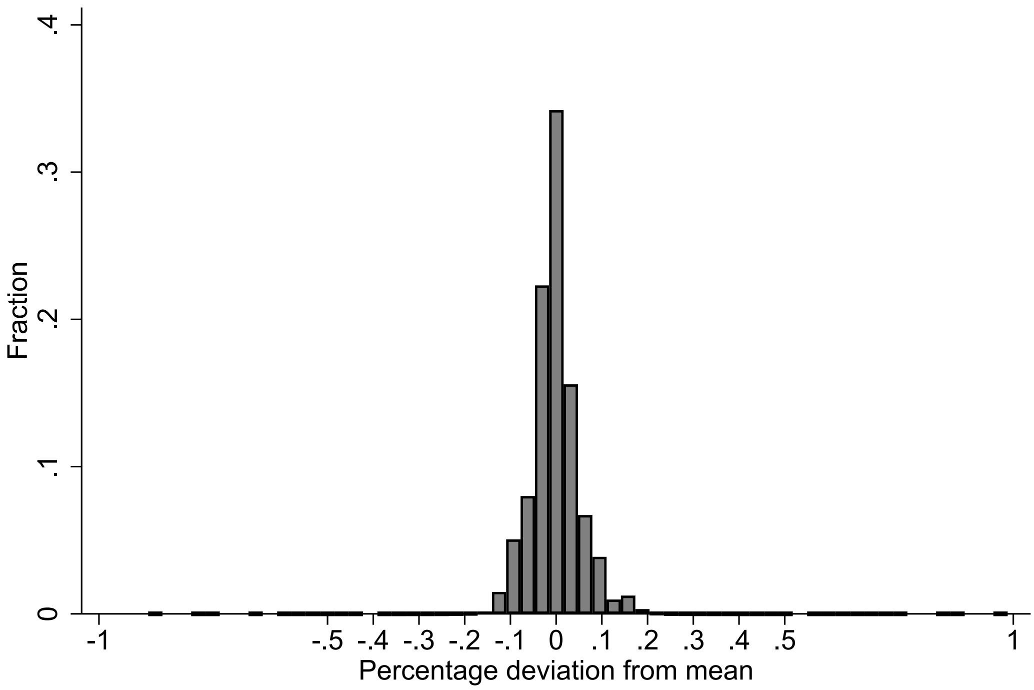
(b) New Matches



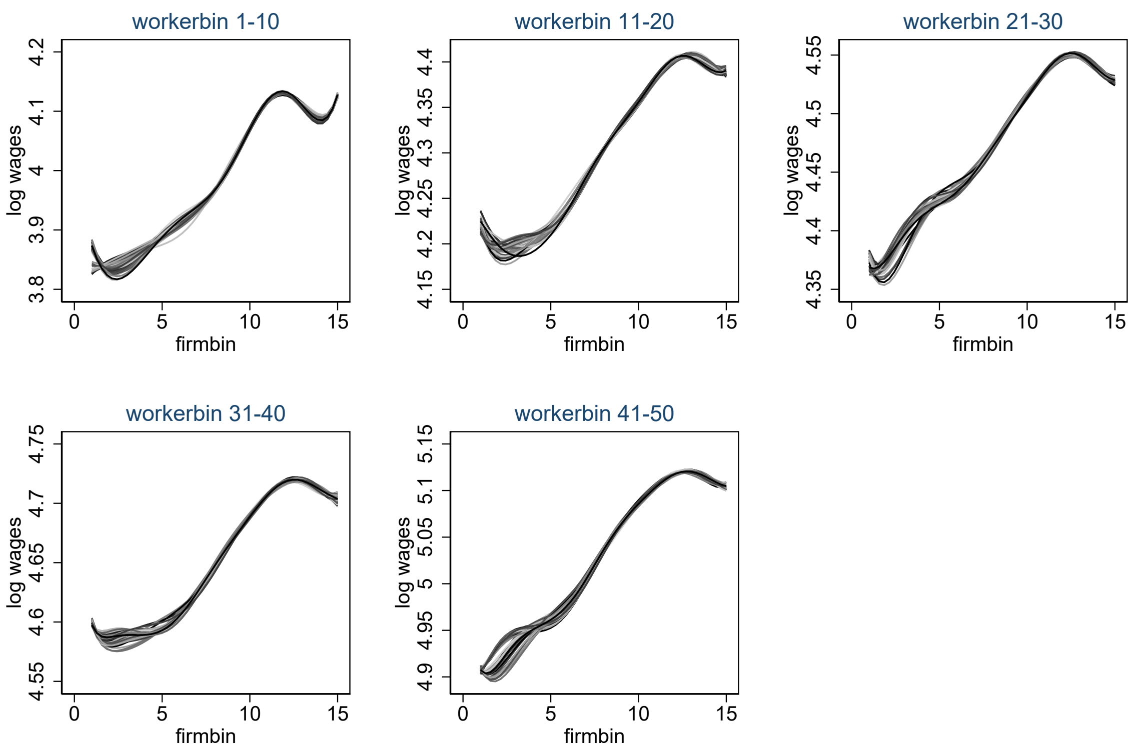
Notes: Plots show the estimated wage–productivity profiles across firm bins constructed from 20 random subsamples (with replacement, clustered by firm-year) from the original sample (size=N). The subsample size M varies randomly between $0.5 * N < M \leq N$. Plots are based on kernel-weighted local polynomial regressions. Kernel: Gaussian. The bandwidth is 2. Data sources: BHP, EP, BeH.

Figure D.11: Robustness to Measurement Error

(a) Histogram of Deviations



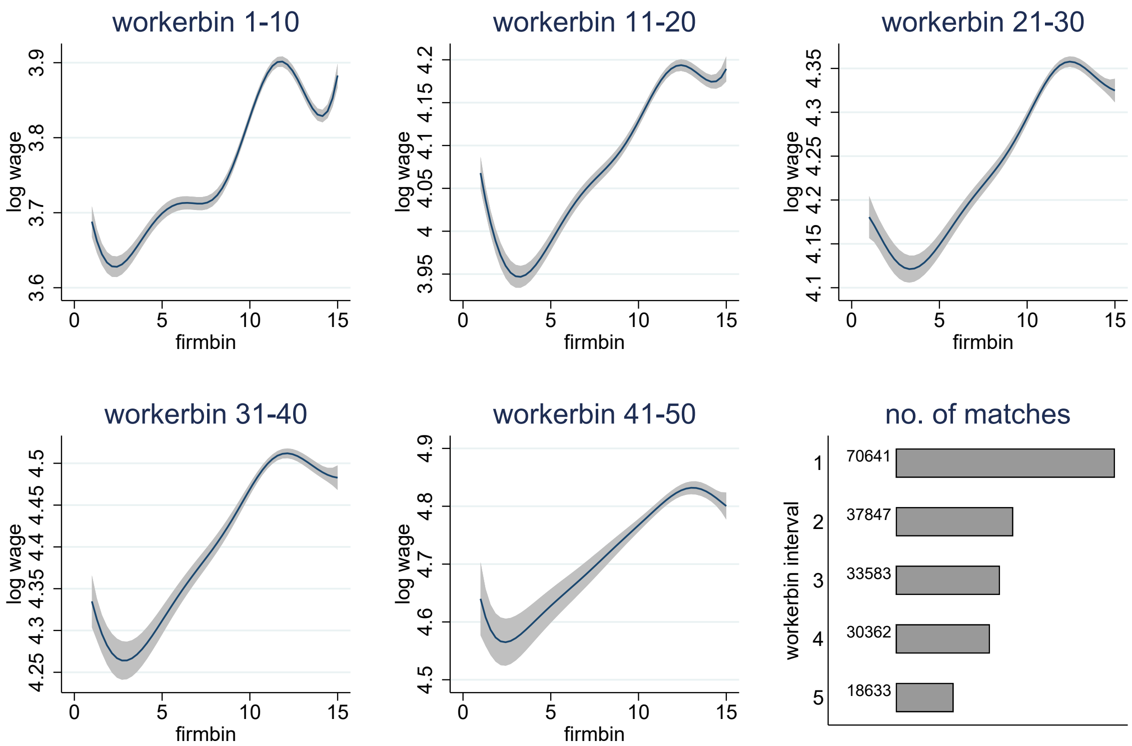
(b) Wage–Productivity Profiles, All Matches



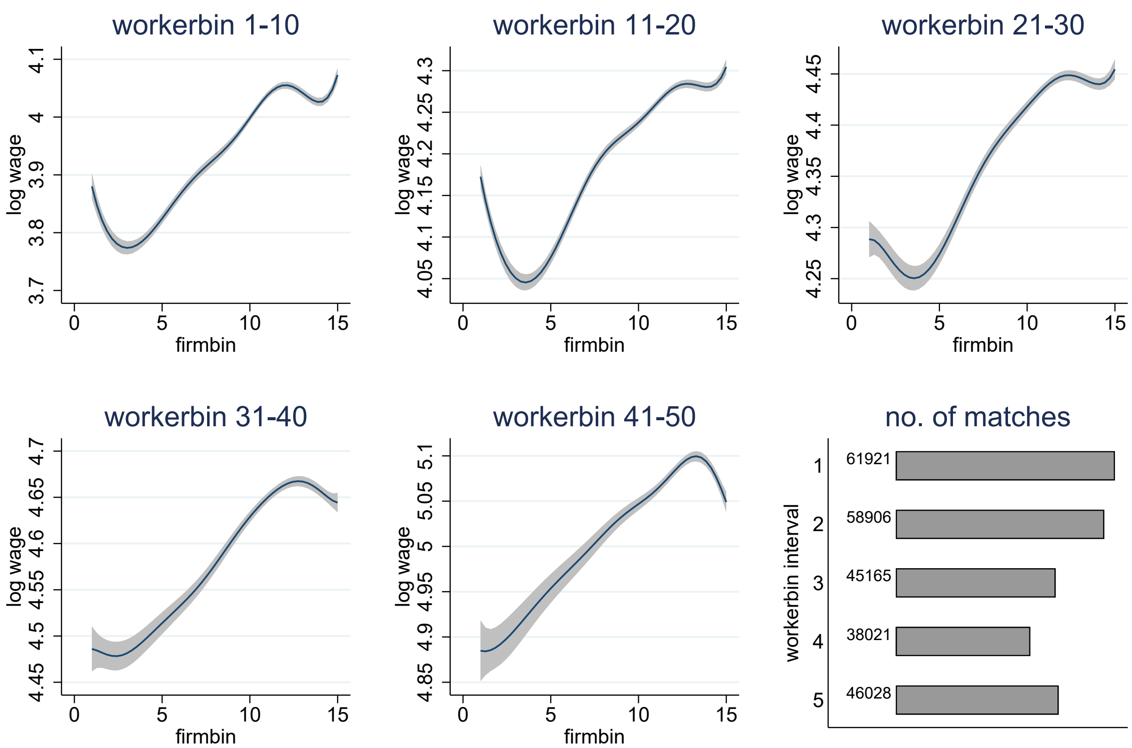
Notes: Panel (a) shows a histogram of deviations from the bootstrapped mean of $\hat{\omega}_{jt}$ measured in logs, so a deviation of 0.1 corresponds to a 10% deviation. Panel (b) shows the estimated wage–productivity profiles for all matches based on firm productivities that are drawn from the confidence band around every $\hat{\omega}_{jt}$ ($\pm$1.96 standard errors, draws based on uniform random numbers). 50 repetitions. Plots are based on kernel-weighted local polynomial regressions. Kernel: Gaussian. The bandwidth is 2. Data sources: BHP, EP, BeH.

Figure D.12: Wage–Productivity Profiles, New Matches, First Match-Year Only

(a) Out of Nonemployment



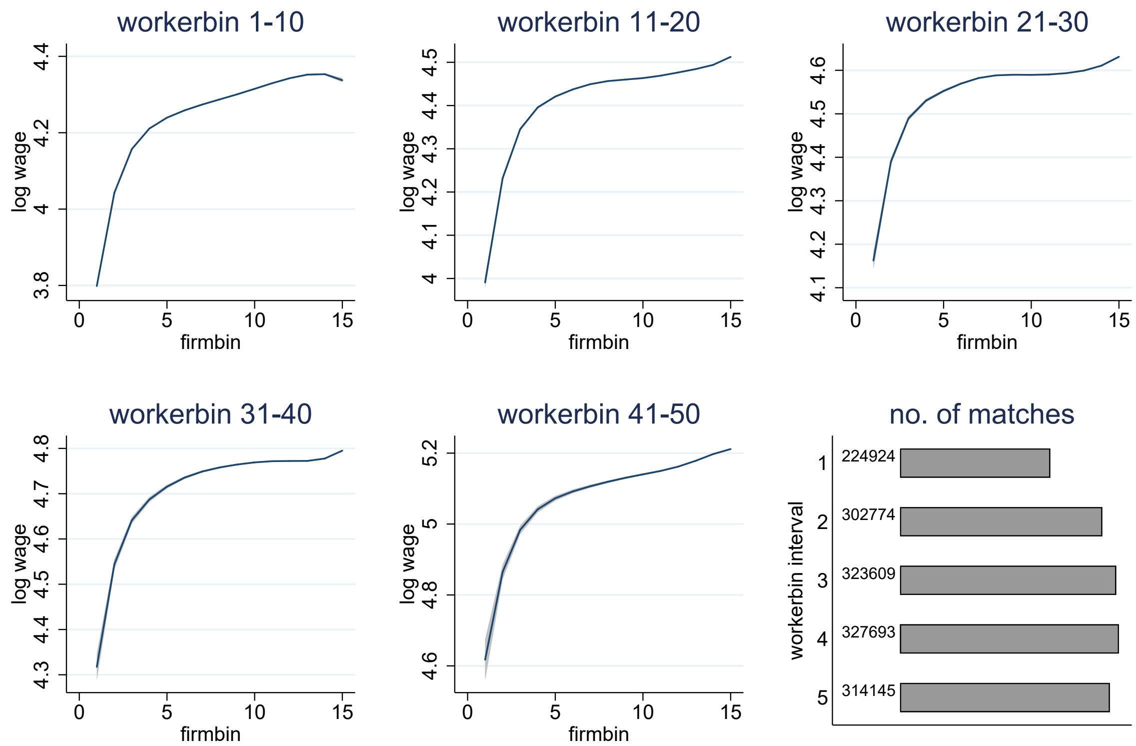
(b) Job-to-Job



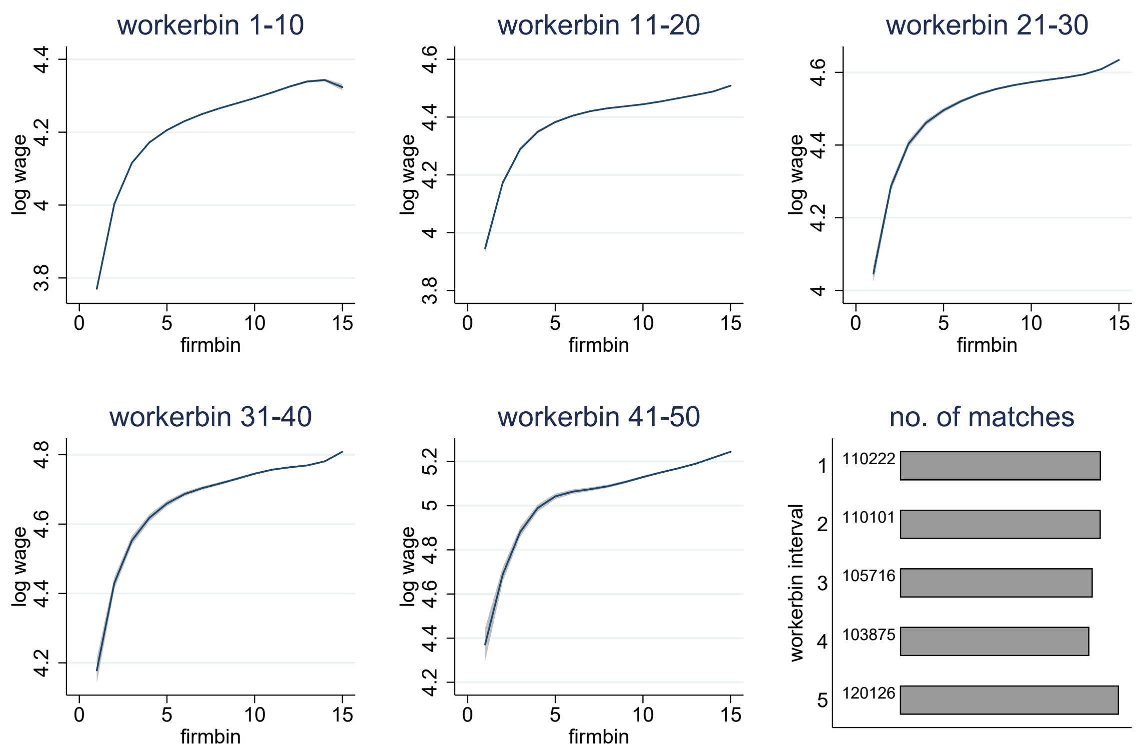
Notes: The plots show the estimated wage–productivity profiles across firm bins constructed using only the first yearly wage observation for all matches and for new matches to remove tenure effects. Plots are based on kernel-weighted local polynomial regressions. Kernel: Gaussian. The bandwidth is 2. 95% confidence bands in gray. Data sources: BHP, EP, BeH.

Figure D.13: Wage Profiles with AKM-Based Firm Types

(a) All Matches



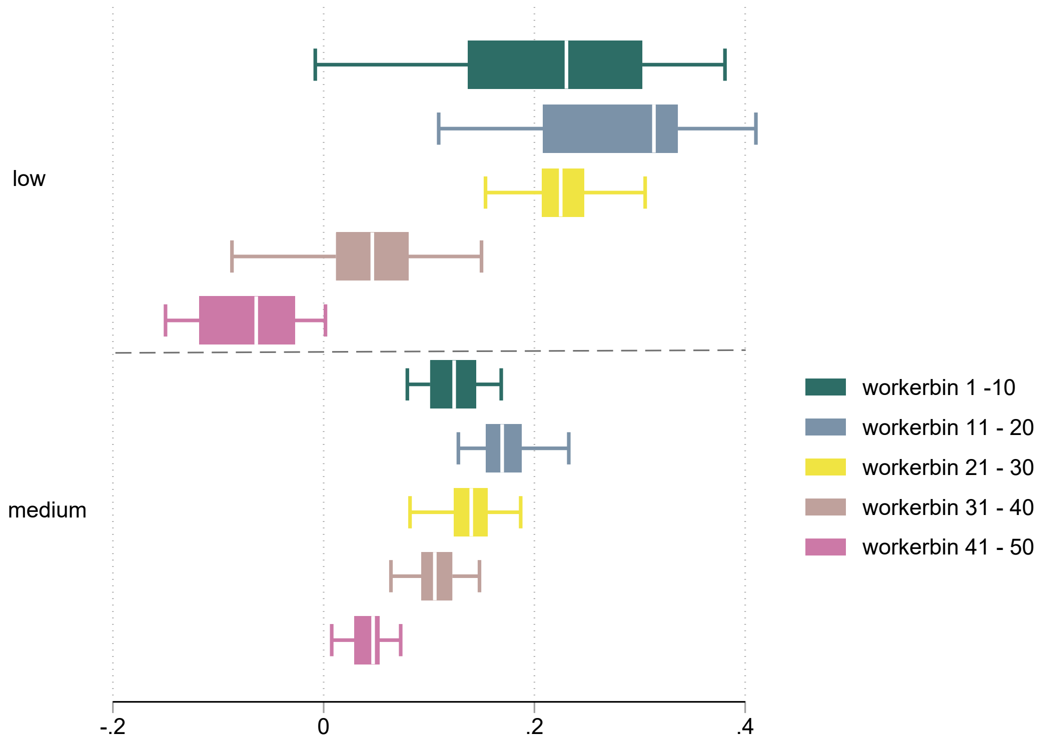
(b) New Matches



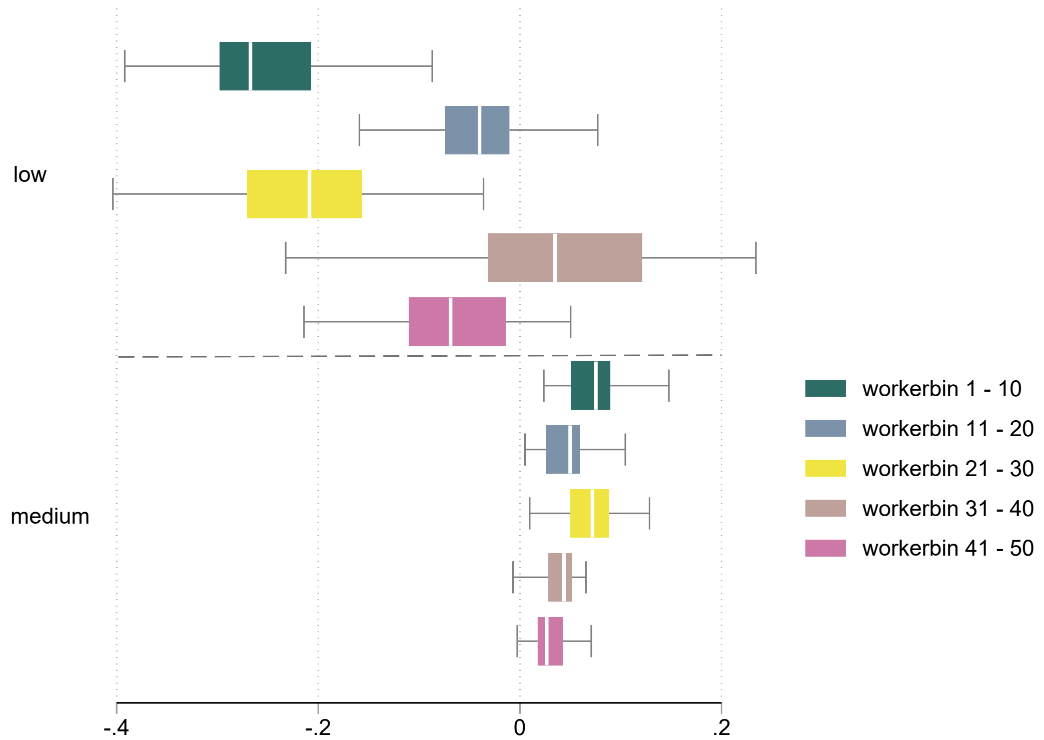
Notes: The plots show the estimated wage profiles across firm bins constructed using the AKM firm effects for all matches and for new matches. Plots are based on kernel-weighted local polynomial regressions. Kernel: Gaussian. The bandwidth is 2. 95% confidence bands in gray. Data sources: BHP, EP, BeH.

Figure D.14: Wage Changes for Observed Transitions, Firm Subsamples

(a) To high

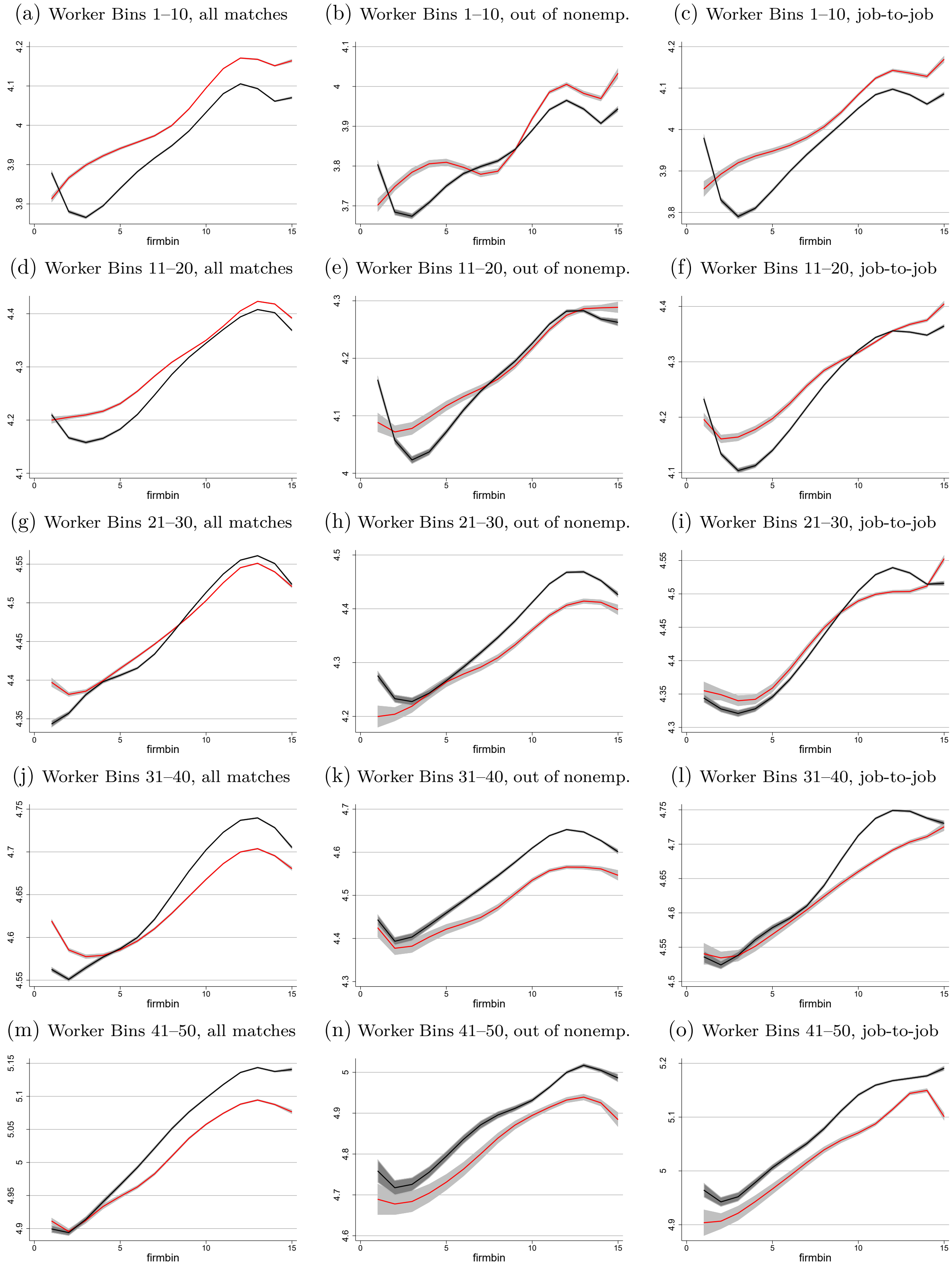


(b) From high



Notes: The plots show the estimated coefficients and 95% confidence intervals (based on robust standard errors) from a linear regression of the individual-level wage differences of transitioning workers on dummies for the origin and destination firm bins using 20 random subsamples (with replacement, clustered by firm-year) from the original sample (size=N). The subsample size M varies randomly between $0.5 * N < M \leq N$. The subsamples consist of new matches (job-to-job switches, no intermediate nonemployment 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 axes capture the destination/origin firm bin groups: low (bins 1–3) and moderate (bins 4–12). Data sources: BHP, EP, BeH.

Figure D.15: Changes in Mean Wages across Worker and Firm Types: 1998–2002 (red) vs. 2003–2008 (black)

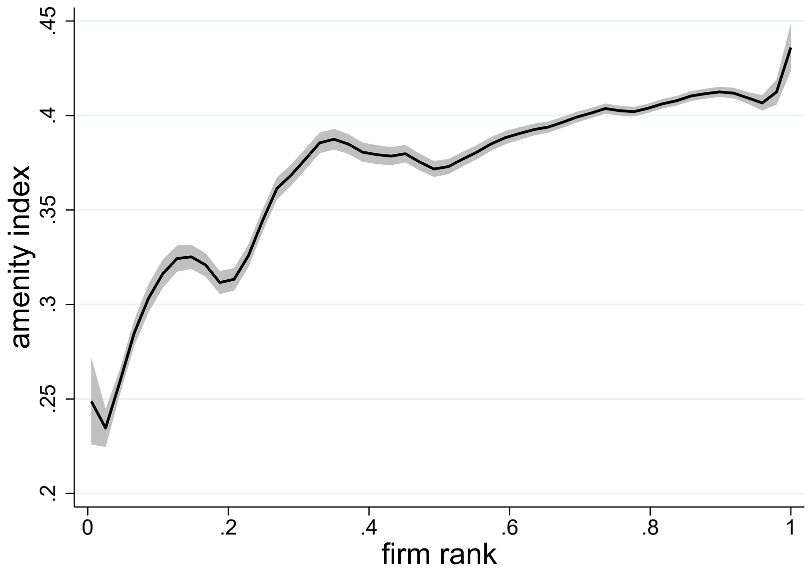


(a) Worker Bins 1–10, all matches
(b) Worker Bins 1–10, out of nonemp.
(c) Worker Bins 1–10, job-to-job
(d) Worker Bins 11–20, all matches
(e) Worker Bins 11–20, out of nonemp.
(f) Worker Bins 11–20, job-to-job
(g) Worker Bins 21–30, all matches
(h) Worker Bins 21–30, out of nonemp.
(i) Worker Bins 21–30, job-to-job
(j) Worker Bins 31–40, all matches
(k) Worker Bins 31–40, out of nonemp.
(l) Worker Bins 31–40, job-to-job
(m) Worker Bins 41–50, all matches
(n) Worker Bins 41–50, out of nonemp.
(o) Worker Bins 41–50, job-to-job

Notes: The plots show the estimated wage profiles across grouped firm bins during two time periods for all matches, new matches out of nonemployment, and job-to-job moves. Plots are based on kernel-weighted local polynomial regressions. Kernel: Gaussian. The bandwidth is 2. 95% confidence bands in gray. Data sources: BHP, EP, BeH.

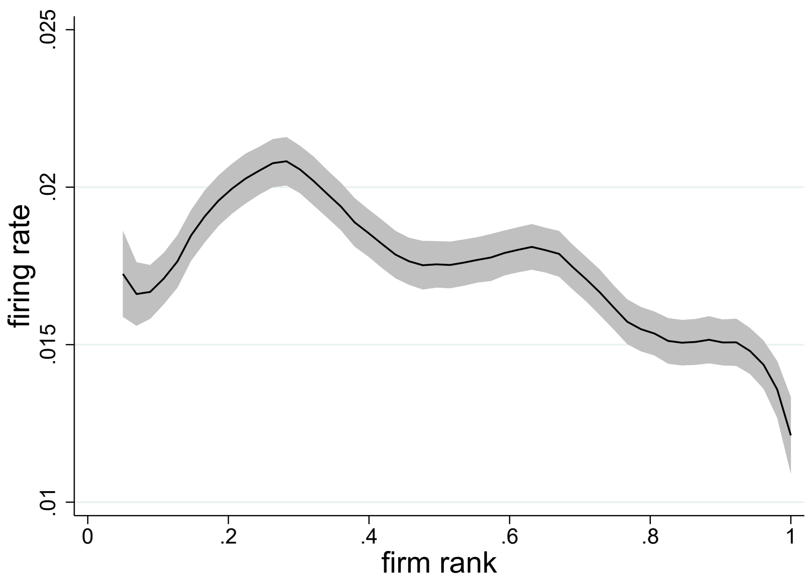
D.7 Extensions

Figure D.16: Amenities



Notes: Estimated univariate kernel density of the amenity index across estimated firm ranks normalized to be between zero and one. Kernel: Epanechnikov. The bandwidth is 0.05. 95% confidence bands in gray. Data sources: BHP, EP, BeH.

Figure D.17: Job security



Notes: Estimated univariate kernel densities for the semiannual rate of employer-initiated separations across estimated firm ranks, normalized to be between zero and one. Kernel: Epanechnikov. The bandwidth is 0.1. 95% confidence bands in gray. Data sources: BHP, EP, BeH.

Online Appendix References

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.

Bagger, Jesper and Rasmus Lentz (2019) "An Empirical Model of Wage Dispersion with Sorting," *Review of Economic Studies*, Vol. 86, pp. 153–190.

Bond, Stephen and Måns Söderbom (2005) "Adjustment costs and the identification of Cobb Douglas production functions," IFS Working Papers W05/04, Institute for Fiscal Studies.

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, 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.

Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg (2009) "Revisiting the German Wage Structure," *The Quarterly Journal of Economics*, Vol. 124, pp. 843–881.

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*, Vol. 126, pp. 405–436.

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.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten (2020) "Leave-Out Estimation of Variance Components," *Econometrica*, Vol. 88, pp. 1859–1898.

Lachowska, Marta, Alexandre Mas, Raffaele Saggio, and Stephen A. Woodbury (2022) "Do firm effects drift? Evidence from Washington administrative data," *Journal of Econometrics*, forthcoming.

Levinsohn, James and Amil Petrin (2003) "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, Vol. 70, pp. 317–341.

Lochner, Benjamin, Stefan Seth, and Stefanie Wolter (2020) "Decomposing the large firm wage premium in Germany," *Economics Letters*, Vol. 194.

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.

Smith, Eric (1999) "Search, Concave Production, and Optimal Firm Size," *Review of Economic Dynamics*, Vol. 2, pp. 456–471.

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.

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