

# Regional Labor Demand, Occupational Persistence and Social Mobility in Germany\*

Lennard Schlattmann <sup>†</sup>

November 12, 2024

## Abstract

Intergenerational persistence in occupational choice is one of the main drivers of income persistence and thus a key determinant of social mobility. In studying this phenomenon, the existing literature has focused exclusively on supply-side mechanisms by assuming that children can freely choose which occupation to work in. This paper, however, quantifies the importance of regional labor demand by merging regional and occupation-specific vacancy shares covering 400 regions and 436 occupations with the German Socioeconomic Panel. The main result indicates that controlling for labor demand reduces occupational persistence by up to 10 percent, the same magnitude as controlling for children's education does. Furthermore, I find that the degree of occupational persistence is relatively homogeneous across parent-child combinations. A child is about 20 times more likely to work in a given occupation if the parent works in that occupation. Finally, children who follow into their parents' occupation experience an initial wage premium of about 6 percent and their annual unemployment risk is halved.

**Keywords:** Social Mobility, Spatial Economics, Occupational Choice, Labor Market Outcomes

**JEL:** J23, J24, J61, J62

---

\*I want to thank Moritz Kuhn, Pavel Brendler, and Christian Bayer for their guidance and support throughout this project. I also would like to thank seminar participants at the university of Bonn and various other institutions for suggestions that have substantially improved this paper. I gratefully acknowledge financial support from the Bonn Graduate School of Economics and the DFG research training group 2281 "The Macroeconomics of Inequality". All errors are my own.

<sup>†</sup>University of Bonn and ECONtribute, lennard.schlattmann@uni-bonn.de, 53113 Bonn, Germany

# 1 Introduction

When entering the labor market, many children choose to work in their parents' occupation.<sup>1</sup> Despite a long tradition of documenting this intergenerational persistence in occupational choice, the underlying mechanisms of this behavior are still not fully understood. Identifying the determinants of this persistence has important implications along several dimensions. First, occupational persistence is a major driver of income persistence and hence social mobility. Therefore, if society wants to improve social mobility, we need to understand the causes of occupational persistence. Second, high levels of occupational persistence may reflect an inefficient allocation of talent, as individuals may not be allocated to occupations based on their skills, but rather based on information networks or even nepotism. So what drives occupational persistence?

To answer this question, the literature has focused exclusively on supply-side effects, such as the transmission of abilities and parental information networks as mechanisms (Kramarz and Skans, 2014; Lo Bello and Morchio, 2022), assuming that children can freely choose which occupation to work in. In reality, however, children only find a job in a particular occupation if there is an employer looking for a worker in that occupation. Thus, the main contribution of this paper is to quantify the importance of demand-side effects. To do so, I use annual vacancy data for each combination of 400 regions and 436 occupations, based on all vacancies which are reported to the German Federal Employment Agency, about 600,000 per year. I show that conditioning on regional labor demand reduces the level of intergenerational persistence in occupational choice<sup>2</sup> by about 5 to 10 percent, the same order of magnitude as conditioning on children's years of education does. The importance of regional labor demand stems from two empirical observations. First, there is a high degree of regional heterogeneity in labor demand across occupations in Germany. Second, workers are highly immobile and many work in the same regional labor market as their parents.

The second contribution is to document that the level of persistence increases with the fineness of the classification of occupations and that the persistence level is similar for different parent-child combinations. At the finest classification, the probability that a son works in a particular occupation increases by factors of 17 and 24, if his mother or father works in that occupation, respectively. The third contribution is to show that children who follow into their parent's occupation experience a wage premium of about 6 percent during the first two years of their first

---

<sup>1</sup>This following behavior is well documented in the literature. Examples include studies of specific occupations, such as physicians (Lentz and Laband, 1989) or pharmacists (Mocetti, 2016), aggregated occupational groups, such as liberal professions (Aina and Nicoletti, 2018), and entire countries, such as the United States (Blau and Duncan, 1967) or the United Kingdom (Lo Bello and Morchio, 2022).

<sup>2</sup>Throughout the paper, persistence is defined as the percentage increase in the probability of working in a given occupation if one's parent works in that occupation relative to someone whose parent does not work in that occupation. A more formal definition is provided in Section 2.

full-time job. Further, their annual unemployment risk is halved from about 7 to 3.5 percent during the first five years, and both effects disappear with job tenure.

Methodologically, this paper uses regional and occupation-specific vacancy shares as proxies for regional labor demand, by merging them with the German Socio-Economic Panel (SOEP). In total, I consider annual vacancy shares for each combination of 400 regions and 436 occupations, defined by the ISCO08 4-digit classification. This results in 174,400 different shares, calculated based on all vacancies reported to the German Federal Employment Agency, about 600,000 per year. Assuming a logistic model, I first compute persistence levels without and then with conditioning on vacancy shares. Comparing both values gives a measure of the importance of regional labor demand for persistence in occupational choice. Intuitively, a miner's son growing up in Bottrop, a city in the industrial Ruhr area where 9.4 percent of full-time workers were employed as miners and quarriers in 2013, might become a miner not because of his father's information networks or the transmission of abilities but simply because the labor demand for miners in Bottrop is particularly high. This labor demand channel seems to be particularly relevant because this paper documents first a high degree of heterogeneity in regional labor demand for different occupations in Germany and second a low labor market mobility, as in my sample between 82 and 86 percent of full-time working children still live in the same region as their parents. To put the importance of regional labor demand into context, I repeat the same exercise but calculate the level of persistence conditional on the children's education level. In this way, I compare the regional labor demand channel with the education channel, which has been shown to be important for social mobility (Caucutt and Lochner, 2020; Heckman and Mosso, 2014). The finding that regional labor demand is a relevant channel for persistence is further supported by proxying regional labor demand with regional, occupation-specific employment shares, which is a more robust measure.

In addition to quantifying the importance of regional labor demand for persistence, this paper examines the impact of following into the parent's occupation on wages and unemployment risk. Evaluating these effects for different subgroups with respect to the child's gender, education level, and tenure allows me to identify heterogeneous effects along these dimensions. First, the wage premium and the reduced unemployment risk for occupational followers fade out 2 and 5 years after labor market entry, respectively. Second, both impacts are mainly driven by those individuals without a high school degree, suggesting that this subgroup is particularly benefiting from following, at least in the short run. Third, while daughters seem to benefit slightly more in terms of the initial wage premium, sons benefit more in terms of the reduced unemployment risk.

There are two branches of literature this work relates to. First, there are papers showing that heterogeneity in local labor markets has important implications along several economic dimensions. It results in workers in different areas earning different wages even after adjusting for costs of living (Moretti, 2010) or facing different unemployment risks and employment

opportunities (Bilal, 2021; Kline and Moretti, 2013; Kuhn et al., 2021). Typically, these local differences are explained by agglomeration effects of productive firms in certain areas (Bilal, 2021) and by more efficient worker-firm matching (Dauth et al., 2022). This paper contributes to this literature by showing that local labor markets also play an important role for intergenerational persistence in occupational choice and hence for social mobility.

Second, there is a literature on how networks impact occupational decisions. Starting from the seminal work by Granovetter (1973) the literature has shown that networks (Brown et al., 2016; Dustmann et al., 2016; Glitz and Vejlin, 2021) and especially strong social ties, like parents (Kramarz and Skans, 2014), are important for occupational choice. Consequently, a substantial share of children follows into their parent’s occupation (Aina and Nicoletti, 2018; Blau and Duncan, 1967; Lentz and Laband, 1989; Lo Bello and Morchio, 2022; Mocetti, 2016). While this following behavior has been well established, the key underlying channels are far less clear. My paper contributes here on the labor demand as well as on the labor supply-side. For the former, it is the first to document and quantify the importance of regional labor demand for intergenerational persistence in occupational choice. For the latter, it extends the supporting evidence in favor of a popular theory developed by Dustmann et al. (2016), which suggests that information networks rather than the transmission of abilities are the main driver of occupational persistence on the labor supply-side, in two dimensions. First, I show that their model predictions of an initial wage premium for occupational followers which fades out with tenure are supported when defining following based on finely classified occupations instead of firms. This definition is arguably more suitable for studying the intergenerational transmission of abilities and skills as they are most likely not firm but rather occupation-specific. Second, I further show that their predictions still hold for the type of network which has been shown to be most important for labor market outcomes: immediate family members.

The remainder of this paper is structured as follows: Section 2 introduces the data and provides descriptive results before Section 3 describes the empirical strategy applied. Thereafter, Section 4 provides the empirical results, which are discussed, related to the literature, and supported by robustness checks in Section 5. Finally, Section 6 concludes.

## 2 Data and descriptive results

This section first introduces the different data sets employed and the most important definitions used in this paper. Thereafter, the data cleaning process is described and descriptive results are shown.

## 2.1 Data and Definitions

This paper makes use of two data sources. First, it applies data from the German Socioeconomic Panel (SOEP), Wave 37, which spans a time period from 1984 to 2020. The SOEP provides information on labor market outcomes such as wages, occupations, and unemployment spells as well as a large set of other socioeconomic variables. While being representative on the national level, one drawback of the SOEP is that it is not representative at the county level. Thus, in order to obtain representative regional, occupation-specific vacancy shares, which will serve as proxies for regional labor demand, this paper additionally uses data from the German Federal Employment Agency (Bundesagentur für Arbeit, BA). This BA data set includes annual vacancy and employment shares according to the ISCO-08 classification up to the 4-digit level. Vacancy shares are based on all vacancies reported to the BA between 2011 and 2020, around 600,000 each year, and employment shares are computed based on all employees in Germany who are subject to mandatory social insurance contributions from 2013 to 2020, around 30 million employees each year. Merging this rich data set with the SOEP allows to control for regional and occupation-specific labor demand when studying intergenerational persistence in occupational choice.

In order to quantify the degree of occupational persistence, the most natural measure would be to simply compute the percentage shares of children who work in the same occupation as at least one of their parents.<sup>3</sup> Only looking at percentage shares for quantifying persistence, however, has at least two important drawbacks. First, it neglects occupational concentration. If everyone worked in one occupation, the percentage share of followers would be equal to one, even though this high persistence might not be driven by actual following behavior. Second, it does not account for differences in the distributions of children's and parents' occupations. For a given following behavior percentage shares of followers mechanically decrease when the distributions of both subgroups differ substantially.

Both drawbacks are not present when using likelihood ratios instead of percentage shares. Hence, following the definition typically used for diagnostic inference in medicine (Doi et al., 2022), I define the likelihood ratio of occupation  $j$  as

$$LR_j = \frac{P(o = j \mid o^P = j)}{P(o = j \mid o^P \neq j)}, \quad (1)$$

where  $o$  and  $o^P$  refer to the child's and parent's occupation, respectively.<sup>4</sup> Thus, the likelihood ratio of occupation  $j$  is given by dividing the probability of a child working in occupation  $j$  if the parent is working in this occupation by the probability of doing so if the parent is not

<sup>3</sup>Table 9 in the Appendix depicts these statistics.

<sup>4</sup>Dal Bó et al. (2009) use a similar measure which they call *dynastic bias* and which is defined by dividing the conditional probability of a father working in a given occupation if the son works in this occupation by the unconditional probability of a father working in this occupation.

working in this occupation. Making the connection to the medicine literature, one can think of children with a parent working in occupation  $j$  as being treated, while those without a parent working in occupation  $j$  are untreated. Hence, the goal is to compare the "risk" of working in occupation  $j$  between these two subgroups. Intuitively, a likelihood ratio of one would indicate that children's occupations are independent from their parents' occupations, while a ratio of 2 implies that the probability of working in a given occupation doubles if one has a parent working in this occupation compared to someone whose parent does not work in this occupation. This definition slightly differs from the one applied by Lo Bello and Morchio (2022) who divide the conditional probability of working in a given occupation if the parent works in this occupation by the unconditional probability of working in this occupation.

In order to obtain an aggregate measure of persistence, likelihood ratios are weighted by the number of children working in each occupation. Formally, the weighted likelihood ratio is defined as

$$WLR = \sum_{j=1}^J w_j LR_j \quad (2)$$

where  $J$  is the number of different occupations and the weights  $w_j$  are based on the number of children working in occupation  $j$  and defined as

$$w_j = \frac{n_j}{N} \quad (3)$$

where  $N$  and  $n_j$  refer to the total sample size and the number of children working in occupation  $j$ , respectively.

Next to the likelihood ratios, this paper will refer to the excess probability as measure of persistence which is simply defined by subtracting one from the likelihood ratios. Intuitively, the excess probability indicates the percentage increase in the probability of working in a given occupation if one's parent works in that occupation relative to someone whose parents does not work in this occupation.

## 2.2 SOEP - Descriptive Results

Out of the total sample size of 742,822 person-year observations, I drop those who are not in the relevant age group, hence those below 16 and above 65 years, and those who are not in the labor force, like students or individuals caring for children or elderly, which leaves 490,535 person-year observations.<sup>5</sup> Since I am interested in intergenerational links, I only keep those individuals who have a parent in this subsample which leaves 107,524 observations. Finally, I exclude all individuals whose parent does not have any information on their occupation which generates the baseline data set including 78,089 observations. Furthermore, for the analysis of

---

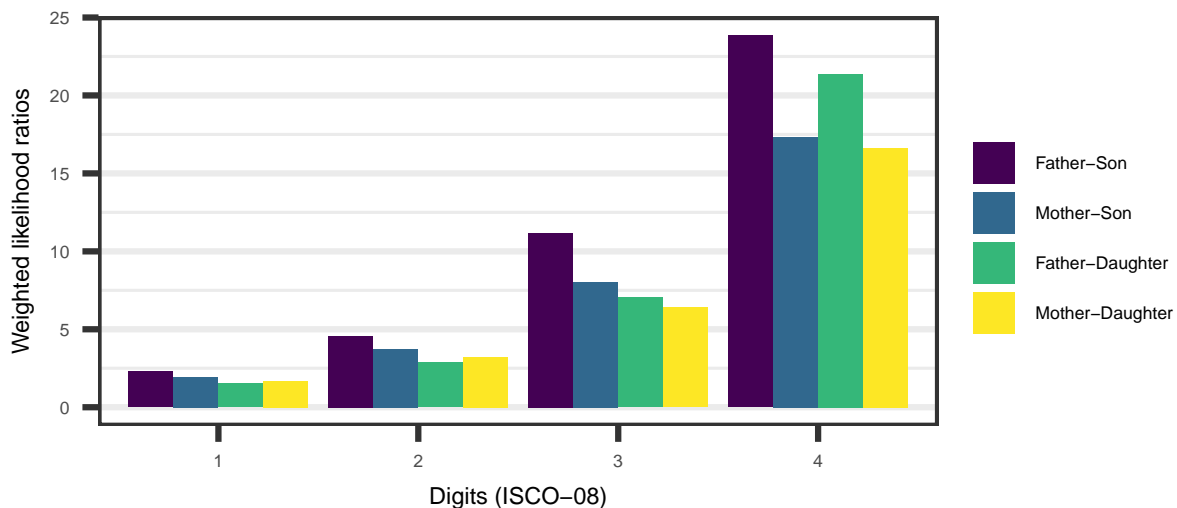
<sup>5</sup>Tables 7 and 8 in the Appendix list the different steps in a table.

regional labor demand, I can only employ SOEP data from 2011 onward because vacancy shares are only available from 2011 to 2021, which sums up to 28,620 person-year observations. For studying aspects which do not require regional information the baseline SOEP sample is used.

Splitting the baseline SOEP sample by gender reveals that sons make up around 56 percent and are thus slightly overrepresented. Note that children are followed up until their parents retire. Hence, some children might start their own family during this period, which might be a reason of daughters being underrepresented, as they leave the labor market after starting an own family to a larger share. Robustness checks are conducted only focusing on those children who still live in the same household as their parents because parent-child ties are arguably stronger for this subgroup.<sup>6</sup>

Figure 1 depicts the weighted likelihood ratios following the ISCO-08 classification<sup>7</sup> for all four parent-child combinations and levels of occupational disaggregation, respectively.<sup>8</sup> The ISCO-08 classification includes 10, 43, 130, and 436 different occupations for the digit levels 1 to 4.

Figure 1: Weighted likelihood ratios of occupational followers



*Notes:* Weighted likelihood ratios are computed according to Equation (2) including the period from 1984 to 2020.

This figure clearly indicates that persistence increases in the fineness of classifying occupations. The purple bar on the very left for example shows that sons were around 2.5 times as likely to work in a given occupation, defined at the 1-digit level, if their father also worked in this

<sup>6</sup>This subgroup amounts to 43,823 children-year observations and thus includes around 55 percent of the baseline sample.

<sup>7</sup>The SOEP data provides occupational information according to the ISCO-08 classification only from 2012 onward. Before 2012 it provides this information based on the ISCO-88 classification which I translated to the ISCO-08 classification based on schemes provided by the International Labour Organization.

<sup>8</sup>The exact numbers are shown in Tables 14 and 15 in the Appendix.

occupation relative to a son whose father worked in another occupation.<sup>9</sup> Examples of this broad level of disaggregation are the categories of *Legislators, senior officials, and managers* or *Craft and related trades workers*. For the finest level of classification, sons were more than 23 times as likely to work in a given occupation if their father worked in this occupation. Examples for this level of disaggregation are occupations like *Dentists, Roofers* or *Miners and Quarriers*. So moving from classifications of 1-digit to 4-digit increases persistence measured by weighted likelihood ratios by a factor of 10.

While this increase in persistence for finer occupational classifications is not too surprising, what, however, might surprise is the relatively high level of homogeneity in persistence across parent-child combinations. For the finest classification the probability for a son to work in a given occupation increases by factors of 24 or 22 if his father or his mother works in this occupation, respectively. For daughters the respective numbers are 18 and 17. Furthermore, while persistence is highest for father-son pairs for all levels of occupational disaggregation, there is no clear ordering for the remaining three pairs. the differences are not substantial. Interestingly, these levels of persistence are not driven by self-employment as Table 11 in the Appendix shows.

In order to compare these values with those documented by Lo Bello and Morchio (2022) for fathers and sons in the United Kingdom between 1991 and 2008, there are three adjustments to make. Next to changing the definition of likelihood ratios by dividing by the unconditional probability of a son/daughter working in a given occupation, I follow their assumptions and exclude soldiers and classify unemployed individuals according to the occupation of their next job. Doing so yields a weighted likelihood ratio of 2.39 for father-sons pairs at the 1-digit level which is still considerably larger than the 1.72 documented in Lo Bello and Morchio (2022), suggesting a relatively high persistence in occupational choice in Germany.<sup>10</sup>

## 2.3 BA data - Descriptive Results

In order to include information about regional labor market characteristics, I merge the SOEP data with regional and occupation-specific employment and vacancy shares provided by the German Federal Employment Agency (Bundesagentur für Arbeit, BA). This data set provides annual employment and vacancy shares for 400 German regions and occupations based on the ISCO-08 classification up to 4-digit. This finest level of disaggregation includes 436 different occupations and thus provides me with 174,400 shares. Annual vacancy shares are computed based on all vacancies reported to the BA, around 600,000 each year, while employment

---

<sup>9</sup>Tables 12 and 13 in the Appendix depict the likelihood ratios for each occupation defined at the 1-digit level separately.

<sup>10</sup>Table 10 in the Appendix shows all weighted likelihood ratios after these adjustments. Unfortunately, the value for father-son pairs at the 1-digit classification is the only number that can be compared as Lo Bello and Morchio (2022) only look at father-son pairs, define the 2-digit level slightly finer, and do not have data for finer classifications.



shares are computed based on all employees who are subject to mandatory social insurance contributions, around 30 million individuals each year. As vacancy shares are potentially a very volatile proxy for regional labor demand, I conduct checks using employment shares which are a more robust measure.

In order for regional labor demand to play a role for intergenerational persistence in occupational choice and hence for social mobility, two conditions have to be fulfilled. First, there has to be some heterogeneity in regional labor demand for different occupations across Germany because otherwise regions would be irrelevant for occupational choice. Second, a large share of children need to live in the same regions as their parents do, such that they face the same labor demand.

Figure 2: Occupation with highest vacancy rate in 2020 (ISCO-08, 2-digit)

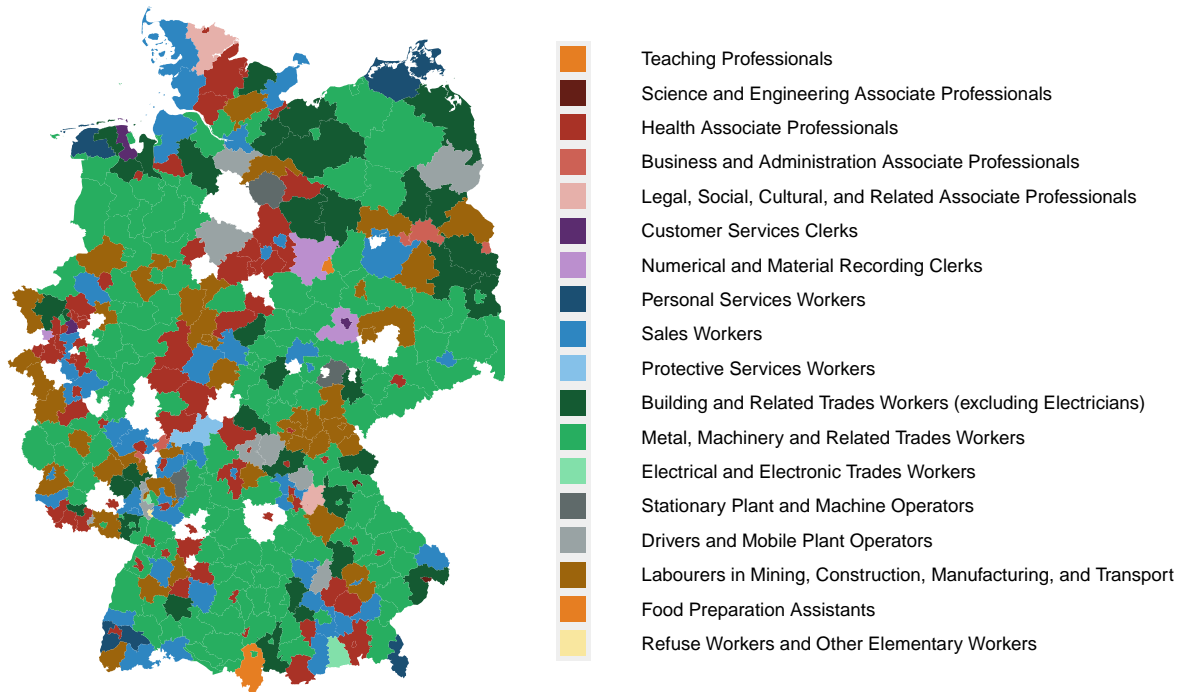


Figure 2 reveals a substantial heterogeneity in regional labor demand across Germany and hence provides evidence that the first condition is satisfied. It depicts the occupation with the highest number of vacancies, defined at the ISCO-08 2-digit level, for every region in Germany in 2020. Out of all 43 occupations of this classification, 18 have the highest number of vacancies in at least one region. One aspect becoming visible in this figure are the different mining locations in Germany, being colored in brown. For example, in the very West the Rhenish coal district is

easily detectable. Hence, depending on the region one is living in, the regional labor demand for a given occupation can differ substantially.

Regarding the second condition, I find that depending on the parent-child combination between 82 and 86 percent of children in my sample live in the same region as their parents. This finding of workers being immobile, especially when having family ties in a given region, is consistent with what the literature has shown (Huttunen et al., 2018). Combining the high level of regional heterogeneity in labor demand across Germany with the immobility of workers suggests that regional labor demand potentially plays an important role for occupational choice and hence for social mobility. Intuitively, a son of a miner in a coal district might not necessarily become a miner himself because of his father’s occupation but simply because it is the predominant occupation in this region. The next section describes how I am planning to quantify this channel.

### 3 Empirical Strategy

This section outlines the empirical strategies used for quantifying the relevance of regional labor demand for occupational persistence as well as for studying the impact of following into parents occupations on wages and unemployment risk .

#### 3.1 Importance of regional labor demand for persistence

In order to quantify the importance of regional labor demand for intergenerational persistence in occupational choice, I compute weighted likelihood ratios conditioning on regional and occupation-specific vacancy shares using a logistic model. More formally, the regression reads

$$\text{logit}(\mathbb{1}(o_{i,t} = j)) = \alpha + \beta_1 \mathbb{1}(o_{i,t}^P = j) + \beta_2 \theta_{i,t} + \epsilon_{i,t}, \quad (4)$$

where  $\mathbb{1}(o_{i,t} = j)$  and  $\mathbb{1}(o_{i,t}^P = j)$  are functions indicating whether child  $i$  and its parent are working in occupation  $j$  in year  $t$ , respectively.  $\theta_{i,t}$  is the vacancy share of occupation  $j$  in the region children  $i$  lives in year  $t$ .<sup>11</sup> Running these regressions, I am able to compute the predicted values of working in a specific occupation, conditional on the parent’s occupation and the regional occupation-specific vacancy share.

Thus, the likelihood ratio of occupation  $j$  conditioning on the regional labor demand, formally

$$LR_j^{vac} = \frac{P(o = j \mid o^P = j, \theta)}{P(o = j \mid o^P \neq j, \theta)}, \quad (5)$$

---

<sup>11</sup>As I am computing likelihood ratios separately for each of the four parent-child combinations, I am also account for potential gender differences.

isolates the part of persistence which is independent of regional labor demand. A prominent example would be parental information networks or the transmission of abilities. Comparing the conditional and unconditional weighted likelihood ratios gives a measure of how important regional labor demand is for persistence. Formally, the percentage reduction in persistence after conditioning on regional labor demand is given by

$$\theta^{share} = \frac{WLR - WLR^{vac}}{WLR - 1}. \quad (6)$$

Thus,  $\theta^{share}$  describes the percentage reduction in persistence after conditioning on regional labor demand.

In order to put this number into context, I repeat the same exercise but now quantify the importance of children's education level on persistence. Education was chosen as comparison because the literature has shown that it is a key determinant for intergenerational persistence. Similarly as before, I defined the share of persistence generated by education as

$$educ^{share} = \frac{WLR - WLR^{educ}}{WLR - 1}. \quad (7)$$

where  $WLR^{educ}$  is computed by weighting

$$LR_j^{educ} = \frac{P(o = j \mid o^P = j, e)}{P(o = j \mid o^P \neq j, e)}, \quad (8)$$

where  $e$  describes children's education level measured in years of education.

### 3.2 Impact of following on wages

Besides quantifying the impact of regional labor demand on persistence, I further document how following into the parent's occupation impacts labor market outcomes. In particular, I provide evidence on the impact on actual hourly gross wages and the annual risk of becoming unemployed for all parent-child combinations and years since the first full-time job. In order to identify the impact of working in the same occupation on wages, I run the following regression

$$\log(wage_{i,t}) = \alpha + \beta follower_{i,t} + \gamma X_{i,t} + \epsilon_{i,t},$$

where  $wage_{i,t}$  is the actual hourly gross wage and  $follower_{i,t}$  is an indicator function being equal to one if child  $i$  works in the same occupation as at least one of its parents in period  $t$ . Since the SOEP data provides occupational information according to the ISCO-08 classification only for 2013 onward, while it provides information according to the ISCO-88 classification for 1984 to 2017, occupations are defined according to the latter using the 4-digit classification.  $X$  is a set

of control variables consisting of a third-degree age polynomial, gender, marital status, years of education, region, migration background as well as a second-order tenure polynomial, the firm size, and survey year. Importantly,  $X$  further includes the occupation.<sup>12</sup> Thus, the estimated coefficients hold within occupations. As I do not employ the BA data for this regression, I can apply the whole baseline SOEP sample.

In order to focus on children’s first major job and abstract from potential smaller vacation or college jobs, I follow a similar strategy as proposed by Topel and Ward (1992). While they define labor market entry as the first quarter in which earnings exceed 70 percent of the minimum wage for full-time work given that the following four quarters do also satisfy this condition, I have to slightly deviate as Germany only introduced its first statutory minimum wage in 2015. Thus, I define labor market entry as the first year in which individuals work full-time given that they continue to do so in the following year.

### 3.3 Impact of following on unemployment risk

Next to differences in wages, I also study whether following into the parent’s occupation impacts the probability of becoming unemployed. For doing so, I run the following regression

$$\mathbb{1}(u_{i,t+1} = 1) = \alpha + \beta \text{follower}_{i,t} + \gamma X_{i,t} + \epsilon_{i,t} \quad \text{if } \mathbb{1}(e_{i,t} = 1),$$

where  $\mathbb{1}(u_{i,t+1} = 1)$  and  $\mathbb{1}(e_{i,t} = 1)$  indicate individuals who are full-time employed in period  $t$  but become unemployed in the period thereafter,  $t + 1$ . As before  $\text{follower}_{i,t}$  is equal to one if children  $i$  works in the same occupation as at least one of its parents and  $X_{i,t}$  is the same set of control variables as in the wage regression. Hence,  $\beta$  describes the impact of working full-time in the same occupation as at least one of your parents on the likelihood of becoming unemployed next period relative to someone who works full-time in another occupation.

## 4 Results

This section presents the empirical results. First, it quantifies the relevance of regional labor demand for intergenerational persistence in occupational choice before studying how children with and without parents working in the same occupation differ with respect to hourly wages and unemployment risk.

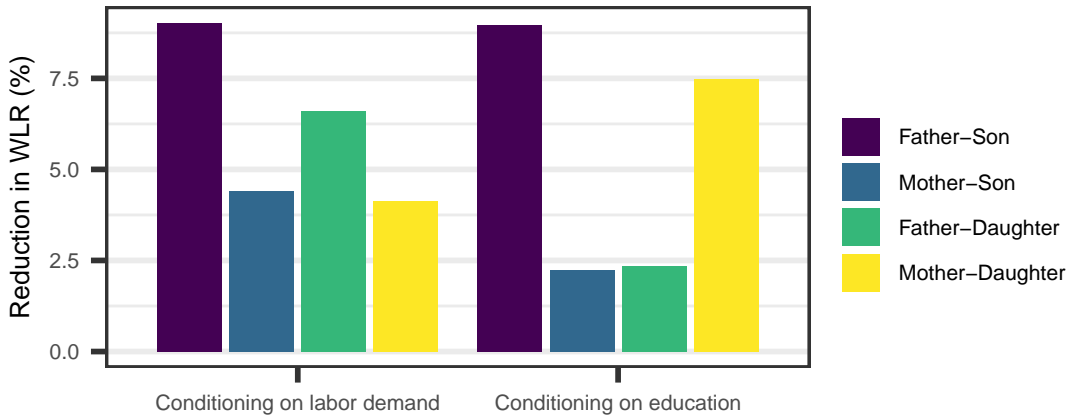
---

<sup>12</sup>This set of control variables is the same as the one used in Lo Bello and Morchio (2022), except that I exclude smoking behavior due to data limitations.

## 4.1 Importance of regional labor demand for persistence

Figure 3 depicts the percentage reductions in persistence after conditioning on regional labor demand and education, as defined in equations (6) and (7). Conditioning on regional labor demand, proxied by regional and occupation-specific vacancy shares, reduces persistence by around 5 to 10 percent. Conditioning on education, measured in years of schooling, generates persistence levels of a similar magnitude. While for father-son pairs it seem to be particularly important and explains around 13 percent, it is less important for cross-gender following. This result suggests that regional labor demand is a relevant factor for persistence in occupations and thus for social mobility. But it also shows that the largest share of persistence results from other characteristics, potentially information networks.

Figure 3: Reduction in WLRs after conditioning on regional labor demand and education



*Notes:* For this figure, occupations are defined at the 4-digit level. As vacancy data is only available from 2011 onward, I also restrict the analysis for education on this subsample to guarantee comparability.

## 4.2 Impact of following on wages

Table 1 shows the impact of working in the same 4-digit occupation as at least one parent on the actual gross hourly wage depending on the time horizon since the first full-time job. Occupational followers have a significant wage premium of around 5.6 percent during the first two years, which fades out thereafter. As the mean actual gross hourly wage for this period and subgroup was around 10 Euros, this premium translates into an increased hourly wage of around 50 Cent. Thus, monthly and yearly gross income for occupational followers is around 80 and 1,000 Euros higher, respectively. When splitting the sample by education into those with and without high school degree, the coefficients indicate that the wage premium is almost entirely driven by the lower educated subsample, which has a highly significant initial wage

premium of around 8 percent. Finally, looking at sons and daughters separately indicates that the impact for daughters seems to be slightly stronger.

Table 1: Impact of following on gross hourly wages

Specification	N	Coefficient				
		All	No HS	HS	Sons	Daughters
First 2 y. after first FT job	11,551	0.056** (0.025)	0.082*** (0.031)	0.032 (0.038)	0.052 (0.032)	0.068* (0.039)
2 - 5 y. after first FT job	9,468	-0.018 (0.026)	0.011 (0.060)	-0.016 (0.027)	-0.035 (0.035)	0.046 (0.041)
6 - 10 y. after first FT job	9,644	-0.020 (0.025)	0.121 (0.076)	-0.041 (0.027)	-0.048 (0.034)	0.047 (0.034)
> 10 y. after first FT job	11,056	0.011 (0.017)	0.069 (0.065)	0.004 (0.017)	0.021 (0.020)	0.027 (0.026)

*Notes:* This figure shows the effect of working in the same occupation as at least one of the parents on actual hourly gross wages. Occupations are defined according to the ISCO-88 4-digit and the ISCO-08 4-digit classifications for the time spans 1984 to 2012 and 2013 to 2020, respectively. Standard errors are clustered on the individual level and written in parentheses.

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

### 4.3 Impact of following on unemployment risk

Table 2 depicts the results with respect to unemployment risk. Occupational followers have a significantly lower risk of becoming unemployed during the first five years after their first full-time job. Considering that the average probability of becoming unemployed for this group is 5 percent, the reduction by 2.2 percentage points almost halves children's unemployment risk.

Table 2: Impact of following on unemployment risk

Specification	N	Coefficient				
		All	No HS	HS	Sons	Daughters
First 2 y. after first FT job	9,708	-0.021*** (0.007)	-0.032*** (0.012)	-0.017** (0.009)	-0.027*** (0.009)	-0.013 (0.012)
2 - 5 y. after first FT job	7,814	-0.022*** (0.006)	-0.039* (0.020)	-0.021*** (0.006)	-0.033*** (0.008)	-0.007 (0.012)
6 - 10 y. after first FT job	8,461	-0.006 (0.008)	0.000 (0.019)	-0.005 (0.009)	-0.019* (0.011)	0.002 (0.014)
> 10 y. after first FT job	9,552	-0.006 (0.009)	-0.043 (0.034)	0.001 (0.009)	-0.006 (0.012)	-0.008 (0.014)

*Notes:* This figure shows the effect of working in the same occupation as at least one of the parents on the probability of becoming unemployed the following year. Occupations are defined according to the ISCO-88 4-digit and the ISCO-08 4-digit classifications for the time spans 1984 to 2012 and 2013 to 2020, respectively. Standard errors are clustered on the individual level and written in parentheses.

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

When splitting the sample by education, we again see that those without a high school degree do also benefit most from following, as their unemployment risk decreases by 3.2 to 3.9 percentage points. The average unemployment risk for this group is around 7 percent. While in terms of wages daughters seem to benefit more from following, sons benefit more with respect to unemployment risk. While for daughters the risk only decreases insignificantly, for sons the annual unemployment risk is reduced by around 3 percentage points.

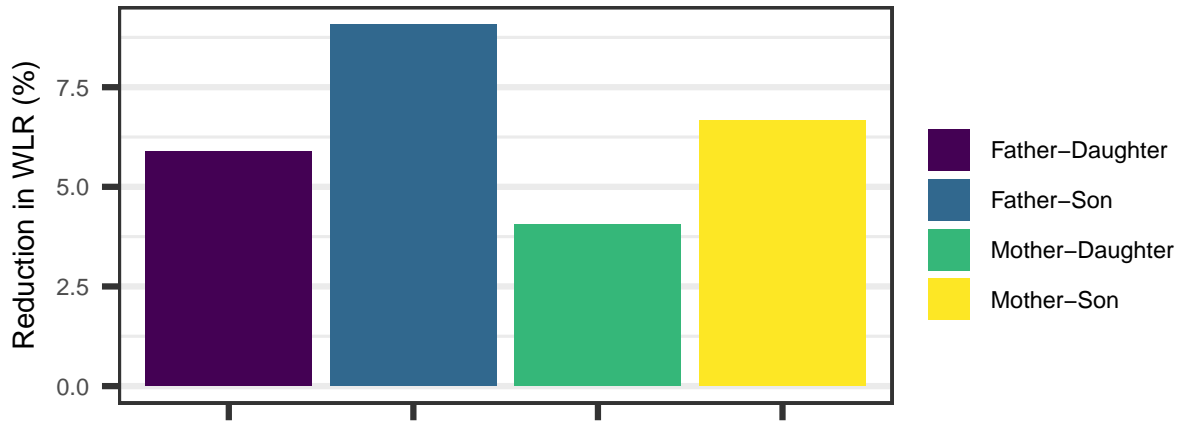
## 5 Robustness checks and discussion

In order to examine the robustness of my results, this section conducts several checks. Thereafter follows a short discussion about the implications of these results.

### 5.1 Robustness checks

In order to ensure the plausibility and robustness of my main results, I conduct several checks. First, with respect to quantifying the importance of regional labor demand for persistence, I repeat the same exercise described in Section 3.1 but now I condition on employment instead of vacancy shares. The idea is that vacancy shares might be an imperfect measure for contemporaneous labor demand, as they are volatile and might thus not mirror actual labor market opportunities in a given occupation. Employment shares, however, are a much more stable measure and hence complement vacancy shares as proxy for labor demand. Figure 4 shows that resulting percentage reductions on persistence. For all four parent-child pairs, the reductions are very similar to the ones found using vacancy shares supporting the robustness of my main result.

Figure 4: Reduction in WLRs after controlling for employment shares



*Notes:* For this figure, occupations are defined at the 4-digit level. As employment data is only available from 2013 onward, this analysis is restricted to this subsample.

In order to further investigate the impact of following into the parent's occupation, I conduct two checks. First, in order to check the robustness of the results, I exclude all children of self-employed parents as the wage premium and the reduced unemployment risk might be driven by self-employed who employ their own children. Tables 3 and 4, however, show that this is not the case. When excluding children of self-employed, the initial wage premium for followers even increases to 7.1 percent and becomes significant to the one-percent significance level. Interestingly, the impact for children without a high school degree is slightly reduced, while the one for those individuals with a high school degree is stronger. This observation might indicate that lower educated children earn higher wages when working for self-employed parents, while the opposite is true for higher educated children. With respect to the reduction in unemployment risk, there are no substantial changes observable.



Table 3: Impact of following on wages without children of self-employed

Specification	N	Coefficient				
		All	No HS	HS	Sons	Daughters
First 2 y. after first FT job	10,261	0.071*** (0.025)	0.065** (0.032)	0.071* (0.038)	0.065* (0.034)	0.089** (0.038)
2 - 5 y. after first FT job	8,145	-0.002 (0.024)	-0.032 (0.050)	0.021 (0.026)	-0.012 (0.034)	0.046 (0.036)
6 - 10 y. after first FT job	7,975	0.018 (0.021)	0.064 (0.064)	0.006 (0.022)	0.013 (0.028)	0.049 (0.032)
> 10 y. after first FT job	9,127	0.020 (0.016)	0.075 (0.050)	0.015 (0.016)	0.044** (0.019)	0.020 (0.025)

*Notes:* This figure shows the effect of working in the same occupation as at least one of the parents on actual hourly gross wages excluding children of self-employed. Occupations are defined according to the ISCO-88 4-digit and the ISCO-08 4-digit classifications for the time spans 1984 to 2017 and 2018 to 2020, respectively. Standard errors are clustered on the individual level and written in parentheses.

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 4: Impact of following on unemployment risk without children of self-employed

Specification	N	Coefficient				
		All	No HS	HS	Sons	Daughters
First 2 y. after first FT job	8,460	-0.017** (0.008)	-0.028** (0.014)	-0.013 (0.010)	-0.028** (0.011)	-0.008 (0.014)
2 - 5 y. after first FT job	6,618	-0.020*** (0.007)	-0.022 (0.024)	-0.023*** (0.006)	-0.029*** (0.010)	-0.007 (0.013)
6 - 10 y. after first FT job	6,884	-0.002 (0.010)	0.008 (0.023)	-0.001 (0.013)	-0.007 (0.014)	0.003 (0.017)
> 10 y. after first FT job	7,718	0.004 (0.011)	-0.030 (0.040)	0.009 (0.011)	0.008 (0.014)	-0.001 (0.017)

*Notes:* This figure shows the effect of working in the same occupation as at least one of the parents on the probability of becoming unemployed next year excluding children of self-employed. Occupations are defined according to the ISCO-88 4-digit and the ISCO-08 4-digit classifications for the time spans 1984 to 2012 and 2013 to 2020, respectively. Standard errors are clustered on the individual level and written in parentheses.

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Second, in order to check the plausibility of the result, I restrict the sample to only those children who still live in their parents' household. This subgroup amounts to more than 52,000 children-year observations and thus includes around 55 percent of the baseline sample. The idea is that the relationship of these children to their parents is arguably closer and, if rather parental information networks than the transmission of innate abilities are the main driver of persistence, we would thus expect stronger results. Tables 5 and 6 show that this is indeed the case. The initial wage premium rises to 6.8 percent and becomes significant at the one-percent level, supporting the plausibility of the presented results. While the impact on unemployment risk during the first two years hardly changes, it becomes stronger during years two to five and

risers from 2.2 to 3.2 percentage points. Interestingly, the difference in impacts for sons and daughters also increases suggesting that the unemployment risk for the former decreases by 5.1 and for the latter by 1.4 percentage points during this period. This result that the impacts of following into the parent's occupation become stronger when parent-child ties are closer also supports rather a theory of parental information networks than of transmitting innate abilities.

Table 5: Impact of following on wages - only children living in parents' household

Specification	N	Coefficient				
		All	No HS	HS	Sons	Daughters
First 2 y. after first FT job	10,294	0.068*** (0.023)	0.079*** (0.028)	0.061 (0.038)	0.056* (0.031)	0.097** (0.038)
2 - 5 y. after first FT job	6,413	-0.001 (0.030)	0.033 (0.056)	-0.002 (0.035)	-0.028 (0.041)	0.059 (0.046)
6 - 10 y. after first FT job	3,506	-0.054 (0.038)	0.038 (0.077)	-0.081* (0.046)	-0.056 (0.046)	0.046 (0.047)
> 10 y. after first FT job	2,680	-0.009 (0.043)	0.032 (0.081)	-0.023 (0.053)	-0.006 (0.054)	-0.041 (0.080)

*Notes:* This figure shows the effect of working in the same occupation as at least one of the parents on actual hourly gross wages for the subgroup of children who live in their parents' household. Occupations are defined according to the ISCO-88 4-digit and the ISCO-08 4-digit classifications for the time spans 1984 to 2017 and 2018 to 2020, respectively. Standard errors are clustered on the individual level and written in parentheses.

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 6: Impact of following on unemployment risk - only children living in parents' household

Specification	N	Coefficient				
		All	No HS	HS	Sons	Daughters
First 2 y. after first FT job	8,567	-0.020*** (0.008)	-0.033*** (0.012)	-0.014 (0.011)	-0.027*** (0.010)	-0.010 (0.014)
2 - 5 y. after first FT job	5,179	-0.032*** (0.007)	-0.067*** (0.021)	-0.023*** (0.008)	-0.051*** (0.009)	-0.014 (0.017)
6 - 10 y. after first FT job	2,955	-0.005 (0.016)	-0.041 (0.028)	-0.000 (0.022)	-0.013 (0.017)	0.033 (0.043)
> 10 y. after first FT job	2,124	0.005 (0.038)	-0.046 (0.069)	0.027 (0.048)	-0.033 (0.043)	0.115 (0.085)

*Notes:* This figure shows the effect of working in the same occupation as at least one of the parents on the probability of becoming unemployed next year for the subgroup of children who live in their parents' household. Occupations are defined according to the ISCO-88 4-digit and the ISCO-08 4-digit classifications for the time spans 1984 to 2012 and 2013 to 2020, respectively. Standard errors are clustered on the individual level and written in parentheses.

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

## 5.2 Discussion

My main result, indicating that regional labor demand is an important driver of intergenerational persistence in occupational choice and thus of social mobility, has potentially important implications on the discussion about the substantial decline in interstate mobility in the United States over the last decades (Kaplan and Schulhofer-Wohl, 2017; Molloy et al., 2011; Molloy et al., 2017). While in 1990 around 3 percent of Americans moved between states each year this number has fallen by half in the following 30 years (Kaplan and Schulhofer-Wohl, 2017). At least two prominent causes for this decline have been proposed in the literature, which have different implications for my main result on this discussion. On the one hand, Kaplan and Schulhofer-Wohl (2017) argued that interstate mobility fell because geographic specificity of returns to occupations declined and workers' ability to learn about other locations before moving both increased. Following this argument lower regional mobility does not imply lower social mobility because with labor markets becoming more equal social advancement is possible even without regional mobility. On the other hand, Cooke (2013) found that the fall in migration is correlated with more people living in dual-earner households which potentially made households less mobile. If this channel dominates and my results also hold for the United States, the reduction in regional mobility does translate into lower social mobility. Hence, a fruitful path for future research could be to study whether my results are generalizable for other countries, in particular for the United States, and what the consequences of lower regional mobility on social mobility in the United States are.

The results on the impact of following into the parent's occupation are qualitatively as well as quantitatively very similar to what has been found in the context of referrals more general (Brown et al., 2016; Burks et al., 2015; Dustmann et al., 2016; Glitz and Vejlin, 2021). For example, Glitz and Vejlin (2021) apply comprehensive Danish administrative data and show that workers who found their jobs through former coworkers' referrals earn initially 4.6 percent higher wages and are 2.3 percentage points less likely to leave their firm during the first year. They rationalize this finding with a learning model build in Jovanovic (1979) and further developed in Simon and Warner (1992) and Dustmann et al. (2016). In this model, firms can either hire new employees via the external or the referral market. The core idea is that for those hired via the latter uncertainty about match-specific quality is lower as referrals provide information which the firm and the employee otherwise would not have. As workers are partially insured against low realizations of their productivity by simply leaving the firm, potential wage growth is higher for employees hired via the external market. Consequently, referral hires have higher reservation wages and turn down wage offers that identical external hires would accept, resulting in an initial wage premium for those hired via referrals. Furthermore, this subgroup is on average better matched. As uncertainty about match-specific quality resolves over time, workers with low productivity realizations separate from the firm resulting in higher unemployment risk

for individuals hired via the external market, and both the wage premium and differences in unemployment risk fade out with tenure.

Not only do my results replicate these model predictions, the finding that the impact of following is particularly strong for children without a high school degree does also fit in this theory very well. As lower educated individuals have fewer school degrees and certificates as signals of their abilities, the uncertainty about the match quality is arguably higher and the impact of parental information thus stronger. Hence, my results are in line with the model predictions and therefore expand the supporting evidence in two dimensions. First, differently to the supporting literature so far, which defines following based on working in the very same firm (Brown et al., 2016; Burks et al., 2015; Dustmann et al., 2016; Glitz and Vejlin, 2021), this paper defines it based on working in the same finely classified occupation. This broader definition of followers has two advantages. First, it allows accounting for parental networks beyond a particular firm, which might be especially relevant for Germany with its large number of small craft businesses. Second, it is more suitable for studying the intergenerational transmission of abilities and skills as they are arguably not firm but rather occupation-specific. The second and probably more important dimension in which this paper extends the empirical evidence in favor of the theory developed in Dustmann et al. (2016) is that it looks at intergenerational occupational following behavior, and thus studies the type of social networks which has been shown to be most important for individual labor market outcomes: immediate family members (Eliason et al., 2022). The fact that my results are not only qualitatively but also quantitatively very similar to what has been found by Glitz and Vejlin (2021) for former coworkers further suggests that genes and thus the transmission of innate abilities are not of first-order importance for the following behavior. But even when agreeing on parental information networks being the main driver behind the persistence, the welfare consequences are not clear. On the one hand, the literature on referrals generally models this channel as welfare improving by reducing initial uncertainty about match-specific quality. For example, Dustmann et al. (2016) estimate that referrals, through the provision of additional information, increase total welfare in the economy by 0.62 percent. On the other hand, Lo Bello and Morchio (2022) model parental information networks rather as a form of nepotism, which results in children not working in the occupation in which they have a competitive advantage in. Hence, they find that shutting down parental networks increases welfare by 0.3 percent in consumption equivalent variation. Therefore, an interesting path for future research could be to combine both channels in a structural model and study the resulting impact of persistence on welfare.

## 6 Conclusion

In this paper, I study the underlying mechanisms behind the empirical observation that many children follow into their parent’s occupation. In particular, I quantify how much of the

persistence can be accounted for by a channel which the literature has abstracted from so far: regional labor demand. For that purpose regional and occupation-specific annual vacancy shares, computed based on all vacancies reported to the German Federal Employment Agency, are merged with the German Socioeconomic Panel (SOEP).

I find that controlling for regional labor demand, proxied by these vacancy shares, decreases the persistence by up to 10 percent when classifying occupations based on the finest level. This impact is of the same magnitude as controlling for children’s education does and suggests that regional labor demand is of substantial importance for occupational persistence and thus for social mobility. Furthermore, I document that the degree of occupational following is relatively homogeneous across parent-child pairs. A child is around 20 times as likely of working in a given occupation if one’s parent works in that occupation compared to a child whose parent works in another occupation. Lastly, I show that children who follow into their parent’s occupation experience a wage premium of around 6 percent during the first two years of their first full-time job. Also their annual unemployment risk halves during the first five years. Both results are primarily driven by those individuals without a high school degree. These results are qualitatively as well as quantitatively very similar to what has been found for referrals based on minority groups (Dustmann et al., 2016) and former coworker (Glitz and Vejlin, 2021), supporting the learning model developed in Dustmann et al. (2016). Consequently, the transmission of innate abilities does not seem to be of first-order importance but, next to regional labor demand, rather parental information networks are the major driver behind the following behavior of children.

For future research, there are at least two promising paths to take. First, it would be interesting to see whether the empirical results found in this paper also hold in other countries. In particular the case of the United States would be interesting as it could provide new implications for the fall in interstate migration observed over the last 40 years. Second, since, next to regional labor demand, parental networks seem to be the main driver behind the persistence in occupational choice, modeling this channel in greater detail and particularly focusing on its impact on the allocation of talent may yield new insights. So far, it has either been exclusively modeled as efficiency improving by reducing match-specific uncertainty, as in Dustmann et al. (2016), or as exclusively deteriorating efficiency as a form of nepotism, as in Lo Bello and Morchio (2022). Incorporating both channels in a structural model can thus yield new insights into the relative importance of both effects.

## References

- Aina, Carmen and Cheti Nicoletti (2018). “The intergenerational transmission of liberal professions”. In: *Labour Economics* 51, pp. 108–120.
- Bilal, Adrien (2021). *The geography of unemployment*. Tech. rep. National Bureau of Economic Research.
- Blau, Peter M and Otis Dudley Duncan (1967). “The American occupational structure.” In: *ERIC*.
- Brown, Meta, Elizabeth Setren, and Giorgio Topa (2016). “Do informal referrals lead to better matches? Evidence from a firm’s employee referral system”. In: *Journal of Labor Economics* 34.1, pp. 161–209.
- Burks, Stephen V, Bo Cowgill, Mitchell Hoffman, and Michael Housman (2015). “The value of hiring through employee referrals”. In: *The Quarterly Journal of Economics* 130.2, pp. 805–839.
- Caucutt, Elizabeth M and Lance Lochner (2020). “Early and late human capital investments, borrowing constraints, and the family”. In: *Journal of Political Economy* 128.3, pp. 1065–1147.
- Cooke, Thomas J (2013). “Internal migration in decline”. In: *The Professional Geographer* 65.4, pp. 664–675.
- Dal Bó, Ernesto, Pedro Dal Bó, and Jason Snyder (2009). “Political dynasties”. In: *The Review of Economic Studies* 76.1, pp. 115–142.
- Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum (2022). “Matching in cities”. In: *Journal of the European Economic Association* 20.4, pp. 1478–1521.
- Doi, Suhail AR, Polychronis Kostoulas, and Paul Glasziou (2022). “Likelihood ratio interpretation of the relative risk”. In: *BMJ Evidence-Based Medicine*.
- Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker (2016). “Referral-based job search networks”. In: *The Review of Economic Studies* 83.2, pp. 514–546.
- Eliason, Marcus, Lena Hensvik, Francis Kramarz, and Oskar Nordström Skans (2022). “Social connections and the sorting of workers to firms”. In: *Journal of Econometrics*.
- Glitz, Albrecht and Rune Vejlin (2021). “Learning through coworker referrals”. In: *Review of Economic Dynamics* 42, pp. 37–71.
- Granovetter, Mark S (1973). “The strength of weak ties”. In: *American Journal of Sociology* 78.6, pp. 1360–1380.

- Heckman, James J and Stefano Mosso (2014). “The economics of human development and social mobility”. In: *Annu. Rev. Econ.* 6.1, pp. 689–733.
- Huttunen, Kristiina, Jarle Møen, and Kjell G Salvanes (2018). “Job loss and regional mobility”. In: *Journal of Labor Economics* 36.2, pp. 479–509.
- Jovanovic, Boyan (1979). “Job matching and the theory of turnover”. In: *Journal of Political Economy* 87.5, Part 1, pp. 972–990.
- Kaplan, Greg and Sam Schulhofer-Wohl (2017). “Understanding the long-run decline in interstate migration”. In: *International Economic Review* 58.1, pp. 57–94.
- Kline, Patrick and Enrico Moretti (2013). “Place based policies with unemployment”. In: *American Economic Review* 103.3, pp. 238–43.
- Kramarz, Francis and Oskar Nordström Skans (2014). “When strong ties are strong: Networks and youth labour market entry”. In: *Review of Economic Studies* 81.3, pp. 1164–1200.
- Kuhn, Moritz, Iouri Manovskii, and Xincheng Qiu (2021). *The geography of job creation and job destruction*. Tech. rep. National Bureau of Economic Research.
- Lentz, Bernard F and David N Laband (1989). “Why so many children of doctors become doctors: Nepotism vs. human capital transfers”. In: *Journal of Human Resources*, pp. 396–413.
- Lo Bello, Salvatore and Iacopo Morchio (2022). “Like father, Like Son: Occupational Choice, Intergenerational Persistence and Misallocation”. In: *Forthcoming Quantitative Economics*.
- Mocetti, Sauro (2016). “Dynasties in professions and the role of rents and regulation: Evidence from Italian pharmacies”. In: *Journal of Public Economics* 133, pp. 1–10.
- Molloy, Raven, Christopher L Smith, and Abigail Wozniak (2011). “Internal migration in the United States”. In: *Journal of Economic perspectives* 25.3, pp. 173–96.
- (2017). “Job changing and the decline in long-distance migration in the United States”. In: *Demography* 54.2, pp. 631–653.
- Moretti, Enrico (2010). *Local labor markets*. Tech. rep. National Bureau of Economic Research.
- Simon, Curtis J and John T Warner (1992). “Matchmaker, matchmaker: The effect of old boy networks on job match quality, earnings, and tenure”. In: *Journal of Labor Economics* 10.3, pp. 306–330.
- Topel, Robert H and Michael P Ward (1992). “Job mobility and the careers of young men”. In: *The Quarterly Journal of Economics* 107.2, pp. 439–479.