I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15154

Sources of Wage Growth

Jérôme Adda
Christian Dustmann

MARCH 2022

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15154

Sources of Wage Growth

Jérôme Adda
Bocconi University and IZA

Christian Dustmann
University College London

MARCH 2022

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Sources of Wage Growth*

This paper investigates the sources of wage growth over the life cycle, where individuals have the possibility to acquire vocational training at the start of their career. Wage growth is determined by sectoral and firm mobility, unobserved ability and the accumulation of human capital. Workers may move between two occupational sectors that require cognitive-abstract (CA) and routine-manual (RM) skills, and job mobility is induced by non-pecuniary job attributes and persistent firm-worker productivity matches. Estimating this model using longitudinal administrative data over three decades, we show that RM skills are a key driver of early wage growth while CA skills become important later on. Moreover, job amenities are an important determinant of mobility decisions. Vocational training has long term effects on career outcomes, affecting the type and quality of matches, with substantial internal rates of return both to the individual as well as society.

JEL Classification: J2, J3, J6

Keywords: wage determination, learning by doing, job mobility, apprenticeship training

Corresponding author:
Jérôme Adda
Department of Economics
Università Bocconi
Via Roentgen 1
20136 Milano
Italy

E-mail: jerome.adda@unibocconi.it

* We thank Jim Heckman, Courtney Brell, David Zentler-Munro, four anonymous referees and participants in seminars at many universities and conferences for helpful comments and discussion. We are particularly grateful to Costas Meghir and Jean-Marc Robin for their input into early version of this paper (Adda et al. (2006)). Dustmann acknowledges funding by the European Research Council (ERC) Advanced Grant (MCLPS 833861), the Norface Welfare program and the DFG Grant 1024/1-2 AOBJ:642097. Edited by Jim Heckman.

1 Introduction

This paper investigates the sources of wage growth for workers over their entire careers, from entry to the labor market onwards. The dynamic model that we develop allows for wage growth to be determined by unobserved ability, mobility, and the accumulation of human capital. In addition, it allows for the choice of education level (and in particular a choice of whether to undertake vocational training before labor market entry), introduces task complexity as a feature that distinguishes occupational sectors, and permits mobility to be driven by both monetary and non-monetary job characteristics.

Our framework therefore combines a number of novel features that allow us to investigate how acquisition of different types of skills and job-to-job mobility drive wage growth at key stages over the lifecycle. We consider occupational sectors characterized by tasks, and as such the accumulation of sector-specific human capital, as well as its returns in different jobs, has a natural interpretation. To study how educational choices interact with the process of human capital acquisition and affect workers' future labor market transitions, we consider an initial choice of labor market entrants whether or not to undergo firm based apprenticeship training, and we allow sector specific returns to work experience and job offer probabilities to depend on this choice. Hence, in our model vocational training shapes future wage profiles not only through initial acquisition of sectoral skills, but also through how these skills aid in on-the-job learning and in inviting future job opportunities. This allows us to assess the return of structured vocational training for both individual workers and society, to investigate what the advantages of this training are both in the short and long run, and to determine whether these advantages come from particular skills that are learned, stronger labor market attachment, firm-to-firm job mobility, or are simply artefacts of selection.

Our analysis is based on administrative data drawn from social security records that allows us to track the careers and wages of individuals from their first entry into the labor market for up to three decades. The data provides precise information on the initial training choices individuals make, as well as the type of occupation they choose. It accurately records all wages, movements between different jobs and occupations, and transitions between non-employment and work, enabling us to precisely track workers' choices and career progression. Our sample covers men working in West Germany, born between 1960 and 1972 and observed between 1975 and 2004, a period that encompasses three decades and many entry cohorts. We combine this data with detailed information about the task content of different jobs, which we use to classify occupations into cognitive-abstract (CA) and routine-manual (RM) sectors, and to compute the amount of experience workers accumulate in each sector over their careers. To estimate choices made prior to labor market entry (including educational choices and initial occupational sector), we use exogenous variation caused by local economic shocks, fluctuations in the availability of training positions and service sector jobs, and the uneven expansion of universities over time and local areas.

The estimated model allows us to understand the dynamic implications of early vocational training, and we identify several advantages of this training that reveal themselves gradually over the lifecycle. Trained workers (i.e. workers who chose apprenticeship training at labor market entry) are more attached to the labor market, especially early on in their careers, allowing them to accumulate work experience faster. From the beginning of their career, they also tend to accumulate more CA experience, which helps sustain wage growth later in the lifecycle (as we find that RM experience contributes to wage growth only in the early years). CA experience reduces the risk of layoff and improves hiring prospects in any sector, as well as incentivizing transitions into CA occupations where this type of experience is rewarded further. Hence, as in Cunha and Heckman (2007), we find that dynamic complementarities through sector specific experience and occupational upgrading are more important sources of advantage for trained workers than the immediate initial wage benefits, although they do not yield returns until a later stage of the lifecycle. We also find that workers sort into vocational training programs according to their unobserved productivity, but even accounting for this reveals that these programmes generate positive returns both for individuals and society, at an internal rate of about 10 percent.

Within our model framework, we revisit several key issues to better understand the different sources of wage growth. As in Topel and Ward (1992), we find that between-job wage growth is an important determinant of overall wage growth for young workers. However, looking more closely we find that the largest single driver of early wage growth is skill accumulation on the job, especially in the RM sector. We also show that gains from marginal mobility (i.e. one additional job change over the career cycle) are small, in comparison to the large wage growth associated with mobility early on in workers' careers. We further illustrate that job-to-job mobility is not driven solely by income maximizing behavior, but that a significant share of these transitions are due to individuals moving towards jobs with (subjectively) superior non-pecuniary attributes. This stresses the importance for future research of better understanding the nature of such attributes, such as working conditions and commuting distances (see e.g. Guglielminetti et al. (2018), Le Barbanchon et al. (2019)).

We also investigate the process of occupational upgrading and mobility across sectors. Sector choice depends on the worker's (unobserved) sector specific "ability" and (observed) educational choices, but also on (observed) sector specific skills that are differentially accumulated through experience in different occupations. This can generate "lock-in" effects: with workers having incentives to remain in sectors where they have previously accumulated more sector specific experience, they will be even less likely to change sector in the future. This effect explains the low mobility of workers across occupational sectors, which remains unexplained by models that distinguish only between general and firm specific skills. Our model is therefore able to explain wage losses following a sector change as described in Neal (1995). Moreover, we find that while sorting explains a sizeable part of the initial sectoral wage difference, the main explanatory factor later in the life cycle is heterogeneous skill accumulation and differential returns to RM and CA experience.

Our paper draws on, combines features from, and extends several literatures examining wage determination and career paths. A literature beginning with Keane and Wolpin (1997) uses structural models to examine the importance of occupational or sectoral mobility for workers' careers in a life cycle setting. Our work builds upon a similar framework to many of these models (notably including Belzil et al. (2017), Hoffmann (2018), and Roys and Taber (2016)), but extends them in several key ways. We incorporate worker mobility not just across sectors but also across firms, a feature our paper shares with Sullivan (2010). This allows for firm mobility as an important driver of wage growth, as stressed by Topel and Ward (1992) and the subsequent literature.

Drawing on insights from the search and matching literature, particularly that pioneered by Postel-Vinay and Robin (2002), Bagger et al. (2014) and Taber and Vejlin (2020), where workers experience permanent wage shocks due to renegotiations after outside wage offers, we allow for time-varying but persistent firm-worker match productivity. The worker-firm match allows us to identify a "search capital" held by workers, which is a potentially important driver of wage growth.[1] Moreover, besides match productivity, our framework also allows for firm specific non-pecuniary match values, so that mobility of workers is determined not just by monetary motives, but also by non-pecuniary incentives. This is an important extension, particularly in light of a recent literature that discusses the reasons for worker mobility and sorting across firms (e.g. Card et al. (2013), Bonhomme et al. (2019), Di Addario et al. (2020), and Bonhomme et al. (2020)).

Another important departure from prior literature is in how we distinguish skills and treat the portability of these skills across sectors. While for instance Keane and Wolpin (1997) define sectors corresponding to white and blue collar jobs, we build on the concept of "tasks" characterising occupations and the type of skills they require (see Autor et al. (2003) and Goos and Manning (2007)), and define sectors according to the mix of tasks they employ, not dissimilar to Dix-Carneiro (2014) and Roys and Taber (2016).[2] This allows us to define task specific work experience that is accumulated differently across sectors according to the intensity with which a sector utilizes it. Moreover, in our model task specific experience is priced differently in the two sectors, so that skills acquired in one sector can be used in another, but with different returns to the worker. By estimating these skill prices we are able to speak to the specificity or portability of human capital (see e.g. Gathmann and Schoenberg (2010)). Thus, our model extends the prototypical Roy (1951) model by allowing for multiple sectors and heterogeneity in multiple skill dimensions, both observed and unobserved, as in Willis and Rosen (1979) and extensions

[1] The literature quantifying the contributions of workers and firms to earnings inequality, pioneered by Abowd et al. (1999), has recently focused on the importance of taking into account the persistence of the wage process at the firm level, see Bonhomme et al. (2019). A different literature has also characterized search capital in the context of equilibrium search models as in Menzio and Shi (2011), Lise and Robin (2017) or Gertler et al. (2020).

[2] Cainesa et al. (2017) also draw attention to the effects of differences in task complexity across occupations on wage growth.

by Heckman and Sedlacek (1985) and Heckman and Honoré (1990).

The paper is organized as follows. In the next section, we provide a discussion of our data and sample, as well as some background information on the vocational training choice available to workers. In Section 3, we specify a dynamic and generalized Roy model for career progression, including the initial training choice. In Section 4 we detail our estimation method and discuss the identification of our model. Section 5 presents the results and decomposes wage growth into different components. Finally, Section 6 concludes.

2 Background and Data

2.1 The Apprenticeship System

The German Apprenticeship System is a vocational training program which combines on-the-job training, provided by the firm, with school education, provided and funded by the state. Apprenticeship training typically starts after secondary school, which tracks children after the age of 10 into lower, intermediate, and upper secondary schools, based on the student's ability and preferences. Those who enroll in lower or intermediate secondary schools typically enroll in blue or white collar apprenticeship training schemes after school graduation (at around the age of 16), or enter the labor market as untrained workers. Those who attend upper secondary schools (which we will sometimes refer to as the "academic" track) are entitled to enroll directly into university.[3] For the cohorts we study, about 20% are assigned to this track. In our analysis, we focus on young men who enter the labor market after lower or intermediate secondary schooling, to either work full time, or to first enroll into an apprenticeship program (about 75-80% of a birth cohort, see Table 1).

The apprenticeship system offers training in more than 500 white and blue collar occupations.[4] In practice though, individuals typically choose from a fairly small number of training professions. For instance, in our data, 70 percent of all male apprentices are concentrated in 20 three digit occupations, with slightly more than two-third of those

[3] See Dustmann (2004) for a detailed description of the German school system.

[4] See http://berufenet.arbeitsagentur.de/berufe/index.jsp for details.

being blue collar. Apprenticeship training is highly structured, with a well-defined curriculum, and lasts between 2-3 years. It takes place at the workplace for 3-4 days a week, where practical and workplace-related knowledge is acquired under the supervision of qualified instructors, and at vocational state schools for 1-2 days a week, where more general and academic knowledge, as well as theoretical knowledge specific to the chosen occupation is obtained. Both the practical and the academic components are examined at the end of the training period, and successful candidates obtain a professional qualification. Apprenticeship wages are substantially below that individuals could obtain as untrained workers. We refer the reader to Steedman et al. (1998) for more details.

2.2 Data and Sample

2.2.1 Registry Data

Our main data is a two percent sample of individual administrative social security records over three decades (1975 to 2004), made available by the German Institute for Employment Research (IAB). Due to our sample period, we focus on former West Germany (to which we refer as "Germany" in what follows), and construct a longitudinal sample of young men whom we observe from entry to the labor force onwards for up to 30 years. In addition to unique worker and establishment identifiers, our data contains detailed information on workers' wages, education, age, occupation and industry, as well as regional information about workplaces.[5] The data records all employment spells in the private and public sectors (except for civil servants and the self employed), with exact dates when each job started and ended. We observe the average daily pre-tax wage at the end of each calendar year for ongoing employment spells. For individuals who change firms within a calendar year, the data provides the average wage from the beginning of the year or the employment spell (if it started after the beginning of the year) until the end of that spell (or calendar year if the spell finishes after the end of the year). Thus wages are not averaged across different firms, a crucial feature for studying job to job mobility.

[5]The wage data is top coded at the earnings limit for social security contributions. For the sample we consider, this concerns only about 1.6 percent of all wage spells. We take top coding into account in our estimation procedure, and we describe details below.

The data also contains information on the apprenticeship training period, and whether a worker holds an apprenticeship qualification or not, as well as their overall educational qualifications. Merged to this data set is information from the unemployment registers, which allows us to observe transitions from and to unemployment.

In our analysis, we focus on West Germany and select all German men born in the period between 1960-1972, who enter the labor market with a lower or intermediate secondary degree, typically obtained by the age of 16. We select these cohorts to ensure that we only include individuals whom we can observe at the start of their labor market career, so that we avoid initial conditions problem. We distinguish between individuals who, after secondary school and before joining the labor market as full time workers, enroll in an apprenticeship scheme for between 2-3 years and successfully complete their training (in what follows we refer to these individuals as "trained"), and individuals who enroll for a shorter period but do not graduate, or do not enroll and enter the labor market directly after secondary school (we refer to these as "untrained").[6]

From this data, we construct a data set of quarterly spells. Whenever multiple spells during a quarter are present (e.g. an employment and an unemployment spell), we assign to that quarter the spell of the longest duration.

Our final sample contains 44,286 individuals who enroll in an apprenticeship training scheme after secondary school, and 4,858 individuals who join the labor market directly and without further training. Following these individuals up until 2004 results in a total of 3,920,492 quarterly observations. To identify the determinants of school track choice at age 10, we use a separate extract from the IAB data of 71,472 individuals who follow either the lower/intermediate or the academic track. We provide more detail on the sample selection in Appendix A.

There is a large overlap in terms of occupations of trained and untrained workers. In our sample individuals are employed in 292 3-digit occupations after labor market entry.

[6]As an alternative to firm-based apprenticeship training, about 6 percent of our sample undertakes qualifying training in vocational schools, which offer classroom training for two to three years, with unpaid work experience, and lead to a certificate equivalent to a firm-based apprenticeship (see Parey (2009) for details). Wage profiles of those who went through firm based training and vocational schools are almost identical. We include these individuals in the group of "trained" workers.

Out of those, 19 occupations employ only untrained workers (and these employ only around one percent of all untrained workers), and 53 occupations employ only trained workers (and these employ just 1.4 percent of all trained workers).

In our analysis, we distinguish two occupational sectors, characterized by occupations that mainly use Cognitive and Abstract skills (CA occupational sector), and occupations that mainly use Routine and Manual skills (RM occupational sector). To classify occupations into these two groups, we use information from a survey data set that measures the precise task content of three digit occupations, and that we describe next.

2.2.2 The Qualification and Career Survey

We use detailed information from the 1991/1992 wave of the German BIBB/IAB Qualification and Career Survey on 19 activities/tasks performed at work to categorize 3-digit occupations into those involving mainly routine/manual (RM) or cognitive/abstract (CA) tasks, similar to Autor et al. (2003) and Spitz-Oener (2006). We follow Antonczyk et al. (2009) and proxy the task intensity of a job in tasks of type j for each survey individual i, $Task_{ij}$, $j \in \{RM, CA\}$, by $Task_{ij} = \frac{T_{ij}}{T_i}$ where T_{ij} is the number of tasks performed by individual i of type j, and T_i the total number of tasks performed by that individual.

Thus, if an individual carries out 6 tasks in total, 4 of which are routine/manual tasks, the task index is 2/3. We then aggregate the individual task indices at the 3-digit occupational level, using the maximum mean task index to classify the occupation as RM or CA, and finally merge this information with the registry data for each 3-digit occupation. We consider an occupation with a CA task share above 50% as belonging to the CA sector. Overall, at labor market entry, about 15% of all individuals work in the CA sector, increasing to 20% after 20 years of potential experience.

2.2.3 Additional Data and County Level Information

To obtain exogenous variation that help in identifying pre-labor market educational and training choices, as well as occupational sector choice, we use regional information either at the district or commuting zone level, drawn from official statistics.[7]

[7] On average, a district (commuting zone) in West Germany has about 200,000 (370,000) inhabitants respectively. In 2000, there were 326 districts and 135 commuting zones in West Germany.

First, as we only consider individuals who attend lower secondary schools that do not grant direct entry to university, we need to account for that choice of track. The three-track system in place during our observation window required individuals to choose a track (lower, intermediate, or upper secondary) around the age of 10. That decision is likely to depend on the availability of university places in the immediate environment, which would motivate children (and their parents) to enroll in an upper secondary school.[8] We draw on data provided by the German Statistical Office (Destatis) which reports yearly numbers of students enrolled at each institute of higher education in the first (winter) semester for the period 1972-1999 (Statistisches Bundesamt, various years).

From this information, we construct a variable that measures the ratio of the number of university places to the total population in a radius of 50 kilometres around the county of residence when the individual is 10 years old (the age of secondary school track choice). Germany underwent an educational expansion planed in the 1960's, which saw the total number of students increase from 155,000 in 1958 to 1 million in 1990. A key objective of the expansion policy was to allocate post-secondary educational institutions more evenly across local areas, and new universities and colleges were mainly opened in rural and semi-rural areas without pre-existing higher education facilities. Municipalities and local governments had very little influence on the location of the new colleges. All this led to exogenous increases in supply of university places, leading to the increased enrolment

During 1972 and 1999, the number of university enrollments in West Germany doubled, though this rapid expansion was far from uniform - in some districts enrollments remained constant, while in others they increased by a factor of 6. It is this large variation in available university places over time and across regions that we use for identification.[9]

Second, we construct two variables that partly drive the decision of young men to either enter the labor market directly as full time workers after finishing (lower or intermediate) secondary school, or to enroll in a 2-3 year apprenticeship training scheme to further improve their skills. The first of these variables is the cyclical component of

[8]In stark contrast to the US, the university system in Germany is (with very few exceptions) publicly provided, free of charge, and not considered to provide degrees of very different quality, so that quality considerations hardly affect enrollment decisions, in particular over the period we consider here.

[9]More precisely, we use the changes across cohorts in the availability of university places within regions, as we condition on commuting zone fixed effects, time fixed effects, and state linear time trends.

the annual log of gross value added (GVA) at the district level (see Volkswirtschaftliche Gesammtrechnung der Laender, various years). This variable reflects regional variation in GDP growth, and may influence the decision of individuals whether or not to acquire additional training before entering the labor market as full time workers. The second variable is the annual number of available apprenticeship training positions at the commuting zone level, which we obtain from official data that report these numbers at commuting zone level (see Berufsbildungsberichte, various years). We scale the number of positions by the relevant cohort size, which provides an indication of the supply of apprenticeship places available in the area.[10]

Third, when enrolling in apprenticeship training, or when entering the labor market directly, individuals have to choose an occupational sector. To model that decision, we extract from the IAB data the share of employed workers in the service industry (banking, hospitality and other services) in the commuting zone for the years 1975-2010. Variation in the size of this industry may affect the decisions of workers whether to start their careers (or training period) in an occupation with more CA or RM skill requirement, as the share of CA occupations is higher in the service industry. We measure the regional industry composition at the time of lower secondary school graduation (around age 15), when individuals typically make their initial occupational choice. Table A1 provides descriptive statistics. We analyze the relationship between these instrumental variables and the initial choices of education and training, and occupational choice in Section 4.

2.3 Descriptive Analysis

Wage Profiles and Labor Market Transitions. Figure 1 displays the log real wage profile as a function of years of potential labor market experience (defined as time since entry to the labor market or beginning of training) for trained workers (those with an apprenticeship qualification, denoted as "Trained Wage"), for those currently training as apprentices ("Wage in training"), and for untrained workers ("Untrained Wage"), as well as the difference in wages between the trained and untrained (right-hand axis). The

[10]Mobility of apprentices is low, as apprenticeship wages are very modest, resulting in the vast majority of apprentices living at home during their training period to avoid additional housing costs.

figure shows that real wages of the untrained increase rapidly during the first five years in the labor market, by 11 percent per year on average. Over the next twenty years however, overall real wage growth is just below 9 percent, or 0.4 percent per year. This is reminiscent of the situation in the US, where most of the wage growth occurs in the first ten years in the labor market (Murphy and Welch, 1992). Those who enroll in apprenticeship training schemes are paid a low wage during their training period, but wages increase sharply afterwards and overtake those of the untrained. In addition, following this the wages of the trained continue to increase slightly faster than the untrained, by about one percent per year on average. As a result, twenty years after graduating secondary school, wages of trained workers are about 15 percent higher than those of the untrained.

Wages are only one dimension along which outcomes of trained and untrained workers may differ. Another important dimension is labor market attachment. Figure 2 shows the proportion of individuals who are in work as a function of potential experience.[11] Labor market attachment of trained workers is stronger than that of the untrained, with a higher fraction of the trained working at any age. Untrained workers also have a higher probability of transitioning into unemployment, as seen in Table 2 (9 percent of employed untrained, but only 3 percent of employed trained workers exit each quarter during the first five years in the labor market), and a lower probability to return to work from unemployment each period (19 percent of trained unemployed individuals with 5-10 years of potential experience, but only 7 percent of the equivalent untrained individuals, find a job from one quarter to the next). All this leads to the untrained spending less time working: over a 25 year period, they work on average a total of 17.4 years, compared with 20.1 years for trained workers.

Figure 3 plots the average number of firms in which an individual has been employed, where the horizontal axis carries potential experience. It shows that the untrained are more mobile during the first few years in the labor market. As such, one might expect that job shopping could be an important source of the large initial wage growth for untrained workers seen in Figure 1. To investigate this further, we decompose wage growth into within and between firm wage growth and plot it against potential experience

[11]Germany had a compulsory military draft system during the period we consider, and we have eliminated interruptions that are due to military service while constructing the figure.

(see Figure 4). Between firm wage growth appears to be substantial, between 20 and 40 percent for the untrained during the first 2-3 years in the labor market, at a time when trained workers are still in the training phase. The gain in wages due to these movements between firms falls over time, but is still large for both trained and untrained workers until about 5-7 years in the labor market, with returns being close to zero after about 15 years. Within firm wage growth for the untrained is likewise very high early on in the career reflecting the rapid learning that takes place on the job.

Occupation and occupational specific skills. Considering now the tasks at which workers are employed, we plot in Figure 5 the amount of exposure to CA tasks over the life-cycle for trained and untrained workers, computed as the average CA task intensity of the employed (see Section 2.2.2), by potential experience. It suggests that untrained workers perform on average about 20 percent CA tasks, with little change as they age. Trained workers, on the other hand, are already more exposed to CA tasks during their training period (about 24 percent), and this share increases as they age, eventually reaching above 30 percent.[12]

In Figure 6 we decompose wage growth due to RM and CA experience, for trained and untrained workers, as a function of potential experience.[13] The figure reveals some interesting patterns. There is a steep increase in log wages as a result of accumulating RM experience for the first 3-5 years for both trained and untrained workers, which levels out after about 5 years. On the other hand, wages increase steadily due to the accumulation of CA experience, for both trained and untrained workers. While these figures are not causal, they suggest that RM experience that matters for productivity is accumulated early on, while the accumulation of CA experience creates productivity gains throughout workers' careers.

[12]The employment shares in our sample in RM and CA jobs are given in Appendix Table A1, where we report among other characteristics, the shares in CA jobs at labor market entry and after 10 years, for trained and untrained workers.

[13]The figure is computed by first regressing log wages on a flexible function of both RM and CA experience separately for trained and untrained workers. We then plot the return to a specific experience multiplied by the average amount of RM and CA experience at a given potential experience.

3 The Model

At labor market entry, individuals choose whether to enroll into vocational training, or to enter full time work directly. They also choose the occupational sector they want to begin working in. After labor market entry, they decide on labor market transitions between jobs, sectors and between work and unemployment. Among other determinants, individuals condition these choices on their work experience, firm tenure, and current wage. There are two types of productive experience, CA and RM, which are differently productive and accumulated with different intensity in each sector in a learning by doing way. We also allow for individual specific unobserved heterogeneity in CA and RM productivity, so that both observed sectoral experience and unobserved productivity determine the sector choice of individuals.[14]

The model is set in discrete time, and one period lasts one quarter. All choices are made in order to maximize the present value of future utility. Individuals derive utility from wages, benefits when unemployed, and leisure. Jobs also have non-pecuniary attributes (e.g. work conditions), so that workers pursue both higher wages and non-pecuniary job attributes through job mobility. Wage growth is determined by the initial choice of individuals whether to train in a structured apprenticeship scheme, and by sector-specific and firm-specific experience. Wages also grow through job mobility, where workers accumulate search capital, as we allow for heterogeneous and persistent worker-firm productivity matches.

3.1 Skills

Each individual is characterized by three types of skills. First, experience acquired in the CA and RM sectors, which we denote by the vector $X_{it} = \{X_{it}^{CA}, X_{it}^{RM}\}$. Second, firm specific experience, denoted by Ten_{it}. Third, education acquired through vocational

[14] Keane and Wolpin (1997), Arcidiacono et al. (2007) or Sauer (2004) develop models with differences in time preferences. Using data from the German Socio-Economic Panel, we do not find any evidence for differences in measures of time preference between individuals who are trained and untrained, neither for differences in desired working hours, risk aversion, impatience, and taste for leisure. Results are available upon request.

training, which can influence the return to sector specific skills as well as transitions in and out of the labor force or across sectors. We denote the educational choice by $T_i \in \{NT, TR\}$ for not trained and trained, respectively.

Each individual enters the labor market with zero sector experience ($X_{it} = \{0, 0\}$). The accumulation of CA and RA skills starts when individuals begin to work or enroll in vocational training. Individuals accumulate RA and CA skills in both sectors, but the accumulation of RM skills is faster in the RM sector, and the accumulation of CA skills is faster in the CA sector. The per period increase in sector specific experience of type j, in sector o, with $j, o \in \{CA, RM\}$ for a worker i, $x^{o(j)}$, varies by training status T_i, and by how much experience the worker has accumulated in the CA and RM sector. The quarterly increase in both stocks of experience is then given by [15]

$$\begin{cases} X_{it}^{CA} = X_{it-1}^{CA} + x^{o(CA)}(T_i, X_{it-1}^{CA}, X_{it-1}^{RM}) \\ X_{it}^{RM} = X_{it-1}^{RM} + x^{o(RM)}(T_i, X_{it-1}^{CA}, X_{it-1}^{RM}) \end{cases} \quad s.t \quad \begin{array}{l} x^{o(CA)} + x^{o(RM)} = 1/4 \\ x^{o(CA)}, x^{o(RM)} \in [0, 1/4]. \end{array} \tag{1}$$

Firm specific tenure Ten_{it} increases by a quarter per period and is reset to zero when the individual starts a job in a new firm.

3.2 Unobserved Heterogeneity

Individuals differ ex-ante in their unobserved productivity in the RM and CA sectors, denoted by $s_i^{RM}, s_i^{CA} \in \{0, 1\}$, respectively. We assume that these two random variables follow a bivariate discrete distribution, each with two points of support. Thus, there are four types of individuals: those with no unobserved sectoral productivity, those with unobserved productivity in either the RM or CA sector, and those with unobserved productivity in both sectors. We denote the proportion of these types in the sample by π_l, $l = 1, \ldots, 4$, and allow the two types of unobserved productivity to be correlated.

[15] To compute $x^{o(j)}$, we assign CA and RM experience to workers by computing the exposure to CA and RM tasks at each point of accumulated CA and RM experience, separately by training status. This accounts for possible compositional changes/differences in 3 digit occupations in workers' careers within the two sectors we defined, and by training status. The per period increase $x^{RM(CA)}$ ranges between 0.03 and 0.07, while $x^{CA(CA)}$ ranges between 0.15 and 0.17, depending on accumulated CA and RM experience. Total experience gained by an employed worker each year is normalized to one.

Hence this specification allows for selection on unobservables as in Heckman and Sedlacek (1985). As explained below, each unobserved ability receives a price in either sector.

3.3 Aggregate Shocks

We characterize the macroeconomic fluctuations of the economy around the steady-state growth trend by de-trended quarterly GDP. The macro state variable G_t is modeled as a discrete two-state Markov process of order one, which allows for persistent aggregate shocks.[16] Individuals observe the current state of the economy and form correct expectations of future shocks. The macro shock affects wages, lay off and job offer rates, as explained below.

3.4 Wages and Matches

We start with the pricing of the unobserved abilities (s_i^{RM}, s_i^{CA}). We define a set of four ability-occupation specific prices $\{p^{o(j)}\}$, where $p^{o(j)}$ is the price of ability j in sector o, with $o, j \in \{RM, CA\}$. We take those prices as constant during our sample period. Unobserved abilities are therefore priced depending on the sector o_{it} in which the individual works:

$$\begin{aligned} \alpha_i^S(o_{it} = RM) &= s_i^{RM} p^{RM(RM)} + s_i^{CA} p^{RM(CA)} \\ \alpha_i^S(o_{it} = CA) &= s_i^{RM} p^{CA(RM)} + s_i^{CA} p^{CA(CA)}\,. \end{aligned} \tag{2}$$

To separately identify individuals with abilities $(0,1)$ from $(1,0)$, we impose a restriction on prices so that $p^{CA(RM)} < p^{RM(RM)}$ and $p^{RM(CA)} < p^{CA(CA)}$.[17] Hence, those with only CA (RM) unobserved ability have an ex-ante absolute advantage in the CA (RM) sector. However, over the life-cycle, productivity also depends on the (observable) vector of accumulated sector experience, X_{it}, so that, ex-post, an individual of a given type may nonetheless be more productive in another sector.

When a worker and firm first meet in period t, they draw a match-specific productivity, κ_{it}, and a non-pecuniary value of the match (e.g. work conditions), μ_{it}.[18] The match-

[16] The transition probabilities are presented in Appendix B in Table A2.

[17] In practice, after several trials, we impose $p^{CA(RM)} = p^{RM(CA)} = 0$.

[18] Keane and Wolpin (1997) consider a non-pecuniary attribute. In contrast to our our specification it is not modeled at the firm level but at the sectoral level and constant across individuals. Sullivan (2010)

specific productivity captures the residual heterogeneity in wages when individuals start a new job. We allow the distributions of these matches to differ by training status and occupation and to follow a normal distribution with mean zero and standard deviations $\sigma_\kappa(T_i, o_{it})$ and $\sigma_\mu(T_i, o_{it})$ respectively. This permits us to estimate the extent to which job opportunities vary between trained and untrained workers, and between the two sectors. For subsequent periods *within* the firm, the two components evolve as

$$\begin{aligned} \kappa_{i,t+1} &= \kappa_{i,t} + u_{i,t+1}, \qquad u_{i,t+1} \sim iid \quad \mathcal{N}(0, \sigma_u^2(T_i, o_{it})) \\ \mu_{i,t+1} &= \mu_{i,t}. \end{aligned} \tag{3}$$

The two components are re-initialised at each transition between firms. The firm-worker match in (3) allows for the value of a match and the contracted wages to change across jobs, while permitting persistence over time within jobs, a pattern that we observe in our data.[19]

The non-pecuniary attribute μ_{it} accounts for why some individuals move from a high paying job to a lower paying one. It also allows for the possibility that individuals are not pure income maximizers, but pursue things other than income in their choice of jobs, as in Heckman and Sedlacek (1985). We model the non-pecuniary attribute of the job as fixed for the duration of a match as we do not observe it, nor do we observe the reason for the dissolution of a match. In contrast, we observe wages within the firm, which allows the identification of a time-varying match productivity. This productivity contribution κ_{it}, together with the non-pecuniary attribute μ_{it}, is part of a "search capital" held by the worker. By selectively moving across firms, this search capital can be accumulated over time and can be an important factor in determining wage growth or worker welfare (as defined below).

Quarterly log earnings $\log w_{it}$ depend on the unobserved ability component $\alpha_i^S(o_{it})$,

introduces a firm specific nonpecuniary match, but in the estimation its variance is set to zero so that it plays no role.

[19] Although our model does not have strategic wage bargaining as in Cahuc et al. (2006) or Dey and Flinn (2005), the random walk feature of the match allows for shocks to be permanent and can rationalize wage increases due to renegotiation in the face of external job offers (among other factors). The model can also accommodate negative shocks to wages, a feature that is present in the data as we consider real wages, with inflation being sizeable during the period of analysis.

the initial training choice T_i, the current occupation sector o_{it}, the macroeconomic indicator G_t, the match-specific component, κ_{it}, and (in a piecewise linear way) on the experience stocks X_{it}^{CA}, X_{it}^{RM} and tenure Ten_{it}:

$$\begin{aligned} \log w_{it} \quad = \quad & \alpha_i^S(o_{it}) + \alpha^G G_t + \alpha^{TR} I_{T_i=TR} + \alpha^{inTR} I_{inTR_{it}} + \alpha^{PCA} I_{o_{it}=CA} \\ & + \alpha^{CA}(X_{it}^{CA}; T_i, o_{it}) + \alpha^{RM}(X_{it}^{RM}; T_i, o_{it}) + \alpha^{Ten}(Ten_{it}; T_i, o_{it}) + \kappa_{it}, \end{aligned} \tag{4}$$

where $I_{T_i=TR}$, $I_{o_{it}=CA}$, $I_{inTR_{it}}$ are indicator variables equal to one if the individual has undergone apprenticeship training, works in the CA sector, and is currently in training, respectively. To define the piecewise linear functions $\alpha^j(x; T_i, o_{it})$, with $j \in \{CA, RM, Ten\}$, denote by $\{N_k^j\}_{k=1}^K$ a set of nodes and by $\alpha_k^j(T_i, o_{it})$ parameters that are specific to a node k, training status T_i, and to sector o_{it}. For a given experience or tenure level x, the return is specified as:

$$\begin{aligned} \alpha^j(x; T_i, o_{it}) \quad = \quad & \alpha_k^j(T_i, o_{it}) + \frac{\alpha_{k+1}^j(T_i, o_{it}) - \alpha_k^j(T_i, o_{it})}{N_{k+1}^j - N_k^j}(x - N_k^j) \\ & j \in \{CA, RM, Ten\}, \qquad k \text{ such that } x \in [N_k^j, N_{k+1}^j], \\ & \alpha_0^j(T_i, o_{it}) = 0, \qquad N_0^j = 0 \end{aligned} \tag{5}$$

Training directly enhances log earnings with a return α^{TR}, as is standard in the literature. Training has also an indirect effect on wages by affecting the returns to experience and tenure, and the variance of the match-specific productivity (see equation (3)) that drives search capital. This indirect effect manifests itself over the life-cycle through different career paths in terms of labor market attachment and occupational choices. Hence training has both an immediate effect on wages, that is similar for all workers and a long-term effect that is potentially heterogeneous across unobserved skill groups. Those with more unobserved CA skills have potentially more to gain from training.

The wages of workers who are in training are lower than those of otherwise equivalent unskilled workers due to the (presumed negative) factor α^{inTR}. Our model allows higher observed wages in the CA sector to arise either because of a generally higher productivity in this sector (if $\alpha^{PCA} > 0$), because of different returns to experience, or because of positive sorting of workers through the unobserved skill component $\alpha_i^S(o_{it})$. The wage function (4) builds on the extant literature (e.g. Flinn (1986), Keane and Wolpin (1997), Roys and Taber (2016) and Belzil et al. (2017)). Its flexible form is designed to pick up differences by training status along several dimensions, which will result in different incentives being important to trained and untrained workers.

3.5 Labor Market Transitions

Employed individuals in our model are laid off with probability $\delta(T_i, X_{it}, o_{it}, G_t, \kappa_{it})$, which depends on the initial training choice T_i, the vector of sector experience X_{it}, the current occupation, o_{it}, the state of the economy and the current match productivity κ_{it}.[20] Workers who are employed draw an alternative job offer with probability $\lambda^W(T_i, X_{it}, G_t)$, while unemployed individuals draw a job offer with probability $\lambda^U(T_i, X_{it}, G_t)$, both functions of the initial training choice, the vector of sector specific experience and the macroeconomic state. We therefore assume a random search environment.

3.6 Occupational Mobility

As well as receiving new job offers, individuals receive offers to change occupational sector with a probability $\lambda^O(T_i)$. These offers may come from the existing firm, in which case the worker would retain their existing job characteristics as they change sector, or they may come together with a new job offer, in which case the occupation change would occur concurrently with a job change. Whether or not it is optimal to accept an occupation change offer will depend, among other things, on the individual's relative productivity in the two sectors and the amount of RM and CA experience accumulated up to that point. As in Dix-Carneiro (2014), workers who switch sectors pay a mobility cost, denoted

[20] Our model distinguishes between exogenous layoffs and voluntary quits. Separate identification is achieved from the assumption of symmetry of the density function of the innovation of the match specific productivity u_{it}, defined in equation (3). Voluntary quits occur when workers receive a large negative shock to wages, which is unobserved by the econometrician. However, the data on wages allow us to identify the density of positive shocks to wages, which, by symmetry, pins down the frequency and magnitudes of negative shocks. This in turns identifies the proportion of workers who will prefer to leave employment in a given period. The exogenous layoffs are identified from the excess flow of workers into unemployment over these expected voluntary quits. In practice, we assume that the layoff probability increases by δ_κ when the match-specific productivity is below a threshold $\underline{\kappa}$.

c_O, which depends on potential experience. This cost has two implications in terms of sectoral mobility. First, older workers will find it less profitable to move, as they have a shorter remaining career span to recoup the cost. Second, those with higher productivity in the other sector will be more likely to move, inducing a selection across occupational sectors. We return to this point in more detail in Section 5.4. Experience in either sector does not depreciate upon an occupational change, but it is priced in different ways. This can lead to a dynamic selection, with an upfront decrease in wages, followed by higher wage growth in future periods.

3.7 Dynamic Choice

Using the notations defined above, we denote the current state vector as:

$$\Omega_{it} = \{X_{it}, Ten_{it}, T_i, o_{it}, G_t, \kappa_{it}, \mu_{it}, s_i^{RM}, s_i^{CA}\}. \tag{6}$$

The value function for individual i in period t is given by:

$$V_t(\Omega_{it}) = \max_{L_{it}} \log(y_{it}(L_{it})) + \mu_{it} I_{L_{it}=emp} + \gamma(1 - I_{L_{it}=emp}) + \beta E_t V_{t+1}(\Omega_{it+1|L_{it}}) \tag{7}$$

where L_{it} is the decision to work (combining labor supply, firm and sectoral choices) and y_{it} is equal to wages or unemployment benefits, depending on work status. We define $I_{L_{it}=emp}$ as an indicator variable for employment and γ represents the utility of leisure.The parameter β is a discount factor, and E_t is the expectation operator conditional on information in period t. The expectation of the individual is over the vector of shocks to income and labor supply, as well as over arrival of job and occupation change offers, and layoffs. $\Omega_{it+1|L_{it}}$ is the updated state space conditional on the choice made in period t. Labor market choices are taken until age 60, at the age at which retirement occurs. Individuals live an additional 20 years, deriving utility from retirement benefits, that depend on the last earned wage. Choices are made under the constraints detailed above and made explicit in Appendix B.2.

3.8 Initial Choices

The choice to enroll in apprenticeship training is based on the comparison of the value of a career with and without the training, allowing for foregone earnings. This decision is taken around the age of 16. In practice, we start modeling from the point we observe individuals joining either their first job or an apprenticeship scheme. As well as the training decision, individuals also choose whether to take their first job in the RM or CA sector. For any choice of training T and initial occupation o, the individual receives a utility that depends on a vector Z_i of local market characteristics (including local GVA, the share of jobs in service industries, and the local availability of apprenticeship training schemes, see Section 2.2.3 for details), a preference shock $\eta_{i,T,o}$, (assumed to be iid and following an extreme value distribution) and the future continuation value of that particular choice as defined in equation (7) and Appendix B.2:

$$V_{t=16}^{Init}(Z_i, s_i^{RM}, s_i^{CA}, \eta_{i,T,o}) = \max_{T,o}\{Z_i\zeta_{T,o} + \eta_{i,T,o} + \beta E_t V_{t+1}^{W}(\Omega_{i0|T,o})\} \tag{8}$$

The variables in Z_i with coefficients $\zeta_{T,o}$ only drive the initial choice of training and occupational sector and not the subsequent labor market decisions. We rely on this exclusion restriction for the non-parametric identification of the model. The specification in (8) implies that in areas and periods where there is a shortage of training slots, some individuals that would benefit from training (net of the utility costs arising from that shortage) end up choosing not to undertake training.

Moreover, and as described in Section 2, the two lower tracks of secondary school prepare their students for apprenticeship training, while the higher track prepares for university attendance. We only consider individuals who left school after graduating in the lower two tracks. Track choice takes place at age 10, based on merit, family decisions and local determinants. To identify the role of unobserved ability in determining this choice, we use variation in university slots in the commuting zone between cohorts. We provide details in Appendix B.2.

4 Estimation and Identification

4.1 Estimation method

Due to the complexity of the model, estimation by standard maximum likelihood is not feasible, and instead we estimate the model using the simulated method of moments. We choose model parameters to minimize the distance between a set of chosen moments from the data and the moments implied by the model using simulated careers (see McFadden (1989) for an early example).

We simulate careers starting from the point when - at 10 years of age - individuals are allocated to the lower, intermediate, or secondary school track. We follow individuals up to age 60 at a quarterly frequency. Using the resulting simulated data we then construct moments that correspond to those we obtain directly from the observed data.[21] We choose the weighting matrix $\hat{\Sigma}$ to be a diagonal matrix which contains the inverse sample variances of the observed moments. Our model contains both continuous outcomes such as wages and discrete ones (labor supply or occupational sector). The estimation method relies on semi-aggregate moments, detailed below, that smooth out the individual discrete decisions, such as in Blundell et al. (2016) or Adda et al. (2017). The standard errors are estimated as in Gourieroux et al. (1993). Estimation is based on the simulation of 24,000 individual careers. This number has been chosen so that the criterion function does not change when increasing the number of simulations further.

4.2 Identification of Life-cycle Outcomes

The model comprises a total of 88 parameters, describing twelve outcomes (wages, occupational sector, transitions from non-work to work, from work to non-work, and from work to work, all both for trained and untrained workers; and education/training choices at age 10 and 16) as a function of several variables and their interactions. To identify those parameters, we use a total of 358 static, conditional and dynamic moments, which are listed in Table 3.[22] These can be categorized into three types: those relating to wages, wage growth, and the variance of wages, those relating to labor market status and transitions, and those characterizing initial choices such as training decisions, initial

[21] For instance, for workers employed for a full calendar year in the same firm, the administrative data report an average of the wage over the year, even if there were wage changes. In the simulations, we also average wages for workers who stay with the same firm. We deal with top coding of wages by imposing the same coding rules in the simulated data as in the observed data. This procedure is similar to a Tobit model, given the normality assumptions for the shocks to wages.

[22] Eisenhauer et al. (2015) emphasize the importance of dynamic moments for the finite sample properties of the estimator.

occupation, and school tracks.

4.2.1 Wages and Transitions

To identify the various determinants of wages, we use moments based on wage equations, estimated both in levels and in first differences. The challenge is to disentangle the effects of different types of experiences and tenure as well as ability bias through positive sorting into training, occupations, and work, on wages. To do this, we rely on several sources of information. First, we rely on differences in coefficients between wage regressions in levels and in changes that provide information on the extent of ability bias. We also use as moments the mean wage conditional on potential experience for unemployed individuals, where the wage consists of the last earned wage while working. To the extent that unemployed individuals are negatively selected, this wage profile will be lower than the average wage of those who work. Second, information on wage growth within and between firms provides information on the importance of skills in determining wages and on the role of search capital.[23] Third, information on unobserved ability in RM or CA occupations is partly obtained from moments based on estimated wage fixed effects. We use panel wage regressions conditioning on CA and RM experience, tenure, occupational sector, training choices and the number of job to job transitions (to proxy for search capital). We then average the wage residual in each of the two sectors for each individual, to obtain a proxy for the ability parameters. We target the variance of these

[23]To further pin down the role of work experience, tenure, and search capital, we exploit information in our data on *exogenous* displacement of workers due to firm closure. We do not model firm closure directly, but our model has both voluntary and involuntary transitions to unemployment (noted δ, see section 3) that we can track. We therefore compare the coefficients of log wage regressions as a function of work experience for all workers who are starting a new job, and for those who are starting a new job post displacement. In both cases, workers have zero firm tenure, and their wages depend on their work experience and search capital. However, under the assumption that firm closure is an exogenous event conditional on observables, the latter group is a random sample of the workforce and workers are not selected into new jobs on the basis of their past choices (see Gibbons and Katz (1991)). Hence, contrasting the OLS returns to experience between those two groups provides information on the importance of job search for wage growth. In practice, we use for our data moments the estimates in Dustmann and Meghir (2005) (Tables 3 and 4, columns 2 and 3), who exploit the same data and firm closures, although for a shorter period.

ability fixed effects for both the trained and untrained workers. Using the subsample of workers who switch across occupational groups, we also compute the correlation between the ability fixed effects in the two sectors. The variances are functions of the skill prices $p^{i(j)}, i, j = \{RM, CA\}$ and the proportions of the unobserved "types" in the sample (π_l), as is the covariance between CA and RM fixed effects. We refer the reader to Appendix C.1 for derivations. Fourth, the variance of wages as a function of age and training status contributes to the identification of the variance of the wage match effects, as well as the distribution of unobserved ability. As pointed out by Heckman and Sedlacek (1985), the variance of wages across occupational sectors decreases with selection, which provides additional moment conditions to characterize unobserved heterogeneity.

We proceed in a similar way to obtain parameters related to labor market status and occupational transitions and we refer the reader to Appendix C for further details.

4.2.2 Initial Conditions

Individuals make three choices before entering the labor market. First, the school track choice between the lower two tracks and the upper track, made at age 10, where only the upper track qualifies for university entry. We model this choice by exploiting geographical information on the availability of university slots at the time when the track choice is made. Second, whether or not to acquire further training when entering the labor market for those who chose the lower two tracks (which is the population that we consider here). We instrument the training choice with local GDP measured as gross value added, and the number of available slots in apprenticeship schemes normalized by cohort size, both at commuting zone level and computed when individuals are 14 years old. Third, which occupational sector to choose. We instrument the initial occupational sector choice using the share of jobs in services (which are more concentrated in the CA sector) within the commuting zone. We provide more detail on these variables in Section 2.2.3. In all cases, we rely on a difference-in-difference design, controlling for time and area effects. Table 4 displays the first stage results for the three initial choices. In the first column we display the results for the school track choice, where we regress an indicator variable for upper track secondary school choice at age 10 on the (log) number of students attending universities located in a radius of 50 kilometers around the county of the individual, normalized by county population size, conditional on county and year fixed effects and state level linear trends, and where we cluster the standard errors at county level. The estimates are significantly different from zero, with an F statistic equal to 25, and imply that a one standard deviation increase in the ratio of university students to population when the focal individual is 10 years old raises the probability of choosing the upper track at secondary school by 1.2 to 1.6 percentage points.[24] Column 2 of Table 4 displays the first stage results for the choice of training or direct labor market entry, and column 3 for the occupational sector choice. Both available training slots and better local economic conditions significantly increase the probability that individuals choose apprenticeship training, with the joint F statistic being 11.6. Moreover, the share of service industry jobs in the commuting zone has a significant and positive effect on the choice of individuals to choose a CA occupation either for training, or as a first job when entering the labor market without further training, with an F statistic of 15.8.

We use these three first stage regressions as additional moments in our estimation. We construct the model counterparts of these regressions by assigning each simulated individual to a particular year of birth and to a commuting zone at ages 10-16, in such a way that the proportions of each group are similar to the proportions observed in the data. To further strengthen the identification of the model, we add regressions of log wages on either training choices or the initial sector, instrumented as explained above (see Table 3). This constitutes a "second stage" that identifies the selection into training or sector.

We further assess the local identification of our parameters and we also conduct a sensitivity test of our parameters with respect to some of the moments we use. Details are shown in Appendix C.4.

4.3 Goodness of Fit

To provide a first visual impression of the fit of the model, we display in Figure 7 the data moments against the simulated ones, with circles with areas proportional to the precision

[24]See Table 1 for descriptive statistics on the share of upper track secondary students, and Table A1 for the ratio of university students to population.

of the data moments. The moments are demeaned and scaled by the standard deviation. Overall, the circles align well with the 45 degree line, showing that the model is globally able to fit the many moments. A regression of the data moments on the simulated ones yields an R^2 close to 0.99, with a slope very close to 1.

Figure 8 depicts the fit of the model for some key outcomes over the life-cycle. The model tracks well the life-cycle profiles for wages, conditional on working.[25] The model also captures well the last earned wage of the unemployed. The latter has a lower profile than the former, especially later on in the life cycle, due to positive selection into work. The model also captures the non-linear profile of the proportion of individuals working, labor market mobility (number of jobs), duration of jobs (firm tenure) and experience in the RM and CA sectors.

We refer the reader to the tables in Appendix D for further results on the fit of the model broken down by training status, such as the relationship between wage levels or wage growth and work experience in different sectors, within and between firm decomposition of wage growth, the accumulation of skills over time and the number of firms an individual has worked for.

5 Results

5.1 Returns to Skills

Our parameter estimates in Table 5 correspond to the parameters in the wage equations (4) and (5). Panel A details the estimated annual increases in log wages for an additional year of experience and tenure. It should be noted that on average, it takes about 1.5 years (5 years) of potential experience to obtain one year of RM (CA) experience, due to spells of unemployment and the fact that experience is divided between

[25] Table A4 in the Appendix shows that the fit is also good when we distinguish individuals by training status. As in Figure 1, wages of trained and untrained workers increase steeply in the first years and flatten out after 5 to 10 years of potential experience. Trained workers' wages begin at lower levels than untrained workers, since they are still in training in the initial periods. However, after 4 to 5 years of potential experience we observe trained workers overtaking their untrained peers.

RM and CA skills (see equation (1)).[26] Given the slower accumulation of CA experience, we choose different intervals for the presentation of estimation results, reporting RM experience returns for the first 4 years of sectoral experience, years 5 to 10, and 10 to 30 years, and CA experience for the first two years of sectoral experience, years 2-5, and years 5-30.

The first four sets of results in Panel A are the returns to RM and CA experience in RM and CA occupations, for trained and untrained workers. The returns to RM experience in RM occupations is highly non-linear. Trained (untrained) workers experience a wage increase of 0.16 (0.09) log points for each year of RM experience of the first 4 years of RM experience. For trained workers, this figure includes the training period, but excludes the discrete jump when they obtain their qualification, which is parameterized separately and discussed below. Beyond that, the effect of RM experience to wage growth is considerably reduced, and close to zero. This suggests that accumulation of RM experience is most important early on in a worker's career (such as plumbing and carpeting skills), leading to rapid productivity increases. However, once the basic skills are acquired, additional experience does not lead to further improvement in productivity.

Returns to the first two years of CA experience in RM jobs are around 0.02 log points per year for trained workers, and close to zero for untrained workers. Interestingly, these returns increase for untrained workers as accumulated CA experience grows, and similarly trained workers experience their highest marginal returns to CA experience when they have accumulated more than 5 years worth. Thus, it seems that workers who have relatively little CA experience are likely to be in roles that have little use for CA skills. Once a worker has sufficient CA experience to take on more complex roles, such as supervisory or white collar positions, then further accumulation of CA experience becomes more useful to them. As it takes more years to accumulate CA experience in RM jobs, this explains a large part of the wage growth for workers in RM occupations in later years (and at higher levels of potential experience). This is qualitatively similar to

[26] For workers starting in the CA sector, it takes on average two years to accumulate one year of CA experience and 3 years to accumulate one year of RM experience. For those initially in the RM sector, it takes about 7 years to accumulate one year of CA experience and an average of 1.5 years for a year of RM experience.

the results displayed in Figure 6 and discussed in Section 2.3.

The next two panels describe the returns of CA and RM experience in CA occupations. Returns to CA experience in CA occupations is high initially, increasing wages by 0.09 (trained) and 0.08 (untrained) log points per year for the first two years. Returns slightly increase for trained workers over the next 2 to 5 years (corresponding to a period of 20 years of potential experience), but are low for untrained workers. Beyond that period, returns are close to zero for both groups. Finally, each year of RM experience leads to 0.06 (0.07) log points of wage growth in the CA sector for trained (untrained) workers for the first 4 years, and then drops close to zero for both group.

Overall, these findings provide a number of important insights. First, there are large differences in the returns to RM and CA experience, with the latter leading to more sustained wage growth over workers' entire careers, while wage growth induced by the former is concentrated to the early period of workers' careers. This suggests that RM skills, while very productive in RM occupations, are learned early on, while accumulation of new CA skills is increasing productivity throughout the career. Second, both RM and CA skills are partly portable when a worker changes sector, although each type of experience is valued more in its own sector. Third, and as a consequence of this, the differing returns may lead to lock-in effects, providing an important disincentive for workers to move across sectors. We come back to this below when we discuss the role of unobserved skills.

The last two rows in Panel A report the returns to firm tenure for both trained and untrained workers, which are close to zero, in line with much of the literature (see e.g. Altonji and Shakotko (1987), Neal (1995) and Dustmann and Pereira (2008)).[27] Small returns to tenure do not mean that there is no firm specific human capital - only that workers do not receive the returns in form of higher wages, as it enhances the worker's productivity only in a single firm, and there is therefore no competition for such human capital, as pointed out by Harris and Felli (1996).

Panel B of Table 5 displays estimates of the log wage intercept for both trained and untrained workers. The intercept is higher for untrained workers but that does not

[27] Larger returns are found by e.g. Topel (1991), Dustmann and Meghir (2005) and Buchinsky et al. (2010).

mean that trained workers obtain lower wages after the training period, as they have accumulated work experience and possibly search capital during training, and may also be positively selected on unobservable skills, a point we discuss below. On average, the wages of trained workers surpass those of the untrained after four years of potential experience. The log wage penalty during training (denoted α^{inTR} in equation (4)) is equal to -0.87 log points, which means that trainees get less than half of a full time wage. This parameter accounts for the discrete jump in wages when training is completed. The wage premium in CA occupations (α^{PCA}) is close to 0.04 log points. We return to the issue of the wage differential across sectors in Section 5.6. Estimates of the business cycle parameter (denoted α^{G} in equation (4)) show that wages are procyclical, with a difference between recessions and booms of about 5 percent.

The last panel of Table 5 presents estimates of the standard deviations of the initial match-specific productivities and their corresponding innovations, σ^j_κ and σ^j_u, $j \in \{RM, CA\}$, the dynamics of which we describe in equation (3). The estimates show that trained and untrained workers draw matches from a similar distribution in the RM sector, but untrained workers draw productivities with a larger initial variance in the CA sector. The standard deviation of the innovation of match-specific productivity (σ^j_u) is larger in the CA sector, which means that matches are less stable, as a bad draw may lead the worker to quit.

While the estimates in Panel A of Table 5 represent the causal effects of an additional year of working, they do not entirely explain the differential wage growth between individuals who made different training choices, as trained and untrained workers accumulate different levels of work experience and search capital over the years. To precisely evaluate the contribution of different stocks of experience to wage growth, we need to also take into account labor force attachment and job mobility. We address this in a comprehensive way in Section 5.4 below, using simulations to construct the appropriate counterfactuals. We first discuss the mobility decisions of workers, and the role of unobserved heterogeneity.

5.2 Labor Market Mobility

Estimates in Table 6 describe the dynamics of labor market mobility. Panel A displays the job destruction rate (δ), which is parametrically specified as a linear probability model where the conditioning variables are both CA and RM experience, a dummy for being in a CA occupation, a dummy for a match below a threshold quality (denoted $\underline{\kappa}$) and a business cycle indicator.

The estimates show that untrained workers have a higher job destruction rate at labor market entry (see estimate "intercept"). This rate diminishes quickly for both groups as they accumulate CA and RM experience. After 5 years of potential experience, the quarterly job destruction rate is on average 3 percent (6 percent) for trained (untrained) workers. The accumulation of CA experience decreases the rate of job destruction far quicker than the accumulation of RM experience.

Workers with poorer match-specific productivities, but potentially with high non-pecuniary match values, are more likely to be laid off, as indicated by the positive coefficient on the dummy for a match below the threshold quality $\underline{\kappa}$. The threshold $\underline{\kappa}$ itself is estimated as -0.08 (last row of Panel A), equal to about one standard deviation of the distribution of initial draws for the match-specific productivity. This implies that there is an increased risk of the job being destroyed for the bottom quarter of the match-specific productivity distribution, which introduces dynamic complementarities as accepting a job with a low match-specific productivity leads not only to immediate lower wages but also a greater likelihood of future unemployment, which would in turn lead to less accumulation of experience and lower future wages. Finally, we find that job destruction rates are pro-cyclical, with a lower probability of job destruction during economic expansions.

Panels B and C display the job offer arrival rate when the worker is either in employment or unemployed. As for the job destruction rate, these are specified as linear probability models with the coefficients estimated as shown. In both cases, the likelihood of a job offer increases with CA and RM experience. Given the differential accumulation of RM and CA experience, the rate of arrival of offers on the job is slightly higher for trained than for untrained workers over the whole life cycle. In contrast, we find a higher rate of offers when unemployed for untrained workers, although the difference becomes small later on in the career. Job arrival rates both when working and unemployed are procyclical.

Mobility is also determined by non-pecuniary job attributes (Panel D). The distribution of the non-pecuniary job attribute has a slightly higher variance in the RM than in the CA sector but is similar for trained and untrained workers. The higher the variance, the more scope there is to search for better jobs, independent of wages. We evaluate in Section 5.5 the relative importance of wages and non-pecuniary attributes for job mobility.

Panel E of the table presents the parameters describing mobility across occupational sectors, which depends on the probability of receiving an offer in the other sector (λ^O), and on the cost of moving, c_O. We find that trained individuals are far more likely to receive an offer to change sector, with a probability of 13 percent per quarter compared to less than one percent for untrained workers. The estimated cost of moving increases with potential experience, being about 0.07 log points of the quarterly wage at labor market entry, but about 15 percent of the annual wage after 10 years of potential experience. This implies a selection of those who switch sector, a point we return to in Section 5.6. As the fixed cost increase, cross-sector mobility is more advantageous earlier in workers' careers. This is in line with the fact that mobility is highest in the first 20 years of potential experience, and before age 35-40. Further, sectoral mobility will depend on productivity, as only those with a high expected return (as a result of observed and/or unobserved characteristics) will be willing to pay the cost to move.

5.3 Unobserved Heterogeneity

We allow for 4 different types of workers, where type 1 has neither unobserved CA or RM productivities, types 2 and 3 have only RM or CA productivities, and type 4 has both unobserved productivities (see section 3.2). The estimated proportions of these types, displayed in the first row of Table 7, show that about 13 percent of workers have only RM skills, 10 percent have only CA skills, and 18 percent (59 percent) have neither (both) productivities. This distribution of unobserved skills affects initial choices, with those who have CA skills or both RM and CA skills being more likely to opt for training at the start of their career (see row 2 of Table 7). Further, those with only CA skills and therefore with a comparative advantage in the CA sector are also more likely to work in that sector (row 3).

Panel B of the table displays estimates of the prices of unobserved skills in either sector (see equation (2)).[28] The price of RM skills is 0.09, while the price of CA skills is 0.27, meaning that those who have unobserved RM skills earn about 10 percent higher wages in the RM sector, while those who have CA skills earn 31 percent higher wages in the CA sector. These differences also imply initial sorting into the two sectors based on unobserved productivity, which we study in more detail in Section 5.6.

5.4 Decomposition of Wage Growth

We now investigate the dynamic implications of the vocational training choice, by decomposing the growth of wages into its different components, separately for those who did and did not choose to undergo training. This allows us to distinguish the contributions of different types of skills and search capital to wage growth.

Presenting simple decompositions of simulated wage growth into within and between firm components as in Topel and Ward (1992), Panel A of Table 8 shows that over the first 5 years of potential experience wages grow on average by 0.34 log points, of which 0.13 log points (38%) are due to between-job wage growth. In these early years, between-job wage growth is a more important driver of overall wage growth for untrained than for trained workers. Over the next 5 years, wage growth slows considerably to 0.1 log points, of which 0.04 log points are due to between-job wage growth. Overall, these figures are similar, although with slightly lower growth rates, to those reported by Topel and Ward (1992), who find for the US that between-job wage growth accounts for around 46 percent of the total wage growth over the first 5 years in the labor market (see their table VII). However, job mobility in the US is higher than in Germany.

We next turn to the decomposition of wage growth to determine the importance of different channels over the life cycle. This decomposition uses the estimates for the wage

[28] We set prices of unobserved CA (RM) skills in the RM (CA) sector to zero, as we do not have enough variation in our data to estimate these cross prices for both observed and unobserved skills.

equation in Table 5 as well as simulated life cycle profiles to determine the amount of sectorial experience, firm tenure, firm-worker matches or business cycle effects at each age. To determine wage growth not only is the return to a specific skill important but also how much of that skill is accumulated at each point in time.

Panel B of Table 8 reports the contribution to overall wage growth of three components: RM experience, CA experience and search capital. In our model, search capital is represented by the growth in the match-specific productivity κ_{it}, while the contributions of experience are captured by the terms involving X_{it}^{CA} and X_{it}^{RM} in equation (4). We ignore other factors that have little effect on wage growth in the first years such as firm tenure, the business cycle, or changes in occupational sector (which we investigate later on). [29]

In the first five years, the largest part (43 percent) of wage growth is due to the accumulation of RM experience, while search capital contributes 30 percent, and CA experience 25 percent.[30] During the next 5 years, the relative contribution of these three factors changes, with search capital becoming more important (46 percent), followed by CA experience and RM experience (respectively 38 and 11 percent). This is in line with the non-linear and strongly decreasing return to RM experience and the increasing return to CA experience that we illustrate in Table 5. When we distinguish wage growth over the first 5 years by training status, we find that the accumulation of RM experience is more important for untrained than for trained workers, as for the latter we start measuring wage growth at the end of their training where they have already accumulated RM and CA experience.

[29] The contribution of the firm-worker match within firms to wage growth is minor, even though we model it as a random walk, which allows us to identify search capital throught the term κ_{it}. This minor contribution of the firm-worker match to within firm growth is due to two compensating forces: workers with a negative shock do not quit the firm immediately, as frictions imply that new and better offers take time to arrive, and non-pecuniary attributes of matches that may compensate for the fall in the productivity match. By setting the variance of the innovation to zero, wage growth decreases by about 0.005 log points for the first 5 years, and by the same amount for the next five years.

[30] The generosity of unemployment insurance can play a role in the importance of search capital. As German unemployment benefits are proportional to earnings in the previous job, unemployed individuals with a previously high match have also higher reservation wages. Countries with a less generous replacement rate or a flat rate would have a lower role for search capital in fostering wage growth.

While this decomposition provides important insights into the drivers of wage growth, it does not immediately indicate the implications for labor market policy. For instance, while suggesting that search capital is an important factor for overall wage growth, this simple decomposition does not address the question of whether mobility should be encouraged to improve wage growth, as it does not define clear counterfactuals. In order to analyze policy implications, we must compare the outcomes resulting from a change in e.g. the job offer rate to those in a status quo baseline scenario, something our model allows us to do. To gain more insight into the role of mobility in wage growth, we therefore study how an increased rate of job offer arrivals while employed (λ^W, see Section 3.5) will affect wage growth in the longer term. We increase λ^W by 0.1, from the baseline value of 0.14 for trained and untrained workers (see the intercept in Table 6, Panel B).[31] We assume that individuals do not anticipate this departure from the baseline, which means that we solve the model and the optimal decisions for the baseline parameter values only.[32]

As might be expected, the effect is an increase of 37 percent in the number of job-to-job transitions in the first five years in the labor market (first row in Table 9), which reduces to 24 percent when instead calculated over a 20 year period. This corresponds to an increase of 0.34 jobs held by each individual in the first 5 years, and an additional 0.58 job-to-job transitions over a 20 year period. Row (3) of Table 9 shows that the increased mobility increases wages by about 1.6 (1.0) percent over a 5 (20) year horizon. To put these numbers in context with the standard between-within decomposition, we report in row (4) the average baseline between-job wage growth scaled by the number of job-to-job transitions, showing that each realized job transition lead to a 0.15 log point increase in wages. In contrast, row (5) shows that the *marginal* effect of additional labor market mobility induced by an increase in λ^W is considerably smaller, at about 0.02 log points per additional job. The reason for this disparity is that while the first job move in a worker's career generates large gains, these gains quickly decline, so that inducement

[31]While the choice of an increase of 0.1 is arbitrary, the results in row (5) of Table 9 are similar with other increments to the parameter λ^W.

[32]This implies that we keep individual behavior constant between scenarios, which abstracts from selection effects on unobservables.

of additional mobility does not produce large returns (see also the left panel of Figure 4).

In summary, the literature following Topel and Ward (1992) has emphasized the importance of search capital for wage growth, and our findings are largely in line with their results. However, while mobility contributes to the observed wage growth, this does not mean that inducing additional mobility through the provision of more job offers will lead to substantially larger wage growth, especially in the longer run, as the marginal return to additional transitions is far smaller than the average return. Finally, a large part of wage growth comes through the accumulation of both RM and CA experience, which play distinct roles over the life cycle, with the former instigating wage growth early on, while the latter is important later on in the career.

5.5 Are Workers Income Maximisers?

Our analysis shows that workers are neither pure income maximizers, nor do they choose jobs solely for their non-pecuniary attributes. To evaluate how important each factor is, we calculate how lifetime income (in present value terms at the start of workers' careers) would change if individuals maximized either only wages or only jobs' non-pecuniary attributes. We find that if job choice was solely driven by non-pecuniary attributes, workers would face a present value loss of about 17 percent in lifetime earnings. In contrast, if workers were only pursuing higher wages, they would gain about 6 percent in lifetime earnings. Using that metric, income maximization appears to be a more important driver of job mobility. However, a model ignoring job amenities altogether would fail to describe job mobility and wage growth adequately. In particular, a number of job-to-job transitions do not result in wage increases, a finding that is in line with results for the US (Hall and Mueller, 2018).

5.6 Occupational Sorting and Lock-in Effects

Workers in the CA sector earn on average higher wages than those in the RM sector. As argued in Section 3.4 when discussing equation (4), this could be for different reasons, such as the occupational premium α^{PCA} or sorting on observable or unobservable characteristics. We show in Table 7 that indeed, workers sort themselves into training and the CA sector based on unobserved ability. We now explore to what extent sorting can account for sectoral differences in wages.

To this end, we compare the log wage differential between the CA and RM sectors for the baseline specification to one in which there are no returns to unobserved ability (i.e. $\alpha_i^S(o_{it}) = 0$). This is shown in Figure 9 as a function of potential experience. At labor market entry, the wage differential is about 0.05 log points, which falls to 0.03 log points if we remove the unobserved ability component. After 5 years of potential experience the wage differential turns negative, due to some workers switching from the RM sector to the CA sector. These individuals have little CA experience and their RM experience is less rewarded in the CA sector, so there is a temporary fall in relative wages. After that period, the wage differential across sectors grows rapidly, due to the faster accumulation and higher return of CA experience in the CA sector. The figures indicate that the contribution of sorting on unobserved ability across sectors to life-cycle wage differences across the RM and CA sectors is quite modest. The main part of the wage differences across sectors stems from the differential accumulation of experience, as well as from different returns to RM and CA experience.

Differential skill acquisition, relative returns across sectors combined with search frictions can generate lock-in effects, where e.g. workers in the RM sector who may be inclined to switch to the CA sector but do not receive an offer to do so at an early stage in their career will accumulate a larger stock of RM experience, so that it becomes increasingly less attractive to change sector. To assess the magnitude of this effect, we construct a counterfactual situation in which offers to move to the CA sector arrive only after several years in the labor market. When young workers entering the labor market cannot make a transition to the CA sector within the first 4 years, about 6 percent of them stay locked in the RM sector. On average, this leads to a 0.3 percent decline in lifetime wages. However, the average treatment effect on the treated is considerably larger, representing a loss of about 10 percent of lifetime wages.

Similar lock-in effects occur e.g. in recessions where the variety of job offers is decreased, and may particularly hurt young workers due to the intense accumulation of sector specific skills early on in the career. Such lock-in effects may therefore be an important explanation for persistent wage disadvantages of cohorts who enter the labor market during recessions (see e.g. Oreopoulos et al. (2012) and Altonji et al. (2016)).

5.7 Internal Rate of Returns to Apprenticeship

Figure 1 suggests that wages of trained workers, while substantially lower than those of untrained workers during the training period, increase sharply once training is complete, and surpass those of untrained workers. Thus, while during the training phase apprentices incur a relative loss, there seems to be a sustained advantage afterward. Figure 1 however is purely descriptive, and ignores both selection of workers into training, and how training affects the type of jobs obtained, unemployment spells, and job finding probabilities. Moreover, to train workers is costly both for firms and society, costs that are ignored in the figure.

To explore this further, we display in Table 10 estimates of internal rates of return to individuals and society, following Heckman et al. (2010). We present results without (first column) and with (second column) controlling for selection into training. We account for selection by first computing the internal rate of return for each individual type, and then re-weighting the returns using the estimated population weights, π_l (see Section 3.2). The first row of the table shows the rates of return from the perspective of those individuals undertaking the training. Calculations using only wages when working are displayed in the first line. The internal rate of return equals 14 percent, but this decreases to 10 percent when controlling for selection, which suggests positive sorting of workers into training. Given that the training period in our sample is on average about 2.5 years, these numbers are similar to estimates by Fersterer et al. (2008) for Austria, who find returns of 2.5 to 4 percent per year, but smaller than earlier OLS estimates for Germany, which are on the order of 15-20 percent per year (see e.g. Krueger and Pischke (1995) and Winkelmann (1996)).

When we take into account unemployment benefits, calculating the costs and gains in terms of total income rather than just wages of the employed, the rate of return (corrected for selection) declines to 9 percent. The reason is that unemployment is more frequent for untrained workers (see Table 2 and Figure 2). We next evaluate the returns to apprenticeship training taking also account of utility derived from leisure when out of work, γ, and from non-pecuniary job attributes, μ_{it}. As untrained individuals have a lower labor market attachment, accounting for the utility of leisure reduces the internal rate of return to training (to 7.7 percent when taking selection into account). On the other hand, non-pecuniary work attributes favor those in work, which increases the rate of return to 9.1 percent. Accounting for all of these factors together, the internal rate of return is about 8 percent.

As we point out above, provision of training is costly for firms, and apprentices attend state sponsored schools during the training period, incurring costs for the tax payer. To assess the internal rate of return of training provision from a societal perspective, we need to account for such costs. We obtain estimates of the costs of state run schools for vocational training of firm-based apprentices from Bundesinstitut für Berufsbildung (2013) for 2017, who report a total of about 3.1 billion euros for the 1,391,886 individuals in firm based training schemes in Germany that year (Destatis, 2013). Hence, on average, each apprentice costs the state about 2,227 euros per year. Moreover, Jansen et al. (2015) conduct an extensive survey to assess the costs of training for firms. They estimate that each apprentice costs firms about 5170 euros per year. As shown in Table 10, taking into account the cost of training decreases the returns from the viewpoint of society to about 8 percent. Next we take into account unemployment benefits, seen here as a cost to society. This raises the return to 12 percent, as trained individuals are less likely to be unemployed. Finally, when we combine both the cost of training and unemployment, the overall rate of return from the societal point of view is equal to about 10 percent.

As such, we can see that correctly accounting for selection is important in quantifying the returns to training, and we find that however we calculate it, this training provides large returns to both the individual and society in general.

6 Conclusion

The analysis in this paper provides important insights into the determinants of workers' career trajectories. Our finding that non-monetary factors play an important role in workers' mobility decisions supports results of earlier papers and emphasizes job amenities as a key determinant of workers' choices. Our distinction between occupational sectors with different possibilities for accumulation of and returns to RM and CA experience reveals RM skills as a key driver of early wage growth, while CA skills become important later on in the lifecycle. CA skills may thus provide advantages for pursuing career possibilities that require supervisory, white collar, or technical competencies, that may materialize at later career stages. These findings stress the importance of different forms of skills whose returns are reaped at different stages over the career cycle. Moreover, they also highlight lock-in effects as a form of labor market inefficiency that exacerbates the presence of frictions, where workers who would, ex ante, be more suited to one sector but are initially allocated to the other, subsequently choose not to change sector in order to avoid losing the benefit of accumulated sector specific skills. Such occupational lock-in effects may be aggravated by recessions that restrict sector choices, in particular for workers who enter the labor market in downturns.

We confirm the importance of search capital for workers' careers. We also show that gains from marginal mobility (i.e. one additional job change over the career cycle) are small, in comparison to the large wage growth associated with mobility early on in workers' careers. Despite the importance of mobility for wage growth early on in workers' careers, we find that the accumulation of CA and RM skills together are a more important driver. Our analysis of the initial training choice and its returns suggests that firm based vocational training induces stronger labor market attachment and eases access to the CA sector, with substantial benefits in the longer run. Hence we uncover important dynamic complementarities of training as in Cunha and Heckman (2007), but which manifest at a later stage of the life-cycle, as training compensates workers by shaping the dynamics of future job search and human capital accumulation. The internal rates of return to apprenticeship training we compute reveal considerable benefits both for the individual and society (at around 8 and 10 percent respectively). However, our analysis emphasizes that assessment of vocational training based on immediate returns alone may be inadequate, and that it is subtler longer term effects on career outcomes that bring gains.

References

Abowd, J., F. Kramarz, and D. N. Margolis (1999). "High wage workers and high wage firms. 67, no. 2:251–333." *Econometrica*, 67(2), 251–333.

Adda, J., C. Dustmann, C. Meghir, and J.-M. Robin (2006). "Career progression and formal versus on-the-job training." IZA Discussion Papers 2260.

Adda, J., C. Dustmann, and K. Stevens (2017). "The Career Cost of Children." *Journal of Political Economy*, 125(2), 293–337.

Altonji, J. and R. Shakotko (1987). "Do Wages Rise with Job Seniority?" In *Unemployment, trade unions, and dispute resolution*, edited by O. Ashenfelter and K. Hallock, volume 47, pages 219–241. International Library of Critical Writings in Economics.

Altonji, J. G., L. B. Kahn, and J. D. Speer (2016). "Cashier or Consultant? Entry Labor Market Conditions, Field of Study, and Career Success." *Journal of Labor Economics*, 34(S1), S361–S401.

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017). "Measuring the Sensitivity of Parameter Estimates to Estimation Moments." *Quarterly Journal of Economics*, 132(4), 1553–1592.

Antonczyk, D., B. Fitzenberger, and U. Leuschner (2009). "Can a Task-Based Approach Explain the Recent Changes in the German Wage Structure?" *Journal of Economics and Statistics*, 229(2/3), 214–238.

Arcidiacono, P., H. Sieg, and F. Sloan (2007). "Living Rationally Under the Volcano? An Empirical Analysis of Heavy Drinking and Smoking." *International Economic Review*, 48(1), 37–65.

Autor, D. H., F. Levy, and R. J. Murnane (2003). "The skill content of recent technological change: An empirical exploration." *Quarterly Journal of Economics*, 118(4), 1279–1333.

Bagger, J., F. Fontaine, F. Postel-Vinay, and J.-M. Robin (2014). "Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics." *American Economic Review*, 104(6), 1551–96.

Belzil, C., J. Hansen, and X. Liu (2017). "Dynamic skill accumulation, education policies, and the return to schooling." *Quantitative Economics*, 8, 895–927.

Blundell, R., M. Costa Dias, M. C., and S. J (2016). "Female Labor Supply, Human Capital, and Welfare Reform." *Econometrica*, 84(5), 1705–1753.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2020). "How Much Should we Trust Estimates of Firm Effects and Worker Sorting?"

Bonhomme, S., T. Lamadon, and E. Manresa (2019). "A Distributional Framework for Matched Employer Employee Data." *Econometrica*, 87(3), 699–739.

Buchinsky, M., D. Fougère, F. Kramarz, and R. Tchernis (2010). "Interfirm Mobility, Wages and the Returns to Seniority and Experience in the United States." *The Review of Economic Studies*, 77(3), 972–1001.

Bundesinstitut für Berufsbildung (2013). "Datenreport zum Berufsbildungsbericht 2013. Informationen und Analysen zur Entwicklung der beruflichen Bildung." page 277, Bonn.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006). "Wage Bargaining with On-the-Job Search: Theory and Evidence." *Econometrica*, 74, 323–364.

Cainesa, C., F. Hoffmann, and G. Kambourov (2017). "Complex-task biased technological change and the labor market." *Review of Economic Dynamics*, 25, 298–319.

Card, D., J. Heining, and P. Kline (2013). "Workplace Heterogeneity and the Rise of West German Wage Inequality." *The Quarterly Journal of Economics*, 128(3), 967–1015.

Cunha, F. and J. J. Heckman (2007). "The Technology of Skill Formation." *The American Economic Review*, 97(2), 31–47.

Destatis (2013). "Berufliche Bildung." Number Reihe 3, page 23.

Dey, M. S. and C. J. Flinn (2005). "An Equilibrium Model of Health Insurance Provision and Wage Determination." *Econometrica*, 73(2), 571–627.

Di Addario, S., P. Kline, R. Saggio, and M. Sølvsten (2020). "It ain't where you're from, it's where you're at: hiring origins, firm heterogeneity, and wages." Mimeo.

Dix-Carneiro, R. (2014). "Trade Liberalization and Labor Market Dynamics." *Econometrica*, 3, 825–885.

Dustmann, C. (2004). "Parental background, secondary school track choice, and wages." *Oxford Economic Papers*, pages 209–230.

Dustmann, C. and C. Meghir (2005). "Wages, experience and seniority." *Review of Economic Studies*, 72(1), 77–108.

Dustmann, C. and S. Pereira (2008). "Wage Growth and Job Mobility in the UK and in Germany." *Industrial and Labor Relations Review*, 61(3), 374–393.

Eisenhauer, P., J. J. Heckman, and S. Mosso (2015). "Estimation of Dynamic Discrete Choice Models by Maximum Likelihood and the Simulated Method of Moments." *International Economic Review*, 57(3), 331–357.

Fersterer, J., J.-S. Pischke, and R. Winter-Ebmer (2008). "Returns to Apprenticeship Training in Austria: Evidence from Failed Firms." *Scandinavian Journal of Economics*, 110(4), 733–753.

Flinn, C. (1986). "Wages and Job Mobility of Young Workers." *Journal of Political Economy*, 94, S88–S110.

Gathmann, C. and U. Schoenberg (2010). "How General is Human Capital? A Task-Based Approach." *Journal of Labor Economics*, 28, 1–49.

Gertler, M., C. Huckfeldt, and A. Trigari (2020). "Unemployment Fluctuations, Match Quality, and the Wage Cyclicality of New Hires." *Review of Economic Studies*, 87(4), 1876–1914.

Gibbons, R. G. and L. F. Katz (1991). "Layoffs and lemons." *Journal of labor Economics*, 9(4), 351–380.

Goos, M. and A. Manning (2007). "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain." *Review of Economics and Statistics*, 89(1), 118–133.

Gourieroux, C., A. Monfort, and E. Renault (1993). "Indirect Inference." *Journal of Applied Econometrics*, 8, S85–S118.

Guglielminetti, E., R. Lalive, P. Ruh, and E. Wasmer (2018). "Home Sweet Home? Job Search with Commuting and Unemployment Insurance." Mimeo.

Hall, R. E. and A. I. Mueller (2018). "Wage Dispersion and Search Behavior: The Importance of Nonwage Job Values." *Journal of Political Economy*, 126(4), 1594–1637.

Harris, C. and L. Felli (1996). "Learning, Wage Dynamics, and Firm Specific Human Capital." *Journal of Political Economy*, 104, 838–868.

Heckman, J. and G. Sedlacek (1985). "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market." *Journal of Political Economy*, 93(6), 1077–1125.

Heckman, J. J. and B. E. Honoré (1990). "The Empirical Content of the Roy Model." *Econometrica*, 58(5), 1121–1149.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz (2010). "The rate of return to the HighScope Perry Preschool Program." *Journal of Public Economics*, 94, 114–128.

Hoffmann, F. (2018). "An Empirical Model of Life-Cycle Earnings and Mobility Dynamics." Mimeo.

Jansen, A., H. Pfeifer, G. Schönfeld, and F. Wenzelmann (2015). "Apprenticeship training in Germany remains investment-focused – results of BIBB Cost-Benefit Survey 2012/13"." BIBB Report 1/2015.

Keane, M. and K. Wolpin (1994). "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation: Monte Carlo Evidence." *Review of Economics and Statistics*, 76(4), 648–672.

Keane, M. and K. Wolpin (1997). "The Career Decisions of Young Men." *Journal of Political Economy*, 105(3), 473–522.

Krueger, A. and J. S. Pischke (1995). "A Comparison of East and West German Labor Markets Before and After Unification." In *Differences and Changes in Wage Structures*, edited by R. Freeman and L. Katz, pages 405–445. University of Chicago Press, Chicago.

Le Barbanchon, T., R. Rathelot, and A. Roulet (2019). "Gender Differences in Job Search: Trading off Commute Against Wage." Mimeo.

Lise, J. and J.-M. Robin (2017). "The Macro-dynamics of Sorting between Workers and Firms." *American Economic Review*, 107(4), 1104–1135.

McFadden, D. (1989). "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration." *Econometrica*, 57(5), 995–1026.

Menzio, G. and S. Shi (2011). "Efficient Search on the Job and the Business Cycle." *Journal of Political Economy*, 119(3), 468–510.

Murphy, K. M. and F. Welch (1992). "The Structure of Wages." *The Quarterly Journal of Economics*, 107(1), 285–326.

Neal, D. (1995). "Industry-Specific Human Capital: Evidence from Displaced Workers." *Journal of Labor Economics*, 13(4), 653–677.

Oreopoulos, P., T. Von Wachter, and A. Heisz (2012). "The Short- and Long-Term Career Effects of Graduating in a Recession." *American Economic Journal: Applied Economics*, 4(1), 1–29.

Parey, M. (2009). "Vocational Schooling versus Apprenticeship Training. Evidence from Vacancy Data." Mimeo, University of Essex.

Postel-Vinay, F. and J. M. Robin (2002). "Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica*, 70(6), 2295–350.

Roy, A. D. (1951). "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers*, (2), 135–146.

Roys, N. and C. Taber (2016). "Skills Prices, Occupations and Changes in the Wage Structure." Mimeo, University of Wisconsin-Madison.

Sauer, R. M. (2004). "Educational Financing and Lifetime Earnings." *Review of Economic Studies*, 71(4), 1189–1216.

Steedman, H., H. Gospel, and P. Ryan (1998). "Apprenticeship: A Strategy For Growth." CEP discussion paper No.CEPSP11.

Sullivan, P. (2010). "A Dynamic Analysis of Educational Attainment, Occupational Choices, and Job Search." *International Economic Review*, 51(1), 289–317.

Taber, C. and R. Vejlin (2020). "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market." *Econometrica*, 88(3), 1031–1069.

Topel, R. (1991). "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority." *Journal of Political Economy*, 99(1), 145–176.

Topel, R. and M. Ward (1992). "Job Mobility and the Careers of Young Men." *Quarterly Journal of Economics*, 107(2), 439–479.

Willis, R. and S. Rosen (1979). "Education and Self-Selection." *Journal of Political Economy*, 87(5), S7–S36.

Winkelmann, R. (1996). "Training, Earnings and Mobility in Germany." *Konjunkturpolitik*, 42, 275–298.

Table 1: Proportion with Different Education Statuses, by Year of Birth

	Birth Cohorts		
	1960	1965	1970
Upper secondary track	20%	21%	24%
Apprenticeship training	64%	67%	65%
No post-secondary education	16%	12%	11%

Source: IAB Social Security data, 1975-2004. See Sections 2.2 and A for details.

Table 2: Observed Quarterly Labor Market Transitions

	Potential Experience (Years)					
Labor market transitions	Untrained			Trained		
	0-5	5-10	10-20	0-5	5-10	10-20
Unemployed to unemployed	0.88	0.92	0.95	0.73	0.81	0.92
Unemployed to work	0.12	0.07	0.05	0.27	0.19	0.08
Work to unemployed	0.09	0.05	0.03	0.03	0.04	0.02
Work to new work	0.04	0.03	0.02	0.02	0.04	0.03
Work to same work	0.87	0.92	0.94	0.95	0.92	0.95

Notes: IAB Social Security data, 1975-2004, aggregated at a quarterly frequency. See Sections 2.2 and A for details. Potential experience starts at first entry into labor market and/or start of training period.

Table 3: Moments Used in Estimation

Moments	Number of moments
Wages:	
Wage by age and training choice	20
Variance of wages by age and training choice	12
Mincer wage regression in levels including experience, tenure, business cycle, sector and training choice	32
Regression of squared wage residual from Mincer wage regression in levels, on experience, tenure, sector and training choice	16
Mincer wage regression in changes including experience, tenure, job to job change and training choice	19
Percentiles of wage changes, by training choice	8
Regression of squared wage residual from Mincer wage regression, in changes on experience, tenure and training choice	11
Variance of wage fixed effects by sector from wage regressions on experience, tenure, training choice and mobility	4
IV and OLS regressions of log wage on potential experience and instrumented education	7
IV regression of log wage on potential experience and instrumented initial sector	2
Correlation between wage fixed effects in RM and CA sector from wage regressions on experience, tenure, training choice and mobility	4
Wage difference post displacement between voluntary and involuntary quits	7
Profile of last wages of unemployed, training choice, age	15
Total wage growth, by training choice, age, within and between firm	20
Proportion with negative job to job wage growth, by education and skills	5
Proportion of top-coded wages, by training choice	2
Labor market status and transitions:	
Proportion working by age and training choice	20
CA experience by age and training choice	14
RM experience by age and training choice	14
Tenure by age and training choice	14
Number of jobs by age and training choice	14
Proportion in CA sector by age and training choice	14
Probability of sectoral change, by experience and training choice	8
Probability of labor market transition, by experience, sector, training choice and past labor market status	52
Probability of job to job transition, by age, experience and training choice	8
Probability of working as a function of wage fixed effects, training and age	4
Probability of transition to CA sector as a function of RM wage fixed effect	2
Choice of training and initial sector:	
Proportion trained	1
Proportion trained as a function of subsequent wage fixed effect obtained from wage regression on experience, tenure, sector	2
Probability of training as a function of offered training positions and local Gross Value Added	3
Probability of initial sector as a function of proportion of firms in service sector in local area	2
Probability of lower track education at age 11 as a function of local number of university students	2
Total	358

Table 4: First Stage: Initial Determinants of Choice of Upper Track at Secondary School, Apprenticeship Training, and Sector

Variable	Secondary School Upper Track (1)	Apprenticeship Training (2)	CA Sector (3)
Log uni places in vicinity	.01645		
	(.0033)		
Local log number of training positions		.042	
		(.0194)	
Local log GVA		.508	
		(.120)	
Share of jobs in services			.826
			(.208)
Time fixed effects	Yes	Yes	Yes
Commuting zone fixed effects	Yes	Yes	Yes
State linear trends	Yes	Yes	Yes
F test	25.3	11.6	15.6
p-value	0.0007	<0.0001	0.0001
Obs	71,472	37,619	51,545
R^2	0.052	0.023	0.013

Notes: Regressions performed on IAB data using a linear probability specification. In (1), the dependent variable is equal to one if the individual went to a higher track at secondary school. The explanatory variable is the log of university places when the individual is 10 years old, in a radius of 50 kilometers around the county of residence and divided by total population. In (2), the dependent variable is equal to one if the individual graduated from apprenticeship training. The explanatory variables are the cyclical component of the log of annual gross value added at commuting zone level and the log of the annual number of apprenticeship training positions scaled by cohort size at commuting zone level. In (3), the dependent variable is the initial choice of CA sector and the explanatory variable is the share of jobs in the service industry (banking, hospitality, and other services) in the commuting zone. The regression includes time fixed effects, commuting zone fixed effects, and state linear trends. All variables are measured when the individual is 14. Standard errors are clustered at commuting zone level.

Table 5: Estimated Parameters: Wage Equations

Parameter	Trained		Untrained	
Panel A: Annual returns to experience and tenure				
RM experience, in RM sectors, $\alpha^{RM}(o_{it} = RM)$:				
RM experience $\in$ [0,4]	0.16	(0.0008)	0.092	(0.0003)
RM experience $\in$ [4,10]	0.0017	(0.0002)	0.0011	(0.0008)
RM experience $\in$ [10,30]	0.0019	(0.0001)	0.0096	(0.0007)
CA experience, in RM sector, $\alpha^{CA}(o_{it} = RM)$:				
CA experience $\in$ [0,2]	0.022	(0.005)	0.0052	(0.0005)
CA experience $\in$ [2,5]	0.0033	(0.0004)	0.037	(0.003)
CA experience $\in$ [5,30]	0.074	(0.0002)	0.074	(0.0001)
CA experience, in CA sector, $\alpha^{CA}(o_{it} = CA)$:				
CA experience $\in$ [0,2]	0.091	(0.0005)	0.079	(0.002)
CA experience $\in$ [2,5]	0.11	(0.00033)	0.0017	(0.0017)
CA experience $\in$ [5,30]	0.0004	(0.0001)	0.0024	(0.0012)
RM experience, in CA sector, $\alpha^{RM}(o_{it} = CA)$:				
RM experience $\in$ [0,4]	0.06	(0.0003)	0.0033	(0.002)
RM experience $\in$ [4,10]	0.003	(0.0002)	0.0013	(0.003)
RM experience $\in$ [10,30]	0.051	(0.0002)	0.0031	(0.006)
Firm tenure (α^{Ten}):				
tenure $\in$ [0,5]	0.00069	(0.0002)	0.0003	(0.0008)
tenure $\in$ [5,30]	0.0015	(0.0001)	0.00087	(0.0001)
Panel B: Wage intercept, sector, and business cycle effects				
Log Wage Constant	3.7	(0.001)	3.84	(0.001)
Wage penalty when training	-0.87	(0.0001)		
Wage premium in CA sector	0.036	(0.0001)	0.036	(0.0001)
Business cycle, high	0.05	(0.0001)	0.05	(0.0001)
Panel C: Standard deviation of firm-worker productivity match				
Std dev of initial match, RM sector (σ_κ)	0.23	(0.0001)	0.23	(0.0001)
Std dev of initial match, CA sector (σ_κ)	0.23	(0.0001)	1.3	(0.0002)
Std dev of innovations, RM sector (σ_u)	0.025	(0.0001)	0.013	(0.0001)
Std dev of innovations, CA sector (σ_u)	0.034	(0.0001)	0.021	(0.0001)

Notes: Log wage is the dependent variable. The experience clock starts at the start of apprenticeship for trained individuals and at the entry into the labor market for untrained workers. Asymptotic standard errors in parentheses. It takes respectively about 1.5 and 4.5 years of labor market experience to reach one year of routine-manual or cognitive-abstract experience. This is due to spells of unemployment and the fact that workers accumulate less than a unit of each type of experience per unit of time (see equation (1)).

Table 6: Estimated Parameters: Labor Market Mobility

Parameter	Trained	Untrained
Panel A: Quarterly job destruction rate (δ):		
Intercept	0.07	0.09
	(0.001)	(0.004)
CA experience (years)	-0.021	-0.025
	(0.024)	(0.01)
RM experience (years)	-0.004	-0.003
	(0.056)	(0.025)
Employed in CA sector	0.001	0.001
	(0.017)	(0.017)
Firm match productivity < $\underline{\kappa}$	0.1	
	(0.004)	
Business cycle high	-0.04	
	(0.039)	
Firm match productivity threshold ($\underline{\kappa}$)	-0.08	
	(0.001)	
Panel B: Quarterly offer arrival rate when employed (λ^W):		
Intercept	0.14	0.14
	(0.003)	(0.018)
CA experience (years)	0.006	0.002
	(0.053)	(0.056)
RM experience (years)	0.005	0.0003
	(0.084)	(0.002)
Business cycle high	0.13	
	(0.0005)	
Panel C: Quarterly offer arrival rate when unemployed (λ^U):		
Intercept	0.116	0.32
	0.022)	(0.001)
CA experience (years)	0.056	0.092
	(0.008)	(0.031)
RM experience (years)	0.021	0.013
	(0.037)	(0.001)
Business cycle high	0.028	
	(0.049)	
Panel D: Non-pecunary job attribute:		
Std dev of non-pecunary job attribute (σ_μ), RM sector	0.05	0.05
	(0.049)	
Std dev of non-pecunary job attribute (σ_μ), CA sector	0.04	0.04
	(0.0245)	(0.00035)
Panel E: Change of sector:		
Quarterly arrival rate of sector change offers (λ^O)	0.13	0.006
	(0.015)	(0.067)
Cost of mobility, intercept (c_O)	0.07	
	(0.004)	
Cost of mobility, potential experience effect	0.05	
	(0.007)	

Note: Potential experience is counted from the beginning of apprenticeship for trained individuals and from entry into the labor market for untrained workers. The job destruction rate and offer arrival rates are modelled as linear probability models with the arguments and parameters displayed in this table. Asymptotic standard errors in parentheses.

Table 7: Unobserved Heterogeneity

Panel A:	Unobserved ability			
Skill set (RM,CA)	(0,0)	(1,0)	(0,1)	(1,1)
Type	(1)	(2)	(3)	(4)
Proportion in sample (π_l)	0.18	0.13	0.10	0.59
Proportion in Apprenticeship	0.87	0.87	0.95	0.93
Proportion in CA sector	0.14	0.14	0.44	0.15
Correlation between RM and CA ability	0.4			

Panel B: Price of unobserved skills by sector (log wage scale):		
	RM skills	CA skills
In RM sector	0.09 (0.0001)	0 (0)
In CA sector	0 (0)	0.27 (0.0001)

Notes: The model allows for 4 types of individuals, with different combinations of RM and CA abilities, see Section 3 for details. The prices displayed in Panel B are defined in equation (2). Asymptotic standard errors in parentheses.

Table 8: Decomposing the Sources of Wage Growth

Potential experience since end of training	0-5	5-10	0-5	
	All	All	Untrained	Trained
Panel A: Between - Within wage growth decomposition:				
(1) Total wage growth	0.34	0.1	0.43	0.33
(2) Within wage growth	0.20	0.06	0.23	0.22
(3) Between wage growth	0.13	0.04	0.2	0.147
Panel B: Decomposition of total wage growth:				
(4) Search capital	30%	46%	36%	30 %
(5) RM experience	43%	11%	61%	42 %
(6) CA experience	25%	38%	3%	27 %

Notes: All results are obtained from simulated data from the estimated model. We refer the reader to Table A5 in the appendix for a comparison between simulated and observed data. Panel A displays wage changes as in Topel and Ward (1992), Table VII. Panel C. Panel B of the table reports the contribution of different channels to wage growth. Row (4) reports the contribution of the match-specific productivity (κ_{it}, see equation (4)) for between job mobility. Rows (5) and (6) report respectively the contribution of RM and CA experiences (α^{RM} and α^{CA} in equation (4)). Rows (4) to (6) do not add up to 100 percent because wage growth also depends on firm tenure, on the within evolution of the match productivity, sector mobility, and business cycles. Potential experience is counted from entry into the labor market for both trained and untrained workers.

Table 9: Effect of Increased Job-to-Job Mobility

Potential experience (years)	0-5	0-10	0-20
(1) Percent. increase in number of job-to-job transitions	37%	31%	24 %
(2) Increase in number of jobs held	0.34	0.48	0.58
Wage growth			
(3) Percentage increase in wages	1.6%	1.7%	1.0%
(4) Average between wage growth, per job (log point)	0.15	0.14	0.14
(5) Marginal wage growth, per job (log point)	0.02	0.02	0.01

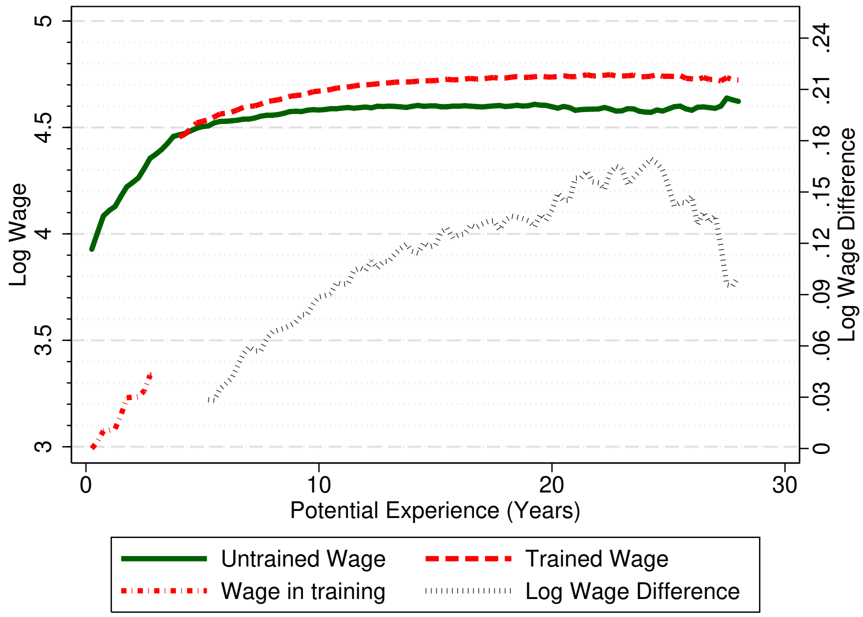
Notes: The table compares simulations of the baseline model with a counterfactual where the rate of arrival of offers while employed (λ^W, see equation (B1)) has been increased by 0.1, from the baseline value (intercept) displayed in Table 6, Panel B. Row (4) divides the between-job wage growth by the number of job-to-job transitions in the baseline. Row (5) displays the difference in the log wage growth between the counterfactual and the baseline, scaled by the increase in the number of job-to-job transitions between the counterfactual and the baseline. Potential experience is counted from entry into the labor market for both trained and untrained workers.

Table 10: Internal Rate of Return to Apprenticeship Training

	Controlling for Selection	
	No	Yes
	(1)	(2)
Individual:		
Only wages	14.3%	10.5%
Wages and unemployment benefits	12.5%	8.8%
Wages and utility of leisure	11.4%	7.7%
Wages and non-pecuniary work attribute	12.9%	9.1%
All	11.8%	8.2%
Society:		
Only wages	14.3%	10.5%
Wages and training costs	11.4%	8.2%
Wages and cost of unemployment benefits	16.4%	12.4%
Wages, training costs, and cost of UB	12.9%	9.6%

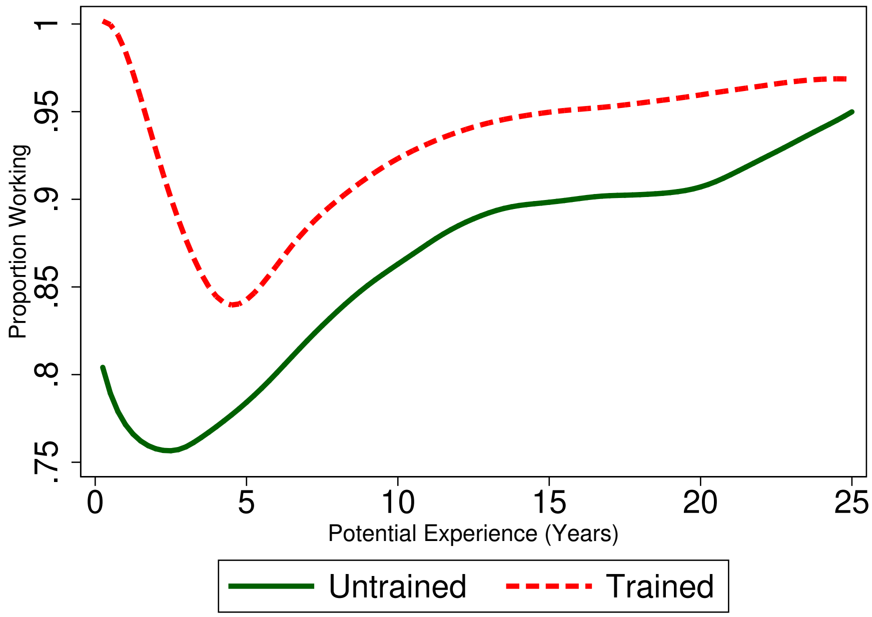
Notes: The table displays the internal rate of return to apprenticeship training, calculated over a period of 40 years following labor market entry. UB: unemployment benefits. Column (1) pools all wage and work spells for all individuals. Column (2) is a weighted average of the internal rate of return for each unobserved heterogeneity type, where the weights are the proportion of types in the sample. Training costs for firms and schools are evaluated at 7,400 euros per year and per individual.

Figure 1: Log Wage by Training Status



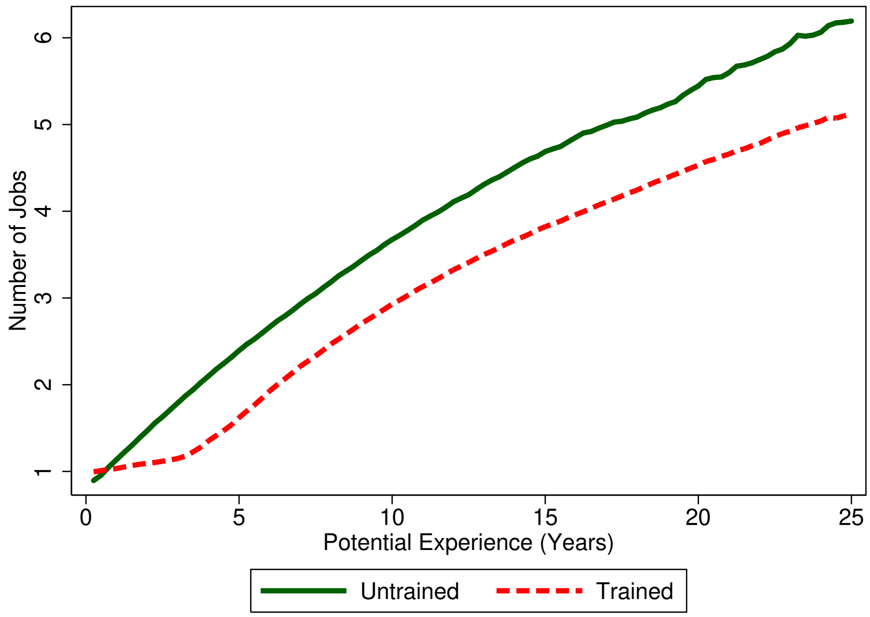
Notes: IAB Social Security data, 1975-2004. See Sections 2.3 and A for details. Potential experience is counted from entry into the labor market for untrained workers and from the start of training for trained workers.

Figure 2: Proportion Working by Training Status



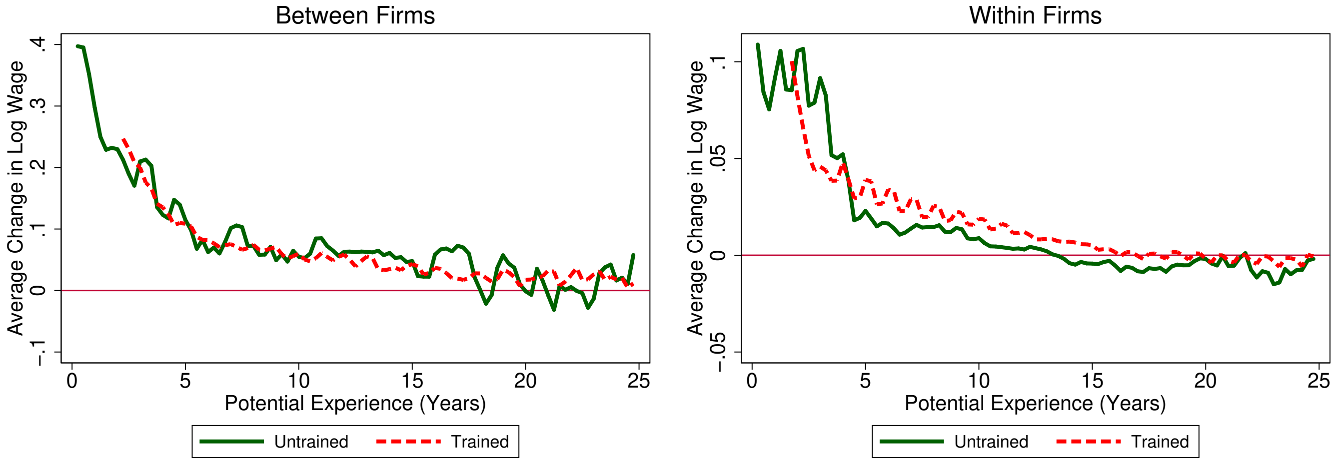
Notes: IAB Social Security data, 1975-2004. See Sections 2.3 and A for details. Potential experience is counted from entry into the labor market for untrained workers and from the start of training for trained workers.

Figure 3: Mobility: Number of Jobs, by Training Status



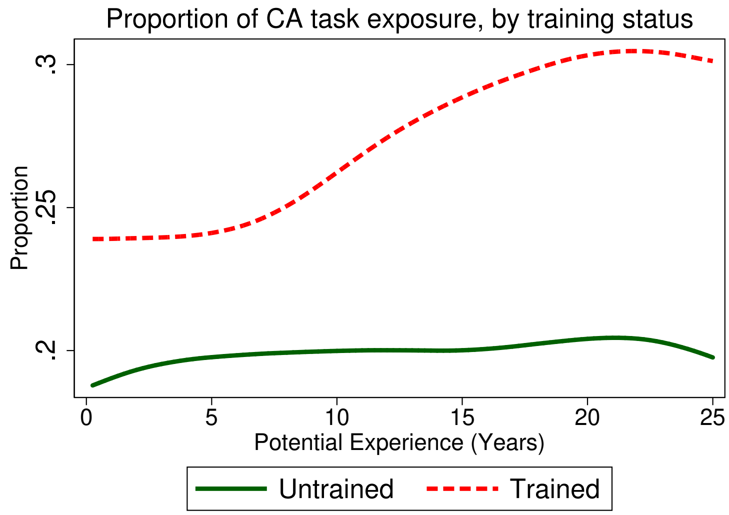
Notes: IAB Social Security data, 1975-2004. See Sections 2.3 and A for details. Potential experience is counted from entry into the labor market for untrained workers and from the start of training for trained workers. For the latter, mobility is low in the first few years as they are in training.

Figure 4: Annual Change in Log Wage



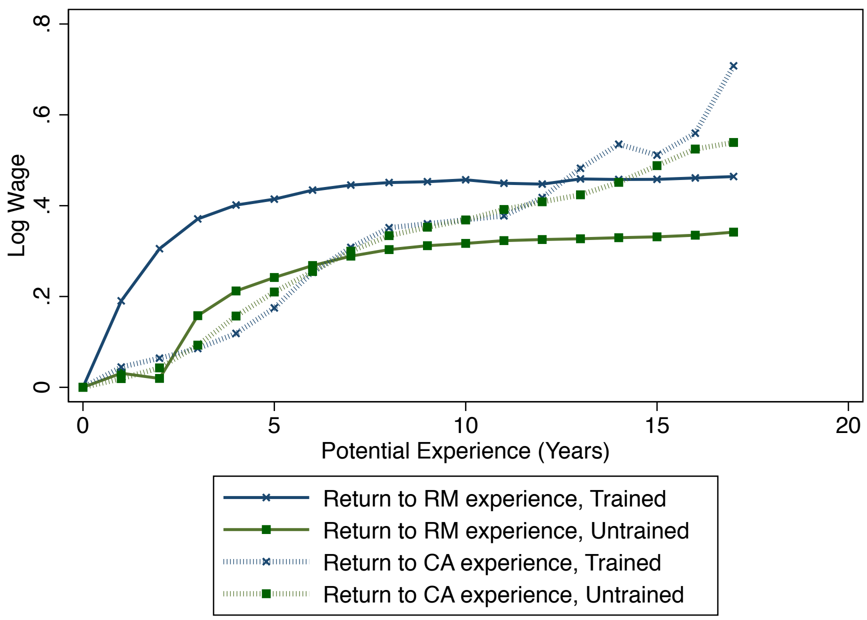
Notes: IAB Social Security data, 1975-2004. See Sections 2.3 and A for details. Potential experience is counted from entry into the labor market for untrained workers and from the start of training for trained workers.

Figure 5: Proportion of Workers in CA Sector and Exposure to CA Tasks



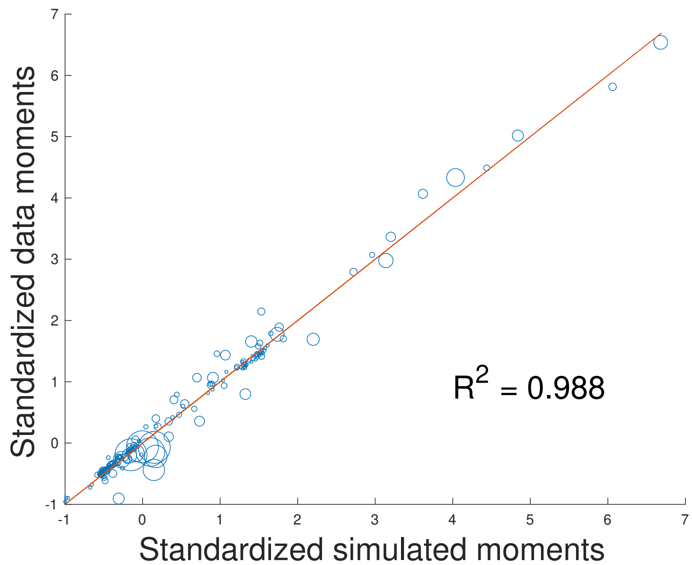
Notes: IAB Social Security data, 1975-2004. See Sections 2.3 and A for details. Potential experience is counted from entry into the labor market for untrained workers and from the start of training for trained workers.

Figure 6: Return to RM and CA Experience, by Training Status



Notes: Estimates obtained from two separate OLS regressions for trained and untrained workers, relating log wages to accumulated CA and RM experience. The regressions also control for firm tenure, a linear trend, a business cycle indicator, and region fixed effects. Experience is counted from entry into the labor market for untrained workers and from the start of training for trained workers.

Figure 7: Overall Fit of the Model



Notes: The figure shows a plot of the data moments against the simulated moments. The moments are rescaled by subtracting the mean and dividing by the standard deviation of the sample of moments. The radius of the circles are determined by the precision of the observed data moments (multiplied by 400 to make them visible). The solid line is the 45 degree line. The reported R^2 is the fit of a regression of observed data moments on simulated moments, weighted by the precision of the data moments.

Figure 8: Fit of the Model: Life-Cycle Profiles

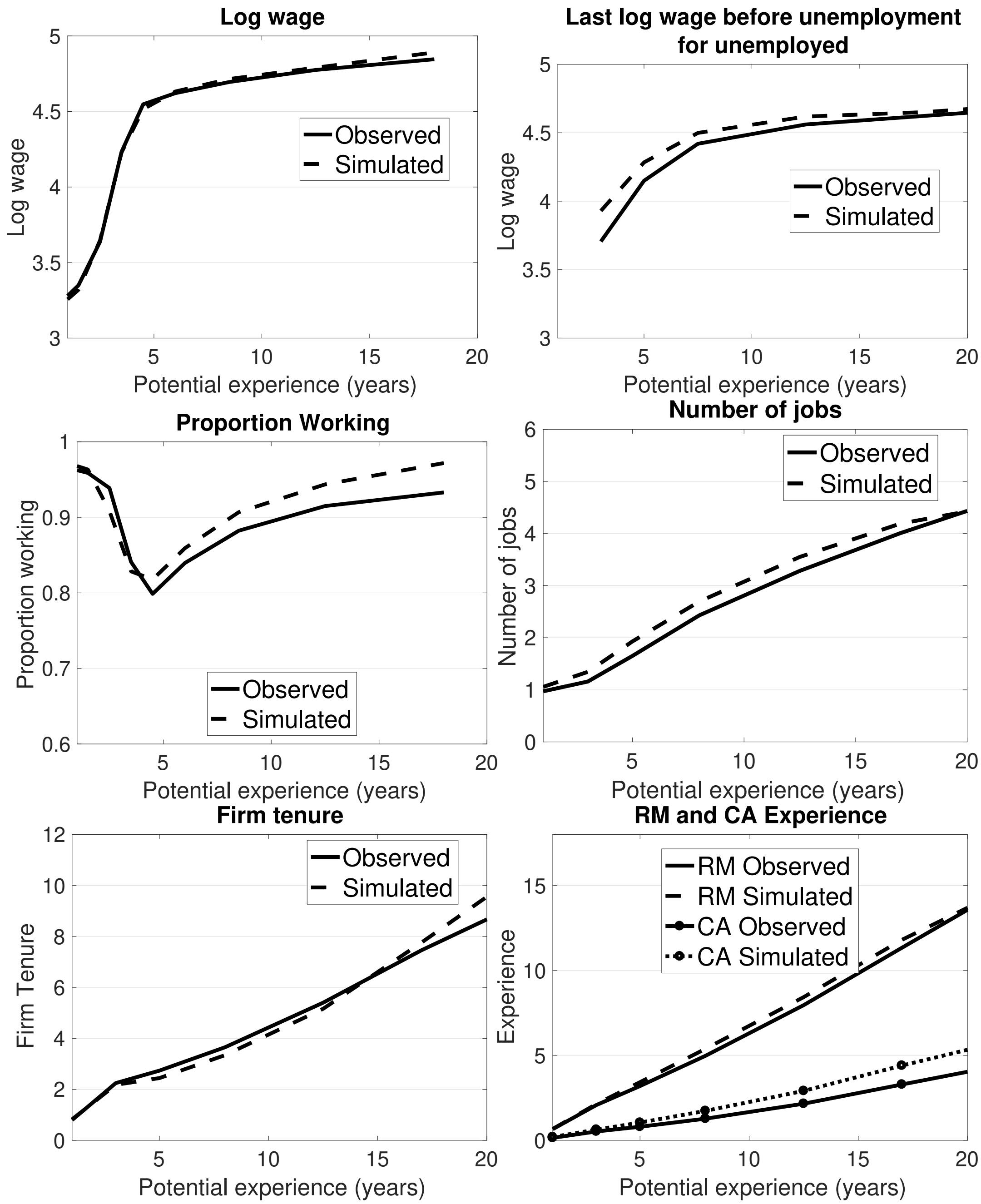
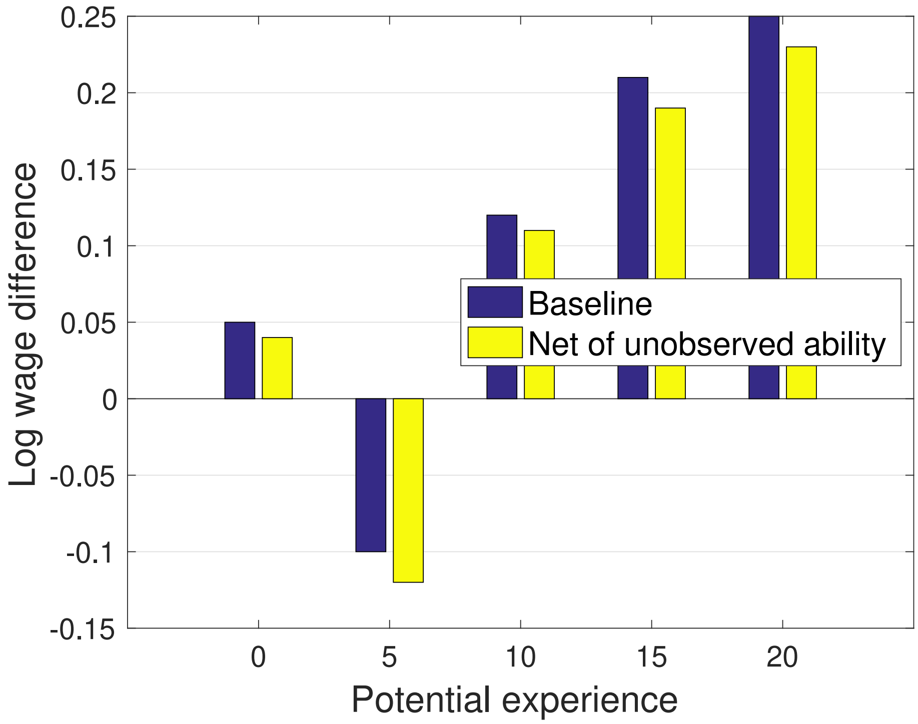


Figure 9: Log Wage Differential between CA and RM Sectors



Notes: Baseline is the simulated difference between log wages in the CA and RM occupational sector. The net of unobserved ability results are obtained by subtracting unobserved ability from the baseline results.

Online Appendix

A Data: Sample Used for Estimation

A.1 Main Sample

From the overall data, we select all male individuals who were born between 1960 and 1972. Thus, we ensure that no individual is older than 15 (the minimum age at which post-secondary labor market entry is possible) in 1975, which is the first year of our data. We consider observations over the entire period between 1975 and 2004. We exclude all individuals who work in East Germany. We also drop individuals who work in the agricultural industry, and individuals who work in family businesses. We restrict our sample to those who are not older than 23 when they enter the labor market for the first time, who enter the labor market with only a lower or intermediate secondary school education, and who either enroll into apprenticeship training directly, or who enter the labor market without further training.[33] [34] We exclude individuals with multiple apprenticeships (which is about 6% of the sample), and workers who are still in training at the end of the observation window, or who have no valid wage spells after apprenticeship training. We also exclude individuals who had a work spell before starting apprenticeship training, and we drop individuals with unreasonably long apprenticeship training periods (we set an upper limit of 1600 days). We restrict our analysis to individuals with German citizenship, as individuals with non-German citizenship are more likely to have acquired (part of) their education abroad.

The wage information in the data is the average daily wage for the length of the working spell. A spell is at most 365 days long if the individual does not change firm, as firms have to report yearly on their employees. If individuals change firm during the calendar year, or exit into unemployment, we observe the average daily wage for each spell of employment. Thus, every wage we observe belongs to one particular worker-firm

[33] In Germany, children enter primary school at the age of about 6. Primary school takes 4 years. After primary school, and at the age of 10, individuals decide between three secondary school tracks: lower secondary school (which takes another 5-6 years), intermediate secondary school (which takes another 6 years), and upper secondary school (which takes another 9 years). For our analysis, we concentrate on individuals who chose lower or intermediate secondary school. These two options do not allow for direct access to university, and after graduation these individuals typically enroll into apprenticeship training, or enter the labor market directly.

[34] As a comparison group of individuals who chose upper track secondary school, which we use to implement our selection correction, we consider all those individuals who enter the labor market either with an upper secondary degree (with or without further training), and before the age of 23, or with college or university education, and before the age of 32.

spell. We compute real wages in 1995 prices and convert them in euros.

We define as "trained workers" all those individuals who entered the labor market with a lower or intermediate secondary school degree, who can be observed after entry on an apprenticeship training scheme for at least 24 months, and who transit to a "trained"status afterwards. We define as "untrained workers" all those individuals who enter the labor market without further training, or who have been observed on an apprenticeship training scheme for less than 7 months, without obtaining a qualification (i.e. dropouts). This group may include individuals who enrolled in one-year preparatory vocational courses (that do not lead to vocational degrees) before entering the labor market. Thus, among our untrained workers some may be individuals who did receive some post-secondary preparatory training.

Another mode of training, as discussed in Parey (2009), is attendance of 2-3 year vocational schools, which provide vocational training with unpaid work experience in specialized schools for a limited number of occupations.[35] These occupations are mainly in female-dominated fields such as caring and health-related occupations. In our sample, these constitute about 6% of individuals.[36] In line with Parey (2009), we find that the wage paths of graduates from these schools are very similar to those of individuals who underwent firm-based training. We thus include them in the group of trained workers, assuming that the choice to undergo training at a full time school is equivalent to choosing apprenticeship training in a firm.

A.2 Area Level Data

Our GVA data until 1996 originates from the archives of Statistik NRW on the national accounts of the German states, and cover the years 1961, 1968, 1970, 1972, 1974, 1976, 1980, 1982, 1984, 1986, 1990, 1992, 1994, 1996. The remaining years until 2000 were extrapolated on the per capita level. Data for population is available for the period 1975-2011, which we extrapolated back to 1970. We then calculated GVA per capita, interpolated the missing years between 1970 and 1996, and extrapolated the remaining years after 1996 on a per capita level.

[35] According to the German Central Labor Office (Bundesagentur fuer Arbeit), firm based apprenticeship schemes train for 541 occupations, while full-time colleges train for only 133 occupations.

[36] The size of this group is smaller than in Parey (2009). One reason for this is that the last year in which individuals in our sample may begin training is 1996, and during these earlier years these school-based vocational schemes were less frequent than in later years.

B Computational Details

This section provides further details on the model presented in Section 3 of the main text.

B.1 GDP and Markov Transition Matrix

To compute business cycles, we use the per capita West German GDP expressed in constant prices at a quarterly frequency, obtained from the OECD for the period 1975-2009. We detrend the series and identify transitions between above trend (representing good times) and below trend (bad times). Table A2 presents the transition matrix for the corresponding first order Markov process, estimated over our sample period.

B.2 Dynamic Choice Model

We denote by $\Omega_{it+1|k}$, $k \in \{W, U, \tilde{W}, \tilde{O}\}$, the state space one period ahead, depending on the labor market choices made at the end of period t. These are respectively: staying employed in the same firm, moving to unemployment, moving to another firm or changing occupational sector. When the individual is working in period t, all skills are updated as explained in Section 3.1. When staying with the same firm, the match-specific productivity contribution is updated as in equation (3). When moving to unemployment the match-specific productivity is not updated, and serves as a basis to determine the unemployment benefits, as they are calculated as a fraction of the previous wage. When the workers opt to join a new firm, they receive the match-specific productivity contribution and non-pecuniary attribute corresponding to the job offer. Additionally, their experience stocks are updated and tenure is reset to zero.

We distinguish two cases for the Bellman equation, the individual being in work, and unemployed. In the first case, the individual derives an instantaneous utility from the log wage and the non-pecuniary job attribute, and a future payoff,

$$
\begin{aligned}
V_t^W(\Omega_{it}) \quad = \quad & \log(w_{it}) + \mu_{it} && \text{(B1)} \\
& +\beta\delta E_t V_{t+1}^U(\Omega_{it+1|U}) \\
& +\beta(1-\delta)(1-\lambda^W)(1-\lambda^O)E_t \max\left(V_{t+1}^W(\Omega_{it+1|W}), V_{t+1}^U(\Omega_{it+1|U})\right) \\
& +\beta(1-\delta)\lambda^W(1-\lambda^O)E_t \max\left(V_{t+1}^W(\Omega_{it+1|W}), V_{t+1}^W(\Omega_{it+1|\tilde{W}}), V_{t+1}^U(\Omega_{it+1|U})\right) \\
& +\beta(1-\delta)(1-\lambda^W)\lambda^O E_t \max\left(V_{t+1}^W(\Omega_{it+1|W}), V_{t+1}^W(\Omega_{it+1|\tilde{O}}) - c_O, V_{t+1}^U(\Omega_{it+1|U})\right) \\
& +\beta(1-\delta)\lambda^W\lambda^O E_t \max\left(V_{t+1}^W(\Omega_{it+1|W}), V_{t+1}^W(\Omega_{it+1|\tilde{O}}) - c_O, V_{t+1}^U(\Omega_{it+1|U})\right).
\end{aligned}
$$

The first line defines the flow utility of working, composed of the utility of wages and of the non pecuniary attribute of the job. The second line corresponds to the future (discounted expected) payoff when the worker is laid off and starts the next period in unemployment with a value V^U, defined below. This happens with a probability $\delta(\Omega_{it})$. For ease of exposition, we have suppressed the dependence of the probability terms on observable variables in (B1). The next lines of equation (B1) correspond to the case when the worker is not laid off, which happens with a probability $(1-\delta)$. In line 3, the worker does not receive an outside job offer, nor an offer from a different sector. The worker then decides whether he stays with the same firm or goes into non employment. Line 4 corresponds to a worker who receives an alternative job offer within the same sector. He must then chose between staying at the same firm, moving to the new firm, or moving to non employment. In line 5 the worker obtains an offer within the same firm to change occupation. The worker chooses between retaining his existing job and occupation, changing occupation but retaining his other job characteristics, or moving into unemployment. The last line corresponds to the case of a worker who gets an offer of another occupation in a different firm.

The Bellman equation for individuals who start the period in unemployment is:

$$
\begin{aligned}
V_t^U(\Omega_{it}) \quad = \quad & \log(b_{it}) + \gamma && \text{(B2)} \\
& +\beta(1-\lambda^U)E_t V_{t+1}^U(\Omega_{it+1|U}) \\
& +\beta\lambda^U(1-\lambda^O)E_t \max\left(V_{t+1}^W(\Omega_{it+1|\tilde{W}}), V_{t+1}^U(\Omega_{it+1|U})\right) \\
& +\beta\lambda^U\lambda^O E_t \max\left(V_{t+1}^W(\Omega_{it+1|\tilde{O}}) - c_O, V_{t+1}^U(\Omega_{it+1|U})\right).
\end{aligned}
$$

The first line includes the utility from unemployment benefits, denoted b_{it}. Unemployment benefits are calculated as a fraction of the last wage when employed, as in the

German unemployment insurance (UI) system that was in place over the period we consider here. When UI is exhausted (after about 18 months), an unemployed worker moves on to the means-tested unemployment assistance. Given the length of time for eligibility and the generosity of social assistance for lower wage individuals, we have made the simplifying assumption that the replacement rate is always 45 percent, which corresponds to the average for our population. In addition, there is a utility of leisure denoted γ, common across all individuals. In line 2, the individual does not receive any job offers (with a probability $1 - \lambda^U$) and continues in unemployment in the next period. In line 3, the individual receives a job offer in the same occupation as he had previously worked. He then chooses between accepting this offer or remaining unemployed. Line 4 corresponds to the case where the individual receives a job offer in a new occupation.

Secondary school track choice: As described in Section 2, it is the two lower tracks of secondary school that prepare for apprenticeship training. The school track is chosen at age 10 based on merit, family decisions, and local determinants, which we denote as $\tilde{Z}_i$. We denote $\{\tilde{\eta}_i^{LT}, \tilde{\eta}_i^{HT}\}$ two idiosyncratic shocks, extreme value distributed, corresponding to the lower two tracks (LT) and upper track (HT) of secondary school and $\tilde{\zeta}^{LT}$ a vector of parameters associated with the lower two track choice. Finally, we denote by V^{HT} the discounted payoff of enrolling in the upper track, and we treat it as a parameter to be estimated. The decision at age 10 is taken to maximize the future payoff:

$$\max\{\tilde{Z}_i\tilde{\zeta}^{LT} + \tilde{\eta}_i^{LT} + \beta EV_{t=16}^{Init}(Z_i, s_i^{RM}, s_i^{CA}, \eta_{i,T,o}), \tilde{\eta}_i^{HT} + V^{HT}\}, \tag{B3}$$

where $V_{t=16}^{Init}$ represents the lifetime continuation value defined in equation (8). Given the distributional assumption on the shocks, the probability of opting for the lower track is:

$$Prob(LT_i = 1|\tilde{Z}_i, Z_i) = \frac{\exp(\tilde{Z}_i\tilde{\zeta}^{LT} + \beta EV_{t=16}^{Init}(Z_i, s_i^{RM}, s_i^{CA}, \eta_{i,T,o}))}{\exp(\tilde{Z}_i\tilde{\zeta}^{LT} + \beta EV_{t=16}^{Init}(Z_i, s_i^{RM}, s_i^{CA}, \eta_{i,T,o})) + \exp(V^{HT})}. \tag{B4}$$

We use this probability to re-weight our simulated sample to take into account potential selection into different tracks. The estimated parameters related to equation (B4) are displayed in Table A3 together with the estimates relating to equation (8).

B.3 Computing the Value Functions

We solve the dynamic model assuming a finite horizon, where individuals work for 40 years and then retire and live for an additional 20 years. We assume a net pension replacement rate of 50 percent of the last wage earned. We allow the value function to depend on age as well as the other state variables.

We analytically integrate out as many state variables and shocks as possible. We approximate the value functions by evaluating them at a number of discrete points in the state space and interpolating multi-linearly in between as in Keane and Wolpin (1994). For experience and tenure the points, denoted $\{N_k^j\}, j \in \{CA, RM, Ten\}$ in equation (5), are taken at 0, 2, 4, 6, 10, and 30 years of experience and 0, 2, 4, 6, and 30 years of tenure. These grids were chosen to optimize the precision, with grid points coinciding with inflection points we could detect from OLS regressions of wage equations. This number of grid points appears to be sufficient as the value functions are smooth. We also discretize the non-pecuniary job attribute on a 12 point grid using linear interpolation when we compute the decision of the workers to change firm. The other state variable is the firm-worker match-specific productivity contribution, which evolves as a random walk while the worker remains in the same job. To discretize this variable, we use 10 points on a grid which depends on skills and tenure to take into account the non-stationary nature of the process. More specifically, given the assumptions made, the match-specific productivity is a normally distributed variable with mean zero and variance $Ten_{it}\sigma_u(T_i)^2 + \sigma_\kappa(T_i)^2$ for an individual with a tenure of Ten_{it} years. We use a quadrature-based method as in Tauchen and Hussey (1991) to generate a grid and transition matrices for this process. To calculate the value functions outside of the grids, we assume that the return to RM and CA skills and tenure is flat after 30 units (years) of these skills. This is motivated by analysis of the data using OLS regressions, which show that the returns to skills are very concave and flatten out with age. For the match-specific productivity, we extrapolate linearly to calculate values outside of the grid. In practice, the values of working and unemployment are smooth with regards to the RM and CA skills, as well as tenure. The model is solved using parallel processing to decrease runtimes.

C Identification

C.1 Relation between Wage Fixed Effects and Unobserved Ability Parameters

In this section, we show how statistics based on wage fixed effects are related to the parameters of the model, and in particular to those pertaining to unobserved heterogeneity. Suppose first that we observe each worker in both sectors during their career. From equation (2), and the definition of the four types with skills $(s_i^{RM}, s_i^{CA}) \in \{(0,0),(1,0),(0,1),(1,1)\}$ and proportions $\{\pi_1, \pi_2, \pi_3, \pi_4\}$, we can write the expected value of the fixed effect in each sector as:

$$\begin{aligned} E(\alpha_i(RM)) &= p^{RM(RM)}\pi_2 + p^{RM(CA)}\pi_3 + (p^{RM(RM)} + p^{RM(CA)})\pi_4 && \text{(C5)} \\ E(\alpha_i(CA)) &= p^{CA(RM)}\pi_2 + p^{CA(CA)}\pi_3 + (p^{CA(RM)} + p^{CA(CA)})\pi_4 && \text{(C6)} \end{aligned}$$

The cross-sectional variance of the wage fixed effect in the RM sector is:

$$\begin{aligned} V(\alpha_i(RM)) &= E(\alpha(RM)_i^2) - E(\alpha(RM)_i)^2 && \text{(C7)} \\ &= (p^{RM(RM)})^2(\pi_2+\pi_4)(\pi_1+\pi_3) + (p^{RM(CA)})^2(\pi_1+\pi_2)(\pi_3+\pi_4) \\ &\quad +2p^{RM(RM)}p^{RM(CA)}(\pi_1\pi_4 - \pi_2\pi_3) \end{aligned}$$

Similarly, the variance in the CA sector can be derived as:

$$\begin{aligned} V(\alpha_i(CA)) &= (p^{CA(RM)})^2(\pi_2+\pi_4)(\pi_1+\pi_3) + (p^{CA(CA)})^2(\pi_1+\pi_2)(\pi_3+\pi_4) && \text{(C8)} \\ &\quad +2p^{CA(RM)}p^{CA(CA)}(\pi_1\pi_4 - \pi_2\pi_3) \end{aligned}$$

The covariance between the RM and the CA sector fixed effects is:

$$\begin{aligned} Cov(\alpha_i(RM), \alpha_i(CA)) &= E(\alpha_i(RM)\alpha_i(CA)) - E(\alpha_i(RM))E(\alpha_i(CA)) && \text{(C9)} \\ &= p^{RM(RM)}p^{CA(RM)}(\pi_1+\pi_3)(\pi_2+\pi_4) \\ &\quad +p^{CA(CA)}p^{RM(CA)}(\pi_1+\pi_2)(\pi_3+\pi_4) \\ &\quad +(p^{RM(RM)}p^{CA(CA)} + p^{CA(RM)}p^{RM(CA)})(\pi_1\pi_4 - \pi_2\pi_3) \end{aligned}$$

The left hand side of equations (C5) to (C9) can be computed with panel data regressions for both trained and untrained workers. This gives us 10 data moments, which are a function of 4 skill prices, as well as 3 proportions of types in the sample (as the π_l sum to 1). Hence, the model is over identified with regards to the unobserved heterogeneity.

In reality, some individuals only work in one sector during their entire career. For those individuals, we can only compute one wage fixed effect. If the choice of sector and the decision to move across sectors is different across unobserved heterogeneity "types", then the formulas above have to be adjusted for selection. In our estimation protocol, these adjustments do not need to be explicitly accounted for, since we rely on indirect inference. The indirect inference procedure consists of matching data and simulated moments, even if they are biased estimators of the left hand side quantities of equations (C5) to (C9). We can still obtain consistent estimates, provided that the bias is the same in the observed and simulated data.

C.2 Parameters Related to Labor Market Transitions

To identify the parameters related to labor market transitions, we rely on a number of linear probability models and OLS regressions relating labor market status, the number of jobs held, the amount of RM and CA skills accumulated, and firm tenure to potential experience, training choice, and a business cycle indicator. Targeting these moments ensure the model reproduces the main lifecycle features in terms of labor market status. On their own, they do not allow separate identification of some key parameters such as the job destruction rate and the arrival rate of offers. Different combinations of destruction and offer rates can generate the same proportion of employed individuals, distributions of tenure, etc. To identify these parameters, we include moments based on the probability of making specific transitions, such as mobility out of work, into work, between jobs, or between sectors. We use linear probability models to relate the likelihood of those transitions to potential experience, training status, and the amount of RM and CA skills. In the case of sector choice we also include a term indicating whether the transition is associated with a change of firm or not.

To identify the role of unobserved heterogeneity in productivity in determining labor market status, we also relate a variable indicating whether the individual is working to potential experience, as well as a wage fixed effect obtained from a prior regression of wages on RM and CA skills, firm tenure, occupation, and the number of jobs (as a proxy for search capital).

Finally, to identify costs of moving between sectors, and how these change over the lifecycle, we use a regression relating the time to the first move from the RM to the CA sector to a wage fixed effect, calculated as explained above. A higher cost of moving discourages individuals for whom the gain of having a job in the CA sector is not large enough. The timing of that move and how it relates to productivity is informative about on how selection operates at different periods over the lifecycle.

C.3 Choice of Training and Initial Sector

To identify the parameters that relate to the initial choices of training and occupational sector, we use linear probability models relating those outcomes to the instruments discussed in Section 2.2.3, but also to the worker's subsequent wage fixed effect. This serves as a proxy for unobserved ability, which contributes to the identification of the sorting into education, together with the first stage regressions.

C.4 Mapping between Parameters and Data Moments

Local identification would fail if the Jacobian of the moment vector is not invertible, which is not the case for our estimation. To gain more understanding on how moments and parameters are related (locally), we investigate how each moment changes with a deviation in each of our parameters.[37]

Figure A1 graphically displays the results. The darker the shade, the larger is the response of the predicted moment to a change in a particular parameter. We first note that there are no rows that are white throughout, which means that there always exists at least one moment that relates to a given parameter, and most parameters are identified by more than one moment. Conversely, each moment relates to at least one parameter. The alternation of darker and lighter shades when moving horizontally along rows is due to the fact that parameters are related to different training groups, and have varying influence on moments pertaining to specific groups. Take for instance the intercept of the wage equation for untrained workers, which is the first row of the figure. A shift in that parameter results in a shift in the entire wage profile of the untrained (labelled moment (1) in the figure), but also changes the incentives to work and therefore changes labor market transitions, and the profiles of experiences and tenure (labelled moments (2, 3, 4 and 12)). Indirectly, it affects the incentives to obtain training, and hence changes the composition of workers in both the trained and untrained groups in terms of unobservables. Figure A1 also serves to demonstrate how intricate the model is. It allows not only for dynamic effects of observable characteristics but also complex sorting patterns across training groups, sectors, and labor market status. All of these effects contribute to the estimation of a simple parameter such as a wage intercept.

C.5 Sensitivity of parameters for first stage variables

We now turn to a sensitivity analysis of the parameters with respect to changes in some of the moments. Andrews et al. (2017) provide a method to assess this sensitivity, at least in a local sense, using the gradient of the objective function. We use this method

[37] Denote by θ the estimated vector of parameters, and define $\breve{\theta}_i = (\theta_1, \ldots, \theta_{i-1}, \theta_i * 1.01, \theta_{i+1}, \ldots, \theta_k)$ the same vector of parameters, where one parameter has been changed by 1 percent. Denote by $g^S(\theta)$ the vector of predicted moments. We calculate the ratios $(g^S(\breve{\theta}_i) - g^S(\theta))/g^S(\theta)$ which measures the departure of the predicted moments due to a deviation of the parameters. This produces a $k \times m$ matrix, with rows pertaining to a specific parameter and the columns to moments. A given parameter would not be identified locally if a change in its value resulted in no change in any predicted moments (and hence a ratio of zero).

to investigate the sensitivity of our parameter estimates with respect to the first stage regressors that identify the choice of secondary school track, vocational training, and initial sector of employment (see Table 4). We display the results in Figure A2 which shows the effect of a one percent change in the variables on the parameters. A sensitivity of 10^{-4} indicates that such a change would affect the fourth digit of that estimate.

As shown in the figure, the sensitivity to changes in most variables is much lower than this level. One variable is an exception, the number of university slots in the vicinity of the individual. The difference between this variable and the other three comes from the fact that it determines the choice of whether to enroll in the upper secondary school track, which in turn allows for university enrolment. Whereas we model explicitly the career paths of both trained and untrained workers who have received lower or intermediate secondary education, we take a reduced form approach for the choice between the lower or intermediate tracks and the upper track, directly estimating the present value payoff for the upper track at age 10 when the choice is made (see equation (B4)). Hence, this part of the model relies solely on the variation of the instrument to determine the sorting into the upper track. In contrast, the sorting between trained and untrained careers is also informed by the repeated information on wages and labor supply. Another way to gauge the sensitivity of the parameters to those "first stage" moments is to compare them with the sensitivity of the parameters to other moments. The 95th percentile of the sensitivity of the parameters to other moments is equal to 0.035, approximately two orders of magnitude higher.

D Additional Evidence on the Fit of the Model

In this section, we present additional evidence on the fit of the model, complementing the results displayed in Figure 8, now conditioning on training status. Tables A4 and A6 provide the results. Each panel is obtained from OLS regressions, either on the observational data or on data simulated from the estimated model, using 24,000 fictitious individuals. The standard errors associated with the coefficients from the observational data are all clustered at the individual level. Even disaggregated by training status, the model is able to reproduce the lifecycle patterns of wages, labor market status and mobility and the accumulation of skills.

Table A7 provides additional information by presenting regressions of wage changes on a number of covariates. These include CA and RM experience, firm tenure, and dummy variables capturing whether the individual changed firm or completed their apprenticeship training over the period in question.

References

TAUCHEN, G. AND R. HUSSEY (1991). "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models." *Econometrica*, 59, 371–396.

Table A1: Descriptive statistics

	At labor market entry	10 years after labor market entry
Average age	16.8 (1.5)	26.6 (1.4)
Birth year	1967.3 (4.9)	1967.0 (4.7)
Proportion trained	0.9 (0.3)	0.9 (0.3)
Proportion working	1	0.91 (0.28)
Average log wage	3.15 (.41)	4.75 (.24)
RM experience	0	6.32 (1.86)
CA experience	0	1.96 (1.58)
Share in CA sector	0.15 (0.35)	0.2 (0.4)
Share in CA sector, trained	0.15 (0.36)	0.2 (0.4)
Share in CA sector, untrained	0.06 (0.24)	0.1 (0.3)
Number of observations	49,144	43,155
	Initial conditions (at age 10 or 15)	
Log number of university students in 50km radius	7.8 (1.1)	
Local number of training positions, per capita	0.76 (0.23)	
Local log gross value added	0 (0.02)	
Share of jobs in services	0.11 (.04)	

Notes: Results in the first panel are based on data from the IAB, and in the second panel on data from Destatis. Area level statistics are calculated at the district ("kreis") level, except for the number of training positions, which are at commuting zone level. University places are calculated at age 10, training positions and value added at age 15. Standard deviations in parentheses.

Table A2: Quarterly transition matrix for business cycle indicator

	$G_{t+1} = 0$	$G_{t+1} = 1$
$G_t = 0$	0.93 (0.04)	0.07 (0.04)
$G_t = 1$	0.07 (0.04)	0.93 (0.04)

Notes: The business cycle indicator G_t is coded as 1 when GDP is above its trend (see Section 3.3). Data source: OECD, quarterly GDP per capita, constant prices, constant PPP, period 1975-2009. Asymptotic standard errors in parentheses.

Table A3: Estimated Parameters: Selection into School Tracks, Training and Sectors

Parameter	Estimate	Std error
Selection into school tracks at age 10:		
Value of university	251.03	(0.002)
Share of university places	17.532	(0.003)
Skill set (RM=1,CA=0)	6.9449	(0.002)
Skill set (RM=0,CA=1)	5.9268	(0.001)
Skill set (RM=1,CA=1)	35.309	(0.005)
Std dev of shock	20.468	(0.003)
Selection into training and sectors at age 16:		
Log GVA	-617.6	(0.001)
Log total population	126.99	(0.015)
Share services	253.92	(0.001)
CA, untrained	-190.84	(0.001)
RM, trained	162.41	(0.002)
CA, trained	44.038	(0.002)
Std dev of shock	54.561	(0.011)

Notes: The table reports the estimated parameters of equations (8) in the main text and (B4) in the Appendix. Asymptotic standard errors in parentheses.

Table A4: Goodness of Fit: Wage Level and Proportion Working

	Trained			Untrained		
	Observed	Std Error	Simulated	Observed	Std Error	Simulated
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Log wage level:						
Potential exp ∈ [0,1]	3.1	(0.0021)	3.1	4.1	(0.0023)	4.1
Potential exp ∈]1,2]	3.2	(0.0021)	3.2	4.2	(0.0023)	4.2
Potential exp ∈]2,3]	3.6	(0.0025)	3.6	4.4	(0.0021)	4.3
Potential exp ∈]3,4]	4.2	(0.0027)	4.2	4.5	(0.0019)	4.4
Potential exp ∈]4,5]	4.5	(0.0023)	4.5	4.6	(0.0018)	4.4
Potential exp ∈]5,7]	4.6	(0.0022)	4.6	4.6	(0.0018)	4.5
Potential exp ∈]7,10]	4.7	(0.0023)	4.7	4.7	(0.0020)	4.6
Potential exp ∈]10,15]	4.8	(0.0028)	4.8	4.7	(0.0025)	4.6
Potential exp ∈]15,30]	4.9	(0.0037)	4.9	4.8	(0.0034)	4.7
Business Cycle high	0.041	(0.00087)	0.054	0.036	(0.0036)	0.05
Panel B: Proportion working:						
Potential exp ∈ [0,1]	0.98	(0.0011)	0.99	0.85	(0.0011)	0.86
Potential exp ∈]1,2]	0.98	(0.0011)	0.98	0.78	(0.0013)	0.78
Potential exp ∈]2,3]	0.96	(0.0012)	0.92	0.78	(0.0014)	0.79
Potential exp ∈]3,4]	0.85	(0.0018)	0.83	0.79	(0.0014)	0.81
Potential exp ∈]4,5]	0.8	(0.0022)	0.82	0.79	(0.0015)	0.83
Potential exp ∈]5,7]	0.84	(0.0020)	0.86	0.82	(0.0015)	0.85
Potential exp ∈]7,10]	0.88	(0.0019)	0.91	0.86	(0.0015)	0.89
Potential exp ∈]10,15]	0.92	(0.002)	0.95	0.9	(0.0018)	0.92
Potential exp ∈]15,30]	0.93	(0.0024)	0.98	0.93	(0.0023)	0.94
Business cycle high	0.024	(0.00079)	0.029	0.038	(0.0037)	0.039

Notes: The observed coefficients (in columns 1 and 4) are derived from IAB data by OLS regressions, with standard errors clustered at individual level. The coefficients in columns 3 and 6 are obtained from OLS regressions based on data simulated from the model.

Table A5: Goodness of Fit: Topel and Ward Decomposition of Wage Growth

Potential experience since end of training	0-5	5-10	0-5	
	All	All	Untrained	Trained
Panel A: Between - Within wage growth decomposition (observed data):				
(1) Total wage growth	0.37	0.07	0.44	0.37
(2) Within wage growth	0.23	0.04	0.21	0.23
(3) Between wage growth	0.15	0.03	0.23	0.14
Panel B: Between - Within wage growth decomposition (simulated data):				
(4) Total wage growth	0.34	0.1	0.43	0.33
(5) Within wage growth	0.20	0.06	0.23	0.22
(6) Between wage growth	0.13	0.04	0.2	0.147

Notes: Panel A and B display wage changes as in Topel and Ward (1992), Table VII. Potential experience is counted from entry into the labor market for both trained and untrained workers.

Table A6: Goodness of Fit: Experience, Tenure, and Job Mobility

	Trained			Untrained		
	Observed	Std Error	Simulated	Observed	Std Error	Simulated
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: CA experience level (in years):						
Potential exp ∈ [0,2]	0.16	(0.016)	0.2	0.019	(0.032)	0.12
Potential exp ∈]2,4]	0.55	(0.016)	0.66	0.29	(0.033)	0.38
Potential exp ∈]4,6]	0.83	(0.017)	1.1	0.54	(0.035)	0.65
Potential exp ∈]6,10]	1.3	(0.019)	1.8	0.96	(0.039)	1.1
Potential exp ∈]10,15]	2.2	(0.023)	3	1.7	(0.049)	1.8
Potential exp ∈]15,20]	3.4	(0.032)	4.6	2.4	(0.071)	2.7
Potential exp ∈]20,40]	4.7	(0.051)	6.2	3.3	(0.12)	3.7
Panel B: RM experience level (in years):						
Potential exp ∈ [0,2]	0.65	(0.017)	0.67	0.7	(0.053)	0.64
Potential exp ∈]2,4]	2.1	(0.018)	2.2	2	(0.055)	2
Potential exp ∈]4,6]	3.2	(0.019)	3.4	3.2	(0.059)	3.4
Potential exp ∈]6,10]	5	(0.021)	5.4	5.2	(0.068)	5.6
Potential exp ∈]10,15]	7.9	(0.026)	8.4	8.4	(0.087)	9
Potential exp ∈]15,20]	11	(0.035)	12	12	(0.12)	13
Potential exp ∈]20,40]	15	(0.056)	15	16	(0.19)	16
Panel C: Firm tenure (in years):						
Potential exp ∈ [0,2]	0.8	(0.025)	0.85	0.76	(0.084)	0.58
Potential exp ∈]2,4]	2.3	(0.026)	2.3	2	(0.087)	1.4
Potential exp ∈]4,6]	2.7	(0.029)	2.5	2.9	(0.097)	2
Potential exp ∈]6,10]	3.6	(0.033)	3.4	4.2	(0.11)	3
Potential exp ∈]10,15]	5.3	(0.04)	5.2	6.2	(0.15)	5.1
Potential exp ∈]15,20]	7.4	(0.054)	7.7	8.3	(0.20)	8.1
Potential exp ∈]20,40]	9.4	(0.083)	11	10	(0.31)	11
Panel D: Number of jobs:						
Potential exp ∈ [0,2]	0.97	(0.013)	1.0	1.0	(0.054)	1.3
Potential exp ∈]2,4]	1.1	(0.013)	1.3	1.6	(0.057)	2.1
Potential exp ∈]4,6]	1.6	(0.014)	1.8	2.1	(0.062)	2.8
Potential exp ∈]6,10]	2.4	(0.016)	2.6	2.8	(0.072)	3.7
Potential exp ∈]10,15]	3.2	(0.021)	3.4	3.6	(0.093)	4.6
Potential exp ∈]15,20]	4	(0.028)	4.1	4.4	(0.13)	5.1
Potential exp ∈]20,40]	4.7	(0.042)	4.5	5.1	(0.19)	5.3

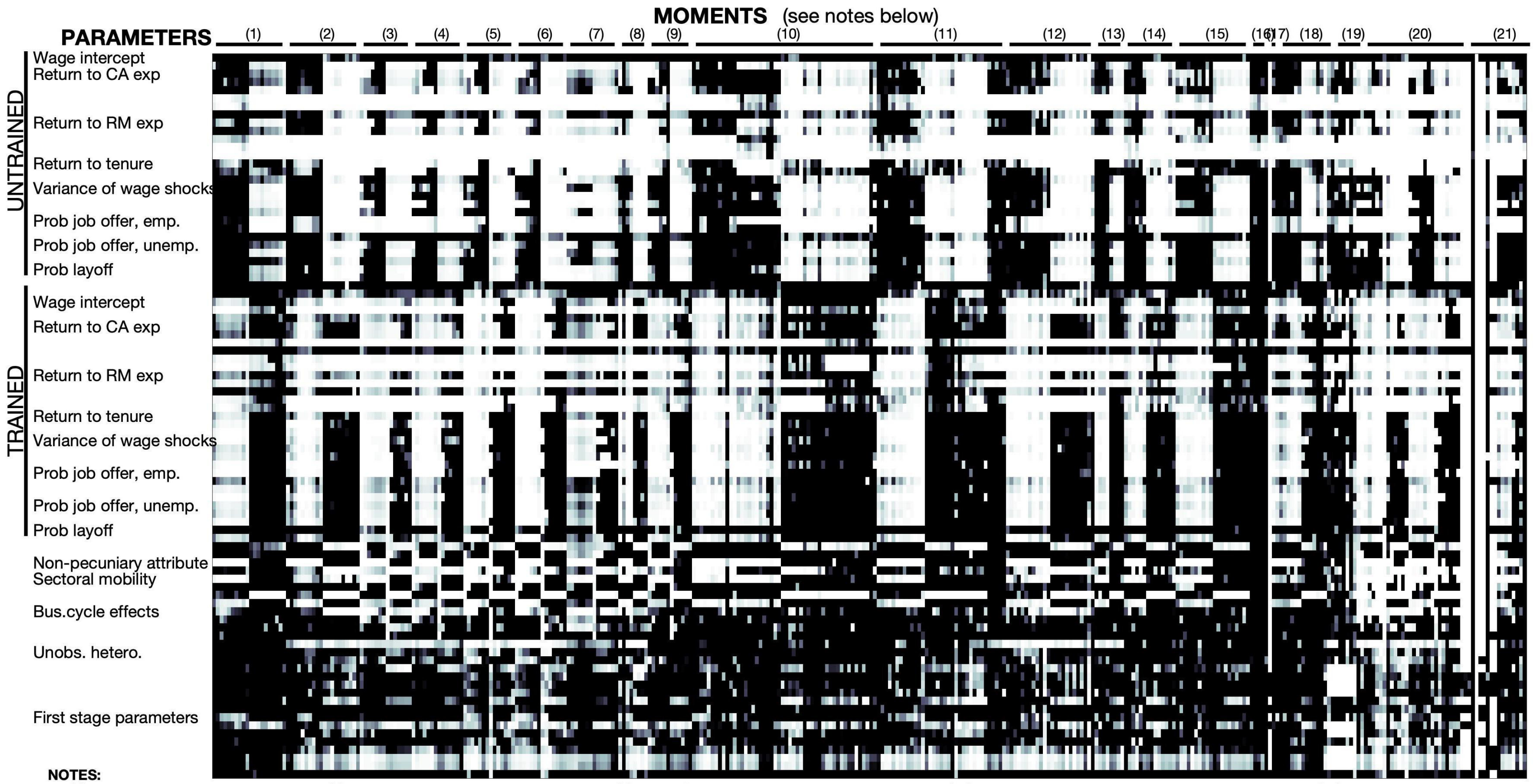
Notes: The observed coefficients (in columns 1 and 4) are derived from IAB data by OLS regressions, with standard errors clustered at individual level. The coefficients in columns 3 and 6 are obtained from OLS regressions based on data simulated from the model.

Table A7: Goodness of Fit: Wage Changes

	Observed (1)	Trained Std Error (2)	Simulated (3)	Observed (4)	Untrained Std Error (5)	Simulated (6)
Firm change	0.062	(0.00085)	0.07	0.097	(0.004)	0.089
CA experience	-0.0026	(0.0025)	-0.0019	-0.0015	(0.00016)	-0.0017
CA experience squared	0.00012	(3.8e-05)	9.3e-05	7.6e-05	(1.4e-05)	0.00011
RM experience	-0.0051	(3.2e-06)	-0.0049	-0.0048	(0.00014)	-0.0037
RM experience squared	0.00018	(2.9e-05)	0.00019	0.00018	(6.2e-06)	0.00017
Firm tenure	0.0004	(1.4e-06)	0.00016	0.00096	(9.6e-05)	0.00011
Firm tenure squared	-1.1e-05	(2.8e-05)	-1.3e-05	-3.3e-05	(4.3e-06)	-9.2e-06
End of apprenticeship training	0.82	(0.002)	0.88	-	-	-
Constant	0.035	(0.00011)	0.034	0.025	(0.00056)	0.022

Notes: The observed coefficients (in columns 1 and 4) are derived from IAB data by OLS regressions, with standard errors clustered at individual level. The coefficients in columns 3 and 6 are obtained from OLS regressions based on data simulated from the model.

Figure A1: Identification: how moments map into parameters



NOTES:

(1): Wage by age and training
(2): Prop working by age and training
(3): Abstract exp by age and training
(4): Manual exp by age and training
(5): Manual exp by age and training
(6): Tenure by age and training
(7): Number jobs by age and training
(8): Prop in abstract by age and training
(9): Prob of sectoral transition by age and training
(10): Wage regressions, level, by training
(11): Wage regressions, first difference, by training
(12): Labor market transitions, by training
(13): Job to Job transitions, by training
(14): Wage fixed effects, by training
(15): Wage regressions, first difference, by training
(16): Wage growth within and between, by training
(17): Education choice
(18): Last wage of unemployed, by training
(19): Quantiles of wage changes, by training
(20): Labor market transitions, by age and training
(21): First stage (school tracks, occupation and training)

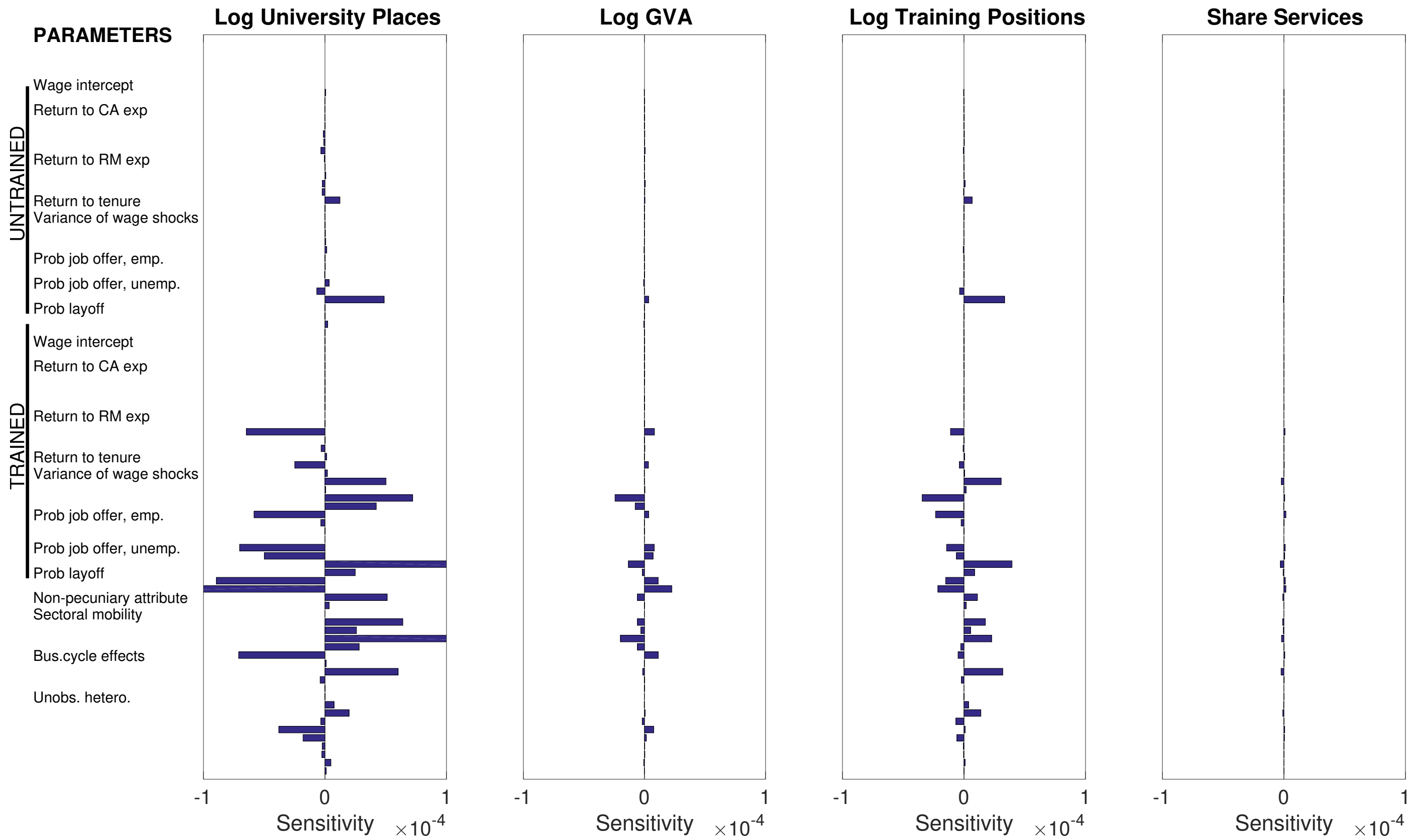


Figure A2: Sensitivity of Parameters to Instrumental Variables

Notes: Each plot shows the plug-in sensitivity of the parameters listed on the left side with respect to the moment associated with the first stage, listed in the title of the plot calculated using the method described in Andrews et al. (2017). The figure shows the effect of a one percent change in one of the instrumental variables on the parameters. For instance, a sensitivity of 10^{-4} indicates that a one percentage change in this variable would affect the fourth digit of that parameter.