

Occupation Growth, Skill Prices, and Wage Inequality, Journal of Labor Economics, 2024

Wenzhi Wang *

January 16, 2025

Böhm et al. (2024)

1. Introduction

During the past decades, occupational employment has changed profoundly across Europe and the United States. A burgeoning literature has established fundamental shifts in labor demand as the most important cause of these changes. Yet it remains puzzling that these demand shifts fail to be clearly reflected in occupational wages or wage inequality. First, occupational employment growth has been decoupled from occupational wage growth. Second, a number of studies suggest that changes in occupation-specific labor demand contributed little to the increase in wage inequality in recent decades.

Making use of individual-level longitudinal information on wages and occupations from German administrative data, we document three important facts. First, individual workers' wage growth is substantially faster within expanding occupations. During our study period (1985-2010), annual individual wage growth is 1 percentage point higher on average in occupations that double in size compared with those whose size is constant. Second, workers who enter an occupation earn 21% less than incumbents on average. Similarly, average wages of workers leaving occupations are 16% lower than those of stayers. Third, both of these effects are increasing in net occupation growth. For a doubling in occupation size, the wage gaps between entrants versus incumbents and between leavers versus stayers widen by 6 and 3 percentage points, respectively. The raw data thus reveal that net growth of an occupation will attenuate average wages within occupations where selection operates in both directions. This explains why wage and employment growth across occupations are uncorrelated in our data.

To interpret these facts, we set up a parsimonious model of occupation choice. Workers have multidimensional occupation-specific skills that evolve heterogeneously across occupations and over the

*This note is written in my pre-doc period at the University of Chicago Booth School of Business.

career. *The central distinction we make is between skill prices – wages paid per constant unit of skill – and average occupational wages. While skill prices are directly affected by shifts in demand, selection may dampen, neutralize, or even overturn their effects on average wages.* Less skilled individuals enter growing occupations, which depress these occupations' average wages. Shrinking occupations also retain the most skilled parts of their workforce, lifting their average wages. Such selection effects imply that between-occupation inequality will underestimate the impact of shifting occupational demand on wage inequality.

We estimate the model using the longitudinal information in our data. Our analysis uncovers three main findings. *First, there is a clear positive relation between skill price and employment growth at the level of detailed occupations.* This indicates that demand shifts were the dominant drivers of both occupational employment and skill-constant wages over the past decades. Characterizing occupations by their task intensities, we find that the patterns are in line with routine-based technical change as one of the important drivers of occupational demand. More generally, the patterns are consistent with polarization, since employment and skill prices of broad occupation groups with high as well as low wages rose compared with those of mid-wage occupation groups.

Since employment growth is uncorrelated with average wage growth and is positively correlated with skill price growth, skills must deteriorate in growing occupations and increase in shrinking occupations. *Our second main finding is that lower-earning workers' net entry into growing occupations and their net exit out of shrinking occupations fully account for the negative relationship between skill changes and employment growth.* We term this the "employment growth selection effect," or simply growth selection. Quantitatively, a doubling of an occupation's workforce translates into a decrease in its average skills by 12%. Viewed through the lens of our model, growth selection stems from both entrants and leavers possessing lower average skills than stayers in any occupation. We exploit the longitudinal dimension of the data to document the nature of growth selection. One component is that growing occupations draw in and shed younger workers who are equally skilled as inframarginal workers but who have had less time to accumulate skills because of their age. We distinguish this age selection from Roy selection, which refers to differences in skills conditional on experience. Because of the fundamental selection problem, we focus on the current spell in a particular occupation. Aggregating across occupation groups, we find that 42% of employment growth selection can be attributed to age selection.

Our third findings is that changes in occupational skill prices have driven much of the increase in wage inequality over the past decades. Inequality between occupations would have been 70% higher at the end of our sample if growth selection had not counteracted the changing skill prices. We break down the quantiles of the wage distribution using our estimates of skill prices and skills along with workers' occupational histories. Our model tracks the wage distribution well. Demographic composition and faster skill accumulation at the median can account for 70% of the rise in the fiftieth to fifteenth percentile wage difference. Controlling for other factors, skill prices alone explain 60% of the increase in the eighty-fifth to fiftieth percentile difference.

2. Data, Literature, and Stylized Facts

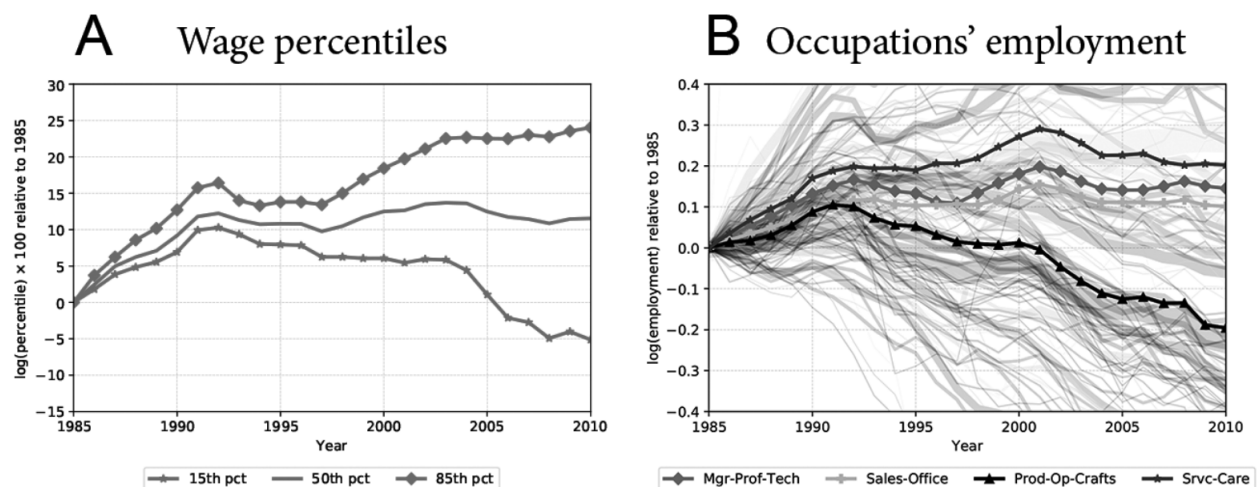
To work with a homogenous sample throughout, we restrict the main sample to German men aged 25-54 years who are working full time (excluding apprentices) in West Germany. We transform the spell structure into a yearly panel by using the longest spell in any given year, adjusting wages appropriately for spells that do not last the entire year. Because of a cap on social security contributions, 12% of wages are right censored at this ceiling; we follow the imputation procedures in past papers. We inflate all wages to 2010 prices using the German consumer price index.

A key strength of the SIAB data is that it provides high-quality longitudinal information on workers' occupations. Until 2010, the SIAB scientific use file contains a consistent set of 120 occupations; we cannot use subsequent years because the classification changes drastically thereafter. Most of our analyses will be based on the raw 120 occupations. To ease interpretation, we also aggregate them into broader groups. These comprise managers, professionals, and technicians (Mgr-Prof-Tech); sales and office workers (Sales-Office); production workers, operators, and craftsmen (Prod-Op-Crafts); and workers in services and care occupations (Srv-Care). See table A.1 for the mapping of detailed occupations into these groups.

2.1. Wage Inequality and Changes in Occupational Employment

Two of the most important trends in developed countries' labor markets over the past decades have been a strong increase in wage inequality and a substantial reallocation of employment across occupations broadly characterized by polarization. Germany is no exception to either phenomenon. Figure 1 reproduces both trends in our dataset.

Figure 1: Evolution of wage inequality and occupational employment



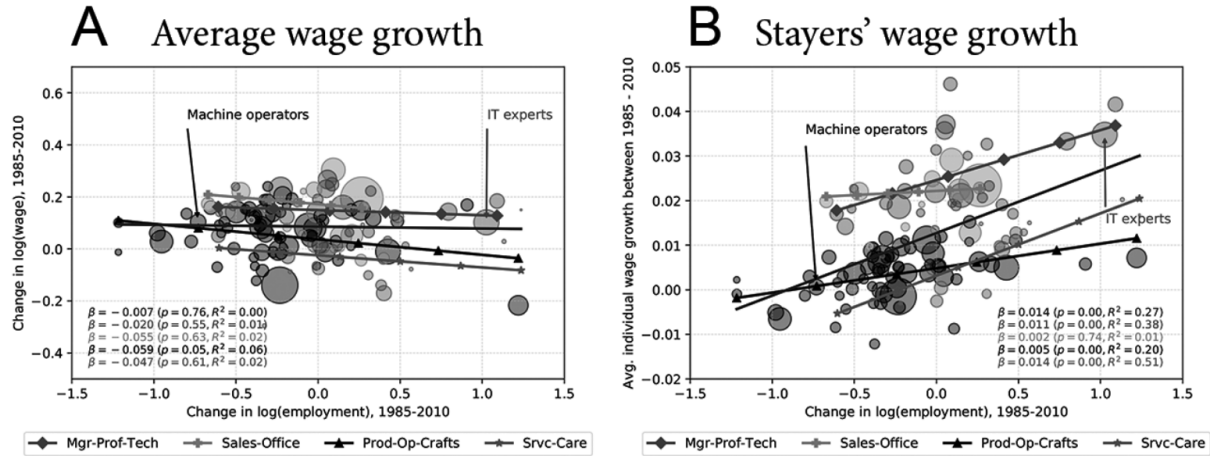
Notes. The vertical axis in A shows the fifteenth, fiftieth, and eighty-fifth log wage percentile over time relative to 1985. The vertical axis in B shows the log change in the number of employed workers within an occupation over time. Shaded lines in the background represent the 120 detailed occupations in the SIAB scientific use file. The four groups show an aggregation of these detailed occupations, as described in table A.1. The thickness of a shaded background line corresponds to the number of employed workers in an occupation averaged across years 1985 until 2010. A color version of this figure is available online.

Figure 1A shows the trends of wage percentiles over the 1985-2010 period normalized to zero in 1985. Inequality increased strongly and steadily both in the upper half, measured by the difference between the eighty-fifth and fiftieth percentile of log wages, and in the lower half (50 - 15 difference). Figure 1B plots the trends in the logs of the detailed 120 occupations' employment (shaded lines) and the four aggregated groups (bold lines with markers). Employment in Prod-Op-Crafts occupations declined by more than 20 log points from a baseline share of more than 60%, whereas the employment share of the other occupation groups increased. This trend has been termed “job” or “employment polarization” because Prod-Op-Crafts workers tend to be located in the middle of the occupational wage distribution. Part of the declining employment in middle-paying occupations appears to be due to changes in technology (affecting codifiable routine-type jobs) as well as international trade and offshoring (affecting manufacturing-type jobs).

One may expect to see such shifts of the demand for different types of occupations directly in the wage distribution, not least because the wage and employment trends occurred largely in parallel. There exists, however, surprisingly little quantitative evidence on the role of occupational change for the evolution of wage inequality: holding occupations' wages fixed at their initial levels and reweighting them with employment in subsequent decades, Goos and Manning (2007) show that composition effects due to employment polarization can account for a substantial part of changing wage inequality.

ity in the United Kindom. Autor (2019) finds that in the United States, a similar exercise explains only small shares of the income differentials across five education groups. For the German case, Dustmann, Ludsteck, and Schönberg (2009) conclude that the rise of lower-half inequality was unlikely to stem from changes in demand. Card, Heining, and Kline (2013) run a set of Mincer regressions and incrementally add occupational identifiers, finding that the role of the latter for rising wage inequality is rather small.

Figure 2: Correlation of changes in employment, average wages, and wage growth



Notes. The vertical axis in A shows the change in average wages between 1985 and 2010. The vertical axis in B depicts individual wage growth averaged across years 1985 until 2010. The horizontal axis in both panels shows the change of the log number of employed workers within an occupation between 1985 and 2010. One bubble represents one of the 120 detailed occupations in the SIAB scientific use file. The four groups show an aggregation of these detailed occupations, as described in table A.1. Bubble size corresponds to the number of employed workers in an occupation averaged across years 1985 until 2010. Regression lines across all occupations (black) and within the four broad groups (symbols) are weighted by the number of employed workers. A color version of this figure is available online.

Figure 2A hints at why these analyses tend to have limited explanatory power, plotting changes in employment against changes in average wages over the 1985-2010 period for each occupation. Variation along the horizontal axis shows that employment changes are very substantial. Many occupations grew or shrank by more than 50 log points. Yet movements of average wages are surprisingly small. *Thus, between-occupation decompositions – such as wage regressions with occupation dummies or reweighting strategies – may attribute little of the trends of in wage inequality to factors like changing skill prices and employment structure and much of its increase to unexplained within-occupation inequality.* In fact, the regression line shows that employment and wage changes are uncorrelated. To pick the two highlighted examples, information technology (IT) experts' employment rose by 102 log points, or 178%. Their average wages grew by 10%, just above the overall average. Machine operators—a prototypical occupation to be hit by RBTC—shrank by 73 log points, or 51%. Yet average wages grew by the same amount as those of IT experts.

Within the broader groups, the noncorrelation between wage and employment growth even turns negative for the lower-earning Prod-OpCrafts and Srvc-Care occupations. This is consistent with the regressions reported by Dustmann, Ludsteck, and Schönberg (2009, sec. IV.D). They conclude that demand shifts were unlikely to have driven lower-end inequality. Finding little or negative correlation between occupational wage and employment growth is not confined to Germany. Hsieh et al. (2019) and Roys and Taber (2019) document very small correlations between the growth rates of occupational employment and wages in the United States. Employment in low-skill occupations increased in the United Kingdom and Canada, while at the same time wages dropped compared with routine occupations (Goos and Manning 2007; Green and Sand 2015).

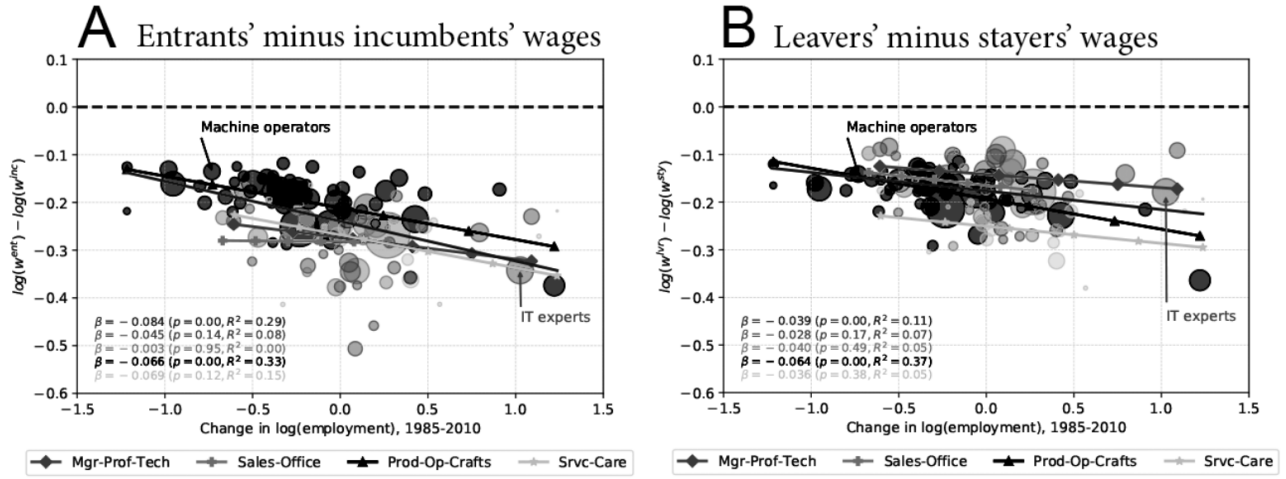
Next to the role that occupations have to play for wage inequality, this begets the more fundamental question of whether, on aggregate, shifts in demand versus supply of labor to different occupations were the dominant factor for the changes of the employment structure. We will find that while the latter may have a role to play, the data strongly suggest that demand changes along the lines of routine-biased technological or international trade are important.

2.2. Individual-Level Wage Growth and Selection

As a first pass, figure 2B shows that there is a strong positive correlation between employment and individual-level wage growth. The horizontal axis is the same as in figure 2A, whereas the vertical axis plots the average annual wage growth of workers who stayed in their occupation for any two consecutive years. Wage growth rates within occupations clearly line up with their employment growth. We will control for other important factors, such as occupation- and age-specific returns to experience, below. But the juxtaposition of figure 2A and 2B already suggests that the key reason for the differences will be selection into occupations. *Put differently, demand shifts will indeed be driving the changes of employment and prices paid for skilled labor across occupations. Negative selection of entrants into growing occupations obscures this relation when looking at average occupation-specific wages.* The underlying occupational prices would be spreading out more than the average occupational wages, which are captured in the above-discussed decomposition analyses.

Data limitations have prevented a more thorough analysis of such selection effects. In particular, the main sources in the United States are repeated cross sections (Current Population Survey [CPS], US Census) or longitudinal data too small in size for investigating individual-level dynamics across detailed occupations (Panel Study of Income Dynamics [PSID], National Longitudinal Survey of Youth). The SIAB data allow us to track occupational biographies over the entire career.

Figure 3: Selection into and out of occupations



Notes. The vertical axis in A shows the average wage of an entrant to an occupation relative to the average wage of incumbents. The average wage is taken across years 1985 until 2010. The vertical axis in B shows the average wage of a worker leaving an occupation next period relative to the average wage of stayers. The average is taken across years 1985 until 2009 to avoid all workers being leavers at the sample end. The horizontal axis in both panels shows the change in the log number of employed workers within an occupation between 1985 and 2010. One bubble represents one of the 120 detailed occupations in the SIAB scientific use file. The four groups show an aggregation of these detailed occupations, as described in table A.1. Bubble size corresponds to the number of employed workers in an occupation averaged across years 1985 until 2010. Regression lines across all occupations (black) and within the four broad groups (symbols) are weighted by the number of employed workers. A color version of this figure is available online.

The vertical axis of figure 3A shows the difference between entrants and incumbents. An occupational entrant is someone newly observed in the occupation in the current period. He could be joining the labor force for the first time, switch from a different occupation, or enter from unemployment or outside the labor force. The wage difference between this group and incumbents is strongly negative.

In principle, the patterns in figure 3A could be generated if occupation choice happened only at labor market entry in combination with substantial returns to experience. If this was the sole effect, however, we would expect that the wages of workers leaving their occupations would be higher than the wages of those who stay on. Put differently, in such a scenario individuals dropping out of our sample after age 54 should dominate the difference between leavers and stayers. Figure 3B and Table 1 show that the opposite is the case. As for entrants, marginal workers have substantially lower wages than those who stay on. This difference is 17.5 log points on average and increases by 3.9 log points per 100 log points higher employment growth. Put differently, only the lowest-skilled workers leave fast-growing occupations. This suggests that the wage gap is not just due to entrants being at an earlier stage of their career than incumbents.

The second specification in Table 1 investigates how much of the wage gaps can be explained with

standard observable skill proxies. *We first run Mincer regressions of individual workers' wages on full interactions of dummies for calendar year, education in three categories, and years of age. We use the residuals to compute the average wage gaps for entrants/leavers and their slopes with respect to occupation growth as before.*

Figure Table 1: Selection Into and Out of Occupations: Regression Results

Specification	Entrants/Incumbents		Leavers/Stayers	
	$\bar{w} X$ (1)	$\hat{\beta}_{\text{occ growth}}$ (2)	$\bar{w} X$ (3)	$\hat{\beta}_{\text{occ growth}}$ (4)
No covariates:				
Coefficient	-.235	-.084	-.175	-.039
SE	.006	.012	.005	.01
Fully stratified by year/education/age:				
Coefficient	-.146	-.069	-.165	-.031
SE	.006	.011	.004	.009
Only occupational switchers:				
Coefficient	-.134	-.026	-.138	-.038
SE	.005	.011	.005	.011
Fully stratified by year/education/age, by age group:				
Ages 25–34:				
Coefficient	-.11	-.054	-.111	-.027
SE	.005	.007	.004	.007
Ages 35–44:				
Coefficient	-.106	-.038	-.112	-.03
SE	.004	.009	.004	.009
Ages 45–54:				
Coefficient	-.091	-.021	-.132	-.009
SE	.004	.008	.004	.008

NOTE.—The numbers are based on regressions of average wage differences between entrants (leavers) and incumbents (stayers) on occupation growth between 1985 and 2010. The average wage is taken across years between 1985 and 2010 and weighted by occupations' employment. In the upper panel, the specification "No covariates" refers to fig. 3. The specification "Fully stratified by year/education/age" controls for the interaction of dummies for calendar year, education (no postsecondary, Abitur or apprenticeship, and university degree), and years of age. The corresponding figure can be found in app. A (fig. A.1). The specification "Only occupational switchers" restricts the samples of entrants and leavers to those who directly switch between occupations (i.e., we drop labor market entrants, sample leavers at age 54, and those with an intermittent spell of unemployment or outside the labor force). See fig. A.2 for the graphical representation of the data and regression line. In the bottom panel, we repeat the specification "Fully stratified by year/education/age," splitting the sample into three age groups.

Observables account for 38% of the gap between entrants and incumbents; they explain much less of the differences between leavers and stayers. The bottom panel shows that coefficients drop by another 30% on average when splitting the sample into 10-year age groups, thereby making the pools of marginal and incumbent workers more homogenous in that dimension. Qualitatively, results are the same. *These patterns underscore that the wage gaps are related to skill selection in education and age. At the same time, the majority of gaps remain unexplained by standard Mincer variables.* The third row of Table 1 investigates the extent to which the gaps persist even among direct occupational switchers (i.e., excluding labor market entrants, sample leavers at age 54, and moves to and from nonemployment). Relative to differences in unconditional wages for the entire sample in the first row, the wage gaps decline by 43% and 21%, respectively. This is natural, as this sample removes the large skill differences due to entrants/leavers moving between nonemployment states or young workers entering the labor market.

The average wage gap remains substantial. This shows that direct occupation switchers will measurably contribute to skill selection below, and it suggests once more that skills are occupation specific.

In all specifications contained in Table 1, the slopes of wage gaps with respect to occupation growth are negative. This underscores that when occupations grow, the composition of workers changes so that average skills become increasingly scarce at all margins – in terms of observables and unobservables and among workers hired directly from other occupations. The last two columns of table 1 are related to Groes et al. (2015) analysis of the Danish labor market. Also using a sample of stayers and leavers, they find that both low and high earners in an occupation are most likely to switch. In our German data, low earners in an occupation have a much higher probability of leaving, dominating the average differences between leavers and stayers (see app. A, sec. A.3).

The prominent models in the literature on occupational changes have difficulty matching the facts in figure 3 and Table 1, as they feature one-dimensional skills. A single dimension leads to a hierarchical ranking of occupations by skill and ensures tractability in general equilibrium. It also implies that upward switchers leave the lower-ranked occupation from above and that downward switchers enter lower-ranked occupations from above. This is hard to square with the finding that even entrants and leavers in low-wage occupations generally earn less than incumbents and stayers, respectively. In principle, the model in Hsieh et al. (2019) could deliver the empirical facts from figure 2, figure 3, and Table 1. To make quantitative predictions, it requires time-constant skills and parametric assumptions about individual skills in all sectors.

The remainder of this paper develops and estimates an empirical model with multidimensional skills that flexibly change over the career. For estimation we only require data on age, occupations chosen, and the corresponding wages. We will use this model to explain and quantify the implications of the new stylized facts shown in this section.

3. Estimating Skill Prices with Longitudinal Data

There are K distinct occupations. At any time t a worker i would earn potential wage $W_{i,t,k}$ in occupation k . This potential wage is the product of the worker's occupation-specific skill $S_{i,t,k}$ and the occupation-specific price for a unit of skilled labor $\Pi_{t,k}$. We use lowercase letters to denote logs. As in Roy (1951), we assume that workers maximize their incomes by choosing the occupation in which they earn the highest wage:

$$w_{it} = \max \{w_{i,t,1}, \dots, w_{i,t,K}\} = w_{i,t,k(i,t)} = \pi_{t,k(i,t)} + s_{i,t,k(i,t)}. \quad (1)$$

The occupation subscript's argument (i, t) indicates that k is i 's choice at time t .

Our goal is to estimate the evolution of skill prices $\pi_{t,k}$ over time. Based on equation (1), we first employ an approximation that allows us to disentangle prices from skills based on observed data. In the empirical formulation, prices evolve at the aggregate level. Skills will grow and be subject to idiosyncratic shocks over an individual's career; we describe how we deal with the challenges that arise from this in a second step.

The Roy model is hard to estimate in its general form (1) and requires strong restrictions. Böhm (2020) derives a computationally simple approximation, which he uses in repeated cross-section data under the assumption that observable and time-invariant proxies for each $s_{i,t,k}$ are available. We adapt Böhm's result to a panel data setting.

For a worker who switches occupations (i.e., $k(i, t-1) \neq k(i, t)$), note from equation (1) that

$$w_{i,t-1,k(i,t)} - w_{i,t-1,k(i,t-1)} \leq 0 \quad w_{i,t,k(i,t)} - w_{i,t,k(i,t-1)} > 0.$$

We also assume the following:

Assumption 1. For workers who switch occupations, one can approximate the wage at indifference, where

$$w_{i,\tau,k(i,t)} = w_{i,\tau,k(i,t-1)}, \tau \in [t-1, t),$$

to be in the middle of the wage interval

$$\left[(w_{i,t-1,k(i,t)} - w_{i,t-1,k(i,t-1)}), (w_{i,t,k(i,t)} - w_{i,t,k(i,t-1)}) \right].$$

That is, we assume the switch point to be halfway between the previous and the current relative wage:

$$\underbrace{w_{i,t-1,k(i,t)} - w_{i,t-1,k(i,t-1)}}_{\text{relative wage in period } t-1} + \underbrace{w_{i,t,k(i,t)} - w_{i,t,k(i,t-1)}}_{\text{relative wage in period } t} = 0. \quad (2)$$

For occupation switchers, the Roy model implies negative relative wages in $t-1$ and positive relative wages in t . *Assumption 1 posits that the absolute values of these relative wages (i.e., the wage rents from the chosen occupation or the wage distance to indifference) are the same in both periods.* From this, we can express realized wage changes as a function of potential wage changes in the occupations worker i chooses in $t-1$ and t .

Result 1. Under Assumption 1, worker i 's observed wage growth between $t-1$ and t can be written in terms of the potential wage growth in the occupations he chooses in those periods:

$$\Delta w_{i,t} = \frac{1}{2} (\Delta w_{i,t,k(i,t)} + \Delta w_{i,t,k(i,t-1)}), \quad (3)$$

where

$$\Delta w_{i,t,k(i,t)} := w_{i,t,k(i,t)} - w_{i,t-1,k(i,t)}, \quad \Delta w_{i,t,k(i,t-1)} := w_{i,t,k(i,t-1)} - w_{i,t-1,k(i,t-1)}.$$

Result 1 is economically attractive as it includes workers' endogenous switches between occupations due to changing potential wages. Naturally, if a worker stays in an occupation during two adjacent periods, his realized wage change is equal to the change in his potential wage in the chosen occupation. If the worker decides to switch occupations and $k(i, t-1) \neq k(i, t)$, half of his wage gain stems from the wage change he would have experienced had he stayed in his previous occupation ($\Delta w_{i,t,k(i,t-1)}$). The other half is the wage change had he been in the current occupation all along ($\Delta w_{i,t,k(i,t)}$). Importantly, only potential wages in the previous and current occupations matter for i 's observed wage

change; potential wages in all other occupations are irrelevant. This stems from the fact that the worker has comparative advantage in both of these occupations and that he chooses accordingly.

Using equation (1), we can decompose result 1 into changes of prices and skills:

$$\Delta w_{i,t} = \frac{1}{2} (\Delta \pi_{t,k(i,t)} + \Delta s_{i,t,k(i,t)}) + \frac{1}{2} (\Delta \pi_{t,k(i,t-1)} + \Delta s_{i,t,k(i,t-1)}) . \quad (4)$$

In the next section, we show how panel data on wages and occupation choices allow us to estimate the evolution of prices and skills imposing minimal yet informative structure on $\Delta s_{i,t,k}$.

3.1. Identification of Skill Prices Changes

The empirical strength of result 1 is that both individuals' realized wage changes $\Delta w_{i,t}$ and occupation choices $k(i, t-1)$, $k(i, t)$ are observed in the SIAB data. This allows linking potential and observed wage changes. However, we need to put more structure on equation (4) in order to separate changes in skill prices from changes in skills. Our framework relies on differences. We thus do not place restrictions on either object's initial levels.

We model the skill accumulation process as learning-by-doing on the job. Its speed is occupation specific and depends on age; working in one occupation k impacts subsequent skills in all other occupations. In particular, we assume the following.

Assumption 2. Individuals' occupation-specific changes are time invariant in expectation. For all $k \in \{1, \dots, K\}$ and $i \in \{1, \dots, N\}$,

$$\mathbb{E} [\Delta s_{i,t,k} \mid a(i, t-1), k(i, t-1), k] = \Gamma_{a(i,t-1), k(i,t-1), k}, \quad (5)$$

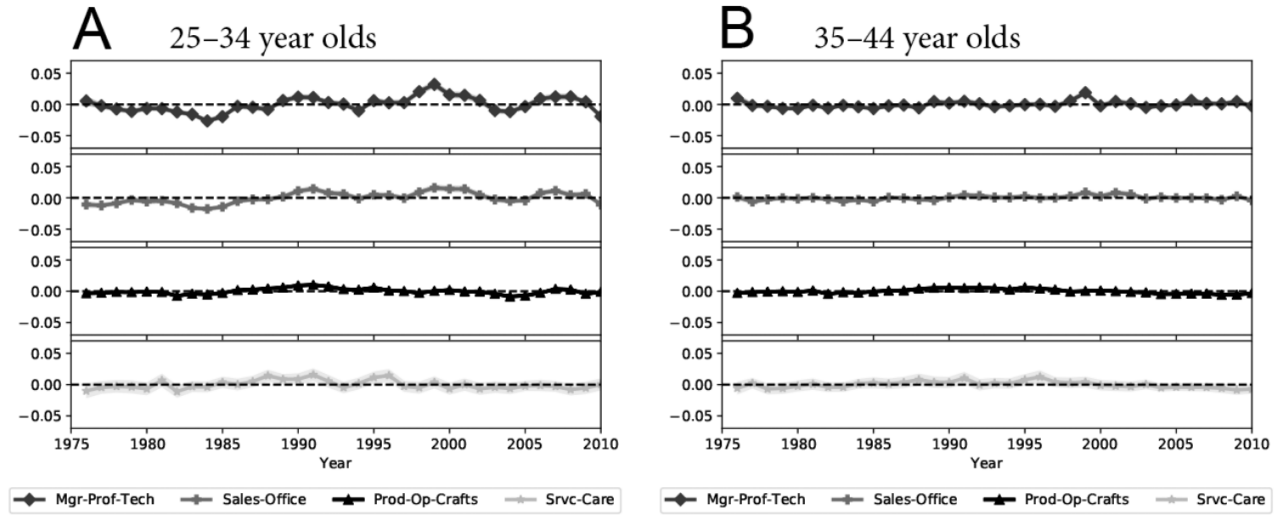
with $a(i, t-1)$ denoting the age group of individual i in period $t-1$.

Assumption 2 restricts reduced-form parameters to be time invariant. It can be interpreted as consisting two restrictions. *The first restriction is time invariance of the structural skill accumulation parameters introduced later.* This means that ex ante skill accumulation within $a(i, t-1) \times k(i, t-1)$ cells for all potential k has not changed over time. This is consistent with the literature studying occupational changes, and more generally with the idea that differences in returns to worker characteristics over time are due to changes in the returns to skills rather than changes in skill endowments.

The second restriction assumes that changing prices do not affect ex post observed skill accumulation among workers of age $a(i, t-1)$ with realized choices $k(i, t-1)$ and $k(i, t)$. This is at odds with the idea of self-selection in the Roy model. The latter would imply, for example, that workers who stay in an occupation during a period of declining prices will be positively selected on skill changes compared with staying workers in periods with constant prices. Therefore, on the one hand, our approach makes progress compared with the literature by incorporating workers' endogenous choices based on changing potential wages. On the other hand, it retains the limitation that skill accumulation parameters are not adjusted for changing self-selection over time.

To examine the plausibility of assumption 2 empirically, figure 4 plots the year-to-year wage growth of 25-34-year-olds and 35-44-year-olds (B) minus average wage growth of 45-54-year-olds. We subtract the overall mean everywhere. Under assumption 2, the ratio of wage growth across different demographic cells should remain constant over time; the eight lines should thus be flat at zero. In figure 4B, all lines come close to it. Figure 4A is somewhat noisier, but there are no systematic trends. These results suggest that within occupations, relative wage growth of different age groups has not substantively changed over time.

Figure 4: Individual wage growth relative to 45-54-year-olds



The lines show average individual wage growth from $t - 1$ to t by year of 25-34-year-olds (A) and 35-44-year-olds (B) minus average wage growth of 45-54-year-olds. Results are centered at zero to show trends over time. The shaded areas around the four lines are 95% confidence intervals. The four groups are based on an aggregation of detailed occupations in the SIAB scientific use file, as described in table A.1. A color version of this figure is available online.

Assumption 2 identifies changes in skill prices from the difference between wage changes and expected skill changes up to a constant. We thus need one further normalization, which we apply to skill price growth during part of our sample.

Assumption 3. Skill prices are constant during a base period, that is,

$$\Delta\pi_{t,k} = 0, \quad \forall k \in \{1, \dots, K\}, t \in \{1, \dots, T_{\text{base}}\}.$$

Under assumption 3, Γ is identified from the base period due to its time-invariant nature. Accordingly, the estimates of $\Delta\pi_{t,k}$ can be interpreted as actual changes of skill prices for $t > T_{\text{base}}$. Distinguishing between price and skill growth is a general challenge of estimation based on panel data; it is often done

rather implicitly. Using the whole decade 1975-84 as the base period, we will abstract from short-term fluctuations. This period covers the entire business cycle.

Another identification strategy is to assume that seasoned workers' average skill growth is zero ("flat spot identification"). Under this assumption, occupation-specific wage growth over the decades is identical to price growth for these ages. Wage growth still incorporates endogenous choices, as in result 1. We will explore this as a robustness check, finding results similar to our specification.

3.2. Estimation and Interpretation of Skill Accumulation Parameters

We combine the equation for wage growth (4) and skill accumulation equation (5) to obtain our baseline estimation equation:

$$\begin{aligned} \Delta w_{i,t} = & \Delta \pi_{t,k(i,t-1)} \cdot \frac{\mathbb{I}(k(i,t-1))}{2} + \Delta \pi_{t,k(i,t)} \cdot \frac{\mathbb{I}(k(i,t))}{2} \\ & + \Gamma_{a(i,t-1),k(i,t-1),k(i,t-1)} \cdot \frac{\mathbb{I}(a(i,t-1)) \cdot \mathbb{I}(k(i,t-1))}{2} + \\ & + \Gamma_{a(i,t-1),k(i,t-1),k(i,t)} \cdot \frac{\mathbb{I}(a(i,t-1)) \cdot \mathbb{I}(k(i,t-1)) \cdot \mathbb{I}(k(i,t))}{2} + \varepsilon_{i,t}, \end{aligned} \quad (6)$$

with $\mathbb{I}(a(i,t))$ and $\mathbb{I}(k(i,t))$ denoting indicator variables for i 's age group and choice of occupation, respectively. In line with assumption 3, we set $\Delta \pi_{k,t} = 0 \forall k \in \{1, \dots, K\}, t \in \{1, \dots, T_{\text{base}}\}$ and estimate equation (6) for the whole period $t \in \{1, \dots, T\}$ by OLS.

Regression (6) is perfectly saturated in age groups, previous occupations, and current occupations. The regression error $\varepsilon_{i,t}$ reflects individuals' skill shocks or idiosyncratic deviations that are different from the average worker in the respective cell spanned by these dummies. That is, we can write $\varepsilon_{i,t} = (1/2) (u_{i,t,k(i,t-1)} + u_{i,t,k(i,t)})$, where $u_{i,t,k} := \Delta s_{i,t,k} - \Gamma_{a(i,t-1),k(i,t-1),k}$. By assumption 2 we have $\mathbb{E}[u_{i,t,k} \mid a(i,t-1), k(i,t-1), k] = 0$, and the error term in equation (6) is uncorrelated with the regressors.

Result 2. Under assumptions 1-3, OLS estimation of equation (6) consistently identifies changes $\{\Delta \pi_{t,k} \mid t > T_{\text{base}}\}$ and average skill changes $\{\Gamma_{a(i,t-1),k(i,t-1),k}\}$. Without assumption 3, OLS identifies these parameters up to the normalization of the base period.

As discussed with assumption 2, average occupation-specific skill changes Γ will not coincide with structural skill accumulation parameters in general. Suppose there exists a learning-by-doing production function, such that ex ante a worker accumulates skills according to

$$\Gamma_{a(i,t-1),k(i,t-1),k}^* = \mathbb{E}[\Delta s_{i,t,k} \mid a(i,t-1), k(i,t-1)] + u_{i,t,k}^*.$$

What we identify in regression (6) is ex post accumulation in the sense that

$$\Gamma_{a(i,t-1),k(i,t-1),k(i,t)} = \Gamma_{a(i,t-1),k(i,t-1),k(i,t)}^* + \mathbb{E}[u_{i,t,k}^* \mid a(i,t-1), k(i,t-1), k(i,t)] \quad (7)$$

is conditional on the current choice $k(i, t)$ and $u_{i,t,k} = u_{i,t,k}^* - \mathbb{E}[u_{i,t,k}^* \mid a(i, t-1), k(i, t-1), k(i, t)]$ is the deviation of the structural shock from its ex post mean. By the nature of the data, $\Gamma_{a(i,t-1),k(i,t-1),k}$ with $k(i, t-1) \neq k$ will be identified from switchers. For those, often

$$\mathbb{E}[u_{i,t,k}^* \mid a(i, t-1), k(i, t-1), k(i, t)] > \mathbb{E}[u_{i,t,k}^* \mid a(i, t-1), k(i, t-1)] .$$

That is, the off-diagonal accumulation parameters tend to be upward biased because of a classic self-selection problem. We thus cannot interpret them as what is learned for $k(i, t)$ while working in $k(i, t-1)$ in the entire population. More subtly, stayers' $\Gamma_{a(i,t-1),k(i,t-1),k(i,t)}$ with $k(i, t-1) = k(i, t)$ will tend to be overestimated relative to Γ^* , too. By self-selection, stayers are likely to have experienced more positive shocks compared with leavers.

Nevertheless, regression (6) succeeds in estimating the skill price changes as long as assumption 2 holds. The price changes are well identified using our method for plausible parameter ranges. It becomes clear, however, that the estimated Γ matrix needs too be interpreted in the above ex post sense. For our decomposition purposes and analysis of skills in occupations, this will be sufficient. The caveats are that we can make statements neither about workers' skill (changes) in occupations that they do not choose nor about the population distribution of the K -dimensional skill vector.

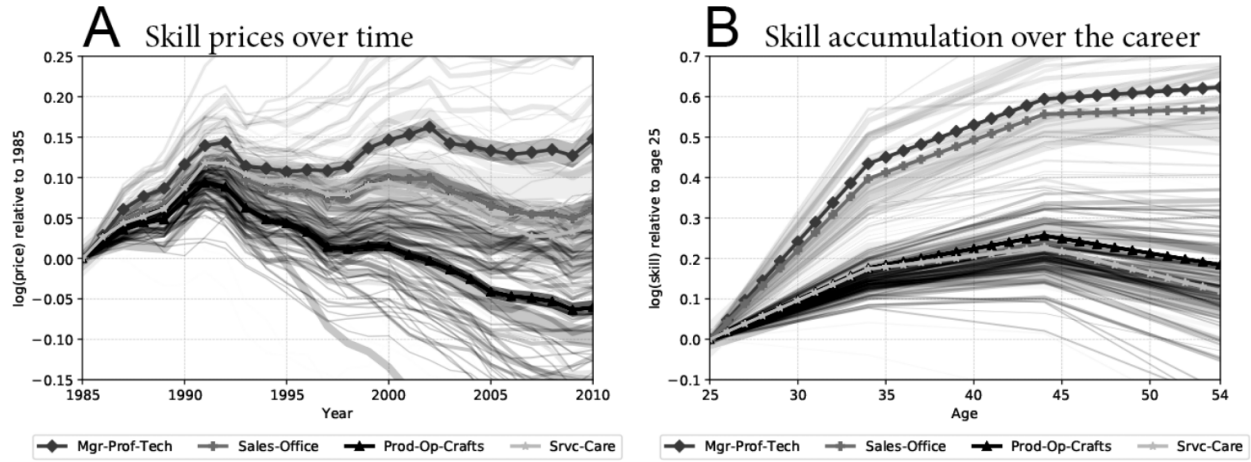
4. Skill Prices and Skill Selection

This section first presents the estimation results. These include the evolution of skill prices, the accumulation of skills over the career, and the relationship of prices and occupations' average skills with employment growth. We then dig deeper into the nature of the implied selection effects, showing that the skill differences between marginal workers and those who remain in their occupations drive the strongly negative association of employment growth with average skill changes in an occupation. We further decompose these in section 4.3 before reporting on various robustness exercises in the final part. One of those is an alternative way to think about unemployment and spells outside the labor force, which are not fully obvious in our setup (in the main estimation, we treat such spells as missing data).

4.1. Estimated Skill Price Changes and Average Skill Accumulation

Figure 5A depicts the evolution of skill prices, normalizing them to zero in 1985 and cumulating the yearly changes until 2010. In the broad occupation groups, skill prices increased strongly among Mgr-Prof-Tech occupations, increased modestly among Sales-Office and Srvc-Care, and decreased among Prod-Op-Crafts. The thin lines in the background show that these broad estimates mask substantial heterogeneity among the 120 detailed occupations. We will explore this in greater detail below.

Figure 5: Estimated skill prices and stayers' average skill accumulation



The graphs show estimated skill price changes over time and occupation stayers' skill accumulation over the life cycle. OLS estimates are as described by equation (6). Shaded lines in the background represent the 120 detailed occupations in the SIAB scientific use file. The four groups show an aggregation of these detailed occupations, as described in table A.1. The thickness of a shaded background line corresponds to the number of employed workers in an occupation averaged across years 1985 until 2010. The shaded areas around the four lines are 95% confidence intervals. A color version of this figure is available online.

Several distinct periods are noticeable. All prices increased in the favorable economic conditions between 1985 and 1991, although this was already less pronounced for the Prod-Op-Crafts occupations. These have experienced a continuous decline thereafter to the point that prices in 2010 were more than 5% below their initial value in 1985. For the other occupations, there was a drop during the 1992-93 recession as well; prices then stayed constant until they rebounded before the turn of the century. This rebound was most pronounced for Mgr-Prof-Tech occupations; prices in this group did not change much for the remainder of our sample period. Skill prices fell by about 5 percentage points for Sales-Office and Srvc-Care occupations between 2000 and 2010. These broad patterns are consistent with the job polarization of figure 1B.

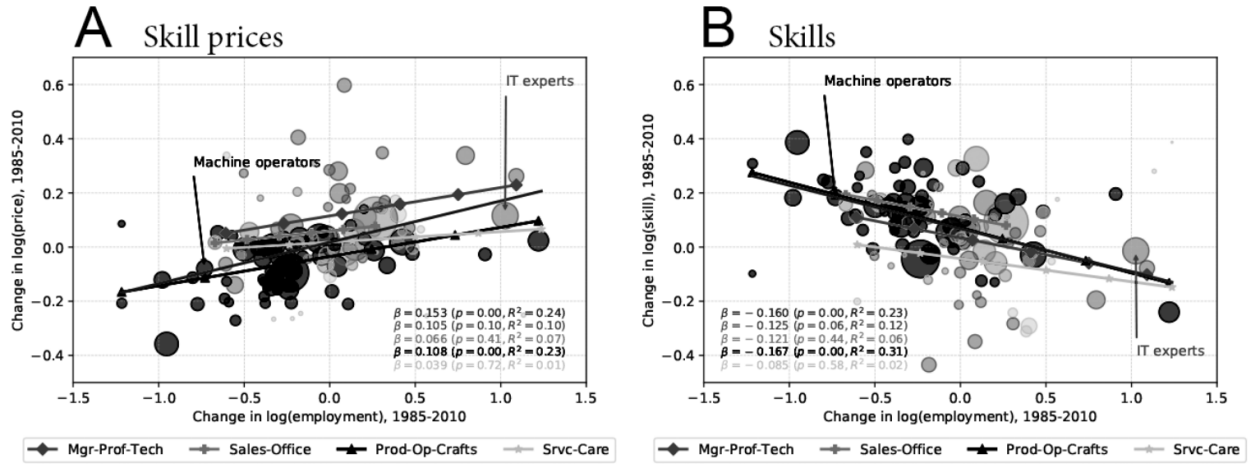
Figure 5B graphs the estimates for the diagonal elements of Γ , that is, the average skill accumulation for stayers. Skill growth in the early years of the career is steep. Absent changes in skill prices, it implies a 20% growth in wages between age 25 and age 34 for Prod-Op-Crafts or Srvc-Care occupations and 50% or more for the other two. It slows down midcareer and flattens out or turns negative toward the end of prime age. This reflects the well-established concavity of life cycle wage profiles. Skill growth differs substantially by occupation. It is initially fast in high-earning Mgr-Prof-Tech and Sales-Office occupation groups and never completely ceases. Growth is flatter and eventually peters out or turns negative in the Prod-Op-Crafts and Srvc-Care groups (i.e., occupations that often require physical labor). Again, the broad groups mask substantial heterogeneity across the detailed occupations. At the

same time, it is the case that the MgrProf-Tech and Sales-Office occupations on the one hand and the Prod-OpCrafts and Srvc-Care occupations on the other hand are almost separate; there are hardly any to be found in the other block. This shows that life cycle wage profiles are decidedly different across occupations, and accounting for this is critical in producing reasonable estimates of prices and skills.

We now hone in our key finding of this section, namely, that employment growth and skill price growth go hand in hand. Figure 2A has shown that the LHS (i.e., the average wage change of professions over the 1985-2010 period) is unrelated to employment growth at the level of the 120 detailed occupations. Figure 6 plots its two components on the right-hand side separately.

$$\begin{aligned} \mathbb{E} [w_{i,t,k(i,t)}] - \mathbb{E} [w_{i,t-1,k(i,t-1)}] &= \underbrace{\Delta \pi_{t,k}}_{\text{price change}} \\ &+ \underbrace{\mathbb{E} [s_{i,t,k(i,t)}] - \mathbb{E} [s_{i,t-1,k(i,t-1)}]}_{\text{mean skill change}} \end{aligned} \quad (8)$$

Figure 6: Correlation of changes in employment, skill prices, and skills



Notes. The vertical axis in A shows the change in skill prices between 1985 and 2010, estimated as detailed in section 3.2. The vertical axis in B depicts the change in skills between 1985 and 2010, estimated as the residual between price and wage changes as shown in equation (8). The horizontal axis in both panels shows the change in the log number of employed workers within an occupation between 1985 and 2010. One bubble represents one of the 120 detailed occupations in the SIAB scientific use file. The four groups show an aggregation of these detailed occupations, as described in table A.1. Bubble size corresponds to the number of employed workers in an occupation averaged across years 1985 until 2010. Regression lines across all occupations (black) and within the four broad groups (symbols) are weighted by the number of employed workers. A color version of this figure is available online.

Figure 6A shows that detailed occupations' log employment changes are positively related to cumulative skill price changes over the same period. The black regression line summarizes this strong relationship for the 120 occupations, which is in marked contrast to the zero correlation for wages in

figure 2A. The relationship holds within occupation groups, showing that the result is more general than a particular demand shifter that predominantly impacts broad occupations. Figure 6A indicates that shifts in demand were the dominant driver of occupational employment changes. Pure shifts in supply would have induced a negative correlation of employment with prices, all else equal. At the same time, there exists heterogeneity of occupations' employment and price growth around the regression line(s). This may, among other things, be due to contemporaneous shifts in supply but also due to variation in the elasticity of skill supply (inelastic occupations would have small employment and large price growth).

Figure 6B shows that implied skill changes are the flip side of the skill price estimates. The regression line indicates that skills of the occupations with the fastest growth declined by 35 log points on average compared with those that shrank the most. These are large effects; we devote the next subsections to examining their components and plausibility.

4.2. Defining and Estimating Employment Growth Selection

We have documented in section 2.2 that entering (leaving) workers' skills on average appear decidedly below those of incumbents (stayers) and that faster-growing occupations draw even less skilled entrants (leavers). Given that growing sectors by definition experience net entry, this could substantially drag down growing occupations' average wages despite rising demand and increasing skill prices. Here we formalize and quantify this effect in the context of our model, showing that it indeed drives the systematic part of the relationship between employment growth and skills.

The change in average skills of an occupation in equation (8) is determined by three mutually exclusive groups of workers: those who leave the occupation after period $t - 1$, those who stay on after period $t - 1$ and are thus incumbent in period t , and those who enter in period t . Denoting the share of leavers after $t - 1$ by $p_{t-1,k}^{lvr}$ and the share of period t entrants by $p_{t,k}^{ent}$, we can decompose the change of skills in occupation k into three terms:

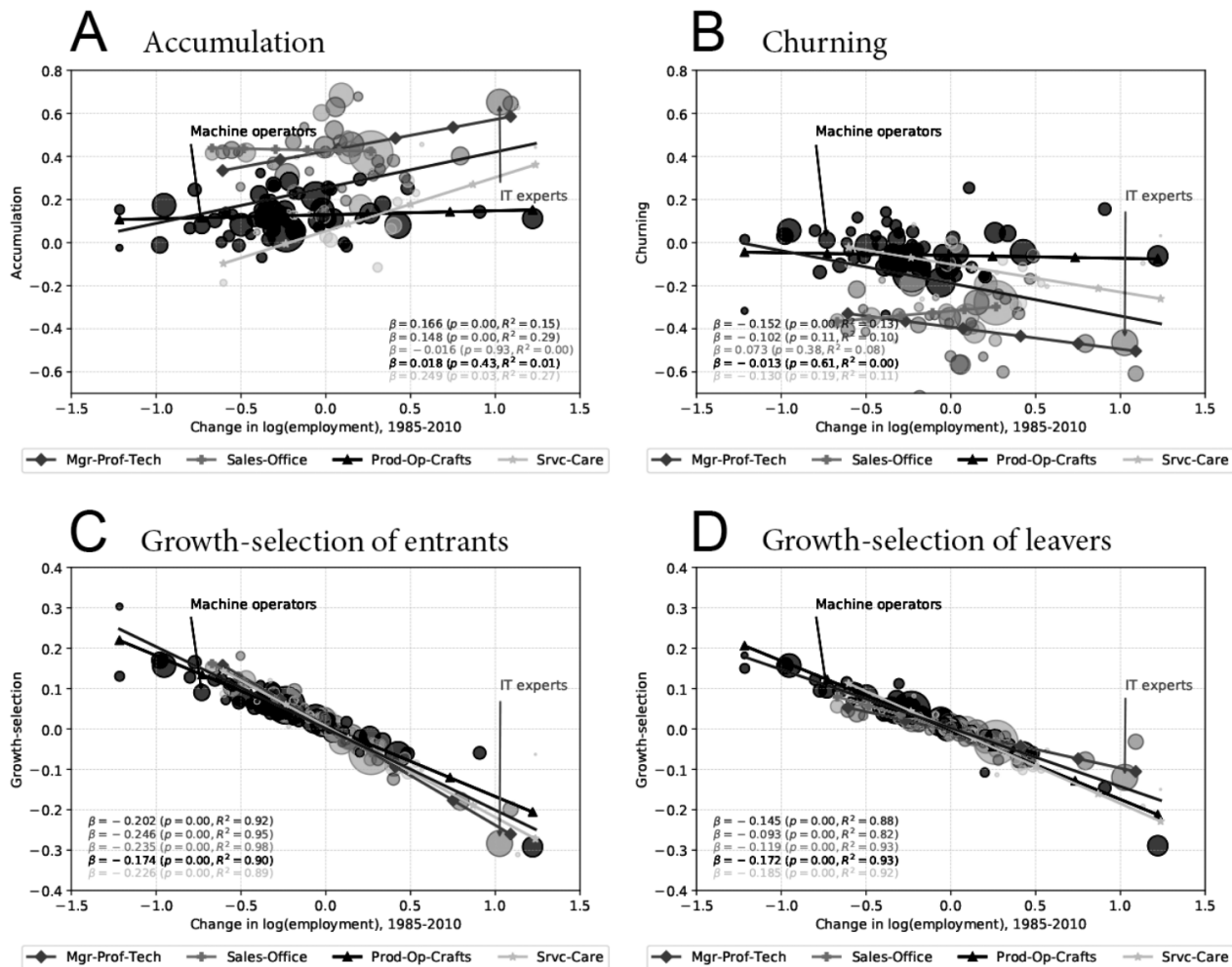
$$\underbrace{\mathbb{E}[s_{i,t,k(i,t)}] - \mathbb{E}[s_{i,t-1,k(i,t-1)}]}_{\text{mean skill change}} = \left(1 - \frac{p_{t-1,k}^{lvr} + p_{t,k}^{ent}}{2}\right) \cdot \mathbb{E}[\Delta s_{i,t,k}^{incumb}] + \frac{p_{t-1,k}^{lvr} + p_{t,k}^{ent}}{2} \cdot (\mathbb{E}[s_{i,t,k}^{ent}] - \mathbb{E}[s_{i,t-1,k}^{lve}]) + \frac{p_{t,k}^{ent} - p_{t-1,k}^{lvr}}{2} (\mathbb{E}[s_{i,t,k}^{ent}] - \mathbb{E}[s_{i,t,k}^{incumb}] + \mathbb{E}[s_{i,t-1,k}^{lvr}] - \mathbb{E}[s_{i,t-1,k}^{sty}]) . \quad (9)$$

See section B.4 of appendix B for the steps of the derivation. Summing the year-to-year changes in equation (9) over the period 1985-2010 gives the components of skill changes we analyze in the following.

The first term of equation (9) reflects just this, that is, the skill accumulation of workers who remain in the occupation. Its impact on occupational skill changes is high if turnover is small and skill accu-

mulation of staying workers is high. Figure 7A shows that this term is positive for the vast majority of occupations and larger among higher-accumulation occupations.

Figure 7: Employment growth versus the components of skill changes



Notes. Results correspond to sample averages following equation (9). The horizontal axis in both panels shows the change in the log number of employed workers within an occupation between 1985 and 2010. One bubble represents one of the 120 detailed occupations in the SIAB scientific use file. The four groups show an aggregation of these detailed occupations, as described in table A.1. Bubble size corresponds to the number of employed workers in an occupation averaged across years 1985 until 2010. Regression lines across all occupations (black) and within the four broad groups (symbols) are weighted by the number of employed workers. A color version of this figure is available online.

The second term of equation (9) is churning, which is composed of average turnover multiplied with the skill differences between period t entrants and $t - 1$ leavers. Churning tends to impact aggregate skills negatively, since leavers will have accumulated more skills than entrants. It becomes more negative for high-turnover occupations and for occupations with high levels of skill accumulation. Hence,

the accumulation and churning effects will often act in opposite directions. Figure 7B shows this to be the case.

Occupation growth does not have a first-order effect on the total of accumulation and churning. By inducing variation in turnover ($p_{t-1,k}^{lvr} + p_{t,k}^{ent}$), occupation growth pushes terms 1 and 2 in opposite directions. It is thus no surprise that the slopes in figure 7A and 7B have opposite signs. Indeed, we fail to find a systematic relation between their sum and occupation growth (see fig. C.1).

Occupation growth directly enters the third term of equation (9), which we call the “*growth selection effect*”. It is the product of net entry and the skill difference between marginal and inframarginal workers. Because that difference is negative for both entrants and leavers everywhere (see figure 3), occupation growth will determine its sign. The growth selection effect is negative for growing occupations, positive for shrinking occupations, and zero when there is no change in size. Growth selection thus formalizes and quantifies the intuition developed in section 2, whereby *the more an occupation grows, the more net entry of less skilled workers it experiences*.

The bottom panels of figure 7 plot employment growth selection. Both for entrants (fig. 7C) and leavers (fig. 7D) the relationship between employment growth and growth selection is strongly negative. Both contribute to the declining skills of growing occupations. Growth selection determines the entire negative slope between changes in employment and changes in skills from figure 6B (the location of the regression line and most of the variation around it stem from accumulation and churning). Therefore, the systematic part of the large selection effects we found in section 4.1 can be traced back to changes in occupations’ sizes multiplied with the skill gaps between marginal and inframarginal workers. An interesting observation is that in the model sketched in section B.1 of appendix B, based on Hsieh et al. (2019), rising skill prices attract workers of exactly offsetting lower quality to an occupation.

4.3. Contributors to Employment Growth Selection

Having established its quantitative importance, we now investigate further the nature of the employment growth selection effect. We decompose equation (9) in a way that allows us to gauge the relative contributions of age and of different worker types. That is, is growth selection primarily determined by entry of young workers who simply have not had a chance to accumulate a lot of skills? Or is it that entrants into a growing occupation have lower skills than incumbents had at the same point in their life cycle? Toward the end of this section, we alternatively consider the contributions of different types of switchers (i.e., other occupations, unemployment, etc).

References

Böhm, Michael J. (2020) “The Price of Polarization: Estimating Task Prices under Routine-Biased Technical Change,” *Quantitative Economics*, 11 (2), 761–799, [10.3982/QE1031](https://doi.org/10.3982/QE1031).

- Böhm, Michael J., Hans-Martin Von Gaudecker, and Felix Schran (2024) “Occupation Growth, Skill Prices, and Wage Inequality,” *Journal of Labor Economics*, 42 (1), 201–243, [10.1086/722084](#).
- Groes, Fane, Philipp Kircher, and Iourii Manovskii (2015) “The U-Shapes of Occupational Mobility,” *Review of Economic Studies*, 82 (2), 659–692, [10.1093/restud/rdu037](#).
- Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow (2019) “The Allocation of Talent and U.S. Economic Growth,” *Econometrica*, 87 (5), 1439–1474, [10.3982/ECTA11427](#).