Skill demand, firm heterogeneity, and wage inequality*

(Preliminary draft - please do not cite/circulate)

Andre Mouton†

August 5, 2021

Abstract

Rising wage inequality is commonly attributed to increasing demand for skilled labor, but empirical wage decompositions find an important role for *wage sorting*: high-earning workers are increasingly likely to work for high-paying firms. In this paper I study how firm heterogeneity interacts with skill-biased technical and structural change. Using matched employer-employee data from West Germany, I show that wage sorting is entirely explained by the occupational and industrial structure of labor markets, and has become more important over time due to changes in industrial composition and rising occupational skill premia. I then develop and estimate a general equilibrium wage accounting model that maintains empirical tractability, while allowing for a rich set of equilibrium relationships between prices and quantities. I estimate that in West Germany, firm heterogeneity amplified by 50% the distributional effects of rising occupational skill premia, and that the entire positive effect of structural change on wage inequality is due to the presence of large industry-employer wage differentials.

1 Introduction

Rising wage inequality in the OECD is widely attributed to growing demand for skilled labor. Advances in information technology and automation are thought to have resulted in a larger productive role for knowledge-intensive work, while substituting for manual and routine tasks. Labor markets have also experienced significant structural changes: employment in manufacturing has generally declined, resulting in fewer of the unskilled and semi-skilled production jobs that once provided broad entry into the middle class. Such technological and structural 'skill-bias' has been widely studied in the macroeconomic literature, and almost always in the context of economic models that assume a direct relationship between skill and wage - models in which the latter is, effectively, the price of the former. These models predict a strong and unambiguous

*I would like to thank Ali Shourideh, Laurence Ales, Rebecca Lessem, and Brian Kovak for their support and advice. Additional thanks to seminar and workshop participants at Carnegie Mellon, and to the Institute for Employment Research (IAB) for their generous provision of data.

†PhD candidate, Tepper School of Business, Carnegie Mellon University (amouton@andrew.cmu.edu)

relationship between skill demand and the wage distribution, and they provide an intuitive and powerful explanation for the observed rise in wage inequality.

Empirical studies nevertheless find that skill is only part of the story, and that a second and important source of wage variation comes from differences between employers. Different firms are observed to pay different wages, even after controlling for differences in labor composition. A typical estimate is that 15% of observed wage variance is due to employer wage differentials (Card, Cardoso, Heining, Kline, 2018). This finding has two potentially important implications for economic models of skill-bias. First, if wages depend not just on one's skill but on one's employer, then there will be no simple and direct relationship between skill demand and wage inequality. The distributional effects of rising skill-bias will depend on the statistical relationship between worker skill and employer pay; there will be a larger effect on wage inequality, for example, when high-paying firms tend also to be skill-intensive firms. Second, the relationship between skill and employer pay is likely to be endogenous to technological and structural change, if these differentially impact high- and low-paying employers. For these reasons firm heterogeneity may play an intermediary role, between skill demand on the one hand and wage inequality on the other.

Evidence that firm heterogeneity is important for explaining the rise in wage inequality is provided by a recent literature on *wage sorting*: in at least some countries, high-earning workers are more likely to work for high-paying employers. Card, Heining, and Kline (2013) estimate that wage sorting is responsible for one-third of the post-1980's rise in West German wage variance, and Song et al. (2019) show quantitatively similar results for the United States. These studies are silent on the origins of wage sorting and its relationship (if any) to structural and technological change, and on the extent to which sorting is associated with skill versus other aspects of person heterogeneity that may influence earnings. The presence of wage sorting suggests, however, that there is a non-trivial relationship between skill and employer pay, and that firm heterogeneity may not only be relevant but also quantitatively important for understanding the effects of changing skill demand.

In this paper I study how firm heterogeneity influences the aggregate relationship between skill demand and wage inequality. I begin with an empirical analysis of matched employer-employee data from West Germany, covering the period 1993-2017. Using decomposition methods, I examine the link between West German wage sorting and observable correlates of structural and technological skill demand. I then develop and estimate a general equilibrium wage accounting model, in order to quantitatively measure (1) the extent to which firm heterogeneity has amplified the effect of rising skill premia on wage inequality, and (2) how firm heterogeneity has influenced the relative contributions of technological and structural change.

The empirical contribution of this paper is to show that West German wage sorting is entirely a feature of the industrial-occupational structure of labor markets. Within broad industry-occupation groups, sorting is absent. I find that wage sorting has contributed to the rise in wage inequality as a result of two broad trends. First, wages have risen faster in skilled occupations, and because these occupations are more common in high-paying employers, the effect on overall wage variance has been greater. Second, employment has shifted away from low-skill, high-paying sectors like materials and crafts manufacturing, and towards low-skill, low-paying sectors like personal services, hospitality, and temp agencies, strengthening the aggregate relationship between skill and employer pay. While these results are descriptive, they suggest that interactions between rising skill demand and employer heterogeneity may have played a fundamental role in West Germany over this time period.

The second contribution of this project is the development and estimation of an equilibrium accounting framework, that allows for endogenous skill premia and employer wage differentials and rent-seeking behavior by workers. Under non-onerous restrictions the model generates a log-additive wage function, allowing it to be non-parametrically estimated from the joint distribution of empirical wage effects obtained from standard wage regressions. Through counterfactual experiments I estimate that industry-firm wage differentials are responsible for one-half of the increase in wage inequality due to changing occupational and industrial demand. These differentials are particularly important for explaining the role of industry, which accounts for 15% of the aggregate trend in wage inequality but would have had a small, negative contribution if employer pay were homogeneous across sectors. I show that the pass-through from labor demand to wage inequality varies substantially across similarly-skilled industries and occupations, and I conduct a comparison with East Germany, where slower wage inequality growth is largely explained by divergent trends in industry composition and industry wage differentials. Taken together, these results indicate that the aggregate relationship between skill bias and wage inequality is sensitive to firm heterogeneity, and for that reason is likely to vary substantially between regions and over time.

The outline of this paper is as follows. In the latter part of this section I give a more detailed discussion of this paper in relation to the literature. The empirical analysis is presented in section 2, and in section 3 I develop the equilibrium model and briefly discuss comparative statics. Estimation and quantitative results are then described in section 4.

1.1 Discussion and related literature

In the past 40 years wage inequality has risen throughout the OECD. Since 1980 the income share of the top 10% of earners has increased in 31 of 38 member nations, rising on average from 32.2% to 37.5% as of 2019[1]. Causal explanations of this increase have focused on changes in the demand for skilled labor, due to declining prices of information technology capital and changes in the trade environment associated with globalization. There exists substantial empirical evidence to support these theories[2], but as the relationship between skill demand and the wage distribution is an equilibrium one, theoretical models of skill-bias have played a critical role in providing intuition and generating testable predictions. Canonical treatments of skill-biased technical change include Krusell et al. (2000), Acemoglu and Autor (2011), and Acemoglu and Restreppo (2018), while prominent models of skill-biased trade include Verhoogen (2008), Costinot and Vogel

[1]Share of pre-tax national income, top 10% of adults, World Inequality Database (https://wid.world/).

[2]See Frey and Osborne (2017) for a comprehensive literature review.

(2010), and Burstein and Vogel (2017). These papers share a common assumption of perfectly competitive labor markets: wage dispersion is solely the result of differences in skill between workers. Although some authors like Autor and Dorn (2013) allow for firm heterogeneity in factor demand, only a handful of papers in this literature (discussed below) have considered any role for employer wage-setting.

The empirical labor literature has nevertheless produced a large body of evidence, accumulated since the late 1990's, on the contribution of firm wage differentials to overall wage inequality. Abowd, Kramarz, and Margolis (AKM, 1999) proposed the use of matched employer-employee data for identifying employers' effects on wages, while controlling for observed and unobserved worker heterogeneity. These authors estimated that slightly more than one-fifth of French wage dispersion was due to employer wage effects, and similar results have since been shown for many other countries. Early studies generally found little or no evidence that high-earning workers sort into high-paying firms, but that has changed as larger, population-level datasets have become available. Card, Heining, and Kline (CHK, 2013) and Song et al. (2019) employ administrative social security data from Germany and the United States, and in both cases find that wage effect correlations have risen over time, reaching values of $.25 - .28$ by the early 2000's. Quantitatively similar results have been shown for Denmark by Bagger, Sorensen, and Vejlin (2013), while Abowd et al. (2012) find significant between-industry wage sorting in France.

Firm wage-setting is important because it breaks the equivalence between skill and wage assumed in models of SBTC. How a rise in skill demand impacts the wage distribution will depend on the relationship between skill and employer pay. When skilled workers are more likely to work for high-paying employers, a rise in skill premia will be more concentrated in the upper wage percentiles, resulting in greater spread of the wage distribution. Figure 2.1 shows this outcome in a numerical example where skill and employer pay are log-normal random variables. When they are positively correlated, skill will tend to increase more quickly in wage, and therefore an increase in the variance of the person effect passes through more strongly to the variance of the wage distribution. It is also important to note that the skill-pay correlation will

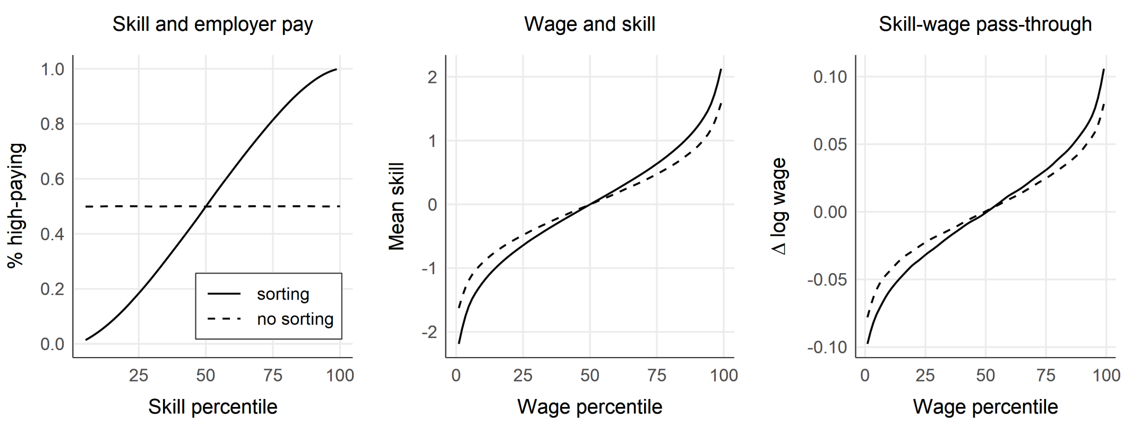

Figure 2.1: Numerical example, firm wage-setting and skill premia pass-through

FIGURE NOTES. Skill and pay are drawn from a joint normal distribution. Correlations in the 'sorting' case are fixed at .5. Panel 3 shows the effects of a 10% increase in the variance of skill premia.

be an endogenous object if rising skill demand is concentrated in high-paying or low-paying firms, if it affects the wages that firms set, or if the composition of agents (firms or workers) is impacted by technical change. Observed structural changes to labor markets make such endogeneity likely. Hence not only will firm heterogeneity intermediate the aggregate relationship between skill demand and wage inequality, but its influence may depend in complicated fashion on the precise way in which demand changes over time

The empirical portion of this paper is a study of the relationship between West German wage sorting and observable dimensions of skill demand. My approach borrows much from CHK, who were the first to document in a comprehensive manner the quantitative importance of wage sorting. The results shown here are consistent with recent work by Haltiwanger and Spletzer (2020) using data from the United States. Those authors find that industrial and occupational classifications explain the majority of rising wage variance over the years 1996-2015, largely the result of changes to occupational wage premia and occupation-industry sorting. The principal difference between the two analyses is that I take an AKM decomposition approach, allowing for a cleaner separation of person and employer wage effects. More broadly, the empirical portion of this paper is related to a literature studying the joint roles of industry and occupation in explaining skill-bias, including Autor and Dorn (2013), Goos, Manning, and Salomons (2014), and Autor, Dorn, and Hanson (2015).

To obtain quantitative results, I develop an equilibrium model that allows for endogenous wage differentials and behavioral responses to prices, while largely maintaining the tractability of the AKM empirical framework. In modeling skill demand I combine elements from Costinot and Vogel (2010) and Autor and Dorn (2013), while employer wage-setting is assumed to arise from a directed search mechanism *a la* Moen (1997). Closely related to this paper is past work by Card, Cardoso, Heining, and Kline (2018), who study a simple model in which monopsony arises due to idiosyncratic preferences for employers, and which has since been extended to an equilibrium setting by Haanwinckel (2021). The main drawback of their approach is that it is empirically tractable only when firms are homogeneous in their demand for skill, and is for that reason poorly suited to the question asked in this paper. Also related is the literature on *assortative matching* between workers and firms due to match production complementarities[3]. These papers feature a wage function that depends on the types of both the worker and the employer, but they abstract from labor differentiation within the firm and product differentiation between firms, and for this reason assortative matching is best thought of as sorting within narrowly-defined labor and product markets. My findings indicate that within broadly-defined industries and occupations, worker and employer wage effects are weakly negatively correlated and this correlation changes little over time, suggesting that assortative matching is not systematically present and is unlikely to explain West German wage sorting.

criticisms of sbtc -¿ wage inequality

[3] See e.g. Eeckhout and Kircher 2010, 2011; Lise, Meghir, and Robin 2016; Hagedorn, Law, and Manovskii 2017; Bagger and Lentz 2018; Bonhomme, Lamadon, and Manresa 2019.

2 Empirical analysis

The principal dataset used in this analysis is the German linked employee-employer dataset (LIAB), provided courtesy of the Institute for Employment Research (IAB). Every year since 1993 the IAB has conducted a stratified survey of German establishments, collecting data on operational, investment, and hiring activities. A matched employee-employer dataset is created by linking surveyed establishments with administrative records for all employees subject to social security taxes. Person-level records include the occupation and daily wage associated with the employment spell, and basic information on demographics and education. In total, the dataset represents a roughly 5% sample of the German workforce, and in any given year contains between 1.5 and 2.5 million observations covering between 4 and 15 thousand establishments.

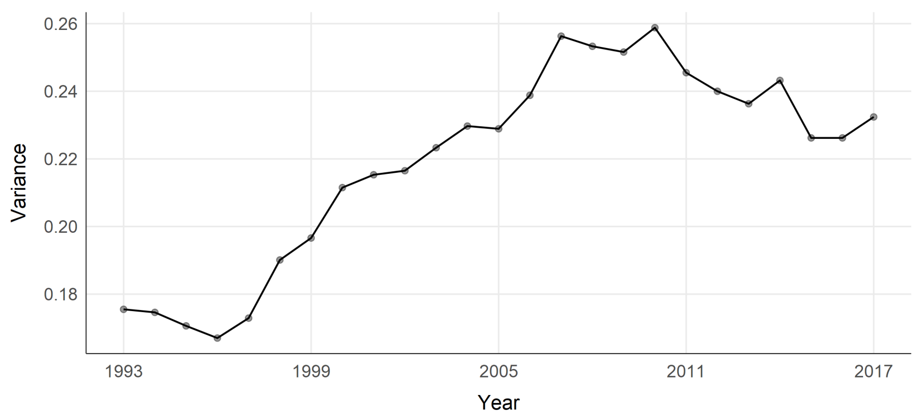

Figure 2.1: West German wage variance, 1993-2017

Figure notes. Variance of log daily wages, full-time West German workers aged 20-60. Earnings above the social security contribution ceiling are imputed using annual Tobit regressions.

The time series of West German wage variance is shown in figure 2.1, where wages are calculated as log daily earnings. Following Card et al. (2013) I impute wages for workers whose earnings exceed the social security contribution ceilings and are therefore top-coded. Between the mid-1990's and 2010, wage variance increased substantially, continuing a trend that is known to have begun in the mid-1980's. Note however that I do not observe workers in "mini-jobs" that are exempt from social security contributions, and that following convention in the wage decomposition literature I exclude part-time workers and apprentices from the sample, who in a typical year represent 3-5% of observation. These exclusions allow for a cleaner comparison of wages earned by different workers, but come with the trade-off that low-earning individuals are to some extent underrepresented in 2.1.

2.1 AKM wage variance decomposition

Following Abowd, Kramarz, and Margolis (1999), it has become common to decompose the variance of wages by first regressing earnings on dummy variables for the person and the employer, optionally including time-varying characteristics in the regression as well; and then by considering separately the variances and covariances of the estimated regression effects. For this project the person and employer effects are provided by IAB, who estimate an AKM regression on the underlying administrative datasets that make up the LIAB. As sample sizes are much larger in the administrative data, this results in better identification of the wage effects than could be achieved by estimating them directly on the LIAB subsample.

Specifically the data provider follows the methodology proposed by Card et al. (2013) and implements the regression equation

$$w_{i,t} = \alpha_i + \psi_{J(i,t)} + x'_{i,t}\beta + \epsilon_{i,t} \tag{1}$$

where $w_{i,t}$ is the log daily wage of person i in year t; α_i is a dummy variable for person i; $\psi_{J(i,t)}$ is a dummy variable for i's current employer j; $x_{i,t}$ is a vector containing year fixed-effects and a cubic polynomial in worker age, interacted with dummies for educational attainment[4]; and ϵ is an I.I.D. error term. Estimation is performed on four partially-overlapping panels that jointly cover the period 1993-2017, with each panel spanning 6-7 years.

Once the regression effects are estimated, wage variance can be decomposed as

$$\begin{aligned} Var(w_{i,t}) &= Var(\alpha_i) + Var(\psi_{J(i,t)}) + Var(x'_{i,t}\beta) + Var(\epsilon_{i,t}) \\ &\quad + 2Cov(\alpha_i, \psi_{J(i,t)}) + 2Cov(\alpha_i, x'_{i,t}\beta) + 2Cov(\psi_{J(i,t)}, x'_{i,t}\beta) \end{aligned} \tag{2}$$

In this section I limit attention to the variances and covariance of the person and employer effects - $Var(\alpha)$, $Var(\psi)$, and $Cov(\alpha, \psi)$ - which together account for virtually all of the trend shown in figure 2.1[5]. The evolution of these moments over time is plotted in figure 4.10. Comparing the 1993-1999 and 2010-2017 periods, 51% and 14% of the total increase in wage variance is explained by the variances of the person and employer effects, respectively, and 28% by their covariance, with the remainder attributable to $Var(\epsilon)$.

These results are qualitatively similar to those shown by Card et al. (2013) for the 1985-2009 period. Quantitatively, the upwards trends in panels 2 and 3 are somewhat smaller when compared to the main results shown by CHK. One reason is that the samples are different: CHK estimate (2) separately for males and females, whereas in this case both genders are included in the sample, and $Cov(\alpha, \psi)$ increases much more quickly over time for males. A second reason is that $Var(\psi)$ falls steeply after 2010. This development is more difficult to explain but appears to be due, at least in part, to statistical bias. Job transition rates are lower in the years following the Hartz reforms (2003-2005), a development that appears to have negatively impacted estimation of the wage effects for the 2003-2010 panel. Because estimation error in (1) tends to bias upwards the variances of α and ψ, and to bias downwards their covariance[6], this likely explains the

[4] Educational categories consists of a lower secondary education, a completed apprenticeship, an upper secondary education, a university degree, and a 'missing' category that represents 12% of the sample in the early 1990's and 6% by the late 2010's.

[5] Additional results and a more detailed discussion of the estimation procedure are presented in the appendix.

[6] See e.g. Andrews et al. (2008) and Bonhomme et al. (2020) for a discussion of so-called 'limited-mobility bias', and the

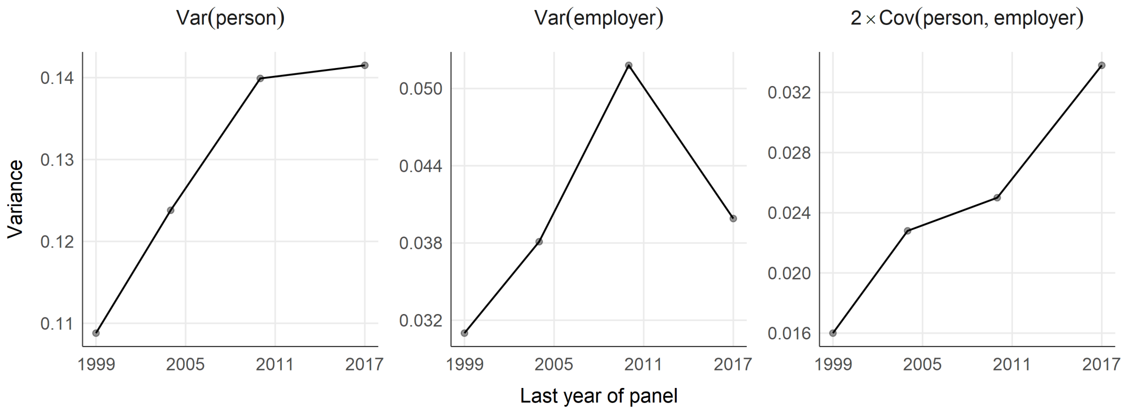

Figure 2.2: AKM wage variance decomposition, 1993-2017

Figure notes. Person and employer wage effects estimated separately in four panels covering 1993-1999, 1998-2004, 2003-2010, and 2010-2017. Weights equalized across years in calculations.

deviations from trend observed in all three panels for the 2003-2010 period. Time-varying bias is a potential concern for this project, and therefore I revisit this discussion in the next section.

2.2 Skill demand and wage sorting

Of key interest is the third panel of figure 4.10, which shows the covariance over time of the estimated person and employer wage effects. Two key questions are (1) whether this covariance is related to worker *skill* versus other aspects of person heterogeneity, and (2) whether this covariance is related to *structural* and *technological* change. These terms require definition, and I define them as follows. I take skill to be variation in the person wage effect α that is explained by the occupation and the industry in which the person is employed. Industry and occupation are obvious choices for labor differentiation both because of their high level of detail, and because they are the variables most studied in the context of technological and structural change. If task content varies across industries and occupations but not within them, then a simple model of occupational choice would predict that between-group variation in α reflects comparative advantage, and within-group variation the absolute advantage of workers; and as shifts in labor demand will only affect relative prices of imperfectly substitutable types of labor, skill-bias should be understood as concerning comparative and not absolute advantage. Given this definition of skill, structural and technological change may then be naturally interpreted as changes over time in the distributions of, and the wages associated with, different industries and occupations.

To answer these questions I employ between-group decompositions of $Cov(\alpha, \psi)$, using the law of total covariance. Given a partition of the sample into G groups of size ω_g where $g \in G$, we may write the covariance

appendix for identification statistics and additional discussion.

as

$$Cov(\alpha, \psi) = \underbrace{\sum_{g \in G} \omega_g \Big(\alpha - \mathbb{E}_g[\alpha]\Big)\Big(\psi - \mathbb{E}_g[\psi]\Big)}_{\text{within-group wage sorting}} + \underbrace{\sum_{g \in G} \omega_g \Big(\mathbb{E}_g[\alpha] - \mathbb{E}[\alpha]\Big)\Big(\mathbb{E}_g[\psi] - \mathbb{E}[\psi]\Big)}_{\text{between-group wage sorting}} \tag{3}$$

The between-group component is shown in figure 2.3 for 15 occupational groups, 12 industry groups, and 180 occupation-industry cells[7]. In all periods, $Cov(\alpha, \psi)$ is mostly a between-occupation and between-industry phenomenon. Within industry-occupation cells, wage sorting is entirely absent; the within-group component of the covariance is negative and largely comparable over time. These results are especially striking given that industry and occupation explain relatively little of $Var(\alpha)$ and $Var(\psi)$ (plots shown in the appendix), and the interaction of industry and occupation accounts for less than 40% of either variance.

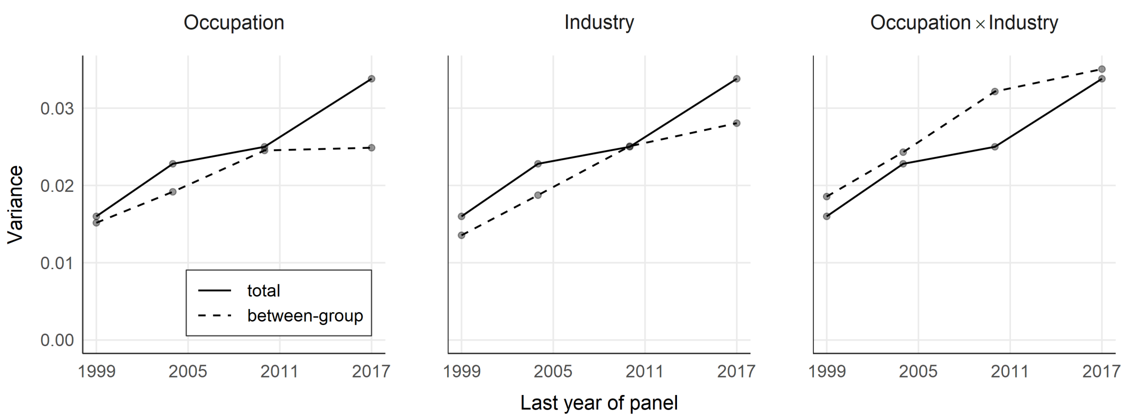

Figure 2.3: Between-group wage sorting, 1993-2017

FIGURE NOTES. Between-occupation component of $Cov(\alpha_i, \psi_{J(i,t)})$ for 15 KLDB 1988 occupational groups, 12 WZ2008 industry groups, and 180 occupation × industry groups.

To more clearly show the sources of between-industry and -occupation wage sorting, in figure 2.4 I plot the group mean person and employer wage effects. Occupational person differentials accord well with the notion of skill: values are particularly high in knowledge-intensive STEM and managerial occupations, and low in jobs associated with unskilled manual labor, and in the appendix I provide results in terms of occupational task content that verify this intuition. Skilled occupations are also strongly and visibly associated with high-paying employers, however, and this is particularly true with technical and engineering jobs. The overall relationship between skill and employer pay is weakened somewhat by unskilled and semi-skilled production jobs, which are associated with much higher employer effects relative to similarly skilled service jobs - a difference that may be immediately rationalized by turning to the second panel of figure 2.4.

Industries exhibit large differentials in terms of person and employer wage effects, and in general manufacturing sectors pay approximately 10% more than service industries with similarly skilled workforces in terms of α. The service sector also exhibits much greater differentiation in terms of employer pay, with

[7] See appendix for details on the industry and occupation groupings.

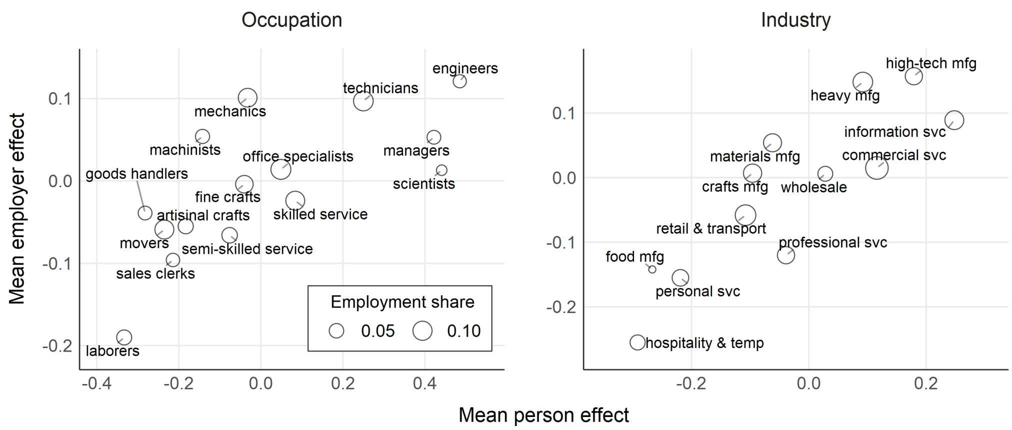

Figure 2.4: Mean AKM wage effects by occupation and industry, 2010-2017

Figure notes. Group mean α and ψ for the 2010-2017 period, calculated for 15 KLDB 1988 occupational groups and 12 WZ2008 industry groups. Point sizes indicate employment shares, and mean effects are labor-weighted intra-industry averages.

personal, professional, and businesses services paying 20% less than commercial and information services. It is not immediately clear whether industry differences in workforce reflect skill or other aspects of person heterogeneity, but in the appendix I show that there is large inter-industry variation in terms of average task content, and that high-α industries are more IT-intensive and engage more frequently in product and process innovation.

2.3 Structural change, technological change, and wage sorting

Results from the previous section show that West German wage sorting is a feature of inter-industry and inter-occupational variation in the estimated AKM wage effects, and that descriptively this variation is associated with knowledge-intensive and cognitively-demanding work, consistent with a skill-based interpretation. But wage sorting has also become quantitatively more important over time, and as discussed in the introduction there are (at least) two potential explanations for the rise in $Cov(\alpha, \psi)$: pre-existing labor sorting patterns may have amplified the pass-through from rising skill premia to the wage distribution, and sorting patterns may have changed as a direct result of technological and structural change. It could be for example that high-paying industries have become relatively more skill-biased, if they are more exposed to skill-biased innovations in technology; or it could be that stable differences in industry skill demand have interact with changes in industry composition, resulting an a stronger aggregate relationship between skill and empoyer pay.

To this end I perform a series of rough counterfactual experiments intended to shed light on the sources of increased wage sorting. I focus on that portion of $Cov(\alpha, \psi)$ explained by 180 interacted industry-occupation groups that, per figure 2.3, successfully capture all of the wage sorting observed in West Germany in each period. In each experiment I hold fixed, individually and in combination, the group mean wage effects and industry and occupation labor shares. Results are shown in figure 2.5. Over the full 1993-2017 time

	93-99 to 98-04	98-04 to 03-10	03-10 to 10-17	93-99 to 10-17
Observed change	**.0030**	**.0038**	**.0014**	**.0082**
Constant means				
Person effect	.0021	.0034	-.0007	.0040
Employer effect	.0019	.0025	.0045	.0079
Both effects	.0011	.0020	.0020	.0039
Constant shares				
Industry shares	.0017	.0020	-.0002	.0039
Occupation shares	.0028	.0033	.0012	.0074
Industry-occupation shares	.0016	.0016	-.0003	.0035
Constant means & shares				
Person effect, industry shares	.0009	.0017	-.0022	.0007
Person effect, ind.-occ. shares	.0008	.0013	-.0023	.0004
Both effects, industry shares	.0000	.0004	.0000	.0001

Table 2.5: Time trend in $Cov(\alpha, \psi)$ under counterfactual scenarios

Table notes. Counterfactual change in $Cov(\alpha, \psi)$, with means and/or weights held fixed at the values of the earlier panel. Shares are in terms of total employment, and means are labor-weighted intra-industry averages.

period, holding constant the group mean person effects reduces the overall trend in half, and holding constant industry employment shares has a similar effect; holding both constant reduces the trend by 90%. Changes in the group mean person effects may reflect both nominal changes - i.e. rising skill premia - or changes to group labor composition, and in general it would be difficult to distinguish between these two explanations. On the other hand a key role for industry composition would suggest that (1) greater occupational sorting across industries does not appear to be a major factor, and (2) the rise in $Cov(\alpha, \psi)$ is at least partly the result of changes to labor market structure.

In figure 2.6 I break down changes to wage premia and employment shares by occupation and industry. The first panel shows that wage growth has been concentrated in skilled occupations, which are in turn associated with high-paying employers. Relative wage declines have been most pronounced in jobs relating to manual labor; wage gains are particularly large in knowledge-intensive occupations and in office administrative jobs, although the latter may be due to changes in composition changes as office jobs are poorly differentiated in the 1988 KLDB coding system. A comparison of figures 2.4 and 2.6 shows that manual labor jobs have also experienced large declines in associated employer wage effects; as a result of changing industry composition, these jobs have become more concentrated in low-paying sectors like temp agencies and food manufacturing. Sectoral shifts have had the opposite effect on engineering jobs, which have are now relatively more likely to take place in high-paying, high-tech manufacturing industries. Changes in occupation mean person and employer effects are strongly correlated overall, suggesting strong similarities between the occupation-level effects of structural and technological change. The second panel of figure 2.6 shows that industry composition has exhibited a polarization pattern over the 1993-2017 timeframe. Middle-skill industries have lost ground to low- and high-skill industries, and because skill and employer pay are so highly correlated across sectors, the distribution of employer effects has tended to polarize as well. The largest

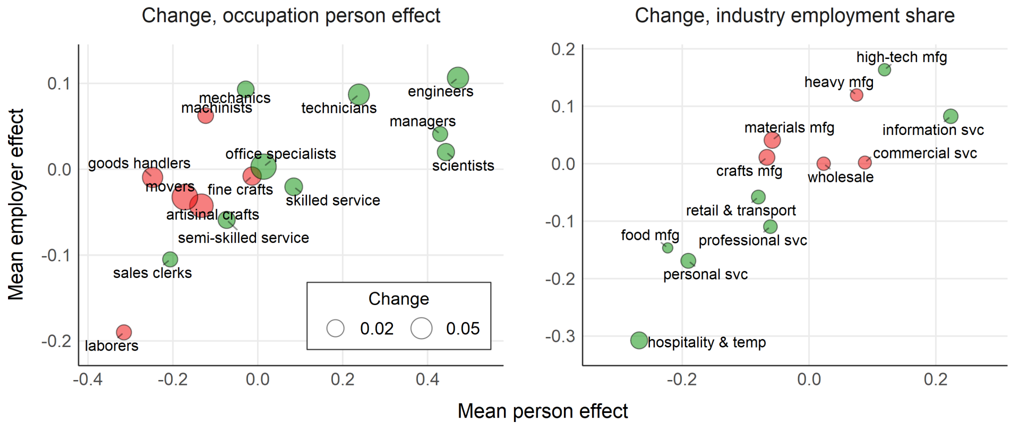

Figure 2.6: Occupational wage and industry share growth, 1993-2017

Figure notes. Change in group mean α and employment shares between 1993-1999 and 2010-2017, calculated for 15 KLDB 1988 occupational groups and 12 WZ2008 industry groups. Axes represent mean log α and ψ, averaged across the 1993-1999 and 2010-2017 periods. Green (red) points indicate growth (decline).

declines in employment shares were experienced by the materials and crafts industry groups, consistent with explanations relating to trade and/or automation. On the other hand German temp agencies have grown substantially over this period, a well-known phenomenon, and one that is particularly significant in light of the low associated employer wage effects. Increased use of temp labor appears to be almost entirely driven by manufacturing; each of the five goods-producing industry groups has seen a rise in temp share of employment from approximately 1% in the 1990's to 5% by the end of the sample. Goldschmidt and Schmieder (2017) show a similar trend in the outsourcing of business services like food catering, which suggests a broader pattern of labor outsourcing from high-paying to low-paying industries.

2.4 Discussion

To summarize the two main empirical findings presented in this section:

1. West German wage sorting is entirely a between-industry, between-occupation phenomenon.
2. Wage variance from sorting has increased over time due to rising occupational skill premia and changing industry composition.

These results are surprising for several reasons. First, although like many past studies I find wage dispersion to be mostly residual - unexplained by observed covariates - wage sorting appears unrelated to this residual component, and completely accounted for by aggregated industry and occupation codes. Second, the growing role of wage sorting over time is not explained by changes in the sorting of occupational labor across industries, but is instead the result of changes to prices and industry composition, with industry skill gradients remaining more or less stable.

These results are consistent with the mechanism studied in this paper, but more importantly they are inconsistent with a number of alternative explanations that might be proposed. The West German experience is not well-described by theoretical models of assortative matching, as these describe sorting between undifferentiated agents, whereas occupation and industry are usually interpreted as proxies for differentiation in labor and product markets. A more general challenge for models of sorting based on match complementarities - whether between workers and firms, or between workers and their peers - is to simultaneously explain rising skill premia and stable industry-occupation sorting patterns; the first suggests that complementarities have become stronger over time, while the second implies that they have not changed.

A key limitation of the empirical analysis in this section is its silence on causality. The results do not prove that rising wage sorting is the result of skill-biased technological and structural change. In addition it is unclear why industry-occupation sorting is present, and whether sorting patterns reflect general economic relationships or merely idiosyncratic and potentially time-varying aspects of the German labor market. Supplemental results provided in the appendix suggest that establishment scale and technological sophistication are strongly correlated across industries, which may provide a natural explanation for the covariance of skill demand with employer pay; but then it is uncertain why a similar relationship does not hold within broad industry groups. In conclusion, the empirical results presented in this section raise more questions than they answer, but I argue that they provide justification for the quantitative analysis that follows, and represent a step forward given our lack of knowledge regarding the phenomena under study.

3 Model

In this section I develop an equilibrium model, with the eventual goal of conducting counterfactuals that will allow me to estimate the contribution of firm heterogeneity to the aggregate relationship between changing skill demand and wage inequality. In particular I wish to know the extent to which firm heterogeneity has amplified the distributional effects of changing demand, and how firm heterogeneity has affected the relative importance of technological and structural changes in demand. A theoretical model is important in this case as it will allow me to disentangle the effects of different changes in the environment, and to take into account the behavioral responses of agents to changes in prices. As the purpose of the model is to allow for quantitative results, empirical tractablility and flexibility are paramount; any failure to accurately predict the empirical joint distributions of wages and employment will directly affect the quality of the results. These goals are challenging in light of the fact that sorting and employer wage-setting are often sources of intractability in theoretical models, and they do not generally yield wage functions that map cleanly into observed wages or estimated AKM wage effects.

I therefore take the following approach. I model skill demand in typical fashion, as arising from a production function defined over differentiated labor inputs. I assume that this function is industry-specific, however, and may vary arbitrarily across industries. Employer wage-setting arises due to search frictions, with tractability coming from the assumption that search is directed. Following the search literature employers are assumed to be atomistic, and they pay different wages because they face different entry costs. In addition to these basic features of the model, I include several elements that aid in structural estimation. First, I assume that production is defined over occupation-specific labor rather than skill, and I model the assignment of skill to occupations. Occupational assignment is important because it generates an equilibrium restriction on the relationship between wages and skill, which in turn allows the model primitives to be identified. I additionally allow for a long-term unemployment state, so as to more straightforwardly capture changes over time in the German unemployment rate. If lower-skill workers are more likely to be unemployed at a given point in time, then a decline in unemployment will act as an increase in the supply of unskilled labor.

I begin the model exposition by describing the environment, after which I formalize the agents' problems. I then describe equilibrium and the optimal assignment; and I close out the section with some simple numerical comparative statics in a 3-industry version of the model.

3.1 Environment

The economy is set in continuous time, and consists of infinitely-lived workers, finitely-lived firms, and a representative household. All agents are risk-neutral and discount the future at rate ρ. Workers exist in unit mass, and are heterogeneous in a variable $s \in [0,1]$ indicating the worker's *skill.* The skill distribution is denoted as $\nu(s)$, which is assumed to be continuously differentiable and strictly positive. I distinguish between two classes of firms: intermediate producers and industry aggregators. Intermediate producers are defined by the *industry* $i \in \{1, ..., I\}$ in which they operate, and by the type of occupation- or *job*-specific good $j \in [0,1]$ that they produce. These goods are then combined by industry aggregators using an i-specific production technology, who then sell an industry composite good to the representative household. In what follows I will refer to intermediate producers as 'firms', and industry aggregators simply as 'aggregators.'

Labor markets operate as follows. At each point in time, existing firms randomly exit the market at an exogenous rate δ. New firms enter by posting vacancies and wage offers in labor submarkets indexed by (i,j). Each entering firm is endowed with a single vacancy, and while entry is free, any firm posting a vacancy must pay an industry- and job-specific vacancy flow cost $C_i(j)$. Workers observe posted vacancies and wages and then choose the submarket to which they will apply. I assume that firms are unable to screen workers, and therefore that posted wages offers are *per unit of worker output.* Due to search frictions not all unemployed workers that apply in submarket (i,j) are matched, and not all vacancies are filled. Following standard assumptions in the search literature, if N unemployed workers search and V total vacancies are posted in a given submarket, then matches are formed at a rate $\zeta V^{\eta} N^{1-\eta}$. Defining market tightness as $\theta = \frac{V}{N}$, vacancies will then be filled at a rate $\zeta\theta^{\eta-1}$ and unemployed workers will find jobs at a rate $\zeta\theta^{\eta}$.

Once matched with a firm, worker flow output is given by the function $m(j,s)$, with output then being sold to industry aggregators at a per-unit price $p_i(j)$. I abstract from voluntary separations, and assume that matches persist until the firm exits. In addition, I impose the condition that $m(\cdot)$ is log-supermodular:

Assumption 1: *$m(j,s)$ is continuously differentiable function and $\frac{m(j',s')}{m(j,s')} > \frac{m(j',s)}{m(j,s)}$ for all $j' > j$*

and $s' > s$.

Log-supermodularity implies that workers with greater skill have a comparative advantage at jobs with a higher j. That $m(\cdot)$ depends only on s and j and not on i is an important assumption, as it implies that relative productivities (i.e. $m(j', s')/m(j, s)$) will be identical across industries conditional on the worker's type and on the type of output being produced.

Unemployed workers who apply to a vacancy in submarket (i, j) receive a flow value $B_{i,j,s}$ representing an unemployment insurance payout, where for the moment B is allowed to depend arbitrarily on the submarket and on the worker's type. Workers who receive the unemployment benefit are required to accept any match with a firm (i.e. there is a perfectly enforced work-search requirement). Workers may also choose not to search for an employer, in which case they receive a fixed social security payment B_{np} and an idiosyncratic home production amenity χ with distribution $F(\chi)$[8]. The worker's flow payoff in this case is $B_{np} + \chi$. Employed workers receive the posted wage $w(i, j)$ multiplied by their productivity $m(j, s)$, together with an industry-specific amenity A_i, which jointly yield the flow utility $(wm)^{\psi} A^{1-\psi}$.

Industry aggregators combine intermediate goods using a CES technology with elasticity σ and intermediate good shares denoted by $\alpha_{i,\chi}(j)$. To keep the model as simple as possible, I abstract from capital and I assume that all output markets are perfectly competitive. Industry composite goods are sold to the representative household at a price P_i, where household utility is also assumed to be CES, with elasticity of substitution τ and industry shares β_i.

For the remainder of this section I limit attention to the steady-state, in which prices and market tightness are constant. To this end I omit time notation, although all endogenous variables should be understood to depend implicitly on time t. I will furthermore limit attention to symmetric equilibria in which all firms choose the same posted wages conditional on i and j, and all unemployed workers choose the same mixed strategy over submarkets when searching for jobs. These assumptions will allow me to abstract from firm- and time-specific notation, and to present a cleaner and more intuitive development of the model.

3.2 Agents' problems

I now formalize agents' choices in this economy. I begin with firms' wage posting and entry decisions, conditional on expected worker behavior. I then examine the industry-/job-search choices of unemployed workers. Finally, I characterize the decisions of industry aggregators and the representative household.

Intermediate producers. The firm's wage posting problem, stated below in system (4), is to choose the wage that maximizes the expected value of the vacancy posting. As stated previously, the flow cost of a vacancy is $C_i(j)$ while the flow profits of a match will be $m(j, s)\big(p_i(j) - w(i, j)\big)$. New matches are formed at rate $\zeta\theta(i, j)^{\eta-1}$, while existing matches dissolve at a rate of δ at which point future profits are equal to zero. All firms are assumed to take as given prices and the decisions of other firms and of workers. Hence

[8] A long-term unemployment state will provide a more tractable way of modeling secular changes in unemployment rates.

the problem of an intermediate good producer can be written as

$$\begin{aligned} w^*(i,j) = \arg\ &\max_w V(i,j|w) \\ &s.t.\ \rho V(i,j|w) = -C_i(j) + \zeta\theta(i,j)^{\eta-1}\big[\mathbb{E}_s J(i,j,s|w) - V(i,j|w)\big] \\ &\rho J(i,j,s|w) = m(j,s)\big(p_i(j) - w\big) - \delta J(i,j,s|w) \end{aligned} \tag{4}$$

where market tightness $\theta(i,s|w)$ will depend upon firms' posted wages. The problem of entering firms, on the other hand, is to compare the costs of posting a vacancy with expected future profits:

$$\text{Enter if } 0 < \rho V^*(i,j) \tag{5}$$

The direct implication of (5) is that at the steady-state equilibrium we must have $V^*(i,j) = 0$.

Workers. Turning to the other side of the market, an unemployed worker must choose the submarket that maximizes the value of job search, where I model the worker's choice as a function $\phi(i,j,s)$ giving the probability that a worker of type s applies to submarket (i,j). As discussed previously, job applicants receive an unemployment flow benefit $B_{i,s}$, and once employed they earn wages $w(i,s)m(j,s)$ and an industry amenity A_i with flow utility given by $\big(w(i,s)m(j,s)\big)^{\psi}A_i^{1-\psi}$. Hence, for any worker searching for employment:

$$\begin{aligned} \phi^*(i,j,s) = \arg &\max_{\phi(i,j,s)} \sum_i \phi(i,j,s)U(i,j,s) \\ &s.t.\ \rho U(i,j,s) = B_{i,s} + \zeta\theta(i,s)^{\eta}\big[W(i,j,s) - U(i,j,s)\big] \\ &\rho W(i,s) = \big[w(i,s)m(j,s)\big]^{\psi}A_i^{1-\psi} + \delta\big[U(i,j,s) - \mathbb{E}_{i,j}W(i,j,s)\big] \\ &\sum_{i,j}\phi(i,j,s) = 1 \end{aligned} \tag{6}$$

If follows directly from this problem that for any submarkets for which $\phi^*(i,j,s) > 0$ we must have $\rho U(i,j,s) = \mathbb{E}_{i,j}\rho U(i,j,s) \equiv U(s)$, where $U(s)$ is the worker's *outside option* and will be an important object in the next section. Works will only search for employment if the flow value of doing so is greater than that of choosing non-participation, and instead receiving the social security flow benefit B_{np} and the home production term χ:

$$\text{Search if } \chi < \rho U(s) - B_{np} \tag{7}$$

From participation decision (7) we will have that the probability a given unemployed worker chooses non-participation is $1 - F\big(U(s) - B_{np}\big)$.

Industry aggregators. Aggregators choose quantities of intermediate goods that minimize costs, taking as given the price P_i of the industry composite good and the prices $p_i(j)$ of the intermediate goods. Production is CES, with an identical elasticity of substitution σ for all industries but heterogeneous intermediate good shares $\alpha_i(j)$. Letting Y_i denote industry output, the profit maximization problem for a representative industry aggregator is:

$$y_i(j)^* = \arg\max_{y(j)} \left(\int \alpha_i(j) y_i(j)^{\frac{\sigma-1}{\sigma}} dj \right)^{\frac{\sigma}{\sigma-1}} - \int p_i(j) y_i(j) dj \tag{8}$$

Perfect competition in industry composite markets implies that $P_i \left(\int \alpha_i(j) y_i^*(j)^{\frac{\sigma-1}{\sigma}} dj \right)^{\frac{\sigma}{\sigma-1}} - \int p_i(j) y_i^*(j) dj = 0$, and hence that the demand for j-goods is proportional to $\alpha_i(j)/p_i(j)$.

Representative household. Finally, the representative household's problem is to choose quantities of industry composite goods that maximize utility, assumed also to take CES form. Prices of the composite goods are therefore assumed to be denominated in units of utility, and so consumption $C(i)$ of industry-i output will be the solution to the problem

$$Y^* = \arg\max_{Y(i)} \left(\sum_i \beta_i Y(i)^{\frac{\tau-1}{\tau}} \right)^{\frac{\tau}{\tau-1}} - \sum_i P_i Y(i) \tag{9}$$

This problem then gives us that the demand for industry output is proportional to $\big(\beta_i/P_i\big)^{\tau}$, where P_i will be equal to industry output divided by industry labor costs.

3.3 Equilibrium

Before defining equilibrium, it remains to describe the market-clearing conditions and laws of motion facing the economy. First, market tightness must be consistent with the optimal job search behavior of workers and firm vacancy policies. Defining $N(s)$ to be the unemployed mass of s-workers and $V_i(j)$ the total number of (i,j)-vacancies in steady-state, it must be that:

$$\theta(i,j) = \frac{V_i(j)}{\int \phi^*(i,j,s) N(s) ds} \tag{10}$$

Second, in steady-state we must have parity between flows into and out of unemployment:

$$\zeta \theta(i,s)^{\eta-1} V_i(j) = \delta\big(\nu(s) - N(s)\big) \tag{11}$$

Finally, market-clearing for intermediate and industry composite goods requires that

$$y_i^*(j) = \frac{\zeta \theta(i,j)^{\eta-1} V_i(j)}{\delta} \times \frac{\int N(s) \phi^*(i,j,s) m(j,s) ds}{\int N(s) \phi^*(i,j,s)} \tag{12}$$

$$Y_i^* = \left(\int \alpha_i(j) y_i^*(j)^{\frac{\sigma-1}{\sigma}} dj \right)^{\frac{\sigma}{\sigma-1}} \tag{13}$$

where the two terms in (12) correspond to total employment and expected labor productivity in submarket (i,j). A steady-state equilibrium in this economy is then wage posting decisions $w^*(i,j)$, output quantities

$y_i^*(j)$ and Y_i^*, and search behavior $\phi^*(i,j,s)$ satisfying problems (4), (6), and (9); endogenous distributions $N(s)$ and $V_j(s)$ satisfying (11) together with firms' entry decisions (5) and workers' participation decisions (7); and prices $p_i(j)$ and P_i such that market-clearing condition (10) and (12)-(13) hold.

The equilibrium cannot be characterized without restricting the unemployment payoff $B_{i,j,s}$, which may have arbitrary effects on worker search behavior and, consequently, on output quantities $y_i^*(j)$ and Y_i^*. To motivate a particular assumption, I begin by noting that in general, the equilibrium (1) is not efficient and (2) does not yield closed-form policy functions. In both cases the reason is that while the value of job search $U(i,j,s)$ has constant elasticity with respect to wages, this is generally *not* true of the elasticity with respect to market tightness. Both wages and market tightness are forms of payment, but the relative values that workers place on them will tend to vary with s. This prevents the derivation of simple policy functions, and it aggravates search congestion in some submarkets which, due to the concavity of the matching function, in turn generates inefficiencies. A natural assumption is therefore that $B_{i,j,s}$ is set so that the elasticity of $U(i,j,s)$ with respect to market tightness is constant, representing the 'optimal' choice by e.g. a utility-maximizing central planner.

I therefore impose the following condition:

Assumption 2: *Unemployment insurance takes the form*

$$B_{i,j,s} = \frac{\rho + \delta + \zeta\theta(i,j)^{\eta} - 1}{\rho + \delta}\zeta\theta(i,j)^{\eta}\big[w(i,j)m(j,s)\big]^{\psi}A_i^{1-\psi} \tag{14}$$

Assumption (14) states that unemployed workers are compensated not only based on wages and amenities associated with their submarket, but based on the job-finding rate as well. Workers are incentivized to seek out submarkets characterized by a higher θ and hence by greater congestion. Although I present (14) as an assumption rather than a theorem, in the appendix I provide an informal proof that this equation must hold (up to a multiplicative constant) for the equilibrium to be efficient.

Before characterizing the policy functions I briefly discuss the assignment of workers to jobs, which is considerably simplified by condition (14). I follow Costinot and Vogel (2010) in defining a correspondence $\lambda : i, s \to j$ that gives the set of jobs j chosen by s-workers in industry i. Provided that $\alpha_i(j)$ and $C_i(j)$ are continuous (which I assume), then given assumption 1, the following result can be shown:

Proposition 1. *If assumptions 1-2 hold, then $\lambda(s)$ is a strictly increasing, differentiable function independent of i, that satisfies the system of equations*

$$\frac{dU(s)}{ds} = \psi\frac{m_s\big(\lambda(s),s\big)}{m\big(\lambda(s),s\big)}U(s) \tag{15}$$

$$\frac{d\lambda(s)}{ds} = \frac{m\big(\lambda(s),s\big)U(s)N(s)}{\sum_i \left(\frac{\psi(1-\eta)}{\eta+\psi(1-\eta)} m\big(\lambda(s),s\big)p_i\big(\lambda(s)\big)\right)^{\psi} A_i^{1-\psi} \frac{\left[\alpha_i\left(\lambda(s)\right)/p_i\left(\lambda(s)\right)\right]^{\sigma}}{\left(\int \frac{\alpha_i\left(\lambda(k)\right)^{\sigma}}{p_i\left(\lambda(k)\right)^{\sigma-1}}dk\right)^{\frac{\sigma}{\sigma-1}}} \frac{\left[\beta_i/P_i\right]^{\tau}}{\left(\sum_k \frac{\beta_k^{\tau}}{P_k^{\tau-1}}\right)^{\frac{\tau}{\tau-1}}} Y} \tag{16}$$

where $U(s) \equiv \rho U(i,j,s)$ is the outside option of an s worker, and where $\lambda(\underline{s}) = 0$ and $\lambda(\overline{s}) = 1$.

Equations (16) and (15) describe the sorting of workers across occupations. Higher-skilled workers are assigned to higher-j jobs, and this assignment is constant across industries when condition (14) holds. As in other models of assignment, the rate at which wages increase in s is determined by the rate at which productivity increases in s, given the worker's job j.

Full derivations of the policy functions are given in the appendix. Two aspects of the model merit discussion, the first being log-additivity of the wage function. Wages can be written as:

$$\begin{aligned}
w^*\big(i,\lambda(s)\big) &= ID_i\big(\lambda(s)\big)PD(s) \\
ID_i(j) &= \left(\frac{\psi(1-\eta)(\rho+\delta)C_i(j)}{\eta\zeta^{\frac{1}{\eta}}A_i^{\frac{(1-\psi)(1-\eta)}{\eta}}}\right)^{\frac{\eta}{\eta+\psi(1-\eta)}} \\
PD(s) &= U(s)^{\frac{1-\eta}{\eta+\psi(1-\eta)}}
\end{aligned}$$

With free entry, match surplus is a function of entry costs and amenities, which jointly determine the 'cost' of a filled vacancy. The greater the profits required to induce firms to enter, the larger the firm wage differential $ID_i(j)$. Worker wage differentials on the other hand are pinned down by the outside option $U(s)$. Note that while wages are separable in employer and skill, they are not separable in *industry* and skill, since $C_i(j)$ may vary with both i and j. Hence the model does not place strong restrictions on how PD and ID covary.

A second key aspect of the model is the behavioral response of workers and firms to changes in prices. Worker search probabilities and total firm output are given by the equations

$$\begin{aligned}
\phi^*(i,j,s) &= \frac{y_i^*\big(\lambda(s)\big)Y_i^* ID_i\big(\lambda(s)\big)^{\psi}A_i^{1-\psi}}{\sum_m y_m^*\big(\lambda(s)\big)Y_m^* ID_m\big(\lambda(s)\big)^{\psi}A_m^{1-\psi}} \\
y_i^*\big(\lambda(s)\big) &= \left(\beta_i\frac{\psi(1-\eta)}{\eta+\psi(1-\eta)}\right)^{\tau}\left(\frac{\alpha_i\big(\lambda(s)\big)m\big(\lambda(s),s\big)}{ID_i\big(\lambda(s)\big)PD(s)}\right)^{\sigma}\left(\int \alpha_i\big(\lambda(k)\big)^{\sigma}\left[\frac{m\big(\lambda(k),k\big)}{ID_i\big(\lambda(k)\big)PD(k)}\right]^{\sigma-1}dk\right)^{\frac{\tau-\sigma}{\sigma-1}} Y
\end{aligned}$$

written for better intuition as functions of ID and PD. Workers are more likely to apply to submarket (i,j) when (1) more (i,j)-output is demanded, and when (2) firms in submarket (i,j) are higher-paying or have relatively greater amenities. Output in submarket (i,j) will be greater, on the other hand, when (1) technological demand is higher due to the parameters α and/or β, or when (2) output substitution across industries and/or occupations results in a higher price of (i,j)-output. For example if $\tau > \sigma > 1$, then employment will tend to be greater in industries and occupations in which unit labor costs $\frac{w(s)}{m\left(\lambda(s),s\right)}$ are smaller.

Hence it is possible to have rent-seeking by both workers and firms in this environment. Worker rent-seeking implies that technological changes (i.e. changes to α) will be more influential when occurring at high-paying firms or high-amenity firms. On the other hand, firm rent-seeking suggests that if technological change affects $PD(s)/m\big(\lambda(s), s\big)$ differently for high-skill and low-skill workers, then this may lead firms to substitute towards or away from skill-intensive firms and industries (i.e. through greater or lesser firm entry). However while both of these channels may be important *in theory*, in practice firm rent-seeking cannot be meaningfully constrained or identified since either quantities $m(j, s)$ or prices $p_i(j)$ must be normalized. Different normalizations will result in different rent-seeking behaviors - a problem more commonly encountered in aggregation problems concerning capital.

3.4 Comparative statics from a 3-industry model

In this section I show, in a simple 3-industry version of the model, how firm heterogeneity influences the relationship between changes in model primitives (α and β) and the wage distribution ($w\big(i\lambda(s)\big) = ID_i\big(\lambda(s)\big)PD(s)$). I assume that $\tau = \sigma = 1$, an assumption that will carry over to the structural estimation in the next section. For this exercise I also impose the functional form assumptions $C_i(j) = C_i$, $\alpha_i(j) = e^{\alpha_i j}$, and $m(j, s) = e^{Bjs}$, and I set parameters so that industries have equal vacancy shares, the variances of the person and employer wage effects are equal, and there is moderate sorting ($Corr(PD, ID) \approx .5$). Across the three industries, I assume that α_i and C_i are monotonically related and hence that industry skill demand and pay are correlated.

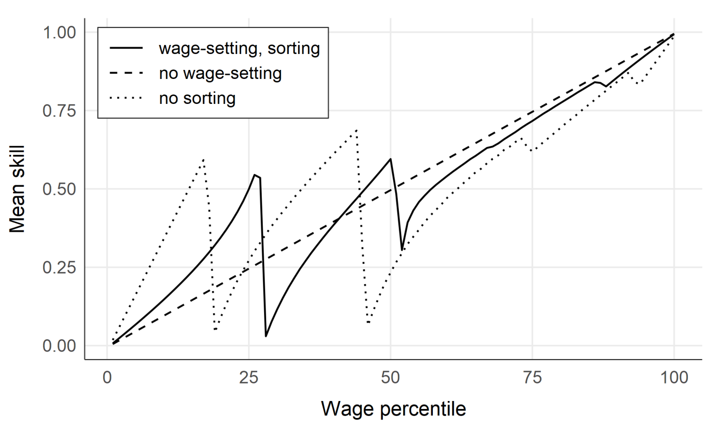

Figure 3.1: Mean skill by wage percentile

I begin by showing, in figure 3.2, how firm heterogeneity affects the relationship between wage and skill. In the absence of wage-setting (i.e. $C_i = C_{i'}$) these are perfectly correlated. The correlation becomes less than one in the presence of industry wage differentials, and it is smallest when there is no sorting (or hypothetically, if PD and ID were negatively correlated). Discontinuities are present due to the assumption of a discrete number of industries and occur at the wage percentiles corresponding to the lowest and highest wages paid in a given industry, which are associated with skill levels $\underline{s}$ and $\overline{s}$, respectively.

Figure 3.2: Distributional effects of skill demand and sorting

Plot 3.2 shows the effects of two changes in technology: an increase in the demand for skilled labor that affects all industries, implemented by increasing $\frac{d}{ds} \log \alpha_i(j)$ identically across industries; and a decrease in the employment share of the middle-skill industry. In the absence of wage-setting, higher skill premia unambiguously increase the wages of high-earning workers, and decreases the wages of low-earners. When the two industries pay different wages, however, wages may also increase in lower percentiles of the distribution, corresponding to high-skill workers in low-paying industries. Industry polarization on the other hand has little effect on the extremes of the distribution, as these represent low-skill workers in the low-skill industry and high-skill workers in the high-skill sector. Rather wage declines are concentrated in the second quartile, and wage gains in the third.

Figure 3.3: Distributional effects by channel

In figure 3.3 I break down these effects by channel: skill premia, wage sorting, and changes to the composition of employer wage effects. In the left-hand panel we see that rising skill demand predominantly interacts with employer heterogeneity through wage sorting. The effect on skill premia is larger when sorting is absent, which merely reflects the fact that a change in the slope of log α will have a larger effect on labor markets when the slope is near zero. Similarly, because the absolute increase in the number of skilled vacancies is greater in the high-paying industry, worker rent-seeking tends to amplify the increase in skill premia. Turning to the second panel, industry polarization has minimal effects on skill premia as it results in greater employment in both the low-skill and the high-skill industries. It does however result in greater wage sorting and greater variance of the employer wage effects. In these examples worker rent-seeking plays a modest role because intra-industry labor distributions are either unchanging or are changing similarly across industries. If skill demand were to increase in only a single industry, for example, then rent-seeking would become more important as it would affect the magnitude of the labor supply response.

4 Quantitative results

4.1 Estimation

I begin by normalizing variables that are not uniquely identified. In a model with assignment, the elasticity of substitution between jobs and the distribution of skill are not separately identified from the labor productivity function γ, so I set $\sigma = 1$ and I assume that s is uniformly distribution in $[0, 1]$. Similarly, it is not possible to identify the utility function exponent ψ independently of industry amenities, and so I impose the condition $\psi = .5$. The discount rate ρ serves only to scale worker utility and I therefore set it to $\rho = 4.2$, corresponding to a discount factor of of .96; all rate variables are assume to reflect annual values.

The remaining aggregate parameters are set as follows. For the (annual) separation rate, I make use of the equilibrium condition that total separations are equal to total hires, where hires and vacancies are calculated from LIAB data. The search exponent η may in principle be identified from aggregate unemployment flows and LIAB data on vacancies and hires, but in practice I obtain estimates that are implausibly high ($\sim .6$), likely due to attenuation bias caused by noise in the LIAB measures. I therefore set $\eta = .35$ following Kohlbrecher et al. (2016). Search efficiency ζ is then obtained by equating predicted with observed hires. The long-term unemployment benefit B_{np} is estimated by matching the decline in West German unemployment rates The industry elasticity of substitution τ is unidentified, both in practice and in principal. It will govern

Parameter	Definition	Value	Source
ψ	utility exponent	.5	normalized
ρ	discount rate	.042	discount factor of .96
δ	separation rate	.169	$\frac{\text{Annual hires}}{\text{Total employment}}$
η	match elasticity	.35	Kohlbrecher et al. (2016)
ζ	match efficiency	2.39	$\frac{\text{hires}}{\text{predicted hires}}$
B_{np}	long-term UP benefit	$\{-.42, -.70, -.88, -\infty\}$	$u_t - u_{t-1}$
σ	job elasticity	1	normalized
τ	industry elasticity	1	see discussion

Table 4.1: Aggregate parameters

the extent to which labor demand responds to α, PD, and γ, but since α and γ are not separately identified from intra-industry cost shares, τ cannot be determined - an instant of the aggregation problem normally associated with differentiated capital inputs. I therefore set $\tau = 1$, which implies that submarket cost shares do not respond to changes in relative wages - a natural assumption given that I have no way of constraining any such response.

Industry-specific parameters are obtained as follows. I estimate industry amenities A_i from the equilibrium equation for industry average vacancy-filling rates, which can be shown to be proportional to $A_i^{\frac{-(-\psi)(1-\eta)}{\eta}} \left(\mathbb{E}_{L_i(j)} PD\big(\lambda^{-1}(j)\big)/ID_i(j)^{\frac{\psi(1-\eta)}{\eta}}\right)^{-1}$ where the term in parentheses is the labor-weighted average within-industry. Once $\alpha_i(j)$ and $\gamma(j,s)$ are obtained (discussed next), industry shares β_i and entry costs C_i are then estimated from empirical employment shares and industry-occupation mean employer wage effects. The overall level of entry costs are set so that model-predicted market tightness is equal to its average empirical counterpart over the 1993-2017 period, which I estimate at .146 using annual unemployment numbers from Destatis. Estimation of the labor share and productivity functions $\alpha_i(j) = \alpha_i^t(j)$ and

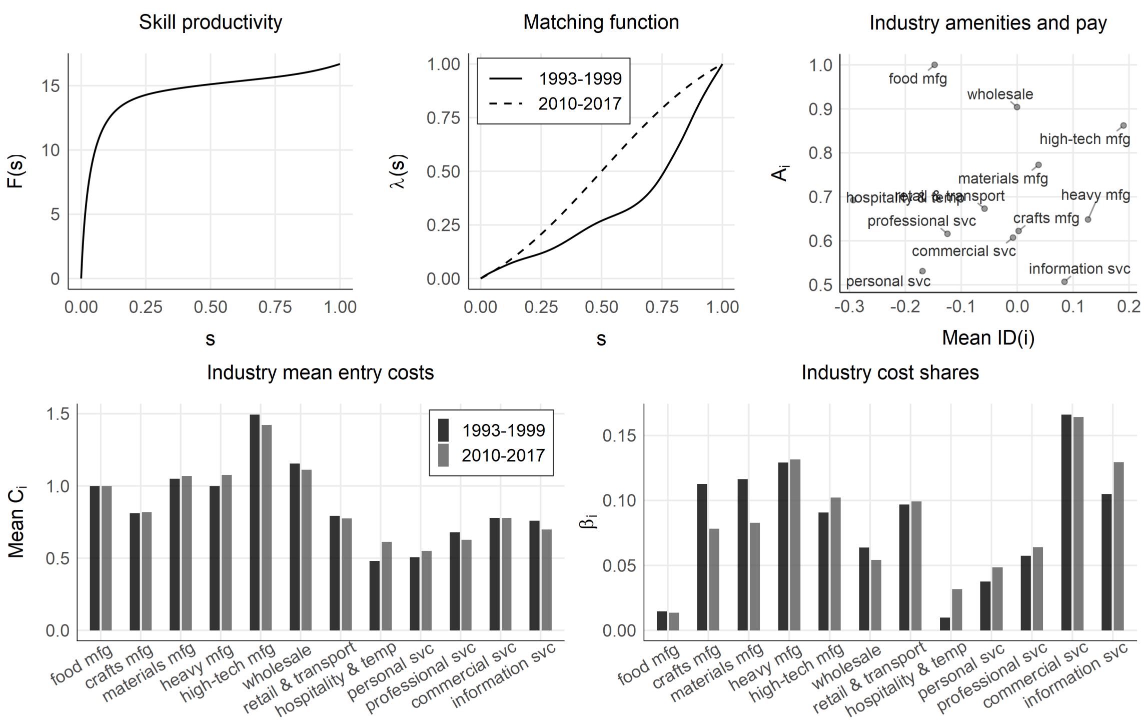

Figure 4.2: Distributional parameters

FIGURE NOTES. Aggregate vacancies calculated by multiplying industry mass times vacancy shares. Industry $ID(i)$ averaged across panels.

$\gamma(j,s) = \gamma^t(j,s)$ is a critical step that requires one of the two parameters to be restricted, as only their joint effect on the wage distribution is identified. In general such a restriction involves a loss of generality, as γ will affect firm profits (and hence wages) whereas while α will not. With $\tau = 1$, however, there will be no loss of generality from imposing the condition

Assumption 3: $\gamma(j,s) = e^{F(s)[\gamma_s + \gamma^t_{j,s} j]}$

where $F(s)$ and γ_s are time-invariant, and $\gamma^t_{j,s}$ is constrained by the assumption that $\lambda(\overline{s}) = 1$. The relationship between skill and productivity is governed by the function $F(s)$, where F is positive but otherwise unrestricted. As F cannot be separately identified from the matching function, I impose the normalization $\lambda(s) = s$ for the 2010-2017 period; F is then estimated over this panel and fixed for the previous 3 panels, allowing λ to be estimated. The term γ_s is included to avoid division by zero in the calculations, and I set it to a value of $\gamma_s = .01$. A final step is to check that the assumption $\lambda'(s) > 0$ holds for all s. A violation of this assumption can in most cases be addressed by changing how λ is normalized when F is estimated.

The functions α^t_i and F are estimated from the aggregate and intra-industry distributions of person effects $PD(s)$ for each of the four panels, where I first take averages of $PD(s)$ conditional on industry and 3-digit occupation in order to capture only that portion of the person wage effect related to skill, as defined in the empirical portion of the paper. For the aggregate wage distribution, I calculate percentiles of the group mean person effects and then truncate the first and last percentiles for better numerical stability. Intra-industry distributions are measured in 20 quantiles, as a finer subdivision is generally not possible given data confidentiality requirements for the LIAB. Because the wage function is linear, all three functions may be obtained non-parametrically from the empirical distribution of AKM person wage effects. The immediate implication that the model can replicate closely the empirical wage and labor distributions, although in practice there is mismatch due to interpolation of the distributions. I therefore iteratively adjust β_i, $\sum_i \alpha_i(j)$, and $C_i(j)$ to ensure that industry employment shares, industry mean employer wage effects, and the simulated matching function are equal to their empirical counterparts.

Moment	1993-1999	1998-2004	2003-2010	2010-2017
Empirical wage variance	**.1620**	**.1919**	**.2240**	**.2264**
$Var(PD)$	.1034	.1178	.1340	.1367
$Var(ID)$	.0307	.0378	.0516	.0397
$Cov(PD, ID)$	.0072	.0104	.0112	.0159
Predicted wage variance	**.0627**	**.0782**	**.0938**	**.1033**
$Var(PD)$	.0369	.0435	.0475	.0566
$Var(ID)$	.0101	.0133	.0182	.0137
$Cov(PD, ID)$	.0079	.0107	.0141	.0165
Model R^2	.387	.408	.419	.456
Wage sorting explained	110%	103%	126%	104%

Table 4.3: Empirical and predicted wage moments

TABLE NOTES. Empirical moments calculated after eliminating topmost and bottom-most percentiles, following the procedure described in the text.

The empirical and model-predicted wage moments are shown in table 4.3. The model captures approximately 2/5 of West German wage variance, with the remainder attributable to other moments in (2) and to within-industry/occupation variance of the person and employer effects. On the other hand wage sorting is entirely replicated by the model, as would be expected given the empirical results in figure 2.3 and the ability of the model to closely replicate the distribution of wage effects conditional on industry and occupation. In addition to the estimation process described above, I estimate the model under two alternative specifications: one in which employer wage differentials are compensating differentials that do not affect market tightness ($A_i \equiv A_i(j) = ID_i(j)^{\frac{-(1-\psi)}{\psi}}$), and one in which workers care only about the employer wage effects and not the amenities ($\psi = 1$). As rent-seeking behavior will in general be sensitive to underlying assumptions, the alternative specifications may help to provide a sense of the contribution and the variability of this channel to the overall results. Nevertheless under the benchmark specification I estimate rent-seeking to only marginally affect worker search probabilities, as industry wage differentials are 'small' relative to wages overall.

4.2 Amplification

In order to quantify the the contribution of employer heterogeneity to the skill-demand-wage inequality relationship, I consider the counterfactual scenario where firms pay the same wage ($C_i(j) = \bar{C}A_i^{\frac{(1-\psi)(1-\eta)}{\eta}}$). Changes to $C_i(j)$ will affect the distribution of vacancies because of free entry, as well as workers' incentives to seek out different industries; hence an important question is how β and α should be adjusted in order for the counterfactual scenario to yield meaningful comparisons. The simplest approach would be to leave β and α unadjusted, but to do so would result in large changes to industry vacancy shares, and is an uninteresting scenario given that entry costs are only identified off of industry wage differentials. More sensible would be to shift the share functions in a time-invariant way, such that the initial (e.g. 1993-1999) vacancy or labor distributions are unchanged under the counterfactual scenario. If the initial vacancy distribution is held fixed, then the initial labor distribution will be different due to worker rent-seeking behavior. Alternatively if the initial labor distribution is held constant, then this affords a much smaller role for rent-seeking, which will only 'matter' in the event that changes over time in the shape of α are substantially different across industries - which, in practice, turns out not to be the case.

As each of these assumptions is reasonable, I consider both throughout this section. I define the two counterfactual experiments as follows:

> **Counterfactual #1**: all firms pay the same wage conditional on skill, and α and β are adjusted so that the distribution of vacancies across submarkets is unchanged for the 1993-1999 period:

$$C_i^{CF1}(j) = A_i^{\frac{(1-\psi)(1-\eta)}{\eta}} \left(\sum_i \beta_{i,t} A_i^{\frac{(1-\psi)(1-\eta)}{\eta}} \int \frac{\alpha_{i,t}(j)}{C_{i,t}(j)} dj \right)^{-1}$$

$$\frac{\alpha_{i,t}^{CF1}(j')}{\alpha_{i,t}^{CF1}(j)} = \frac{\alpha_{i,t}(j')}{\alpha_{i,t}(j)} \frac{C_{i,1993-1999}(j)}{C_{i,1993-1999}(j')}$$

$$\frac{\beta_{i',t}^{CF1}}{\beta_{i,t}^{CF1}} = \frac{\beta_{i',t}}{\beta_{i,t}} \left(\frac{A_{i',1993-1999}}{A_{i,1993-1999}} \right)^{\frac{(1-\psi)(1-\eta)}{\eta}} \frac{\int \frac{\alpha_{i',1993-1999}(j)}{C_{i',1993-1999}(j)} dj}{\int \frac{\alpha_{i,1993-1999}(j)}{C_{i,1993-1999}(j)} dj}$$

Counterfactual #2: all firms pay the same wage conditional on skill, and α and β are adjusted so that the distribution of labor across submarkets is unchanged for the 1993-1999 period:

$$C_i^{CF2}(j) = A_i^{\frac{(1-\psi)(1-\eta)}{\eta}} \left(\sum_i \beta_{i,t} A_i^{\frac{(1-\psi)(1-\eta)}{\eta+\psi(1-\eta)}} \int \frac{\alpha_{i,t}(j)}{C_{i,t}(j)^{\frac{\eta}{\eta+\psi(1-\eta)}}} dj \right)^{-1}$$

$$\frac{\alpha_{i,t}^{CF2}(j')}{\alpha_{i,t}^{CF2}(j)} = \frac{\alpha_{i,t}(j')}{\alpha_{i,t}(j)} \left(\frac{C_{i,1993-1999}(j)}{C_{i,1993-1999}(j')} \right)^{\frac{\eta}{\eta+\psi(1-\eta)}}$$

$$\frac{\beta_{i',t}^{CF2}}{\beta_{i,t}^{CF2}} = \frac{\beta_{i',t}}{\beta_{i,t}} \left(\frac{A_{i',1993-1999}}{A_{i,1993-1999}} \right)^{\frac{(1-\psi)(1-\eta)}{\eta+\psi(1-\eta)}} \frac{\int \frac{\alpha_{i',1993-1999}(j)}{C_{i',1993-1999}(j)^{\frac{\eta}{\eta+\psi(1-\eta)}}} dj}{\int \frac{\alpha_{i,1993-1999}(j)}{C_{i,1993-1999}(j)^{\frac{\eta}{\eta+\psi(1-\eta)}}} dj}$$

Note the implicit use of $\tau = \sigma = 1$. In the event where elasticities of substitution are different from one, the adjustments to α and β will depend on wages and the matching function, and will not have closed-form solutions.

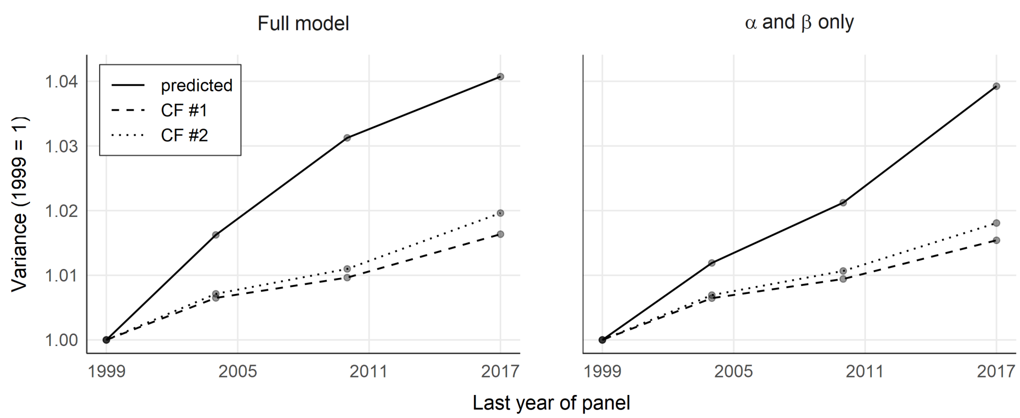

Figure 4.4: Predicted and counterfactual wage variance

FIGURE NOTES. Simulated results for the cases where (1) α, β, C, and B are time-varying, and (2) where C and B are fixed at their 1993-1999 and 2010-2017 values, with the average of the two cases shown.

The predicted trend in wage inequality is shown in figure 4.4. In the first panel entry costs and the long-term unemployment benefit are allowed to vary over time, and in the second panel these are fixed. The technological parameters α and β account for virtually all of the model-predicted trend over the full 1993-2017 sample period, and under each of the counterfactual scenarios this trend is less than half as large. The model only predicts three-fifths of the empirical trend, however, and so there is some ambiguity as to the correct method for quantifying the amplification due to firm wage-setting, depending on whether one attributes residual wage inequality to technical change as well. In table 4.5 I therefore present a range of estimates for the contribution to trend, with the amplification naturally appearing largest when measured against the model-predicted effect of α and μ, and smallest when compared with the empirical trend. For the full sample period this gives a range of estimates between 32.4% and 59.8%, representing an amplification of trend between 50% and 150%.

Scenario	93-99 to 98-04	98-04 to 03-10	03-10 to 10-17	93-99 to 10-17
Counterfactual #1: constant vacancy distribution				
Contribution, α and μ	**.0053**	**.0065**	**.0120**	**.0238**
% total effect of α and μ	46.1	66.1	64.9	59.8
% model-explained trend	34.0	41.6	127.4	58.7
% empirical trend	17.6	20.3	500.2	37.0
Contribution, all parameters	**.0092**	**.0121**	**.0027**	**.0240**
% model-explained trend	59.4	77.3	28.5	59.1
% empirical trend	30.8	37.7	111.9	37.2
Counterfactual #2: constant labor distribution				
Contribution, α and μ	**.0047**	**.0058**	**.0105**	**.0210**
% total effect of α and μ	41.1	58.5	56.6	52.6
% model-explained trend	30.3	36.9	111.1	51.6
% empirical trend	15.7	18.0	436.1	32.5
Contribution, all parameters	**.0086**	**.0112**	**.0011**	**.0209**
% model-explained trend	55.3	71.7	11.1	51.4
% empirical trend	28.7	35.0	43.7	32.4

Table 4.5: Contribution of firm heterogeneity to rise in wage variance

Table notes. Total contribution is the estimated change in wage variance due to α and μ. The model-explained trend is the predicted change in wage variance, and the empirical trend is total increase in West German wage variance.

Comparing the two counterfactuals, there is a larger estimated contribution from firm wage-setting when the initial vacancy distribution is held constant. The reason is that, when industry wage differentials are equalized, there are more applicants per job in low-skill industries and fewer applicants per job in high-skill industries. Consequently growth in the former will tend to have a larger negative effect on skill premia, while growth in high-skill sectors will tend to raise skill premia by less. On net this means that industry polarization increases wage variance by less. Hence, through the lens of the model, one of the reasons that wage inequality has increased is that high-skill industries like information technology "punch above their weight"; job growth in these industries has a larger effect on wage inequality because these jobs pay more and therefore elicit a larger supply response.

4.3 Attribution

I now turn to the relative contributions of α and β, and how changes in these parameters have interacted with firm heterogeneity. In figure 4.6 the effects of changing occupation and industry demand are shown, with entry costs and the long-term unemployment held fixed at both their 1993-1999 and 2010-2017 values and then averaged across the two cases. Of the total change in variance, roughly one-fourth is attributed to changing industry composition and the remainder to rising demand for skilled occupations, with the remaining time-varying parameters having a negligible effect over the full sample period. Firm heterogeneity has substantially different implications for wage variance from industry and occupation demand, amplifying the former by 50% and accounting for all of the effect of the latter. Because declines in low-skill manufacturing sectors have been offset by gains in low-skill services, changing industry composition has on net reduced skill premia, and so the entire contribution of industry demand to the aggregate trend is through wage sorting and the composition of employer wage differentials.

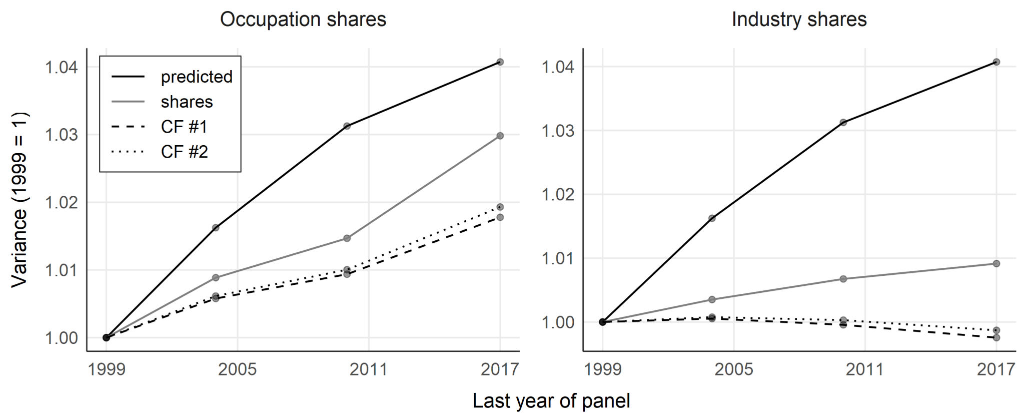

Figure 4.6: Sources of predicted wage variance

FIGURE NOTES. Effects calculated by holding other parameter values at their 1993-1999 and 2010-2017 values and averaging the simulated change in variance.

This is shown in figure 4.7, where the effects of changing demand are broken down by channel. Occupational demand explains all of the rise in the model-predicted variance of the person effect, and three-fifths of the empirical increase. On the other hand industry composition explains all of the increase in the variance of the employer effect predicted by the model; this represents less than half of the empirical trend, however, which is mostly the result of greater dispersion within industry and occupation. Finally, industry and occupation demand have had comparable effects on wage sorting, this representing the one channel in which both types of technical change are important. The model accounts for all of the growth in empirical wage sorting, which is attributed 60% to occupational skill demand and 40% to industry demand. Tabulated results are given in the appendix.

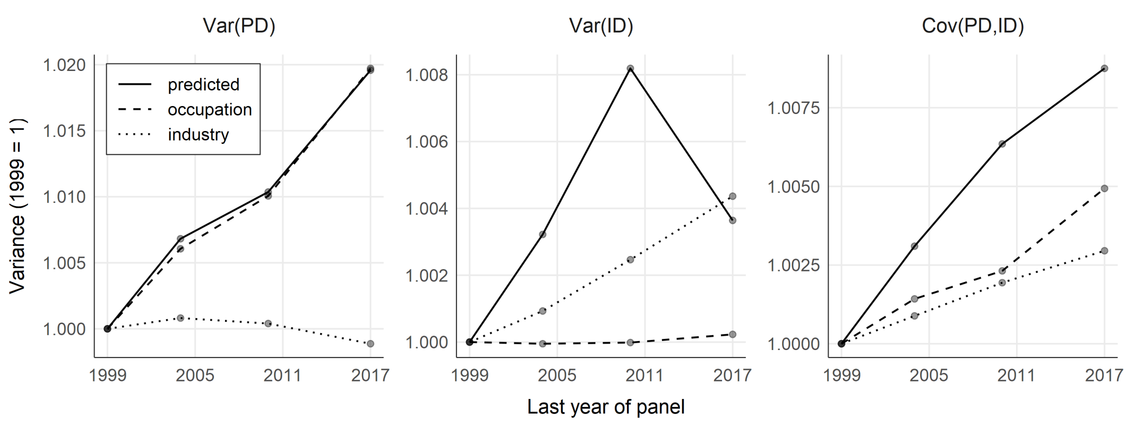

Figure 4.7: Predicted wage variance by channel

Several aspects of these results are surprising. First is the general lack of interaction between occupation and industry demand, and their tendency to affect only two of the three channels. This is not a property of the model, but rather the result of industry composition changing in a way that is neutral with respect to skill demand, and skill demand rising comparably across industries. Technical change has affected aggregate skill demand and industry composition, but not the underlying patterns of association between industry and occupation. Second, there is remarkable consistency over time in the predicted effects of occupation and industry demand, both in the aggregate and in each of the three channels in figure 4.7. This would seem to support the notion that changes in α and β reflect technical change rather than short-lived shocks to German labor markets, and it also suggests that any AKM estimation bias in the middle panels has a negligible effect on results. Finally, the counterfactual results in figure 4.6 raise the possibility that while sectoral change is indeed important from a distributional standpoint, its role has been mis-diagnosed; industry is almost always thought to influence the wage distribution through the skill premia channel. It may be that the West German case is unique, but the fact remains that industry wage differentials appear to an important and potentially dominant channel by which industrial composition affects wage inequality.

4.4 Occupation and industry contributions to wage variance

A benefit of the structural model is that it allows results to be disaggregated to the level of individual occupations and industries. This is straightforward in the case of industry, which is an explicit part of the model; we may adjust the components of β_i individually and then simulate the effect on the wage distribution. For occupations this exercise is more involved as there is no direct correspondence between empirical occupations and the occupational space assumed by the model. I therefore define occupations as distributions over i and j. Because data disclosure restrictions prevent me from measuring this joint distribution in detail, I instead estimate it using the joint industry-occupation distribution and the distribution of occupations across deciles of the AKM person effect. Some detail is lost in this process, and when comparing different points in time I am forced to assume that occupational distributions are fixed at their initial or final values; hence the results for occupations should be regarded as rough.

In figure 4.8 two results are shown: the estimated historical contributions to the change in wage variance, and the effects of a 1% increase in vacancy share beginning from the 2010-2017 parameter estimates. In terms of low-skill jobs, the largest contributions are from craft and production occupations, as well as office work, all of which are associated with industries and regions of the task space that see declining demand over this period. By far the largest contributions come from rising demand for skilled labor, particularly engineers and managerial professionals. The second panel gives a somewhat clearer picture as to the role played by firm heterogeneity, and the differences between occupations in terms of their associated with employer pay. Production jobs and unskilled labor jobs are strongly associated with high-paying and low-paying industries, respectively, and so while they are roughly equal in terms of skill, the latter has an effect on wage variance roughly three time as large. Similarly, engineering and scientific jobs are comparable in terms of their skill level but differ substantially in how they affect wage variance; increased demand for engineers generates an

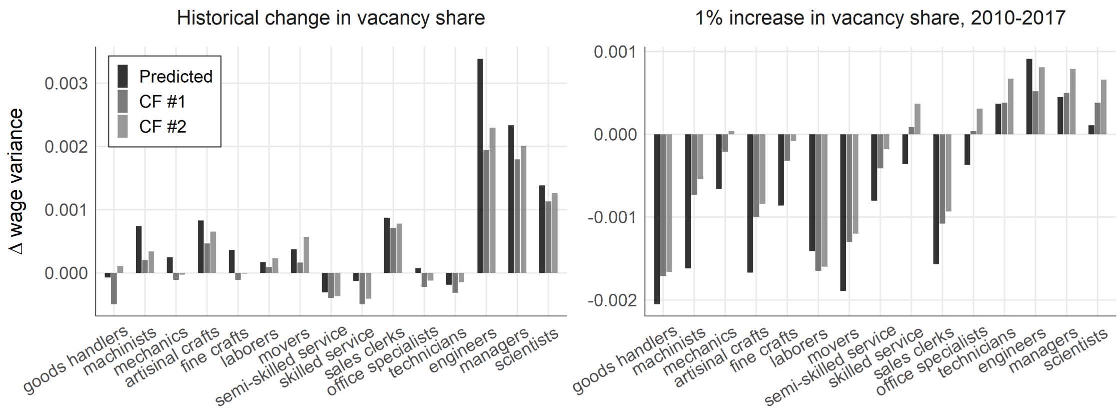

Figure 4.8: Occupational contributions to change in wage variance

Figure notes. The historical change is the change in occupation vacancy share between 1993-1999 and 2010-2017; effects are measured with other parameters fixed at their 1993-1999 and 2010-2017, with the average of the two cases shown. The effect of a 1% increase in vacancy share is estimated with all other parameters at their 2010-2017 values.

effect twice as large, due the concentration of engineering jobs in large scale, high-paying manufacturing sectors.

Comparable results for industries are shown in figure 4.9. The historical contributions are dominated by the crafts and materials manufacturing sectors, the hospitality and temp agency industry group, and information services. The impact of employer heterogeneity varies substantially. Declines in manufacturing employment predominantly have an effect three times as large as would be the case otherwise; rising employment at temp agencies generates large increases in wage variance, whereas it would otherwise tend to reduce wage variance; and the effect of rising employment at information services firms is only marginally affected by shutting down firm wage-setting. The right-hand panel presents similar contrasts. Commercial services and heavy manufacturing are equally skill-intensive, but the second of these two exerts a distributional influence five times as large. Firm heterogeneity tends in general to dampen the effects of growth in service industries, and to amplify the effects of growth in manufacturing, due to the much higher wages paid by goods-producers.

One aspect of figure 4.9 that deserves further attention is the growth of the hospitality/temp industry group. Because LIAB contains establishment-level data on the use of temp labor, I am able to simulate wage inequality under the counterfactual scenario where outsourcing to temp agencies does not increase over the sample period. To do this I first use temp agency employment to estimate the change in $\beta_{\text{hosp/temp}}$ associated with increased outsourcing between 1993-1999 and 2010-2017. Based on observed changes in industry-level use of temp labor, I then re-distribute this demand across artificial "industries" with the same occupational labor demand as the hospitality and temp sector, but with entry costs and amenities associated with the outsourcing industries. This experiment predicts that wage variance would have risen by about 6% less (.039 log points) in the absence of greater outsourcing to temp agencies, due to to smaller increases in wage

Figure 4.9: Industry contributions to change in wage variance

Figure notes. The historical change is the change in industry vacancy share between 1993-1999 and 2010-2017; effects are measured with other parameters fixed at their 1993-1999 and 2010-2017, with the average of the two cases shown. The effect of a 1% increase in vacancy share is estimated with all other parameters at their 2010-2017 values.

sorting and the variance of the employer wage effect. This is offset to some extent by a greater increase in the variance of the person effect, since reallocating low-skill vacancies to high-cost industries reduces the number of vacancy postings.

In light of these results, the quantitative implications of rising skill demand, and in some cases even the qualitative effects, are likely to depend on which industries and occupations are most impacted. If changes to demand have manifested differently in different countries or at different times, or if there is substantial geographic variation in the employer wage effects associated with different sectors, then the aggregate relationship between skill demand and the wage distribution is also likely to vary. On the one hand this suggests that caution should be taken when extrapolating the effects of skill-bias from historical data. On the other hand the results in this paper suggest that it may be possible to explain some of the variation in trend observed across countries, if sufficiently rich matched data is available.

4.5 Comparison with East Germany, 1993-2017

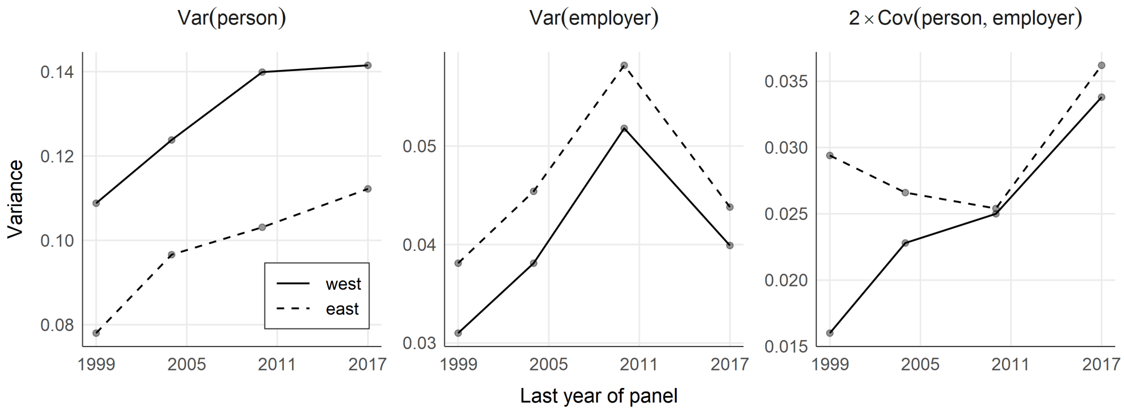

Figure 4.10: AKM wage variance decomposition for East Germany, 1996-2017

Figure notes. Person and employer wage effects estimated separately in four panels covering 1996-1999, 1998-2004, 2003-2010, and 2010-2017. Weights equalized across years in calculations. Note: West German results include the years 1993-1995.

5 Conclusion

In this paper I study how employer heterogeneity influences the relationship between skill-biased technical change and wage inequality, focusing on trends in West Germany over the period 1993-2017. My empirical analysis builds on past work by Card, Heining, and Kline (2013) who find that one-third of the upward trend wage variance is due to sorting of high-earning workers into high-paying employers. I show two main results. First, West German wage sorting is a feature of the industry-occupation structure of labor markets; conditional on broad industry and occupational classifications, sorting is absent. Second, wage sorting has contributed to growing wage variance due to a combination of rising skill premia for occupations associated with high-paying employers, and changes to industry composition that on net have tended to concentrate low-skill labor in low-paying service industries.

I then develop and estimate a model of labor markets and, through counterfactual exercises, quantify the role of firm heterogeneity. Results indicate that more than half of the effect of changing occupational and industry labor demand on wage variance is due to the presence of employer wage differentials, and between 32% and 37% of aggregate empirical trend. Although changing industry shares are found to explain roughly one-sixth of the rise in West German wage variance, this is entirely due to interactions with employer wage premia, as structural change has tended to favor unskilled labor and to *reduce* the dispersion of skill premia. Finally, I show that the quantitative implications of firm heterogeneity differ substantially across occupations and industries and between East and West Germany, creating variation in the aggregate relationship between skill demand and wage inequality.

These results are potentially important for several reasons. First, they indicate that the linkage between SBTC and wage inequality is non-trivial, and that models of technical change that do not explicitly account for firm wage-setting may yield counterfactual predictions about the wage distribution. Second, firm heterogeneity may provide traction in explaining observed variation in the relationship between technical change and wage inequality, both temporal and geographic. There is a large literature that attempts to understand why, for example, wage inequality has declined in Brazil as it has risen elsewhere; firm heterogeneity may be critical for understanding these differences. Finally, while past empirical research has found that domestic outsourcing is associated with wage decreases, the results presented here indicate that such outsourcing has sizable distributional effects, as a result of large employer wage gaps between manufacturing and service sectors. This is particularly relevant in the West German case, given ongoing controversy regarding the use of temp agency labor and regulatory efforts to close the pay gap between temporary and full-time workers.

At the same time this paper is subject to important limitations that may hopefully be addressed by future research. The empirical results are purely descriptive. This may be unavoidable given the macroeconomic and somewhat complicated nature of the relationships being studied; but an immediate and important question is whether similar results hold for other countries. Availability of matched employer-employee data is a key challenge in answering this question. The quantitative results in this paper should be regarded as a first step, given the stylized nature of the equilibrium model that I estimate. Past work has tended either to rely on purely empirical models of the wage distribution, or on theoretical frameworks that generate counterfactual wage distributions and are difficult to take to the data in a quantitatively meaningful way. In attempting to bridge the gap between these two approaches I have prioritized transparency and tractability, at the cost of strong restrictions on the search environment and a simplistic model of inter- and intra-industry market structure. Relaxation of these assumptions represents a natural direction for future study.

References

[1] Abowd, John, Francis Kramarz, Paul Lengermann, Kevin L. McKinney, and Sébastien Roux. "Persistent inter-industry wage differences: rent sharing and opportunity costs." IZA Journal of Labor Economics 1, no. 1 (2012): 7.

[2] Abowd, John, Francis Kramarz, and David Margolis. "High wage workers and high wage firms." Econometrica 67, no. 2 (1999): 251-333.

[3] Acemoglu, Daron, and David Autor. "Skills, tasks and technologies: Implications for employment and earnings." In Handbook of labor economics, vol. 4, pp. 1043-1171. Elsevier, 2011.

[4] Acemoglu, Daron, and Pascual Restrepo (2018). "The race between man and machine: Implications of technology for growth, factor shares, and employment." American Economic Review 108, no. 6: 1488-1542.

[5] Ales, Laurence, Musab Kurnaz, and Christopher Sleet. "Technical change, wage inequality, and taxes." American Economic Review 105, no. 10 (2015): 3061-3101.

[6] Andrews, Martyn, Len Gill, Thorsten Schank, and Richard Upward. "High wage workers and low wage firms: negative assortative matching or limited mobility bias?." Journal of the Royal Statistical Society: Series A (Statistics in Society) 171, no. 3 (2008): 673-697.

[7] Autor, David, and David Dorn. "The growth of low-skill service jobs and the polarization of the US labor market." American Economic Review 103, no. 5 (2013): 1553-97.

[8] Autor, David, David Dorn, and Gordon Hanson (2015). "Untangling trade and technology: Evidence from local labour markets." The Economic Journal 125, no. 584: 621-646.

[9] Bagger, Jesper, and Rasmus Lentz. "An empirical model of wage dispersion with sorting." The Review of Economic Studies 86, no. 1 (2019): 153-190.

[10] Bagger, Jesper, Kenneth L. Sorensen, and Rune Vejlin. "Wage sorting trends." Economics Letters 118, no. 1 (2013): 63-67.

[11] Bellman, Lisa, Ben Lochner, Stefan Seth, and Stefanie Wolter. "AKM effects for German labour market data." FDZ-Methodenreport, IAB FDZ. 2020.

[12] Bonhomme, Stephane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler. "How Much Should we Trust Estimates of Firm Effects and Worker Sorting?" NBER Working Paper No. 27368, June 2020.

[13] Bonhomme, Stephane, Thibaut Lamadon, and Elena Manresa. "A distributional framework for matched employer employee data." Econometrica 87, no. 3 (2019): 699-739.

[14] Burstein, Ariel, and Jonathan Vogel (2017). "International trade, technology, and the skill premium." Journal of Political Economy 125, no. 5: 1356-1412.

[15] Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. "Firms and labor market inequality: Evidence and some theory." Journal of Labor Economics 36, no. S1 (2018): S13-S70.

[16] Card, David, and John DiNardo (2002). "Skill-biased technological change and rising wage inequality: Some problems and puzzles." Journal of labor economics 20, no. 4: 733-783.

[17] Card, David, Jorg Heining, and Patrick Kline. "Workplace heterogeneity and the rise of West German wage inequality." The Quarterly Journal of Economics 128, no. 3 (2013): 967-1015.

[18] Costinot, Arnaud, and Jonathan Vogel. "Matching and inequality in the world economy." Journal of Political Economy 118, no. 4 (2010): 747-786.

[19] Eeckhout, Jan, and Philipp Kircher. "Sorting and decentralized price competition." Econometrica 78, no. 2 (2010): 539-574.

[20] Eeckhout, Jan, and Philipp Kircher. “Identifying sorting—in theory.” The Review of Economic Studies 78, no. 3 (2011): 872-906.

[21] Goldschmidt, Deborah, and Johannes Schmieder. “The rise of domestic outsourcing and the evolution of the German wage structure.” The Quarterly Journal of Economics 132, no. 3 (2017): 1165-1217.

[22] Goos, Maarten, Alan Manning, and Anna Salomons. “Explaining job polarization: Routine-biased technological change and offshoring.” American Economic Review 104, no. 8 (2014): 2509-26.

[23] Haanwinckel, Daniel. “Supply, Demand, Institutions, and Firms: A Theory of Labor Market Sorting and the Wage Distribution.” Unpublished manuscript (2021).

[24] Hagedorn, Marcus, Tzuo Hann Law, and Iourii Manovskii. “Identifying equilibrium models of labor market sorting.” Econometrica 85, no. 1 (2017): 29-65.

[25] Hall, Anja, and Beermann, Beate (2021). BIBB/BAuA-Employment Survey 2006. GESIS Data Archive, Cologne. ZA4820 Data file Version 1.0.1, https://doi.org/10.4232/1.11072.

[26] Haltiwanger, John, and James R. Spletzer. “Between Firm Changes in Earnings Inequality: The Dominant Role of Industry Effects.” No. w26786. National Bureau of Economic Research, 2020.

[27] Kohlbrecher, Britta, Christian Merkl, and Daniela Nordmeier. “Revisiting the matching function.” Journal of Economic Dynamics and Control 69 (2016): 350-374.

[28] Lise, Jeremy, Costas Meghir, and Jean-Marc Robin (2016). “Matching, sorting and wages.” Review of Economic Dynamics 19 (2016): 63-87.

[29] Moen, Espen (1997). “Competitive search equilibrium.” Journal of Political Economy 105, no. 2: 385-411.

[30] Schmidtlein, Lisa, Stefan Seth, and Matthias Umkehrer (2019): Linked-Employer-Employee-Daten des IAB: Linked Employer-Employee Data from the IAB: LIAB Cross-Sectional Model 2 (LIAB QM2) 1993–2017. FDZ-Datenreport, 06/2019 (en), Nuremberg. DOI: 10.5164/IAB.FDZD.1906.en.v1

[31] Song, Jae, David Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter (2019). “Firming up inequality.” The Quarterly journal of economics 134, no. 1: 1-50.

[32] Spermann, Alexander (2011). “The New Role of Temporary Agency Work in Germany.” IZA Discussion Papers 6180.

[33] Verhoogen, Eric (2008). “Trade, quality upgrading, and wage inequality in the Mexican manufacturing sector.” The Quarterly Journal of Economics 123, no. 2: 489-530.

LIAB Linked-Employer-Employee-Data of the IAB. This study uses the LIAB cross-sectional model 2, version 1993-2017, of the Linked-Employer-Employee Data (LIAB) from the IAB. Data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and subsequently through remote data access. AKM wage estimates also provided by IAB, following the methodology of Card et al. (2013) and as detailed in Bellman et al. (2020). DOI: 10.5164/IAB.LIABQM29317.de.en.v1

BIBB/BAuA-Employment Survey 2006. Appendix results make use of task data from the 2006 BIBB/BAuA employment survey, made available by GESIS as a scientific use file. These data were aggregated by 3-digit KLDB 1988 occupation and subsequently merged with LIAB. DOI: 10.4232/1.11072

A Empirical appendix

A.1 Data description

	1993-1999	1998-2004	2003-2010	2010-2017
Observations	10,681,636	9,228,262	9,558,801	7,109,955
Persons	3,383,086	3,336,966	3,129,852	2,366,668
Establishments	11,934	24,545	26,981	23,031
Identified sample				
Observations	10,645,792	9,185,430	9,511,130	7,080,695
Persons	3,351,623	3,301,958	3,097,052	2,347,597
Establishments	8,151	18,518	19,989	17,684
% weighted sample	97.8	97.1	97.0	97.7
Wage standard deviation	.411	.448	.483	.482

Table A.1: LIAB summary statistics

Table notes. Sample consists of full-time employees (male and female) aged 20-60 in surveyed West German establishments. Person/establishment fixed effects estimated by IAB on the underlying administrative dataset following Card, Heining, Kline (2013). Daily wage in log 1995 euros, censored values imputed.

A.2 AKM wage variance decomposition

	1993-1999	1998-2004	2003-2010	2010-2017
Sample characteristics				
Persons (estimation)	32,645,910	30,598,327	29,865,417	30,787,607
Persons (LIAB)	3,351,623	3,301,958	3,097,052	2,347,597
Establishments (estimation)	2,543,452	2,537,182	2,476,096	2,103,298
Establishments (LIAB)	8,151	18,518	19,989	17,684
Variance components (LIAB)				
Std. dev. daily wage	.411	.448	.483	.482
Std. dev. person effects	.330	.352	.374	.376
Std. dev. estab effects	.176	.195	.228	.200
Std. dev. time-varying effects	.062	.073	.073	.114
Std. dev. residual	.113	.122	.126	.136
Corr. person/establishment	.138	.166	.147	.225
Corr. person/time-varying	.054	.008	.007	-.211
Corr. estab/time-varying	.064	.056	.072	.018

Table A.2: AKM covariance components

Table notes. Estimation sample consists of full-time and apprentice workers aged 20-60 (see Bellman et al. 2020 for details). LIAB sample excludes apprentices and East German establishments. Daily wage in log 1995 euros, censored values imputed. Time-varying effects not provided, so I obtain them by regressing log wage net of fixed effects on year dummies and squared and cubic age terms, each interacted with education.

	1993-1999	1998-2004	2003-2010	2010-2017
% sample unidentified				
Fulltime	.022	.029	.030	.023
Fulltime male	.017	.019	.022	.017
Fulltime female	.034	.048	.049	.037
1-24 employees	.067	.082	.089	.071
25+ employees	.003	.004	.005	.003
Job exit rates				
All	.179	.177	.164	.176
Fulltime	.172	.168	.152	.166
Job entry rates				
All	.166	.167	.152	.173
Fulltime	.151	.150	.133	.154

Table A.3: AKM identification statistics

Table notes. Entry (exit) rates are calculated as percent of workers that enter (exit) the establishment in the previous (following) year. 'All workers' includes apprentices and those in part-time and marginal employment.

A.3 Skill demand and wage sorting

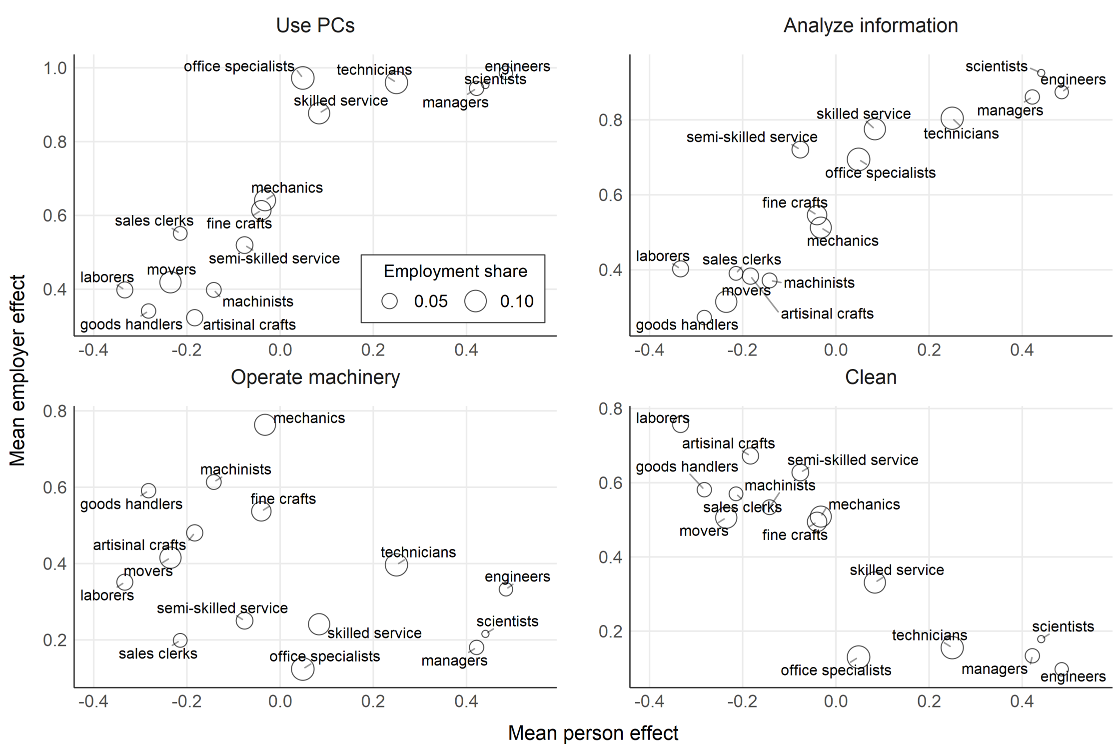

Figure A.4: Occupational person effects and job tasks, 2006

Figure notes. Person effect is the labor-weighted mean AKM person effect. Job tasks are obtained from the 2005-2006 BIBB employment survey, and imputed by 3-digit occupation, educational attainment, and the interaction of 1-digit occupation and educational attainment. Task values represent mean occupation values after merging task data with the LIAB.

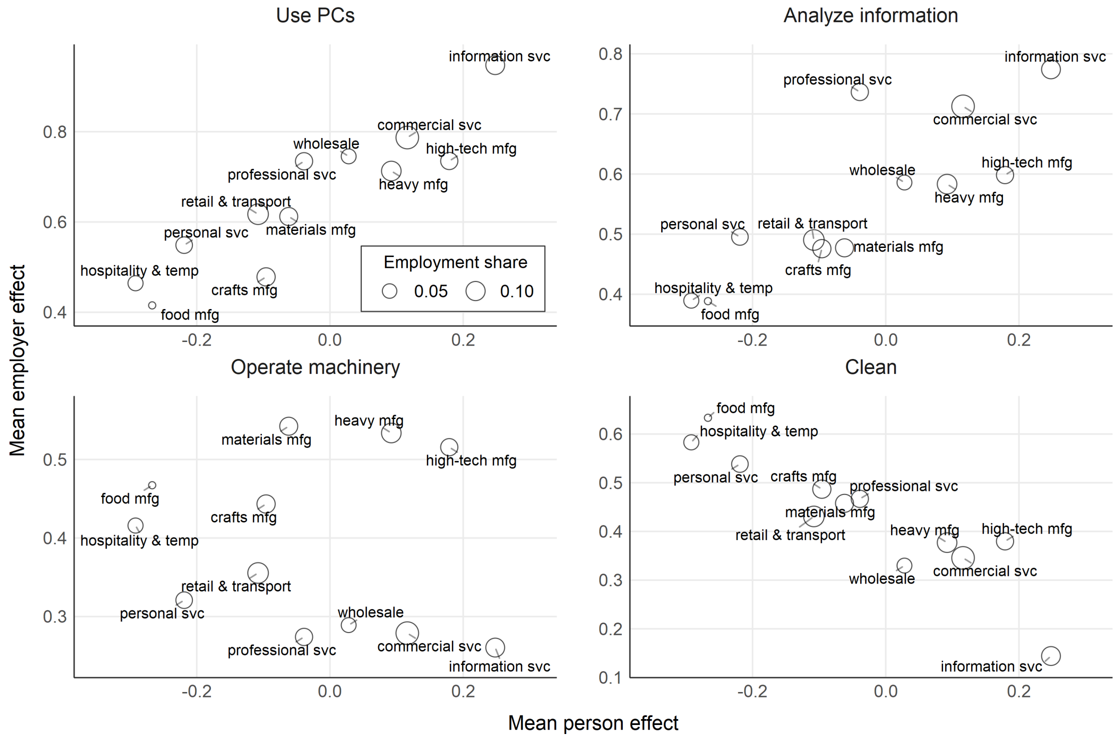

Figure A.5: Industry person effects and job tasks, 2006

FIGURE NOTES. Person effect is the labor-weighted mean AKM person effect. Job tasks are obtained from the 2005-2006 BIBB employment survey, and imputed by 3-digit occupation, educational attainment, and the interaction of 1-digit occupation and educational attainment. Task values represent mean industry values after merging task data with the LIAB.

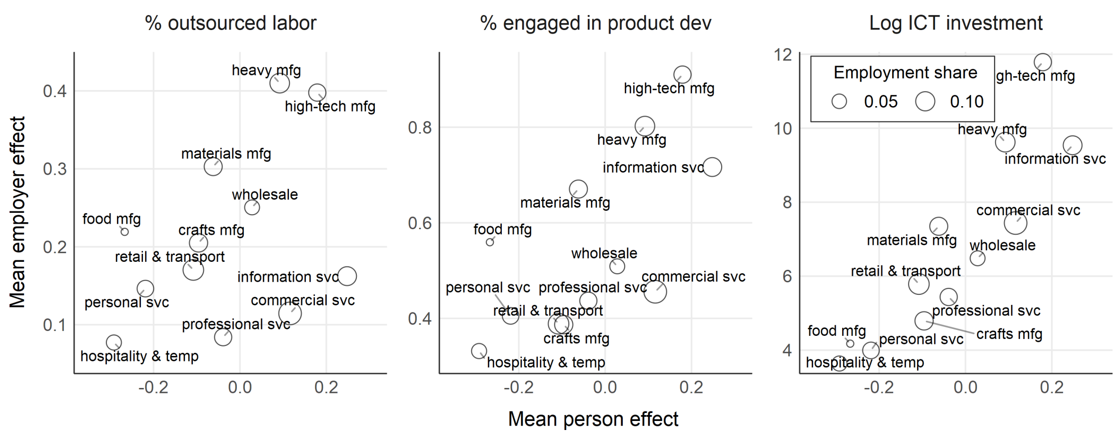

Figure A.6: Industry person effects and skill-related technologies, 2006

FIGURE NOTES. Person effect is the labor-weighted mean AKM person effect. Technology values are obtained from the IAB panel and averaged by industry; missing ICT investment values are first imputed over the 2000-2007 period by regressing log investment on log employment, a dummy variable representing 4 employment size categories, and dummies for year and 45 industry groups.

A.4 Between-group variance decompositions

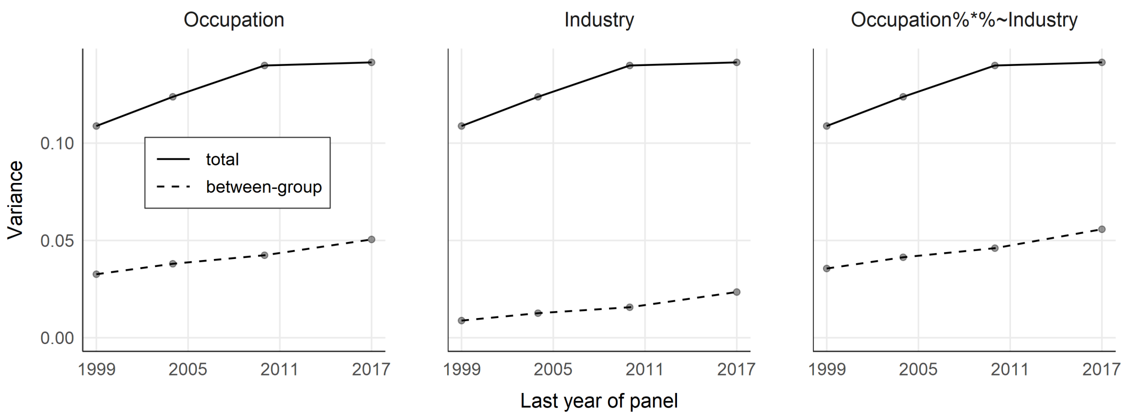

Figure A.7: Between-group variance of person effect

FIGURE NOTES. Between-occupation component of $Var(\alpha_i)$ for 15 KLDB 1988 occupational groups, 12 WZ2008 industry groups, and 180 occupation × industry groups.

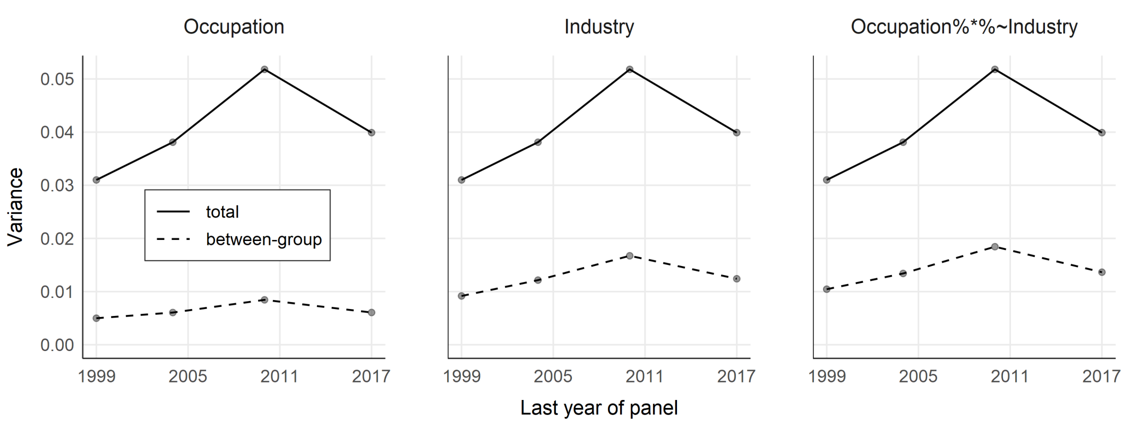

Figure A.8: Between-group variance of employer effect

FIGURE NOTES. Between-occupation component of $Var(\psi_{J(i,t)})$ for 15 KLDB 1988 occupational groups, 12 WZ2008 industry groups, and 180 occupation × industry groups.

	1993-1999	1998-2004	2003-2010	2010-2017
Person characteristics				
Occupation (75 groups)	94.5	84.2	98.4	73.2
Occupation (15 groups)	88.6	80.7	96.0	75.2
Education (5 groups)	23.8	27.2	35.2	35.5
Age (4 groups)	11.2	7.0	8.8	5.3
Gender	43.8	28.1	28.0	9.5
Estab characteristics				
Industry (46 groups)	88.6	85.1	101.6	84.0
Industry (12 groups)	84.8	82.5	100.0	82.8
No. employees (4 groups)	65.8	57.0	64.8	43.2
Sales revenue (4 groups)	82.0	75.6	109.2	69.1
Age (4 groups)	16.3	8.8	13.6	14.8
Mixed characteristics				
Ind (46) $\times$ occ (75)	119.0	108.8	130.6	107.7
Ind (12) $\times$ occ (15)	116.5	106.1	129.0	103.6
Federal state (10)	6.3	3.5	4.8	3.0

Table A.9: Between-group wage effects covariance

Table notes. Between-group component of covariance, as % of total. Number of groups provided in parentheses; see main text for details on grouping variables.

A.5 Aggregated industry and occupation classifications

Industry and occupational classifications are based off of the WZ 2008 and KLDB 1988 classifications, respectively. For years prior to 2008, the data provider imputes industry using extrapolation when possible and, when not, correspondence tables. This is primarily an issue for observations prior to 1999, which were recorded with a much older classification system (WZ 1973). Occupation codes after 2010 are recorded as KLDB 2010 values, for which there exists no direct correspondence with the earlier system. The data provider imputes KLDB 1988 occupation for these years but with substantial inaccuracies; comparing matches observed both before and after 2010, roughly 1/3 experience a change in occupational code, generally involving a 'move' to a similar occupation or from a detailed occupation to a broad N.E.C. category.

To minimize errors from recoding, I conduct the empirical analysis in this paper using aggregated industry and occupational groups. Doing so introduces a second problem: the KLDB 1988 system lacks hierarchical categories, and aggregate WZ 2008 classifications (which follow the NACE system) do an unsatisfactory job of preserving industry wage differentials. In both cases I find it preferable to aggregate classifications manually by combining neighboring industries (occupations) that exhibit similar mean establishment (person) wage effects. This is necessarily an arbitrary approach, but it is successful in producing industry and occupation classifications that are comparable over time while preserving as much as possible of the underlying wage structure.

Industry	Mean person wage effect	Mean estab wage effect	WZ 2008 industry codes
Food mfg.	4.38	-.15	11-32, 161, 101-103, 107
Crafts mfg.	4.53	.01	310-332, 370-439
Materials mfg.	4.54	.04	104-106, 108-152, 162-182, 221-239, 251, 255-259, 292, 331-332, 370-390, 411-429,
Durables mfg.	4.67	.12	51-99, 241-245, 252-254, 264-275, 281-289
High-tech mfg.	4.71	.18	191-212, 261-263, 279, 290-291, 293-309
Wholesale	4.62	.01	461-469
Retail/transport	4.52	-.05	451-454, 471-532
Hospitality/temp	4.34	-.34	561, 563, 781-783
Personal svc.	4.40	-.18	472-473, 476-478, 551-559, 562, 801-822, 829, 920, 931-932, 960
Professional svc.	4.53	-.13	691-692, 741-774, 791-822, 829, 855-856, 862-889, 951-952
Commercial svc.	4.68	.00	661-683, 711, 731, 823, 841-854, 860-861, 900, 910, 941-949
Information svc.	4.82	.09	581-653, 701-702, 712-722, 732

Table A.10: Aggregate industry classifications

Table notes. Mean wage effects calculated first by panel, and then averaged across the four panels.

Occupation	Mean person wage effect	Mean estab wage effect	KLDB 1988 occupation codes
Goods handlers	4.35	-.02	111, 121, 143-152, 181-184, 212, 242, 321-346, 321, 331-344, 361-376, 402-403, 412, 432-433, 492, 531
Machinists	4.47	.07	82-101, 112, 131-135, 161-162, 164, 231-241, 322-323, 422-432, 441-442, 461-466,
Mechanics	4.57	.10	71-72, 141-142, 191-211, 213-226, 263, 273-274, 284-291, 312-314, 521, 541-549, 711
Artisans	4.47	-.05	11-21, 51, 163, 175-177, 391-401, 451-453, 470-491, 501-513, 804, 834
Fine crafts	4.59	-.01	102, 171-174, 251-262, 270-271, 275-282, 301-311, 315, 624, 631, 634-635, 713-716, 725
Unskilled labor	4.28	-.21	41-44, 53, 351-358, 411, 792, 856, 901-902, 912-921, 923-937
Movers	4.43	-.03	522, 712, 714, 723-724, 731-732, 741-744,
Semi-skilled service	4.52	-.07	791, 793-801, 805, 814, 851-852, 854-855, 861, 864, 911
Skilled service	4.67	-.02	681, 683, 701-705, 753, 772, 782-783, 811, 822-833, 835-838, 842-844, 853, 857, 862-863, 875-877, 891-893,
Sales clerks	4.39	-.10	682, 684-686, 688, 706, 733-734, 773, 784,
Office specialists	4.61	.00	781
Technicians	4.84	.09	32, 52, 61, 283, 621-623, 625-629, 632-633 691-694, 721-722, 726, 771, 774, 802-803, 922
Engineers	5.07	.11	601-612
Managers	5.03	.05	687, 751-752, 761-763
Doctors	5.04	.02	813, 821, 841, 871-874, 881-883,

Table A.11: Aggregate occupation classifications

TABLE NOTES. Mean wage effects calculated first by panel, and then averaged across the four panels.

B Theoretical appendix

B.1 Unemployment insurance and efficiency

In this section I take as fixed worker assignment and suppress j-notation. In addition I impose the condition $\psi = 1$, which simplifies but does not substantially affect the analysis. Define $U(s) = \rho U(i, s)$.

Writing the worker's flow value of unemployment as

$$\begin{aligned}\rho U(i,s) &= \frac{(\rho+\delta)\frac{B_{i,s}}{w(i,s)} + \zeta\theta(i,s)^{\eta}}{\rho+\delta+\zeta\theta(i,s)^{\eta}} w(i,s) \\ &\equiv G(i,s)w(i,s)\end{aligned}$$

and defining the elasticity of G with respect to θ as $\epsilon_G(i,s)$ (which may implicitly depend on $\theta(i,s)$), we can show that total (i,s)-vacancies satisfy

$$V_i(j) = \frac{\alpha_i(s)^\sigma \left[m(s)\zeta\theta(i,s)^{\eta-1}\right]^{\sigma-1}\left[\frac{1}{C_i(s)}\frac{\epsilon_G(i,s)}{\epsilon_G(i,s)+\psi(1-\eta)}\right]^\sigma}{\left(\int \alpha_i(k)\left[\alpha_i(s)m(s)\zeta\theta(i,k)^{\eta-1}\frac{1}{C_i(k)}\frac{\epsilon_G(i,s)}{\epsilon_G(i,s)+\psi(1-\eta)}\right]^{\sigma-1}dk\right)^{\frac{\sigma}{\sigma-1}}} Y_i^*$$

Fixing any two labor types s and s', the vacancy ratio $\frac{V_i(s')}{V_i(s)}$ will be equal to

$$\frac{V_i(s')}{V_i(s)} = \left(\frac{\alpha_i(s')C_i(s)}{\alpha_i(s)C_i(s')}\right)^\sigma \left(\frac{\left[\theta(i,s')^{\eta-1}\right]^{\sigma-1}\left[\frac{\epsilon_G(i,s')}{\epsilon_G(i,s)+\psi(1-\eta)}\right]^\sigma}{\left[\theta(i,s)^{\eta-1}\right]^{\sigma-1}\left[\frac{\epsilon_G(i,s)}{\epsilon_G(i,s)+\psi(1-\eta)}\right]^\sigma}\right)$$

Efficiency requires that the term in parenthesis be independent of i, which will only be the case when $\epsilon_G(i,s)$ is a constant.

To see this, consider a central planner that directly allocates vacancies and job applicants across submarkets in order to maximize the utility of the representative household. I simplify by assuming $C_i(j) = C_i$. Writing the planner's Bellman equation:

$$\begin{aligned}\rho V(N,Y) = &\left(\sum_i \beta_i Y_i^{\frac{\tau-1}{\tau}}\right)^{\frac{\tau}{\tau-1}} + \\ &\max_{\phi,v}\left[\left(\delta[\kappa(s)-N(s)] - \sum_i \zeta v(i,s)^\eta \big(N(s)\phi(i,s)\big)^{1-\eta}\right)\frac{\partial V}{\partial N(s)} - \sum_i \int C_i v(i,s) + \right. \\ &\left.\left(\left[\int \alpha_i(s)\left(\zeta v(i,s)^\eta \big(N(s)\phi(i,s)\big)^{1-\eta} m(s)\right)^{\frac{\sigma}{\sigma-1}} ds\right]^{\frac{\sigma-1}{\sigma}} - \delta Y(i)\right)\frac{\partial V}{\partial Y(i)}\right]\end{aligned}$$

As the planner is free to substitute workers across i, we can define a shadow price $\omega(s)$, and by taking the first-order conditions of the planner's problem we may derive the following condition for market tightness:

$$\theta^P(i,s) = \frac{\eta}{1-\eta}\frac{\omega(s)}{C_i N(s)}$$

At the planner's solution, market tightness is multiplicatively separable in industry and worker type. But in the market equilibrium, θ will only be separable if ϵ_G is a constant. More explicitly, if we define $H(\theta) = \frac{G(i,s)}{\zeta\theta^{\eta-1}\epsilon_G(i,s)}$ we can show that

$$\theta^M(i,s) = H^{-1}\left(\frac{U(s)}{(1-\eta)(\rho+\delta)C_i}\right)$$

For this to be separable it must be that H has a constant elasticity with respect to market tightness, which in turn implies the same for G and hence that ϵ_G is constant.

Supposing then that $\epsilon_G(i,s) = D$ for some constant D, the market equilibrium will result in market tightness equal to

$$\theta^M(i,s) = \frac{\sum_k V_k^M(s) C_k^{\frac{1}{D+1-\eta}}}{N(s) C_i^{\frac{1}{D+1-\eta}}}$$

and search probabilities will take the form

$$\frac{\phi^M(i',s)}{\phi^M(i,s)} = \frac{V_{i'}^M(s) C_{i'}^{\frac{1}{D+1-\eta}}}{V_i^M(s) C_i^{\frac{1}{D+1-\eta}}}$$

But from the planner's problem we have that

$$\frac{\phi^P(i',s)}{\phi^P(i,s)} = \frac{V_{i'}^P(s) C_{i'}}{V_i^P(s) C_i}$$

Therefore we must have $D = \eta$ at the efficient allocation. It follows that $G(i,s) = \theta(i,s)^\eta$, and as the proof is not materially affected by setting $\psi < 1$, it is then straightforward to show that $B_{i,s}$ takes the form given in assumption 1.

B.2 Occupational assignment

I begin with the following notation:

- let $\omega(s)$ denote the set of jobs chosen by skill type s in at least one industry: $\omega(s) = \{j \mid \exists\, i \; s.t. \; \phi^*(i,j,s) > 0\}$.
- let $\sigma(j)$ denote the set of skill types s that choose job j in at least one industry: $\sigma(j) = \{s \mid \exists\, i \; s.t. \; \phi^*(i,j,s) > 0\}$.

The initial part of the proof follows Costinot and Vogel (2010), with minor differences due to the presence of firm heterogeneity and search frictions.

Lemma 1. *There exists a continuous and strictly increasing function $\lambda : [\underline{s}, \overline{s}] \to [0,1]$, independent of i, such that $\phi^*(i,j,s) > 0$ if and only if $\lambda(s) = j$, and where $\lambda(\underline{s}) = 0$ and $\lambda(\overline{s}) = 1$..*

Proof. That $\omega(s)$ is non-empty follows from $N(s) > 0$, which will be true if (1) $\delta > 0$ and $\kappa(s) > 0$ and (2) B_{np} is sufficiently small that $F(-B_{np}) > 0$, which I assume. It will then certainly be the case that $F(U(s) - B_{np}) > 0$, since $\rho U(i,j,s) = \zeta \theta^\eta (wm)^\psi A^{1-\psi} > 0$ for strictly positive prices, and we will have $\phi^*(i,j,s) > 0$ for at least one (i,j) pair.

Regarding non-emptiness of $\sigma(j)$ suppose that $\sigma(j)$ is empty for some j. By assumption we have $\alpha_i(j) > 0$, and so it must be that if $\omega(s)$ is non-empty and prices are strictly positive, then $V_i(j) > 0$ for all i and j provided that firms rationally expect that $\phi(i,j,s) > 0$ for at least one s. At the same time $\omega(s)$ non-empty implies that for any s, there exists at least one i' and j' for which $\phi^*(i',j',s) > 0$. Now if $\phi^*(i,j,s) = 0$ for all i

and all s, then $\frac{\int \phi^*(i,j,s)N(s)ds}{\int \phi^*(i',j',s)N(s)ds} = 0$. However, defining $\bar{m}_i(j) > 0$ to be a firm's expected worker productivity, from workers' first-order condition we have that $\frac{\int \phi^*(i,j,s)N(s)ds}{\int \phi^*(i',j',s)N(s)ds} = \frac{V_i(j)\left[(C_i m(j,s)/\bar{m}_i(j))^{\psi} A_i^{1-\psi}\right]^{\frac{1}{\eta+\psi(1-\eta)}}}{V_{i'}(j')\left[(C_{i'} m(j',s)/\bar{m}(i,j'))^{\psi} A_{i'}^{1-\psi}\right]^{\frac{1}{\eta+\psi(1-\eta)}}} > 0$, a contradiction.

Third, $\sigma(j)$ is non-decreasing. Suppose otherwise. From the worker's first-order condition, we must have

$$\begin{aligned}
0 &\geq \zeta^{1-\psi}\theta(i,j)^{\eta+\psi(1-\eta)}\left(\frac{1-\eta}{\eta}(\rho+\delta)\frac{C_i(j)}{\bar{m}_i(j)}\right)^{\psi} A_i^{1-\psi} m(j,s)^{\psi} - U(s) \\
&\equiv R(i,j)m(j,s)^{\psi} - U(s)
\end{aligned}$$

with equality if $\phi^*(i,j,s) > 0$. Supposing that there exist two industries i and i', two jobs $j^+ > j^-$, and two worker types $s^+ > s^-$ such that $\phi^*(i,j^+,s^-) > 0$ and $\phi^*(i',j^-,s^+) > 0$, then we must have

$$\begin{aligned}
0 &= R(i,j^-)m(j^-,s^+)^{\psi} - U(s^+) \\
&\geq R(i,j^-)m(j^-,s^-)^{\psi} - U(s^-) \\
&= \frac{m(j^-,s^-)^{\psi}}{m(j^-,s^+)^{\psi}}U(s^+) - U(s^-) \\
&> \frac{m(j^+,s^-)^{\psi}}{m(j^+,s^+)^{\psi}}U(s^+) - U(s^-) \\
&= -\frac{U(s^-)}{m(j^+,s^+)^{\psi}R(i',j^+)}\Big(R(i',j^+)m(j^+,s^+)^{\psi} - U(s^+)\Big) \\
&\geq 0
\end{aligned}$$

a contradiction for $\psi > 0$. Note that the result holds across industries due to assumption 1, without which there would be a non-separable term depending on (i,j,s).

Fourth, ω and σ are single-valued almost everywhere. The proof is unchanged from CV and so I provide only the intuition: if ω (or σ) has positive measure over a domain with positive measure, then from the previous result the range of the correspondence will have measure greater than the measure of $[0,1]$ (or $[\underline{s},\bar{s}]$), a contradiction.

Fifth, $\sigma(j)$ is single-valued. If this is not the case, then from step 3 there exists a non-degenerate interval $[s,s']$ in which all workers choose job j. Step 4 implies that there exists another job j' that is chosen by a single worker type. From the firm's first-order condition and market clearing it must be that

$$\sum_i \Big(\int \phi^*(i,j,s)N(s)ds\Big)\zeta\theta(i,j)^{\eta} = \sum_i \frac{\left[\alpha_i(j)\frac{\zeta\theta(i,j)^{\eta-1}}{C_i(j)}\right]^{\sigma}\left[\bar{m}_i(j)\right]^{\sigma-1}}{\left(\int \alpha_i(k)\left[\alpha_i(k)\frac{\zeta\theta(i,k)^{\eta-1}}{C_i(k)}\bar{m}(i,k)\right]^{\sigma-1}dk\right)^{\frac{\sigma}{\sigma-1}}}Y_i^*$$

But then

$$\frac{\sum_i \frac{\left[\alpha_i(j')\frac{\zeta\theta(i,j')^{\eta-1}}{C_i(j')}\right]^{\sigma}\left[\bar{m}(i,j')\right]^{\sigma-1}}{\left(\int \alpha_i(k)\left[\alpha_i(k)\frac{\zeta\theta(i,k)^{\eta-1}}{C_i(k)}\bar{m}(i,k)\right]^{\sigma-1}dk\right)^{\frac{\sigma}{\sigma-1}}}}{\sum_i \frac{\alpha_i(j)\left[\alpha_i(j)\frac{\zeta\theta(i,j)^{\eta-1}}{C_i(j)}\right]^{\sigma}\left[\bar{m}_i(j)\right]^{\sigma-1}}{\left(\int \alpha_i(k)\left[\alpha_i(k)\frac{\zeta\theta(i,k)^{\eta-1}}{C_i(k)}\bar{m}(i,k)\right]^{\sigma-1}dk\right)^{\frac{\sigma}{\sigma-1}}}} = 0$$

which violates the assumptions that $\alpha(i,j)$ is strictly positive and continuous (and therefore finite), that $m(j,s) > 0$, and that $C_i(j)$ is finite.

From the last step we have $\sigma(j)$ single-valued; from the third step, weakly increasing; from the first step, continuous and such that $\sigma(0) = \underline{s}$ and $\sigma(1) = \overline{s}$; and from the fourth step, σ is strictly increasing. Hence we have a continuous, strictly increasing bijection $\lambda(s) = \omega(s) = \{j \mid \exists\ i\ s.t.\ \phi^*(i,j,s) = 1\} = \sigma^{-1}(s)$. □

Lemma 2. *Reservation values satisfy the equation*

$$\frac{d\log U(s)}{ds} = \psi\frac{m_s\big(s,\lambda(s)\big)}{m\big(s,\lambda(s)\big)}$$

Proof. From lemma 1, for any $\phi^*(i,j,s) > 0$ we must have

$$0 \geq R(i,j)m(j,s)^{\psi} - U(s)$$

and following CV the following two inequalities must hold:

$$\begin{aligned} R(i,\lambda(s)) - \frac{U(s)}{m(\lambda(s),s)^{\psi}} &\geq R(i,\lambda(s)) - \frac{U(s+ds)}{m(\lambda(s),s+ds)^{\psi}} \\ R(i,\lambda(s+ds)) - \frac{U(s+ds)}{m(\lambda(s+ds),s+ds)^{\psi}} &\geq R(i,\lambda(s+ds)) - \frac{U(s)}{m(\lambda(s+ds),s)^{\psi}} \end{aligned}$$

and therefore

$$\begin{aligned} R(i,\lambda(s))\Big[m(\lambda(s),s+ds)^{\psi} - m(\lambda(s),s)^{\psi}\Big] &\leq U(s+ds) - U(s) \\ &\leq R(i,\lambda(s+ds))\Big[m(\lambda(s+ds),s+ds)^{\psi} - m(\lambda(s+ds),s)^{\psi}\Big] \end{aligned}$$

Regarding continuity in s: U is continuous given continuity of θ, which in turn follows from continuity of α, $\kappa(s)$, and λ. If θ is continuous then continuity of $R(i,j)$ also follows, and hence we can divide by ds and take the limit of the previous inequalities to show that

$$\begin{aligned} U'(s) &= R(i,\lambda(s))\psi m_s(\lambda(s),s)m(\lambda(s),s)^{\psi-1} \\ &= U(s)\psi\frac{m_s(\lambda(s),s)}{m(\lambda(s),s)} \end{aligned}$$

and hence

$$\frac{d \log U(s)}{ds} = \psi \frac{m_s\big(s, \lambda(s)\big)}{m\big(s, \lambda(s)\big)}$$

and the result follows. ☐

Lemma 3. *The matching function satisfies*

$$\frac{d\lambda(s)}{ds} = \frac{m\big(\lambda(s), s\big) U(s) N(s)}{\sum_i \left(\frac{\psi(1-\eta)}{\eta+\psi(1-\eta)} m\big(\lambda(s), s\big) p_i\big(\lambda(s)\big)\right)^{\psi} A_i^{1-\psi} \frac{\left[\alpha_i\big(\lambda(s)\big)/p_i\big(\lambda(s)\big)\right]^{\sigma}}{\left(\int \frac{\alpha_i\left(\lambda(k)\right)^{\sigma}}{p_i\left(\lambda(k)\right)^{\sigma-1}} dk\right)^{\frac{\sigma}{\sigma-1}}} \frac{\left[\beta_i/P_i\right]^{\tau}}{\left(\sum_k \frac{\beta_k^{\tau}}{P_k^{\tau-1}}\right)^{\frac{\tau}{\tau-1}}} Y}$$

where $\lambda(\underline{s}) = 0$ *and* $\lambda(\overline{s}) = 1$.

Proof. Total demand for j-output is given by

$$\sum_i y_i^D(j) = \sum_i y_i(j)$$

while from market-clearing and lemma 1 we have

$$y_j^S(j) = \sum_i m\big(\lambda(s), s\big) \phi^*(i, j, s) \zeta \theta\big(i, \lambda(s)\big)^{\eta} N(s) \delta\big[j - \lambda(s)\big]$$

with δ the Dirac function. Combining the two equations:

$$\sum_i y_i^D(j) = \int m\big(j, \lambda^{-1}(j')\big) \sum_i \left[\phi^*\big(i, j, \lambda^{-1}(j')\big) \zeta \theta(i, j)^{\eta} N\big(\lambda^{-1}(j')\big)\right] \delta\big[\lambda(s) - j'\big] \frac{1}{\lambda'\big(\lambda^{-1}(j')\big)} dj'$$

This simplifies to

$$\lambda'(s) = \frac{m\big(\lambda(s), s\big) N(s) \sum_i \phi^*\big(i, \lambda(s), s\big) \zeta \theta\big(i, \lambda(s)\big)^{\eta}}{\sum_i y_i\big(\lambda(s)\big)}$$

where

$$\phi^*(i, \lambda(s), s) = \frac{\frac{y_i\left(\lambda(s)\right)}{m\left(\lambda(s), s\right)} \left[m\big(\lambda(s), s\big) p_i\big(\lambda(s)\big)\right]^{\psi} A_i^{1-\psi}}{\sum_k \frac{y_k\left(\lambda(s)\right)}{m\left(\lambda(s), s\right)} \left[m\big(\lambda(s), s\big) p_k\big(\lambda(s)\big)\right]^{\psi} A_k^{1-\psi}}$$

$$\theta\big(i, \lambda(s)\big)) = \left[\frac{U(s)}{\zeta \left(\frac{\psi(1-\eta)}{\eta+\psi(1-\eta)} m\big(\lambda(s), s\big) p_i\big(\lambda(s)\big)\right)^{\psi} A_i^{1-\psi}}\right]^{\frac{1}{\eta}}$$

$$y_i\big(\lambda(s)\big) = \frac{\left(\frac{\alpha_i\left(\lambda(s)\right)}{p_i\left(\lambda(s)\right)}\right)^{\sigma}}{\left[\int \alpha_i\big(\lambda(k)\big)\left(\frac{\alpha_i\left(\lambda(k)\right)}{p_i\left(\lambda(k)\right)}\right)^{\sigma-1} dk\right]^{\frac{\sigma}{\sigma-1}}} \frac{\left(\frac{\beta_i}{P_i}\right)^{\tau}}{\left[\sum_k \beta_k \left(\frac{\beta_k}{P_k}\right)^{\tau-1}\right]^{\frac{\tau}{\tau-1}}} Y$$

Substitution gives us

$$\frac{d\lambda(s)}{ds} = \frac{m\big(\lambda(s),s\big)U(s)N(s)}{\sum_i \left(\frac{\psi(1-\eta)}{\eta+\psi(1-\eta)} m\big(\lambda(s),s\big)p_i\big(\lambda(s)\big)\right)^{\psi} A_i^{1-\psi} \frac{\left[\alpha_i\left(\lambda(s)\right)/p_i\left(\lambda(s)\right)\right]^{\sigma}}{\left(\int \frac{\alpha_i\left(\lambda(k)\right)^{\sigma}}{p_i\left(\lambda(k)\right)^{\sigma-1}} dk\right)^{\frac{\sigma}{\sigma-1}}} \frac{\left[\beta_i/P_i\right]^{\tau}}{\left(\sum_k \frac{\beta_k^{\tau}}{P_k^{\tau-1}}\right)^{\frac{\tau}{\tau-1}}} Y}$$

and the result is shown. □

B.3 Equilibrium policy functions

Unrestricted $B_{i,j,s}$. In this section I suppress subscripts and arguments when it is possible to do so without ambiguity. Beginning with (6), it is straightforward to show that workers' search flow value takes the form

$$\begin{aligned} U(s) &\equiv \rho U(i,j,s) \\ &= G(i,j,s)\big[w(i,j)m(j,s)\big]^{\psi} A_i^{1-\psi} \end{aligned}$$

where G is a function of market tightness, wages (inclusive of worker productivity), and $B_{i,j,s}$. From the non-participation decision we have that workers will search whenever $U(s) - B_{np} \geq \chi$, and hence the proportion of unemployed s-workers searching for jobs will be

$$Pr(search) = 1 - F\big(U(s) - B_{np}\big)$$

Firms anticipate worker search behavior when posting vacancies, with the optimal posted wage taking the form

$$w^* = \frac{\psi(1-\eta)}{\epsilon_G(i,j,s) + \psi(1-\eta)} p_i(j)$$

where ϵ_G is the elasticity of $G(i,j,s)$ with respect to market tightness. From free entry we have that

$$p_i(j) = \frac{\epsilon_G(i,j,s) + \psi(1-\eta)}{\epsilon_G(i,j,s)} \frac{(\rho+\delta)C_i(j)}{\zeta\theta(i,j)^{\eta-1}\bar{m}_i(j)}$$

with $\bar{m}_i(j)$ anticipated worker productivity. The first-order condition for industry aggregators gives us the equation

$$y_i(j) = \left(\frac{\alpha_i(j)}{p_i(j)}\right)^{\sigma} P_i^{\sigma} Y_i$$

where the zero-profit condition implies that

$$P_i = \left[\int \alpha_i(k) \left(\frac{\alpha_i(k)}{p_i(k)}\right)^{\sigma-1} dk\right]^{\frac{-1}{\sigma-1}}$$

Finally, utility maximization by the representative household gives us that

$$Y_i = \frac{\left(\beta_i/P_i\right)^{\tau}}{\left[\sum_k \beta_k \left(\beta_i/P_i\right)^{\tau-1}\right]^{\frac{\tau}{\tau-1}}} Y$$

where normalization of prices in terms of household utility implies that

$$\left[\sum_k \beta_k \left(\beta_i/P_i\right)^{\tau-1}\right]^{\frac{-\tau}{\tau-1}} = 1$$

Unless $\epsilon_G(i,j,s)$ is constant, it is not possible to solve analytically for worker search values or market tightness. We can manipulate workers' reservation values to obtain the equation

$$\theta(i,j) = \left(\frac{\epsilon_G(i,j,s)^{\psi}}{G(i,j,s)}\right)^{\frac{1}{\psi(1-\eta)}} \left(\frac{U(s)}{\left[\frac{\psi(1-\eta)(\rho+\delta)C_i(j)}{\zeta} \frac{m(j,s)}{\bar{m}_i(j)}\right]^{\psi} A_i^{1-\psi}}\right)^{\frac{1}{\psi(1-\eta)}}$$

but the first term is in general non-separable, precluding a closed-form solution for $\theta(i,j)$.

With assumption 2. The immediate implication of assumption 2 is that we can write market tightness as

$$\theta(i,j) = \left(\frac{U(s)}{\zeta \left[\frac{\psi(1-\eta)}{\eta+\psi(1-\eta)} \bar{m}_i(j) p_i(j)\right]^{\psi} A_i^{1-\psi}}\right)^{\frac{1}{\eta}}$$

which then allows us to derive workers' relative search probabilities:

$$\frac{\phi^*(i',j',s)}{\phi^*(i,j,s)} = \frac{\frac{p_{i'}(j')y_{i'}(j')}{C_{i'}(j')} \left[\left(p_{i'}(j')m(j',s)\right)^{\psi} A_{i'}^{1-\psi}\right]^{\frac{1}{\eta}}}{\frac{p_i(j)y_i(j)}{C_i(j)} \left[\left(p_i(j)m(j,s)\right)^{\psi} A_i^{1-\psi}\right]^{\frac{1}{\eta}}}$$

Equilibrium wage posting then simplifies to

$$w^*(i,j) = \frac{\psi(1-\eta)}{\eta+\psi(1-\eta)} p_i(j)$$

with the price of an intermediate good given by

$$p_i(j) = \left(\frac{\eta}{\eta+\psi(1-\eta)} \frac{\zeta \bar{m}(i,j)}{(\rho+\delta)C_i(j)} \right)^{-1} \left(\frac{U(s)}{\zeta \left[\frac{\psi(1-\eta)}{\eta+\psi(1-\eta)} \bar{m}_i(j) p_i(j) \right]^{\psi} A_i^{1-\psi}} \right)^{\frac{1-\eta}{\eta}}$$

With optimal assignment. With $\lambda : s \to j$ defined, we will have $\bar{m}(i,j) = m\big(j, \lambda^{-1}(j)\big)$, and prices simplify to

$$p_i\big(\lambda(s)\big) = \frac{\eta+\psi(1-\eta)}{\psi(1-\eta)} \frac{1}{m\big(\lambda(s),s\big)} \left(\frac{\psi(1-\eta)(\rho+\delta)C_i\big(\lambda(s)\big)}{\eta \zeta^{\frac{1}{\eta}} A_i^{\frac{1-\psi)(1-\eta)}{\eta}}} \right)^{\frac{\eta}{\eta+\psi(1-\eta)}} U(s)^{\frac{1-\eta}{\eta+\psi(1-\eta)}}$$

$$P_i = \frac{\eta+\psi(1-\eta)}{\psi(1-\eta)} \left[\frac{\psi(1-\eta)(\rho+\delta)}{\eta \zeta^{\frac{1}{\eta}} A_i^{\frac{(1-\psi)(1-\eta)}{\eta}}} \right]^{\frac{\eta}{\eta+\psi(1-\eta)}} \left(\int \alpha_i\big(\lambda(k)\big)^{\sigma} \left[\frac{m\big(\lambda(k),k\big)}{C_i\big(\lambda(k)\big)^{\frac{\eta}{\eta+\psi(1-\eta)}} U(k)^{\frac{1-\eta}{\eta+\psi(1-\eta)}}} \right]^{\sigma-1} dk \right)^{\frac{-1}{\sigma-1}}$$

All together, the policy functions and market tightness are given by

$$w^*\big(i,\lambda(s)\big) = \left(\frac{\psi(1-\eta)(\rho+\delta)C_i\big(\lambda(s)\big)}{\eta \zeta^{\frac{1}{\eta}} A_i^{\frac{(1-\psi)(1-\eta)}{\eta}}} \right)^{\frac{\eta}{\eta+\psi(1-\eta)}} U(s)^{\frac{1-\eta}{\eta+\psi(1-\eta)}}$$

$$\phi^*(i,j,s) = \frac{\frac{p_i(\lambda(s)) y_i(\lambda(s))}{C_i(\lambda(s))} \left[\big(p_i\big(\lambda(s)\big) m\big(\lambda(s),s\big) \right]^{\psi} A_i^{1-\psi} \Big]^{\frac{1}{\eta}}}{\sum_k \sum_m \frac{p_m(\lambda(k)) y_m(\lambda(k))}{C_m(\lambda(k))} \left[\big(p_m\big(\lambda(k)\big) m\big(\lambda(k),k\big) \right]^{\psi} A_m^{1-\psi} \Big]^{\frac{1}{\eta}}}$$

$$y_i^*\big(\lambda(s)\big) = \left(\frac{\alpha_i\big(\lambda(s)\big)}{p_i\big(\lambda(s)\big)} \right)^{\sigma} P_i^{\sigma} Y_i$$

$$Y_i^* = \frac{\big(\beta_i/P_i\big)^{\tau}}{\left[\sum_k \beta_k \big(\beta_i/P_i\big)^{\tau-1} \right]^{\frac{\tau}{\tau-1}}} Y$$

$$\theta(i,j) = \left(\frac{U(s)}{\zeta \left[\frac{\psi(1-\eta)}{\eta+\psi(1-\eta)} m\big(\lambda(s),s\big) p_i\big(\lambda(s)\big) \right]^{\psi} A_i^{1-\psi}} \right)^{\frac{1}{\eta}}$$

and the equilibrium is characterized.

These equations can be rendered somewhat more intuitive by defining the employer and person wage differentials

$$ID_i(j) = \left(\frac{\psi(1-\eta)(\rho+\delta)C_i(j)}{\eta \zeta^{\frac{1}{\eta}} A_i^{\frac{(1-\psi)(1-\eta)}{\eta}}} \right)^{\frac{\eta}{\eta+\psi(1-\eta)}}$$

$$\begin{aligned} PD(s) &= U(s)^{\frac{1-\eta}{\eta+\psi(1-\eta)}} \\ &= \left(\frac{\frac{\psi(1-\eta)}{\eta+\psi(1-\eta)} \sum_m \alpha_m\big(\lambda(s)\big)^\sigma \beta_m^\tau ID_m\big(\lambda(s)\big)^{\psi-\sigma} A_m^{1-\psi} \left(\int \alpha_m\big(\lambda(k)\big)^\sigma \left[\frac{m\big(\lambda(k),k\big)}{ID_m\big(\lambda(k)\big)PD(k)} \right]^{\sigma-1} dk \right)^{\frac{\tau-\sigma}{\sigma-1}} Y}{N(s)} \right)^{1-\eta} \end{aligned}$$

With these definitions we can re-write prices as

$$\begin{aligned} p_i\big(\lambda(s)\big) &= \frac{\eta+\psi(1-\eta)}{\psi(1-\eta)} \left(\frac{m\big(\lambda(s),s\big)}{ID_i\big(\lambda(s)\big)PD(s)} \right)^{-1} \\ P_i &= \frac{\eta+\psi(1-\eta)}{\psi(1-\eta)} \left(\int \alpha_i\big(\lambda(k)\big)^\sigma \left[\frac{m\big(\lambda(k),k\big)}{ID_i\big(\lambda(k)\big)PD(k)} \right]^{\sigma-1} dk \right)^{\frac{-1}{\sigma-1}} \end{aligned}$$

and policies can be written more or less succinctly as functions of model primitives, the solutions λ and U to the system of differential equations, and the wage differentials ID and PD which are themselves functions of primitives and U:

$$\begin{aligned} w^*\big(i,\lambda(s)\big) &= ID_i\big(\lambda(s)\big)PD(s) \\ \phi^*(i,j,s) &= \frac{\alpha_i\big(\lambda(s)\big)^\sigma \beta_i^\tau ID_i\big(\lambda(s)\big)^{\psi-\sigma} A_i^{1-\psi} \left(\int \alpha_i\big(\lambda(k)\big)^\sigma \left[\frac{m\big(\lambda(k),k\big)}{ID_i\big(\lambda(k)\big)PD(k)} \right]^{\sigma-1} dk \right)^{\frac{\tau-\sigma}{\sigma-1}}}{\sum_m \alpha_m\big(\lambda(s)\big)^\sigma \beta_m^\tau ID_m\big(\lambda(s)\big)^{\psi-\sigma} A_m^{1-\psi} \left(\int \alpha_m\big(\lambda(k)\big)^\sigma \left[\frac{m\big(\lambda(k),k\big)}{ID_m\big(\lambda(k)\big)PD(k)} \right]^{\sigma-1} dk \right)^{\frac{\tau-\sigma}{\sigma-1}}} \\ y_i^*\big(\lambda(s)\big) &= \left(\frac{\alpha_i\big(\lambda(s)\big)m\big(\lambda(s),s\big)}{ID_i\big(\lambda(s)\big)PD(s)} \right)^\sigma \left(\int \alpha_i\big(\lambda(k)\big)^\sigma \left[\frac{m\big(\lambda(k),k\big)}{ID_i\big(\lambda(k)\big)PD(k)} \right]^{\sigma-1} dk \right)^{\frac{-\sigma}{\sigma-1}} Y_i \\ Y_i^* &= \left(\beta_i \frac{\psi(1-\eta)}{\eta+\psi(1-\eta)} \right)^\tau \left(\int \alpha_i\big(\lambda(k)\big)^\sigma \left[\frac{m\big(\lambda(k),k\big)}{ID_i\big(\lambda(k)\big)PD(k)} \right]^{\sigma-1} dk \right)^{\frac{\tau}{\sigma-1}} Y \\ \theta(i,j) &= \frac{PD(s)^{1-\eta}}{\zeta^{\frac{1}{\eta}} ID_i\big(\lambda(s)\big)^{\frac{\psi}{\eta}} A_i^{\frac{1-\psi}{\eta}}} \end{aligned}$$

The equilibrium is then fully characterized conditional on the solution to (15)-(16), which must be obtained numerically.

C Quantitative results

C.1 Estimation

Matching parameters η and ν. In principle it is possible to estimate the match function $H = \zeta V^{\eta} N^{1-\eta}$ from annual data on industry vacancy postings, industry hires, and aggregate unemployment in West Germany. With $\sigma = \tau = 1$, total hires in industry i will be

$$M(V,N) = \int \zeta \left(\frac{\delta y_i^*\big(\lambda(s)\big)}{\zeta\theta\big(i\lambda(s)\big)m\big(\lambda(s),s\big)} \right)^{\eta} \big(\phi^*\big(i,\lambda(s),s\big)N(s)\big)^{1-\eta}$$

I do not observe s-unemployment and so I assume that changes over time in N are in equal proportion for all skill types: that is, $N(s) \equiv N(s)U$ where U is aggregate unemployment and $N(s)$ is assumed constant. The previous equation becomes

$$M(V,N) = SV_i U^{1-\eta} A_i^{\frac{(1-\psi)(1-\eta)}{\eta}} \frac{\int \left(\frac{\alpha_i\big(\lambda(s)\big)}{ID_i\big(\lambda(s)\big)PD(s)} \right)^{\sigma} m\big(\lambda(s),s\big)^{\sigma-1} ds}{\int \left(\frac{\alpha_i\big(\lambda(s)\big)}{ID_i\big(\lambda(s)\big)PD(s)} \right)^{\sigma} m\big(\lambda(s),s\big)^{\sigma-1} \frac{PD(s)}{ID_i\big(\lambda(s)\big)^{\frac{\psi(1-\eta)}{\eta}}} ds}$$

where V_i is total i-vacancies. The last term is one over the observed average of the term $PD/ID^{\frac{\psi(1-\eta)}{\eta}}$:

$$\frac{\int L_i\big(\lambda(s)\big) \frac{PD(s)}{ID_i\big(\lambda(s)\big)^{\frac{\psi(1-\eta)}{d}}} s}{\int L_i\big(\lambda(s)\big) ds} = \frac{\int \left(\frac{\alpha_i\big(\lambda(s)\big)}{ID_i\big(\lambda(s)\big)PD(s)} \right)^{\sigma} m\big(\lambda(s),s\big)^{\sigma-1} \frac{PD(s)}{ID_i\big(\lambda(s)\big)^{\frac{\psi(1-\eta)}{\eta}}} ds}{\int \left(\frac{\alpha_i\big(\lambda(s)\big)}{ID_i\big(\lambda(s)\big)PD(s)} \right)^{\sigma} m\big(\lambda(s),s\big)^{\sigma-1} ds}$$

Neither this nor A_i or known, but I assume these terms are invariant in the short-run and may be removed by differencing. This allows for estimation of η while controlling for annual variation in the industry distribution of vacancies. Match efficiency ν can then be obtained from aggregate vacancies, unemployment, and hires, conditional on the estimate of η.

However, estimates obtained in this fashion result in values of η that are implausibly large (in the .6-.7 range) relative to estimates in the literature, likely due to the short length of the time series and noise in the estimates of V and hires. Attenuation bias will tend to result in low estimates of $1-\eta$ and hence, high estimates of η. I therefore use the estimate $\eta = .35$ obtained by Kohlbrecher et al. (2016) from German labor market data. Matching efficiency ζ is then found by substituting all values into the last equation above and taking the average across industries.

Estimation of A. Industry amenities are identified from industry vacancy-filling rates. Solving fort industry hires and vacancies and taking the ratio, we obtain

$$\frac{M(V,N)}{V_i} = SA_i^{\frac{-(1-\psi)(1-\eta)}{\eta}} \frac{\int \left(\frac{\alpha_i\big(\lambda(s)\big)}{ID_i\big(\lambda(s)\big)PD(s)}\right)^{\sigma} m\big(\lambda(s),s\big)^{\sigma-1} ds}{\int \left(\frac{\alpha_i\big(\lambda(s)\big)}{ID_i\big(\lambda(s)\big)PD(s)}\right)^{\sigma} m\big(\lambda(s),s\big)^{\sigma-1} \frac{PD(s)}{ID_i\big(\lambda(s)\big)^{\frac{\psi(1-\eta)}{\eta}}} ds}$$

where S is a constant term that may be ignored. Hence

$$A_i = \left(\frac{M(V_i,N)}{V_i} \frac{\int L_i\big(\lambda(s)\big) \frac{PD(s)}{ID_i\big(\lambda(s)\big)^{\frac{\psi(1-\eta)}{\eta}}} ds}{\int L_i\big(\lambda(s)\big) ds} \right)^{\frac{\eta}{(1-\psi)(1-\eta)}}$$

Amenities are then estimated from the average observed value of $\frac{PD(s)}{ID_i\big(\lambda(s)\big)^{\frac{\psi(1-\eta)}{\eta}}}$ within each industry, together with vacancies and hires.

I note in passing that the amenity share ψ will independently affect only the model-predicted labor share

$$\text{labor share} = \frac{\psi(1-\eta)}{\eta + \psi(1-\eta)}$$

For empirical labor shares and plausible values of η, however, this will result in $\beta > 1$; hence I treat β as a normalization and set it equal to .5.

Estimation of γ, α, and λ. I begin by normalizing the empirical matching function λ in the initial panel and estimating γ, as the two are not separately identified. I set $\lambda(s) = s$ and assume that labor productivity takes the form $\log \gamma = F(s)(\gamma_s + \gamma_{sj}^t j)$ where $F(s)$ is an unknown and time-invariant function, and $G^t(j)$ an occupational productivity shifter that is allowed to depend on time. I impose the assumption that $\gamma_{sj}^1 = 1$, whereas γ_s is only included so as to prevent division by zero in the estimation of $F(s)$, and so I set this parameter to a small value ($\gamma_s = .01$). It is then possible using proposition 1 to estimate $F(s)$ non-parameterically from the empirical distribution of the person wage effect $PD(s)$. Given $F(s)$, the empirical matching function in panels 2-4 may be estimated. The parameter γ_{sj} is fixed in subsequent panels by the condition that $\lambda(1) = 1$.

Production shares $\alpha_i(j)$ are identified from the intra-industry distribution of person effects $PD(s)$:

$$\frac{w\big(i,\lambda(s)\big)L_i\big(\lambda(s)\big)}{\int w\big(i,\lambda(k)\big)L_i\big(\lambda(k)\big)dk} = \frac{\alpha\big(\lambda(s)\big)^{\sigma} \left(\frac{m\big(\lambda(s),s\big)}{ID_i\big(\lambda(s)\big)PD(s)}\right)^{\sigma-1}}{\int \alpha\big(\lambda(k)\big)^{\sigma} \left(\frac{m\big(\lambda(k),s\big)}{ID_i\big(\lambda(k)\big)PD(k)}\right)^{\sigma-1} dk}$$

using empirical wage effects, intra-industry labor shares, and the estimated matching function.

Estimation of β and C. Industry entry costs are obtained by using the free entry condition

$$C_i(j) = \frac{ID_i(j)^{\frac{\eta+\psi(1-\eta)}{\eta}} A_i^{\frac{(1-\psi)(1-\eta)}{\eta}} \eta\zeta^{\frac{1}{\eta}}}{\psi(1-\eta)(\rho+\delta)}$$

Taking the ratio of $C_{i'}(j')$ to $C_i(j)$,

$$\frac{C_{i'}(j')}{C_i(j)} = \left(\frac{ID_{i'}(j')}{ID_i(j)}\right)^{\frac{\eta+\psi(1-\eta)}{\eta}} \left(\frac{A_{i'}}{A_i}\right)^{\frac{(1-\psi)(1-\eta)}{\eta}}$$

which gives us entry costs up to a multiplicative constant, conditional on amenities, employer wage differentials, and the parameters ψ and η.

To obtain industry shares in final good production β_i, we may integrate over wages $w\big(i, \lambda(s)\big)L_i\big(\lambda(s)\big)$ to obtain

$$\int w\big(i, \lambda(s)\big)L_i\big(\lambda(s)\big)ds = \left(\beta_i \frac{\psi(1-\eta)}{\eta+\psi(1-\eta)}\right)^{\tau} \left(\int \alpha_i\big(\lambda(k)\big)^{\sigma} \left[\frac{m\big(\lambda(k), k\big)}{ID_i\big(\lambda(k)\big)PD(k)}\right]^{\sigma-1} dk\right)^{\frac{\tau-1}{\sigma-1}} Y$$

Rewriting this and dividing by $\sum_i \beta_i$ we obtain

$$\beta_i = \frac{\dfrac{\left(\int w\big(i,\lambda(s)\big)L_i\big(\lambda(s)\big)ds\right)^{\frac{1}{\tau}}}{\left(\int \alpha_i\big(\lambda(s)\big)^{\sigma}\left[\frac{m(\lambda(s),s)}{ID_i(\lambda(s))PD(s)}\right]^{\sigma-1} ds\right)^{\frac{\tau-1}{\tau(\sigma-1)}}}}{\sum_m \dfrac{\left(\int w\big(m,\lambda(s)\big)L_m\big(\lambda(s)\big)ds\right)^{\frac{1}{\tau}}}{\left(\int \alpha_M\big(\lambda(s)\big)^{\sigma}\left[\frac{m(\lambda(s),s)}{ID_M(\lambda(s))PD(s)}\right]^{\sigma-1} ds\right)^{\frac{\tau-1}{\tau(\sigma-1)}}}}$$

which can be estimated from industry cost shares, α, the estimated matching function, and the empirical wage differentials.

C.2 Attribution

alternative specification table: amplification

industry/occupation tables: one each

east german tables: AKM table, estimation plot, amplification table

Parameter	$Var(w)$	$Var(PD)$	$Var(ID)$	$Cov(PD, ID)$
α	.0090	.0204	.0002	.0100
β	.0306	-.0012	.0044	.0059
C	.0000	.0000	-.0001	.0010
B	.0007	.0005	.0000	.0002
	Counterfactual #1: constant vacancy distribution			
α	.0185	.0185	–	–
β	-.0025	-.0025	–	–
C	.0002	.0002	–	–
B	.0004	.0004	–	–
	Counterfactual #2: constant labor distribution			
α	.0202	.0202	–	–
β	-.0014	-.0014	–	–
C	.0001	.0001	–	–
B	.0007	.0007	–	–

Table A.12: Parameter contributions to predicted rise in wage variance

TABLE NOTES. Contributions are estimated by holding other parameters fixed at their 1993-1999 and 2010-2017 values, and taking the average across these two cases.