Career Progression, Economic Downturns, and Skills*

Jerome Adda[†], Christian Dustmann[‡], Costas Meghir,[§] Jean-Marc Robin[¶]

October 31, 2012

Abstract

This paper analyzes the career progression of skilled and unskilled workers, with a focus on how careers are affected by economic downturns, and whether formal skills, acquired early on, are able to shield workers from the effect of recessions. We estimate a dynamic life-cycle model of education choice, labor supply and wage progression using detailed administrative data for Germany for numerous birth cohorts across different regions, following workers from labor market entry onwards. The model allows for labor market frictions that vary by skill group and over the business cycle. We find that sources of wage growth differ, with learning-by-doing being an important component early on for unskilled workers, whereas job mobility is important for workers who acquire skills in an apprenticeship scheme before labor market entry. Economic downturns affect skill groups through very different channels: unskilled workers loose out from a decline in productivity and human capital, whereas skilled individuals suffer mainly from a lack of mobility.

Keywords: Wage determination, Skills, Business cycles, Apprenticeship Training, Job Mobility

*We thank Joe Altonji, David Card, Mariacristina De Nardi, Eric French, Guy Laroque, Uta Schoenberg, Chris Taber and participants at seminars and conferences. We are grateful for funding from the DfES through the Centre for Economics of Education and to the ESRC through CEMMAP and the Centre for Fiscal Policy at the IFS. Costas Meghir also thanks the ESRC for funding through a Professorial Fellowship (RES-051-27-0204)

[†]European University Institute and IFS

[‡]University College London

[§]Yale, University College London and IFS.

[¶]Science Po-Paris, University College London and IFS.

1 Introduction

The early years in a worker's career are essential, not only because wages rise most rapidly, but also because workers are most vulnerable to economic shocks, and make important choices about training and investment into human capital. Since these early choices and events have significant long-term career consequences, it is important to understand their dynamics and effects, and the way they interact with shocks to the overall economy.

An essential part of this early career progression is wage growth. It has been viewed as a consequence of investment into learning and human capital (see e.g. Ben-Porath (1967), Becker (1994), Rosen (1972), Rosen (1976))[1], mobility and job shopping (see e.g Mincer and Jovanovic (1981), Topel and Ward (1992)), or both (see for instance Gladden and Taber (2000), Altonji, Smith, and Vidangos (2009), or Gladden and Taber (2009b)). While this literature has provided us with important insights into worker's wage progression, far less is known about how early career progression is affected by economic shocks, and how wage growth, transitions between jobs and into and out of non-employment, and the economic cycle interact.[2] This is surprising: not only is youth unemployment a major concern, but recent research has also highlighted the potentially harmful effects that economic shocks at early ages may have on workers' careers (see e.g. Oreopoulos, Von Wachter, and Heisz (2012)).

A related question is how the harmful effects of economic shocks on young workers' careers can be minimized. It is possible that skills - not acquired on the job, but in specifically designed training schemes - help shielding young workers from adverse labor market shocks. The observation that the great recession has had a far milder impact on youth unemployment in countries with a well developed firm-based vocational training schemes points indeed at the possibility that this type of training provision may help lessening the impact of economic shocks on young workers.[3] To assess this, it is important to un-

[1]See Lemieux (2006) for an assessment of estimating wage determination equations based on learning models.

[2]See also French, Mazumder, and Taber (2006) who emphasize this point, and address it in a reduced form context.

[3]For instance, while the youth unemployment rate between 2007 and 2011 has increased in most OECD countries, it has remained stable in Austria and Switzerland, and has even decreased in Germany - all countries with a large structured apprenticeship system that trains young workers for particular occupations after secondary school. SOURCE: OECD Labor Market Statistics 2012

derstand better the factors that determine wage growth, mobility and non-employment, and how these are related to economic shocks, in a context where young individuals have the opportunity to obtain vocational training in a structured apprenticeship programme.

In this paper, we address these questions. We first ask how do workers' careers progress after secondary school in a world where wages grow through job shopping and on-the-job learning, and where workers have the initial choice to acquire occupation specific skills in a 2-3 years structured vocational training scheme. We then investigate how career profiles of workers compare that have, and that have not chosen to enroll in a vocational training scheme, and how these are affected by economic shocks that hit individuals at different career stages. Addressing these questions will not only help understanding how early career apprenticeship programmes affect career progression - an issue under renewed scrutiny in the policy debate in many countries - but also how early-career vocational training may help alleviating the effects of economic downturns on employment and career progression of young workers.

To make progress, we develop a life cycle model of career choice and career progression where workers decide whether to acquire occupation specific training after secondary school, and before entering the labor market, or to join the labor market as unskilled workers. We model this in accordance with the institutional features in Germany, where close to four in five workers enter the labor market after secondary school in one of two ways: either directly as unskilled workers, or as apprentices, enrolling in a 3-years structured and firm based training scheme before entering the labor market as skilled workers.[4] The German system is unique in that it allows a precise distinction between skilled (i.e. those who chose apprenticeship training after secondary school) and unskilled workers (i.e. those who decided to join the labor market without further training) in a homogeneous work environment, where training decisions are made at the start of the career, and where skilled and unskilled workers do similar jobs.[5]

[4]Apprenticeship training combines formal classroom teaching with on-the-job training by qualified supervisors who implement a structured curriculum and that leads to certification of skills, within a narrowly defined occupation, such as bank-clerk or plumber. Firm-based apprenticeship training schemes have a number of advantages over vocational schools: craft techniques and customer interaction may be taught more effectively in a work environment than in the classroom, and firms may know better than schools which skills are needed at the workplace. Firm-based training may also allow for smoother transitions of firm-trained apprentices into employment (see Ryan, 2001a and Parey, 2009 for evidence).

[5]There is a large overlap in occupations for workers who enter the labor market directly without

Our model allows for direct job-to-job mobility, as well as transitions into and out of non-employment. We allow the key parameters that characterize search frictions to differ across skill groups, over the experience profile and, importantly, over the business cycle. We model workers' career progression in a framework where wages grow because workers learn on the job, and through job shopping. Our model draws from models of education choice[6] and wage determination.[7] It builds on the empirical labor literature, allowing for a rich stochastic structure of the wage equation, where wages grow with experience and job (firm) specific tenure, and depend on a match specific component as in Wolpin (1992).[8] The wage equations are specific to the two alternative careers (skilled or unskilled) as in a Roy type model. With search frictions, these careers could differ in rates of job arrival, job destruction and mobility. Thus for example, if occupation-specific apprenticeship training reduces flexibility because of the specificity of training, the job arrival rates should be lower for apprentices leading to longer job finding periods.[9] Our framework also draws on the macro labor literature, by allowing aggregate shocks to affect relative wages between the two groups, as well as labor market transitions (see Barlevy (2002), Nagypal (2005), Petrongolo and Pissarides (2008) or Shimer (2012)). Our model thus combines these two strands of the literature and allows us to study the effect of the business cycle on labor market attachment, experience and job mobility, in a way that allows for heterogeneous effects across skill groups and at various stages of a career.

Our analysis is based on unique administrative data drawn from social security records, which allows us to track the careers and wages of individuals from their entry to the labor market onwards. It also records precisely the training choices individuals

further training, and for workers who train in an apprenticeship scheme after school. In our sample individuals are employed in 292 3 digit occupations after labor market entry. Out of those, 19 occupations only employ non-apprentices (but employ only about 1 percent of all non-apprentices), and 53 occupations only employ apprentices (but employ only 1.4 percent of all apprentices).

[6]See Card (1999), Taber (2001), Card (2001), Cameron and Heckman (1998).

[7]See e.g. papers by Willis and Rosen (1979), Heckman and Sedlacec (1985), Altonji and Shakotko (1987), Topel (1991), Altonji and Williams (1998), Altonji and Williams (2005), Parent (1999), Dustmann and Meghir (2005).

[8]For recent contributions on wage dynamics see for example Meghir and Pistaferri (2004), Low, Meghir, and Pistaferri (2009) and Altonji, Smith, and Vidangos (2009). Sullivan (2010) and Pavan (2011) study wages in a structural context where agents can choose between occupations.

[9]See e.g. Heckman (1993). Fitzenberger and Kunze (2005) investigate whether this lock-in effect explains part of the gender wage gap in Germany.

make after labor market entry. This high quality of the data is an important strength of our approach. All transitions between employment and work and between different jobs as well as wages are recorded accurately by the firms thus allowing us to precisely assign wages to firms, and to observe detailed transitions across firms, and into and out of employment. Our sample covers men from what used to be West Germany who are born between 1960 and 1972. We observe these individuals the first time in 1975, and we follow them until 2004, covering therefore three decades and many entry cohorts. Our data offers the opportunity to compare careers of individuals entering the labor market facing effectively different economic conditions and training costs, because of the varying availability of skilled training. This provides exogenous variation that allows us to identify the initial choice of whether to enroll into apprenticeship training, or enter the labor market directly, which we combine with a dynamic structural model that characterizes apprenticeship and non-apprenticeship careers. The data provides also variation in the economic cycle, and where workers are exposed to recessions at various stages of their career.

We find that the careers of individuals who chose, and who chose not to acquire apprenticeship training after labor market entry are markedly different at an early career stage. Those who undergo training enter the labor market with far higher wages, while those who enter as unskilled workers undergo a period of rapid wage growth during the first 5 years in the labor market. Remarkably, this wage growth during the early career phase is mainly due to learning on-the-job, and to a far lesser extend due to job shopping. Interesting are the differences in the fundamental parameters that drive wage progression for these two groups: while unskilled workers have higher job destruction rates than skilled workers, they also have higher job arrival rates, both on the job and off the job. These differences narrow over the career, but never converge. This is surprising, given that individuals are fairly homogeneous before making their training choice, and compete for similar jobs.

These differences in the underlying parameters, which are larger at the early career stages, lead to surprisingly different ways how skilled and unskilled workers respond to economic shocks. Evaluating the long-run effect of a recession, we find that economic shocks have permanent effects on human capital for unskilled as well as for skilled workers. However, it matters at what career stage a recession hits: while an economic shock early on in workers' careers reduces human capital of unskilled workers by twice as much as for skilled workers, these differences tend to become far smaller when the recession hits at later career stages. Do these shocks to human capital translate into wages? We find that exposure to an economic shock early in a worker's career leads to wage reductions that persist 5-10 years. However, the differences in wages between skilled and unskilled workers is far smaller than what the difference in human capital would suggest. This is due to a dramatic reduction in job mobility of skilled workers during a recession, who tend to remain with the same firm, while unskilled workers are more mobile and compensate for the loss in human capital through the accumulation of search capital.

Our model helps us to understand the precise channels through which workers' careers are affected by economic shocks. It thus contributes to an important and growing recent literature that investigates the effect of economic shocks on workers' careers (see e.g. Oddbjorn and Roed (2006), Schmieder and von Wachter (2010) Davis and von Wachter (2011), or Oreopoulos, Von Wachter, and Heisz (2012)). While these papers provide impressive insights into the possibly devastating effects of economic shocks on workers' careers, they do not distinguish between job destructions that are caused by an economic recession, and job destructions that would have happened anyway. Furthermore, in any analysis that is based on DiD type identification strategies, longer term projections may be contaminated by other economic shocks. While supporting the key findings of these papers, our analysis adds to this literature by allowing us to separate effects of a recession from separations that would have occurred anyway, to isolate the impact of a past shock on future careers from other possible determinants, and to compare the career impacts of shocks that hit workers at different career stages.

Our paper also contributes to a better understanding of training schemes that develop workplace related vocational skills. Such schemes are (once again) recognized as a key factor in strengthening competitiveness and growth.[10] A crucial question for assessment

[10]See e.g. President Obama's "manufacturing skill speech" (http://www.whitehouse.gov/the-press-office/2011/06/08/president-obama-and-skills-americas-future-partners-announce-initiatives), or the renewed emphasis on firm-based apprenticeship programmes of the British Government (see the UK 2011 budget (http://cdn.hm-treasury.gov.uk/2011budget complete.pdf), where the government announces an

of such schemes is how they affect workers' productivity and employment patterns. A small literature estimates the effects of apprenticeship training on wages (see e.g. Winkelmann (1996) and Fersterer and Winter-Ebmer (2003) who report OLS estimates for the wage returns to apprenticeship training in Germany and Austria of around 15-20 percent, and Fersterer, Pischke, and Winter-Ebmer (2008) who report IV estimates of 2.5 and 4 percent per year of training). Although these papers provide important insights into the returns to enrollment in vocational training schemes, they focus on the wage component only, and do not consider the role of endogenous experience profiles and the effects of selection into work (and its effect on life cycle earnings through employment). However, these factors are likely to be very important when comparing careers of skilled and unskilled workers. Moreover, they may interact differently with aggregate shocks for skilled and unskilled workers. Our contribution is thus to provide a more detailed understanding of the various channels that lead to higher returns for workers who undergo apprenticeship training - which is a key factor in assessing whether such training schemes should be encouraged in other countries.

The structure of the paper is as follows. Section 2 describes the data set, discusses institutional features and provides descriptive statistics. Section 3 presents the model. We explain our estimation method in Section 4. The results are presented in Section 5. Section 6 concludes.

2 Background and Data

In this section, we give some brief description about training choices and the firm based apprenticeship system we are analyzing in this paper. We then describe our data and sample, and provide some descriptive statistics.

2.1 The Apprenticeship System

The German Apprenticeship System is a vocational training programme which combines on-the-job training, provided by the firm, with school education, provided and funded by the state. Similar systems operate in Austria and Switzerland. The system offers training

---

additional £180 million for up to 50,000 additional apprenticeship places).

in more than 500 white- and blue collar occupations[11]. In practise, individuals 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 occupations, with slightly more than two-third of those being blue collar ones.

Apprenticeship training typically starts after secondary school, at around the age of 16. Germany tracks children after the age of 10 in lower, intermediate and upper secondary schools. Pupils who attend lower and intermediate secondary schools typically enrol in blue or white collar apprenticeship schemes. Pupils who attend upper secondary schools are entitled to enrol directly into university.[12]

Apprenticeship training is highly structured, with a well-defined curriculum. It takes place at the workplace at 3-4 days a week, under the supervision of qualified instructors, where practical and workplace related knowledge is acquired, and at vocational state schools at 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. We refer the reader to Steedman, Gospel, and Ryan (1998) for more details.

2.2 Data and Sample

Our main data is a 2 percent sample of administrative social security records, covering the years between 1975 and 2004, and made available by the German Institute for Employment Research. It records all spells of employed work of workers in the private and public sectors, with exact dates when each job started and ended. The data does not cover civil servants and the self employed. The data set reports 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, we observe the average wage from the beginning of the calendar year or the employment spell (if it started after the beginning of the calendar year) until the end of that spell. Thus wages are not averaged across different firms. The wage data is top coded at the earnings limit for social security contributions.

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

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

For the sample we consider, this concerns only about 2.2 percent of all wage spells. We take top coding into account in our estimation procedure, and we describe details below. The data contains also information on the apprenticeship training period, and whether a worker holds an apprenticeship qualification or not, as well as their overall educational qualifications.

In our analysis, we focus on West-German men born in the period between 1960-1972, who enter the labor market with a lower or intermediate secondary degree, which is not sufficient for attending university directly, and which is 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 any initial conditions problem.

We then define two groups: individuals who enrol in apprenticeship schemes for at least 2 years and successfully complete their training (in what follows we refer to these individuals as "apprentices" or "skilled"), and individuals who enrol for a shorter period, but do not graduate, or do not enrol and enter the labor market directly (we refer to these as "non-apprentices" or "unskilled").[13] There are also one-year vocational courses, which do not lead to vocational degrees. Thus, among the non-apprentices, we may include some who were exposed to some apprenticeship training, or who attended a vocational preparatory classes, without following up further training.

From this data, we construct a data set of quarterly spells, thus assuming that all decisions are made on a quarterly basis. Whenever during a quarter multiple spells are present (e.g. an employment and an unemployment spell), we assign to that quarter the spell that covers the largest proportion of that quarter. When the individual does not move firms and thus the wage we observe is an average over more than one quarter, we treat this as a time aggregated wage where we do not observe the individual constituents of this average. This time aggregation problem is fully accounted for during estimation, as we explain later.

[13] As an alternative to firm-based apprenticeship training, some youth attend 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). About 6 percent of our sample undertakes qualifying training in these vocational schools. Wage profiles of those who went through firm based training and vocational schools are almost identical.We add these to the group of apprentices, but allow their work experience to accumulate at a different rate during the training period.

The data contains 38,018 individuals who enroll in an apprenticeship training scheme after secondary school, and 4,392 individuals who join the labor market directly and without further training. These are followed through time, quarter after quarter up until 2004; we have thus a total of 3,667,223 quarterly observations. Finally, to identify the determinants of choices of school tracks at age 10, we use 69,084 individuals who follow the vocational track and 10,608 who follow the academic track. We provide more detail on the sample selection in the Appendix.

2.3 Descriptive Analysis of the Data

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 age minus the age at the end of compulsory schooling, taken to be 16) for those with an apprenticeship qualification ("Skilled Wage"), for those currently training as apprentices ("Wage in Apprenticeship") and for non-apprentices ("Unskilled Wage") as well as the difference in wages between apprentices and non-apprentices (right hand axis).

The figure shows that non-apprentices have a rapid increase in their wage during the first five years in the labor market, with real wages increasing by 11 percent per year on average. Over the next twenty years however, overall wage growth is just below 9 percent, resulting in a 0.4 percent real average growth per year. Those who enroll in apprenticeship training schemes are paid a very low wage during their training period, covering part of the cost of their training. At the end of the apprenticeship training, however, wages increase sharply and overtake those of non-apprentices. From there on, the wages of those with an apprenticeship qualification increase slightly faster, by about 1 percent per year. After twenty years, wages of apprentices are about 15 percent percent higher than those of non-apprentices.

From this graph it almost seems puzzling that anyone wishes to follow an apprenticeship career, given the large up-front investment in training that lasts about 3 years and the apparently low rate of return in terms of wages. Comparing the net present value of the flow of wages as depicted in Figure 1 between skill groups show that unskilled individuals are better off by about 2.3 percent.[14] Of course these simple figures are mis-

[14]This figure is calculated over a horizon of 25 years using an annual discount rate of 0.95 and assuming

leading, as comparative advantage and other differences between the two career paths may well explain the large participation rates in apprenticeship schemes. This is one of the questions we investigate below, by allowing for such differences in the model that will follow.

Indeed, wages are only one dimension in which education groups may differ. Another important dimension is labor market attachment. Figure 2 shows the proportion of individuals who are in work as a function of age.[15] It is apparent from the figure that labor market attachment of apprentices is stronger than that of non-apprentices, with a higher fraction of apprentices working at any age. The difference in the proportion of individual working narrows from about 10 percent at age 25 to 5 percent around age 40.

In Table 1 we report in more detail the transitions of apprentices and non-apprentices between the different states. The table displays the quarterly transition probabilities by apprenticeship status and time in the labor market, which starts when the individual has found a first job or an apprenticeship training scheme. The figures show that unskilled workers have a higher probability of dropping out of work. During the first five years in the labor market, each quarter, about 3 percent of employed skilled workers exit, while this figure is about 9 percent for the unskilled. This proportion decreases when we focus on more senior workers, and the difference between the two groups narrows. The figures in Table 1 also reveal that apprentices have a higher probability to return to work from non-employment. For instance, for workers with 5 to 10 years of potential experience, 19 percent of skilled unemployed individuals find a job from one quarter to the other. For unskilled this figure is only 7 percent. Further, the probability of job to job transitions is higher at the beginning for non-apprentices, but declines after five years for both groups and becomes marginally higher for apprentices.

To summarize, these figures indicate that - overall - the unskilled spend less time working. Over a 25 years period, they work a total of 21.9 years, compared with a total of 22.5 years for skilled workers. The greater job attachment and the resulting higher earnings contributes to "compensate" apprentices for the lost earnings early on. If we

no selection into education.

[15]Germany has 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.

combine labor market participation and wages, using a replacement rate of 40 percent when unemployed, we find that skilled individuals are two percent better off in terms of net present value when they first enter the labor market; this number increases to 5 percent if we assign zero earnings to unemployed workers. Hence, the decision to obtain apprenticeship training cannot be assessed solely on the basis of the implied earnings advantage as depicted in Figure 1. Another important dimension of this choice is the employment prospect.

Figure 4 plots the number of firms in which an individual has, where the horizontal axis carries potential experience. It is evident from this figure that the unskilled are more mobile during the first few years in the labor market. Thus, job shopping can be an important source of the large initial wage growth for the unskilled, as we illustrate in Figure 1. To investigate this further, we decompose wage growth into within and between firm wage growth and plot it against potential experience (see Figure 3), distinguishing between the two skill groups. Between job wage growth is indeed substantial, between 20 and 40 percent for the unskilled during the first 2-3 years in the labor market, when apprentices are still in the training phase. The gain in wages falls over time, but is still large for both groups until about 5-7 years in the labor market, with returns being close to zero after about 15 years. If we think of wage improvements as being due to better matches, as in our model presented below, the decline is expected because the probability of an improvement will decline as the worker climbs up the job-quality ladder. Within firm wage growth for the non-apprentices is likewise very high early on in the career reflecting the rapid learning that takes place on the job. The equivalent training for the apprentices takes place during the official training period (which we have not shown in the figure).

Figure 5 shows the path of residual wages for both skill groups over time, together with the deviation of GDP from a trend. The residual wage is obtained by projecting wages on age and regional dummies, so as to make individuals comparable across years. We also have shaded the periods when GDP is below its trend, which we define as an economic downturn. Our data encompass three downturns, one in the mid-seventies, a large one in the early eighties and one at the end of our sample, which starts in 2004. The figure shows that wages are procyclical, with a correlation with GDP of 0.4 for unskilled workers and 0.57 for skilled ones. The precise mechanism that leads to such a correlation is difficult to ascertain in a reduce form context. We return to this issue in detail in Section 5.4.

3 The Model

We now turn to describing the key features of our model, which is set in discrete time, and where one period lasts one quarter. It focuses on individuals who leave secondary education at age 16 (and who chose the low or intermediate school track at age 10, see Section 2.1). At that point individuals have the choice either to enroll in an apprenticeship scheme, or to enter the labor market as unskilled workers. Once this choice has been made, individuals start their career. During apprenticeship, individuals may move to a new employer but not to unemployment. Throughout their work career, individuals receive job offers with some probability, which may differ depending on whether they are employed or not. Jobs can end either because of a quit or because of exogenous job destruction. Individual choices include moving between jobs when the opportunity arises and between work and unemployment as well as the initial education choice.

Individuals derive utility from consumption and leisure, and experience a potential loss of utility associated with a job change. When out of work individuals are entitled to unemployment benefits, which (according to the benefit system in Germany) are a function of the wage earned in the last job. Individual i maximizes the flow of utility over the life-cycle:

$$V_{it}(\Omega_{ift}) = \max_{\{L_{if't}|L_{ift-1}\},} E_t \sum_{\tau=t}^{T} \beta^t u(c_{if\tau}, L_{if\tau-1}, L_{if'\tau}) \tag{1}$$

$$\text{subject to} \quad c_{if\tau} = w_{if\tau} L_{if\tau} + (1 - L_{if\tau}) b^U_{i\tau},$$

where Ω_{ift} is a vector of state variables for individual i, in period t, who was in firm f at the end of the previous period. In case the individual was not working in the preceding period, f indicates the identity of the last firm in which the agent was employed. This vector includes the education status, work experience, past labor supply, tenure on the job in the firm, the match quality of the job, the aggregate state of the economy as well as time invariant unobserved characteristics pertaining to that individual. At the start of period t, the individual chooses labor supply in a potentially new firm f', conditional on past labor supply in firm f. We denote labor supply by $L_{if't}$, which takes two values, either work, $L_{if't} = 1$, and non-employment, $L_{if't} = 0$. As detailed below, past labor market status conditions the availability of offers with potentially different arrival rates for employed and unemployed individuals. Employed workers may receive offers from other firms and have to decide whether to stay with the current firm or move to another firm. We denote wages by $w_{if\tau}$ and unemployment benefits by $b^U_{if\tau}$, which is indexed by f as benefits depend on past employment, We denote by β the discount factor. The individual derives utility from consumption $c_{if'\tau}$ and we abstract from savings. Hence, consumption equals wages or benefits. We detail below each component of the model in more detail.

Aggregate shocks: We characterize the macroeconomic fluctuations of the economy around the steady-state growth trend by de-trended GDP. The macro shock is relevant because it potentially affects the relative price of the two skill groups as well as the relative attractiveness of being out of work. It also affects the probability of finding a job as well as the job destruction rate, in a specific way to both skill groups. This allows the model to capture the different effect of business cycles on skilled and unskilled workers along several dimensions, such as unemployment duration or job tenure, and which we will explore later on. The macro state variable G_t is modeled as a discrete two state Markov process of order one. The transition probabilities are presented in Appendix C in Table A1.

Wages and Matches: The central component of the model is the job contract. If a worker i and a firm f form a match at time t, the output is split according to a rule that yields an annual wage w_{ift} to the worker. The way the split is determined is not modeled here.[16] One simple way to think about the wage-setting mechanism is Nash bargaining.

[16]For equilibrium wage determination with shocks to firm productivity and heterogeneous workers see Lise, Meghir, and Robin (2009).

Worker i and firm f negotiate a wage given match output and job amenities. If the worker happens to meet another firm $\widetilde{f}$ while employed, he compares the two bargaining solutions and takes the best offer. Wage contracts are continuously updated following shocks to match productivity, and, as in a standard Mortensen and Pissarides (1994) model, really bad productivity shocks may result in unemployment.

Wages are modeled as follows. Let $Ed_i \in \{0,1\}$ denote the worker's apprenticeship qualification status (1 for apprentices and 0 for non-apprentices). Let X_{it} be the number of quarters spent in work (including the apprenticeship period) since age 16.[17] Let T_{ift} denote the number of quarters spent in the current job ($T_{ift} = 0$ if the job in firm f starts in period t). Let ε_{Wi} be a permanent individual characteristic that is unobserved by the econometrician but is known by the worker and observed by the employer. Quarterly earnings w_{ift} are functions of the macroeconomic shock, G_t, education, Ed_i , experience, X_{it}, tenure, T_{ift}, the unobserved permanent heterogeneity variable, ε_i, and a match-specific component, κ_{ift}:

$$\ln w_{ift} = \alpha_0(\varepsilon_{Wi}) + \alpha_{Ed} Ed_i + \alpha_X(X_{it}, Ed_i) + \alpha_T(T_{ift}, Ed_i) + \alpha_G(Ed_i)G_t + \kappa_{ift}, \quad (2)$$

where α_X and α_T are two education-specific functions of experience and tenure. We use a piecewise linear function, with nodes at 0, 2, 4, 6, 10 and 30 years of experience and tenure. The specification is motivated by the fact that most of the non linearity in wages profiles is early on, so we have a denser grid between 0 and 10 years of actual experience. Unobserved heterogeneity affects the overall level of log wages. This specification is in line with the empirical evidence found in French, Mazumder, and Taber (2006). They show that the return to experience appears to be unrelated to the business cycle. The specification with an additive and separate unobserved productivity term is consistent with findings in Gladden and Taber (2009a).

When the worker and the firm first meet ($T_{ift} = 0$) they draw a match specific effect $\kappa_{ift} = \kappa_{if}^0$ such that

$$\kappa_{if}^0 \sim \mathcal{N}(0, \sigma_0^2\,(Ed_i)) \quad (3)$$

which captures the heterogeneity in wages when individuals start a new job. We interpret

[17] $X_{i,t+1} = X_{it} + 1$ if the worker is working in period t; otherwise, $X_{i,t+1} = X_{it}$. We do not allow for depreciation of skills while unemployed.

this as match specific heterogeneity and we allow it to differ by apprenticeship status allowing us to estimate the extent to which job opportunities vary for skilled and unskilled workers. Flinn (1986) shows the importance of such matches to explain the wage path of young workers. Then, whenever $T_{ift} \geq 1$,

$$\kappa_{ift} = \kappa_{ift-1} + u_{ift}, \tag{4}$$

$$u_{ift} \sim iid \quad \mathcal{N}(0, \sigma_u^2(Ed_i)). \tag{5}$$

This allows for the possibility that the value of a match and the contracted wage can change, while allowing for persistence over time. Contrary to the US and the UK, in Germany, the cross sectional variance of wages does not increase over the lifecycle, which means that a random walk of wages that continued across jobs would lead to counterfactual implications and would be inappropriate. This led us to the above specification, where the random walk component is reinitialized when changing jobs, leading to wages that are stationary over the life-cycle, because jobs have a finite expected life.

The utility of working and being out of work: Utility is assumed to take a log form. In addition, we allow for a mobility cost or benefit μ_{if} when a worker moves between jobs. This allows for the possibility that workers may move to a job that pays lower wages, as is observed in the data. The one off benefit/cost of moving is an iid random variable μ_{if} such that

$$\mu_{if} \sim \mathcal{N}(m_\mu(Ed_i), \sigma_\mu^2(Ed_i)).$$

While unemployed, the individual derives a utility from unemployment benefits; these are calculated as a fraction of the last wage when employed (denoted as $w_{i(-1)}$), 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 40 percent.[18] In addition,

[18]We have taken a replacement rate that is on average correct for our population. Modeling the entire system would imply an increased state space.

there is a utility of leisure which varies across individuals on the basis of education, experience, unobserved heterogeneity ε_{Wi} and a Gaussian white noise η_{it} with variance σ_η^2. Thus, the instantaneous utility of unemployment is:

$$R_{it}^U \equiv R^U(Ed_i, X_{it}, w_{i(-1)}, \eta_{it}) = \ln(\gamma_U w_{i(-1)}) + \gamma(X_{it}, Ed_i, \varepsilon_{Wi}) + \eta_{it},$$

$$\eta_{it} \sim iid \quad \mathcal{N}(0, \sigma_\eta^2(Ed_i)),$$

with $\gamma_U = 0.4$ and where $\gamma(X_{it}, Ed_i, \varepsilon_{Wi})$ is the utility of leisure, which is education-specific, and varies with experience and unobserved heterogeneity. The effect of experience is modeled as a piecewise constant function (with nodes at 0, 2, 4, 6 and 30 years of experience).

Finally, we assume that all shocks $\{\kappa_{if}^0,\ u_{ift},\ \mu_{if},\ \eta_{it}\}$ are jointly as well as serially independent, and independent of the unobserved heterogeneity vector ε_i (see below for a complete description of unobserved heterogeneity).

Transitions: Employed individuals may be laid off with probability $\delta_{it} \equiv \delta(G_t, Ed_i, X_{it})$, which depends on the state of the business cycle as well as experience and apprenticeship status. Exogenously displaced individuals suffer a loss of their match specific effect which will lead on average to lower wages upon re-entry, followed by a catch-up, which is consistent with findings in Bender, Dustmann, Margolis, and Meghir (2002) and von Wachter and Bender (2006). Conditional on not being laid off, they draw an alternative job offer with probability $\pi_{it}^W \equiv \pi^W(G_t, Ed_i)$.

Unemployed individuals draw a job offer with probability $\pi_{it}^U \equiv \pi^U(G_t, Ed_i, X_{it})$ function of the aggregate shock, education and experience. They can choose to take this job, depending on how the value of working compares to the value of unemployment.

Our model has therefore some of the crucial features which are discussed in the macro labor literature, which has emphasized the role of frictions in generating unemployment (see for instance Davis and Haltiwanger (1992), Barlevy (2002), Nagypal (2005), Petrongolo and Pissarides (2008) or Shimer (2012)). An important insight of this literature is the role of job offers when unemployed, as well as job mobility, and how these interact with the business cycle.

Individual decisions to work, to move to a new job or to quit working are carried out by comparing the lifetime values of each of these states. The structure of the value functions is presented in appendix B.

Education decision: The choice to follow an apprenticeship training is assumed to be a one off decision made at age 16 by comparing the value of a career under the two training alternatives allowing for both the direct costs of training and foregone earnings. We assume that both an unskilled job and an apprenticeship position are available immediately. For simplicity, we refer to that decision as "decision at age 16", although there is some heterogeneity ion our sample, and in practise, we start modeling from the point we see individuals joining the first job or an apprenticeship scheme.[19] The choice to become an apprentice is based on comparing the value of this decision with the value of joining the labor market directly, minus the cost of the training decision, which can be expressed as

$$V_{it}\left(\Omega_{ift}|Ed_i = 1\right) - cost_{it} > V_{it}(\Omega_{if't}|Ed_i = 0), \tag{6}$$

where $\Omega_{ift}|Ed = j,\ j \in \{0, 1\}$ is the state vector at age 16, with zero experience and tenure, business cycle state G_t and with an offer from firm f for a training position, consisting of a match effect κ^0_{ift} and mobility cost μ_{if}. The value of unskilled work is conditioned on $\Omega_{if't}$, which is also evaluated at zero experience and tenure, the same business cycle shock, but with an offer from a different firm f' for an unskilled position. This offer consists of a match effect $\kappa^0_{if't}$ and mobility cost $\mu_{if'}$.

The cost of training is modeled as:

$$cost_{it} = \lambda_R(R_i, G_t) + \lambda_0(\varepsilon_{Ti}) + \omega_{it}, \tag{7}$$

where $\lambda_R(R_i, G_t)$, represents the (deterministic) direct costs of apprenticeship training, which we allow to depend on the relative scarcity of apprenticeship training schemes across time and regions (see e.g. Parey (2009) who illustrates the strong variation in training schemes across regions in Germany). We proxy these by including interactions between region of residence, (R_i), and the state of the business cycle, G_t, both measured

[19]On average individuals start apprenticeship at 17.1 and an unskilled job at 17.8 years of age. Delays between graduation age and labor market entry may be due to undertaking short vocational courses, gap periods, or compulsory military service before finally opting for an apprenticeship or a non-apprenticeship career.

when the choice is made at age 16. These interactions reflect how aggregate shocks affect each of the eleven regions of (West) Germany. Such differential effects of GDP shocks across regions will occur because industrial composition differs across regions or because employment in some industries is more procyclical than in others. The availability of data for thirteen birth cohorts observed in eleven states provides sufficient variability for estimation.

We allow for unobserved heterogeneity in the costs of training, $\lambda_0(\varepsilon_{Ti})$, so as to capture the possibility that individuals may differ in their ability to learn in an academic environment. The unobserved heterogeneity term ε_{Ti} is potentially correlated with the other heterogeneity term, ε_{Wi}, which affects wages in equation (2). Hence this specification allows both for selection on unobserved returns to education and for ability bias as expressed in the labor literature.[20]

Finally, we denote by ω_{it} a normally distributed iid shock to the cost of training (capturing for instance travel costs as well as family background) that is revealed to the individual before the training choice is made. It induces a probability for this choice, conditional on all the other shocks, from which it is independent. The shocks ω_{it} and $\lambda_0(\varepsilon_{Ti})$, together with the match specific effects in both alternatives and the non-pecuniary benefits, need to be integrated out because they are not observed.[21]

Unobserved heterogeneity: As detailed above, wages and apprenticeship costs depend on unobserved heterogeneity. As argued by Taber (2001), who also analyses a model of schooling choice and careers, it may be far too restrictive to allow just for one factor of heterogeneity. We thus assume that the vector $\varepsilon_i = \{\varepsilon_{Wi}, \varepsilon_{Ti}\}$ consists of two random variables which follow a bivariate discrete distribution, each with two points of support (Heckman and Singer (1984)). The two elements capture the ability to learn (which thus correspond to the individual specific costs of training), and productivity in the labor market; they may be positively or negatively correlated or possibly not be correlated at

[20]See for example Griliches (1971), Card (2001), Heckman and Vytlacil (2005) and Carneiro, Heckman, and Vytlacil (2006) among many others.

[21]In principle one could estimate a richer model allowing for regional shocks and mobility but this would greatly increase the state space and the choices to be made (see Kennan and Walker (2010) or Dahl (2002)).

all.[22] Education choices depend on the costs of education (observed or not) and on the expected wage gains.

4 Estimation

4.1 The Selection of our Population and Initial Conditions

As we explain above, the population whose labor market behavior we model consists of individuals who at 10 years of age have enrolled in the lower or intermediate secondary school track (a decision that is made by parents, based on primary school teacher's recommendations), but not in the high track school track, who complete secondary schooling at the age of 15 or 16, and who either enroll into an apprenticeship training scheme afterwards, or enter the labor market without further education. [23]

Thus, the population we consider does not cover those who - at the age of 10 - enroll into higher track schools, allowing them to ultimately enter university. This is about 20 percent of each cohort. To address this initial conditions problem we specify a reduced form probability of choosing the academic path, as a function of the region and year of birth of the individual (reflecting the economic conditions at the time) as well as of the two factors of unobserved heterogeneity in the vector ε_i. The key assumption in this approach is that the distribution of unobserved heterogeneity is independent of region and cohort. We estimate the parameters describing the probability of choosing the lower tracks together with the parameters of the model.

4.2 Method of Simulated Moments

The model is estimated using simulated method of moments, by minimizing the distance between a set of chosen moments from the data and the moments implied by the simulated careers from the model (McFadden (1989)). The criterion we minimize takes the following form:

[22] In practice we normalize one point of support to be zero and include a constant in the wage of each sector and in the costs of apprenticeship.

[23] Table 2 shows that for the cohorts 1960, 1965, and 1970, around two in three individuals choose apprenticeship training; the fraction of each cohort entering the labor market without further education decreases slightly, from 16 percent for the 1960 cohort, to 11 percent for the 1970 cohort. The fraction of those who choose an academic career (which typically follows graduation from a high track secondary school) increases slightly, from 20 percent to 24 percent.

$$M(\theta) = (\hat{m} - g^S(\theta))'\hat{\Sigma}^{-1}(\hat{m} - g^S(\theta))$$

where $\hat{m}$ represents a vector of data moments, $g^S(\theta)$ represents the moments implied by the model, based on S simulated careers, and $\hat{\Sigma}$ represents a weight matrix. Here we chose $\hat{\Sigma}$ to be a diagonal matrix which contains the variances of the observed moments. The standard errors are estimated as in Gourieroux, Monfort, and Renault (1993).

Estimation is based on the simulation of 12,000 individual careers, starting from the point when - at 10 years of age - individuals are allocated to the lower, intermediate, or higher (and more academic) track. Using the simulated data we then construct moments that correspond to those we obtain from the observed data. We deal with time aggregation in wages by generating simulated data at the quarterly frequency, imposing the same time aggregation as on the real data, and constructing the moments in the same way. For instance, for workers employed a full calendar year within the same firm, the administrative data we use reports 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 firm.

We deal with top coding of wages in a similar way. We impose the same rules for top coding in the simulated data as in the observed ones. This procedure is essentially similar to a Tobit model, given the normality assumptions we have made for the shocks.

We use a total of 414 moments to estimate a total of 116 parameters. The career paths of apprentices and non apprentices are characterized by 169 moments which we use to estimate 70 parameters; the training choice is characterized by 13 parameters, and we use 124 moments to estimate these; the choice of the academic track is described by 33 parameters, where estimation is based on 121 data moments. A full list of moments can be found in the tables of Appendix D, and we will describe here only the estimation of some of the key parameters of the model. When constructing moments, we always control both for region and aggregate time trends so that identification does not rely on pure cross-sectional or temporal variation.

The career path of individuals is characterized by a number conditional moments, obtained from linear regressions. These are obtained, for instance, by regressing the (log) wage level on a function of experience, tenure and the business cycle for skilled and unskilled individuals. This set of moments helps identifying the return to experience and tenure by skill groups. To identify the variance of wages over the life-cycle, which depends on the distribution of initial matches and unobserved ability, we regress the squared residual of the wage equation on a constant, a function of potential labor market experience, by skill groups. Moments obtained from regressing the changes in log wage on a function of experience, tenure, business cycle and skill group helps to identify match specific heterogeneity, as well as the return to tenure and experience. To identify the innovation to the match specific effect, we use as moments the coefficients from a regression of the squared residual of the wage change equation on skill groups dummies.

We further estimate linear probability models to characterize the proportion of individuals in work and linear regressions to describe the number of jobs held as a function of potential experience and business cycle. When considering business cycle effects, we always allow for separate effects between skill groups and interact it with potential experience. This interaction captures how business cycles affect young and older workers differently.

For the choice of apprenticeship at age 16, we use as moments the proportion of apprentices by region and year. We proceed in a similar way for the choice of the academic track, by matching the proportion of individuals who chose the lower track by region and year in the observed data and in the simulated data. Finally, in constructing the moments we account for heterogeneity due to the initial region of residence at age 16, as well as aggregate time trends by including regional dummies and a quadratic trend.

5 Career Paths across Skill Groups and Economic Shocks

5.1 The Fit of the Model

We start with summarizing how well our model fits the data, by comparing some of its key predictions to those we obtain from the raw data. One important set of moments are the evolution of employment and log wages over the life cycle, for the two groups of workers. These are summarized in Figure 6, comparing the profiles we obtain from the data with those generated by our model. As is apparent from the figures, not only does the model capture the wage profile over the life career cycle very well, but it also matching quite precisely the slightly U-shaped profiles of the proportion of individuals in work. An important moment is the fraction of individuals who enroll in apprenticeship training. Here the overall proportion of non-apprentices is 10.5 percent in the raw data, while the model's prediction is about 9.4 percent. We provide additional assessment of the fit of the model along various dimensions in much detail in Appendix D. Overall, these tests show that the model fits the data moments remarkably well.[24]

5.2 The Parameter Estimates

We now turn to the estimated parameters. Table 3 presents a subset of parameters that are fundamental for understanding differences in the early and later career paths between skilled and unskilled workers. These include parameters that characterize the distribution of innovations to match specific effects, the distribution of match specific effects (first panel) as well as the job destruction rate and the job arrival rates (second panel).

The first panel reports the standard deviations of the initial match specific effects and the innovations to match specific effects, σ_0 and σ_u, the dynamics of which we describe in equations (3) and (4). The estimates show that apprentices and non-apprentices face different match specific distributions. Whereas initial matches are similar across skill groups, the variance of innovations to match specific effects for non-apprentices is larger than for qualified apprentices. The difference is quite sizeable, with the standard deviation of innovations for the unskilled being nearly twice as large as for qualified apprentices. Differences in these parameters, paired with a higher job-to-job mobility due to differences in job destruction and offer rates (to which we turn next), may partly explain the high wage growth for non-apprentices, which is shown in Figure 1. We present below in Section 5.3 a decomposition of wage growth to better understand its determinants.

[24]We do not assess the fit of the model using chi-square tests. Given the large number of observations we use for the estimation of the moments, and given the degree of over-identification, even small deviations from the data moments will be statistically significant.

In the second panel of Table 3 we report the job destruction rates (δ), and the job arrival rates when employed (π_W) and when unemployed (π_U), again separately for skilled and unskilled workers. We do not report estimates for individuals who are in apprenticeship training, as - in accordance with regulations in Germany - individuals cannot be fired during the training period, once enrolled. As we explain in section 3, we allow the job arrival and destruction rates to vary with skill level, time in the labor market, and the business cycle. Inspection of the Table shows that the job destruction rates are markedly higher for unskilled than for skilled workers, particularly in the first four years in the labor market. The difference persists beyond that period, but becomes smaller. Thus, exogenous separations seem to play a far more important role for the mobility of unskilled workers during the first years in the labor market. Unskilled individuals have - on the other hand - higher job arrival rates while on the job, as well as when in unemployment, in booms as well as recessions. These differences between the two groups explain the differences in transitions in Table 1 which we discussed above. They will also be important for our analysis of the way skilled and unskilled workers enter and exit non-employment during recessionary periods. Our estimates indicate, as emphasized by Petrongolo and Pissarides (2008) or Shimer (2012), that variations across the business cycle in separation rates are smaller than the variation in the probability of obtaining job offers.

We now turn to the returns to experience and tenure. Our parameter estimates in Table 4 correspond to the wage equation (2). As we explain in Section 3, we allow for non-linear returns to tenure and work experience, and we allow the tenure and experience profiles to vary by skill group. Notice that we start the experience and tenure clock at the start of the first job for unskilled workers and at the start of apprenticeship training for skilled workers. The wage profiles based on the raw data, and displayed in Figure 1, suggest that the returns to work experience are non linear, steepest during the first 6 years, and basically flat beyond that period. This is reflected by the estimated parameters in the table: during the first six years in the labor market, wages grow faster for non-apprentices. Over a period of 30 years of experience, the average wage gain from experience is 1.5 percent per year for unskilled workers and 1 percent for skilled workers. The lower returns to experience for the skilled is partly due to the return to experience being captured in the education effect, which is substantial (0.98 log points). The estimated return to tenure, on the other hand, are very low for both skill groups, varying between 0.1 to 0.2 percent per year.[25] These estimates represents the causal effect of an additional year on the job. However, they do not explain entirely the differential wage growth across skill groups, as skilled and unskilled workers accumulate different levels of work experience and job seniority over the years. We address this issue directly in section 5.3 below, using simulations to construct the proper counterfactual.

How are wage profiles of apprentices and non-apprentices affected by the business cycle? We address that by allowing the effect of the business between skill groups on log wages to differ (see equation (2) for details). The estimates in the table show that during upturns wages increase by about 2 percent for skilled individuals, and by about 5 percent for unskilled workers. It has long been hypothesized that productivity moves pro-cyclical (see e.g. Basu (1996)). However, our findings provide evidence of procyclical productivity, net of composition effects (induced by both observed or unobserved characteristics) due to differential participation in the labor market. We return to the effect of business cycles in more details below.

As we point out above, we allow for two dimensions of unobserved heterogeneity: first, individuals may differ in their ability to learn, which is important for the decision whether or not to enrol in apprenticeship training. Secondly, individuals may be differently productive at any level of skills accumulated. This formulation recognizes that abilities to perform in the labor market may differ from those required to acquire further education - which we believe is an important distinction in particular when modeling jobs with a high craft and manual component. We find that high ability individuals and those with lower cost of education are more likely to become apprentices. This is because the returns to choosing a skilled career is higher for high ability workers. We also find evidence that the two unobserved ability characteristics are correlated (although not strongly), where high ability individuals are also more likely to have higher training costs. Hence, the selection of individuals into the skilled track is complex, as it draws

[25] See Altonji and Shakotko (1987), Neal (1995) and Gathmann and Schoenberg (2010) who also find low return to firm tenure respectively on US and German data.

both high productivity individuals for whom the return to skilled job is higher *and* low productivity individuals who, on the other hand, have a lower cost of training. We refer the reader to the appendix Table A12 for a presentation of the results.

5.3 Wage Returns and Wage Growth by Skill Group

While in the previous section we discussed the parameters of the wage equation, we now turn to wage returns of career choices, wage growth, and its determinants. We first consider the gains from choosing apprenticeship training over the lifecycle, by comparing it to the lifetime value of joining the labor market as an unskilled worker. We then decompose the wage growth of the two career choices, by holding constant its various determinants, like human capital accumulation or job shopping.

Wage Returns What are the wage returns to choosing an apprenticeship training scheme as opposed to entering the labor market directly? To address this question, we compute the returns to training over a 40 year horizon, by simulating wage profiles for workers and by computing the net present value of earnings. We report here average treatment effects, i.e. the returns to training for the *average* worker. To compute these we allocate workers to both skill groups and compare their net present values under both scenarios.

The figures we present in Table 5 are the ratios of the net present values of earnings with, and without apprenticeship training. We compute these for two scenarios: evaluated before (column *Age 16*), and after (column *Age 19*) the training period. The former will include the apprenticeship period, and thus the foregone wages while in training. Notice that the figures we present in the table are not simply the returns to training while in work, but incorporate all differences in career paths, including non-employment spells and differences in job destruction rates. These numbers are not directly comparable to the parameters estimated in earnings functions, which are, under fairly strong assumptions, interpretable as the *internal* rates of return to training (see e.g. Willis 1986, Card (1999) or Card (2001)).[26]

[26] Among these assumptions are that education and experience profiles are log-additive, and that workers are continuously employed after labor market entry. Further, as these are *marginal* rates of returns, costs of education incurred through reducing the lifespan available for working are not considered.

The first row reports the "OLS" returns, which are calculated by comparing wage (and unemployment benefit) flows, but ignores sorting. The return to apprenticeship is close to 16 percent, or just above 5 percent per year. Evaluated before the training period, this figure is lower, about 7.2 percent. In the next row we display the average treatment effect. We now find lower returns, close to 11 percent (or 4 percent if the training period is included).

As the returns we compute include non-employment spells, the question arises how these should be evaluated. In the figures in row 2, we assign to those spells imputed unemployment benefits, which are rather generous in Germany. An alternative is to allocate zero wages to those spells.[27] As skilled individuals have a higher labor market attachment (see e.g. Figure 2), the returns now increase slightly, from 10.9 percent to 11.6 percent. Finally, in the last column we report the internal rate of return to apprenticeship training, computed at the age of 16, which is at 9.5 percent roughly in line with the ratio of net present values at age 19.[28] Although not directly comparable, our estimates are thus of a similar magnitude than the 2.5 - 4 percent returns per year of apprenticeship training obtained by Fersterer, Pischke, and Winter-Ebmer (2008), who, in a reduced form setting, instrument the length of apprenticeship training using as an instrument for the length of training information about the time to failure of firms that close down during the training period.

Decomposing Wage Growth We now turn to the components of wage growth over the life cycle. This is similar to French, Mazumder, and Taber (2006) who study wage growth for a population of young and low skilled individuals in the US in a reduced form framework. However, while with reduced form techniques, it is difficult to assess the relative magnitude of these alternative sources of wage growth, due to the endogeneity of labor supply and job to job mobility, one strength of our model is that it allows us to construct counterfactual life-cycle profiles, by comparing profiles with and without

[27]This would be more standard, and, for instance, in line with the literature that evaluates the effect of firm closure on wages (see e.g. Sullivan, LaLonde, Jacubson

[28]We have computed this by assuming zero wages when not working. Note that this rate is not the same than the return to education obtained from wage regressions, as it does not impose separability assumptions, and takes account of all career implications when choosing apprenticeship training.

returns to experience, tenure, or job mobility.

We simulate life-cycle profiles of wages and labor supply for both skilled and unskilled workers over their life cycle and report annual wage growth - conditional on working - over many periods. For apprentices, we compute the annual wage growth 5 years after enrollment in apprenticeship training, to avoid capturing the graduation effect (three years after enrollment), which is substantial (see Figure 1). For unskilled workers, we decompose wage growth for the first 5 years, and for all the subsequent years. We then assess the contribution of experience and tenure to wage growth, by simulating wages and labor market transitions when one of these components of wage growth is set to zero. A third channel of wage growth in our model is the evolution of the firm-worker match. This process follows a random walk, and conditional on staying in the same firm, the match quality is likely to rise, as negative shocks would lead to quits. To understand how important this is for wage growth, we simulate wage profiles, setting the variance of these innovation to zero. A final channel of wage growth comes from job shopping. To assess its contribution to wage growth, we simulate an economy where individuals never receive alternative offers while on the job. We assume that individuals do not anticipate any of these departure from the baseline, which means that we solve the model and the optimal decisions for the baseline parameter values. This implies that we keep individual behavior constant between scenarios, and we can therefore abstract from changes in wages because of composition effects.

We present the results of these simulations in Table 6. The baseline results in the first row of the Table show that workers who enter the labor market without further training experience strong wage growth over the first years of their careers, with wages growing at a rate of 11 percent per year. Wage growth slows down considerably after this initial period, to about 0.7 percent. For apprentices, wage growth after the first five years in the labor market is slightly higher at 1.4 percent.

In line with findings by Altonji and Shakotko (1987) and Altonji and Williams (2005), firm tenure plays a minor role for wage growth, as suggested by the estimates in the second row. Likewise, the evolution of the worker-firm match plays a negligible role, except for unskilled workers in the later part of their career. On the other hand, the effect of experience is very important, in particular for workers who enter the labor market without training. Over the first 5 years in the labor market, the annual wage growth decreases from 11 percent to only 0.8 percent if we exclude experience effects. After five years, the returns to experience are far lower for both apprentices and non-apprentices. It is perhaps unsurprising that human capital accumulation through work experience is an important driver for this group of workers, as they are more likely to learn on-the-job what apprentices learn in a more formal training environment. However, the relative magnitude of the contribution of experience to wage growth, in particular during the first half decade in the labor market, is remarkable. This is particularly so as the contribution of job shopping is far lower: Job-to-job mobility increases average annual wage growth from 2.9 percent to 3.8 percent - which is substantial, but far less than the contribution of experience.

At first sight, these relatively low returns to job shopping in the early career phase seems at odds with Figure 3, where workers who move to a new firm have on average large increases in their wages. However, for these increases to contribute to wage growth over several years, workers need to have fairly stable careers, which is not the case for young unskilled workers during their early career stage. After the first five years, mobility becomes relatively more important, with an increase in the growth rate from only 0.1 percent without mobility to 0.7 percent per year.

For apprentices, work experience plays a smaller due to their concentrated human capital accumulation during their training period, and subsequent higher entry wages after training. However, the contribution of general work experience to wage growth is notably higher for skilled than for unskilled workers after the first five years in the labor market. Job mobility plays likewise an important role in explaining wage growth, with a change in wage growth from 0.5 percent per year to 1.4 percent (which is in absolute terms higher than for unskilled workers over the same period).

Thus, the perhaps most interesting result from these decompositions is that - while job shopping contributes importantly to wage growth of young workers who enter the labor market without further training - learning through work experience is by far the most important component of their wage growth in the early career stages.

5.4 Career Effects of Recessions

Young people have most likely been the main victims of the last economic crisis, and have been most severely affected by unemployment in almost all OECD countries. One exception is Germany, where youth unemployment was only 3 percentage points above the overall unemployment rate in 2007, and where this difference has decreased to 2.5 percentage points in 2011. Moreover, Germany's youth unemployment rate has been persistently lower than that in many OECD countries over the last few decades. Some authors suspect this to be a consequence of the apprenticeship training scheme in Germany that facilitates entry into the labor market for young workers (See e.g. Ryan (2001)). But how exactly this should work, and whether these transitions may also help young workers to remain in work during a recession is altogether unclear.

Our analysis allows us to shed light on this question, and to study the effect of business cycles on the careers of skilled and unskilled workers. Furthermore, in contrast to the reduced-form literature, we are able to isolate the longer-run effects of an economic crisis on both future wages and employment prospects, and to identify the factors that may lead to differences in the way an economic shock affects career prospects of young workers. This addresses the question of whether apprenticeship type education schemes help to shield young workers from the consequences of an economic downturn on unemployment.

We explore this by simulating our model, comparing careers of workers who face two situations. First, a baseline scenario, where no recession occurs. Second, a scenario where a recession takes place either early, or later in a worker's career (we set these at 2 and 15 years of potential experience). While workers do not know ex ante when the recession occurs and for how long it will last, they have expectations that are consistent with the history of booms and recessions in Germany over the period we consider. In our simulations, a recession lasts for 3 years, which is consistent with workers' expectation, given the stochastic process described in Table A1. We then compute the differences in labor market status, work experience and firm tenure, and the difference in log wages (assuming zero wages for the unemployed) between each of these two scenarios, and the baseline scenario. The results are displayed in Figures 7 to 11. In each Figure, the period of the recession is indicated by the shaded area.

Employment, Experience and Tenure In Figure 7 we display the change in employment for the two skill groups. A recession early on in a cohort's career (left panel of Figure 7) decreases the proportion of individuals working by about 2 percent.[29] Interestingly, the effect is different for the two skill groups, with non-apprentices experiencing non-employment at a much earlier stage in the recession than apprentices. It takes both groups about 5 years after the end of the recession to return to their baseline employment. When the recession hits workers at a later career stage (after 15 years in the labor market, right panel), the effects are smaller and more more short-lived for both groups. Further, they are now larger for skilled workers.

One channel through which these employment effects lead to lasting career effects is by reducing the accumulation of human capital. We explore that in Figure 8 where we show the effects these shocks have on labor market experience. Experiencing a recession at an early stage of the career leads to a permanent decrease in human capital, in particular for unskilled workers. On average, while skilled workers loose about 0.04 years of experience, the effect for unskilled workers is twice as large. For the older cohort, the effect is smaller, and - as implied by the previous figure - the reduction in experience is more pronounced for unskilled individuals.

Besides affecting labor market experience, economic shocks may also have an effect on job mobility. On the one hand, a recession may reduce mobility, by reducing job offer arrivals while on the job; on the other hand, it may increase job mobility, by increasing job destruction rates. Both these may differ by skill groups. Our model allows for the underlying fundamental parameters to change through a recessionary period, as illustrated by the estimates in Table 3. One way to illustrate how exposure to a recession affects job mobility is to consider its effect on firm seniority, which is what we do in Figure 9. These figures show a distinctively mobility response for the two skill groups. While unskilled workers experience a decrease in their firm tenure during the recession, skilled workers face an *increase*. There are two counteracting processes at work: During recessions there are more transitions from work to non-employment, forcing workers to look for new jobs; on the other hand, those who are in work choose more often to remain

[29] This figure is consistent with the numbers reported by Burda and Hunt (2011), for the recessions that occurred during that period.

with the same firm, which increases firm tenure. While for skilled workers, the latter effect dominates, the opposite is the case for unskilled workers. After the recession, firm tenure decreases, as skilled workers start moving between firms again. It is noticeable that the effect on mobility is quite persistent, especially for skilled workers. When the recession hits older cohorts, the overall response pattern are similar. Thus, it seems that recessions *decrease* mobility for skilled workers, which may have consequences for their earnings - something we will investigate next.

Wages and Workforce Composition Figure 10 shows the effect on earnings, which we set to zero for the unemployed. For a recession striking after 2 years of potential experience, both skilled and unskilled workers suffer a loss in earnings of comparable magnitude. However, as implied by the graphs above, the reason of this drop differs across skill groups. While it is mostly the loss in human capital accumulation through a decrease in experience for the unskilled, it is the lack of accumulating search capital for the skilled. A recession leads to a prolonged decrease in earnings, especially for skilled individuals, which can last for up to 10 years. When the recession hits an older cohort (right panel), the effects are more moderate for skilled workers' earnings, but larger for unskilled workers' earnings. of a recession is similar for unskilled workers, but more moderate for skilled workers. Again, this is a consequence of the effect on job mobility, as the loss of search capital is smaller for older skilled workers. We have also evaluated the total effect of a recession, calculated as the change in the net present value of earnings over a period of 15 years, and starting from the beginning of the recession. For workers hit by a recession early on in their career, the net present value of earnings drops by about 2.3 percent for both skill groups. For a recession that hits workers at a later stage, the effect is 3 percent for unskilled and 2 percent for skilled workers.

A recession changes also the composition of the workforce - something that we have so far ignored. in Figure 11 we illustrate composition effects, by plotting the ratio of high to low productivity individuals who are in work. These figures show that the composition of the workforce in terms of workers' unobserved abilities changes indeed, with low productivity individuals being more likely to exit to non-employment. This is similar to the findings of Solon, Barsky, and Parker (1994) and Lemieux (2006), although these authors emphasize the composition bias in aggregate statistics due to the underweighting of (observed) low skilled individuals. Our focus here is different, as we condition on a population of low skilled individuals, and uncover the change in unobserved ability.

The composition effect is stronger early on in the career, with a change in the ratio of high to low productivity workers of about 2 percent. The effects are also long-lived: it takes about 6 years after the recession has ended to bring the ratio of high to low productivity workers to the pre-recessionary level. This suggests that low productivity workers in both skill groups are harder hit by a recession, and find it more difficult to get back to work, even years after the recession has ended. The magnitude of this selection decrease with the age of the cohort exposed to the recession. This change in composition of the workforce tends to moderate the procyclicality of wages, a phenomenon that has been described in aggregate data (see for instance Stock and Watson (1999)).

What Is the Effect of a Recession? What exactly is the effect of a recession on e.g. workers' earnings? Studies typically compare the earnings of workers who loose their job in a recession with the earnings of workers who do not. While this is certainly an interesting parameter, it does not however answer the question of earnings losses of workers who loose their job *because* of a recession, neither does it answer the question of the effect on earnings losses of workers who loose their job during a recession, compared to the case when no recession hits. This distinction is important to isolate the effect of the recession on outcomes, but difficult to implement in a reduced-form analysis, even for short-run effects. The reason is the two counterfactuals needed to compute these two parameters are not observed. The group of workers who loose their job during a recession includes also workers who may have lost their job anyway; further, wage paths of those who have not lost their job during a recession may nevertheless be affected by the recession, for instance through reduced mobility.

Our model allows us to construct these counterfactuals. We are thus able to disentangle the effect of a recession on those who loose their job *because* of the recession, and those who do not. To implement this, we identify those individuals who loose their job during the recession, but would not have lost their job in the non-recessionary baseline scenario. We also identify those who have not lost their job because of the recession.

Note that in the group of individuals that has not been laid off because of the recession, some individuals may loose their jobs, both in the baseline and in the counterfactual scenario, but not in both. We calculate for these two groups of workers the net present value of their earnings for the baseline scenario, and the recession scenario, and compare the effects. As before, we consider the period from the start of the recession until 15 years after the recession.

Focussing on the effect on earnings, this reveals considerable heterogeneity within skill groups due to the fact that some - but not all - loose their job because of the recession. When the recession hits workers at an early career stage, those who loose their job because of the recession suffer a loss of 23 percent in discounted life-time earnings, where the loss is similar across skill groups. Most interestingly, also those workers who do not loose their job because of the recession, forego about 1 to 2 percent in net present value. The reason is that these workers loose also search capital because of reduced job to job mobility. The latter is especially important for skilled workers who are employed throughout the recession.

Our results therefore illustrate the difficulty of estimating the effect a recession has on workers' wage careers, due to the difficulty to define an appropriate control group. As we demonstrate here, also those workers who do not loose their job in a recession are affected. A second issue, and which we highlight above, is the selection of workers based on their unobserved productivity. Our analysis is able to overcome both issues by using simulations to construct a proper counterfactual.

6 Conclusion

How individuals' careers interact with economic shocks has attracted renewed interest since the great recession. Any analysis of this subject is inherently complex, as recessions impact on the various parameters that govern the career process of workers simultaneously. Thus, a thorough assessment requires - as a first step - to model and to estimate the fundamental mechanisms that drive career paths of workers, which is what we do in this paper. We base our empirical analysis on rare administrative data for Germany, a country that has attracted attention for its performance throughout the last recession.

A distinctive feature of Germany's labor market is the structured vocational training scheme that trains about 65% of each cohort, and which has sometimes being credited for the performance of the German labor market. Thus, besides modeling the career paths of workers and how they interact with economic shocks, we also model the initial choice of individuals whether or not to enroll into an apprenticeship training scheme at labor market entry, and allow for differences between skilled and unskilled workers.

Our results are interesting in many ways. Not only do we find that skilled and unskilled workers have different career paths, but they respond to economic shocks in very different ways. This is related to the way the underlying parameters of the process of human capital accumulation and job mobility differ between the two groups, and are affected by economic shocks. After conditioning on unobserved heterogeneity (where we allow individuals to differ in terms of productivity and their ability to learn), we find that, although vocational training within an apprenticeship scheme offers a higher return, this additional return is quite modest, and corresponds to less than 4 percent per year of training. One reason is that workers who do not enroll in apprenticeship training experience rapid learning on the job during the first years in the labor market. A most interesting finding is that - although job shopping is important for early wage growth for unskilled workers - far more important for these workers is on the job learning.

Another distinctive feature between the two groups is the way they respond to economic downturns, in particular when the recession hits workers at an early stage in their career. While unskilled workers are more likely to transit to non-employment, and suffer larger losses to human capital than skilled workers, they do not experience larger wage losses. The reason is that, during a recession, skilled workers loose out on search capital, as they tend to remain with the same employer. What this illustrates is that the two drivers of wage growth, learning-by-doing and job mobility, are differently affected by economic shocks for skilled and unskilled workers.

One important aspect of our analysis is that it is performed for an country where the vast majority of young people enrolls in apprenticeship training. Thus, young workers who enter the labor market directly, and without further training are exposed to a work environment where a large fraction of co-workers are well trained, having obtained training within a structured three-years apprenticeship scheme, which creates a fertile learning environment. This may be the reason not only for why learning, as opposed to job shopping, plays such an important role for these young workers during their first years in the labor force, but also why the returns to apprenticeship training are relatively low. Therefore, as the learning environment at the workplace changes with the average skill levels of one's peers, so may the returns to experience. In consequence, while the returns to enrollment into a three-year structured apprenticeship training scheme may seem relatively modest in Germany, they may be far larger in countries where only a small fraction of workers has received structured job training, like for instance in the UK. We believe however that the fundamental differences in the way young skilled and unskilled workers respond to economic shocks which our analysis uncovers are likely to generalize to other economic environments.

One important insight our analysis provides is the difficulty of precise assessment of recessions on the careers of workers, which is intrinsically related to the complexity of recessionary effects on individuals' careers. Not only will an economic shock lead to responses through a variety of different channels (like job experience and learning, job shopping, or innovations while in work), but it also changes the composition of the workforce with respect to unobservable characteristics, and it affects all workers, including those individuals who do not loose their jobs as a direct consequence of the recession. While addressing the first two issues is challenging in a reduced form context, identifying appropriate counterfactual scenarios is even more difficult. Our analysis emphasizes that it is important to precisely define which effects are identified in any analysis of the consequences of a recession, and demonstrates the strength of our structural approach not only in isolating the direct long term consequences of an economic shock on individuals' careers, but also its ability to estimate different parameters of interest, by being in the position to create different counterfactual situations.

References

ALTONJI, J., AND R. SHAKOTKO (1987): "Do Wages Rise with Job Seniority?," in *Unemployment, trade unions, and dispute resolution*, ed. by O. Ashenfelter, and K. Hallock, vol. 47, pp. 219–241. International Library of Critical Writings in Economics.

ALTONJI, J., AND N. WILLIAMS (1998): "The Effects of Labor Market Experience, Job Seniority and Mobility on Wage Growth," *Research in Labor Economics*, 17, 233–276.

——— (2005): "Do Wages Rise with Job Seniority? A Reassessment," *Industrial and Labor Relations Review*, pp. 370–397.

ALTONJI, J. G., A. SMITH, AND I. VIDANGOS (2009): "Modeling Earnings Dynamics," NBER Working Papers 14743.

BARLEVY, G. (2002): "The sullying effect of recessions," *Review of Economic Studies*, 69(1), 65–96.

BASU, S. (1996): "Cyclical productivity: Increasing returns or cyclical utilization?," *Quarterly Journal of Economics*, 111, 719–751.

BECKER, G. (1994): *Human Capital*. The University of Chicago Press, 3rd edn.

BEN-PORATH, Y. (1967): "The production of human capital and the life cycle of earnings," *Journal of Political Economy*, 75, 352–365.

BENDER, S., C. DUSTMANN, D. MARGOLIS, AND C. MEGHIR (2002): "Worker Displacement in France and Germany," in *Losing Work, Moving on: International Perspectives on Worker Displacement*. Peter J. Kuhn,, Michigan, Upjohn Institute for Employment Research, Kalamazoo.

BURDA, M. C., AND J. HUNT (2011): "What Explains the German Labor Market Miracle in the Great Recession?," *Brookings Papers on Economic Activity*, pp. 273–335.

Cameron, S. V., and J. J. Heckman (1998): “Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males,” *Journal of Political Economy*, 106(2), 262–333.

Card, D. (1999): “The causal effect of education on earnings,” in *Handbook of Labor Economics*, ed. by O. C. Ashenfelter, and D. Card, vol. 3, pp. 1801–1863. Elsevier.

——— (2001): “Estimating the Returns to Schooling: Progress on some Persistent Econometric Problems,” *Econometrica*, 69, 1127–1160.

Carneiro, P., J. Heckman, and E. Vytlacil (2006): “Estimating Marginal and Average Returns to Education,” mimeo UCL.

Dahl, G. B. (2002): “Mobility and the Return to Education: Testing a Roy Model with Multiple Markets,” *Econometrica*, 70(6), 2367–2420.

Davis, S. J., and J. Haltiwanger (1992): “Gross Job Creation, Gross Job Destruction, and Employment Reallocation,” *Quarterly Journal of Economics*, 107(3), 819–863.

Davis, S. J., and T. von Wachter (2011): “Recessions and the Costs of Job Loss,” *Brookings Papers on Economic Activity.*

Dustmann, C. (2004): “Parental background, secondary school track choice, and wages,” *Oxford Economic Papers*, pp. 209–230.

Dustmann, C., and C. Meghir (2005): “Wages, experience and seniority,” *Review of Economic Studies*, 72(1).

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.

Fersterer, J., and R. Winter-Ebmer (2003): “Are Austrian returns to education falling over time?,” *Labour Economics*, 10(1), 73–89.

Fitzenberger, B., and A. Kunze (2005): "Vocational Training and Gender: Wages and Occupational Mobility among young Workers," *Oxford Review of Economic Policy*, 21(3), 392–415.

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

French, E., B. Mazumder, and C. Taber (2006): "The Changing Pattern of Wage Growth for Low Skilled Workers," in *Working and Poor: How Economic and Policy Changes Are Affecting Low-Wage Workers*, ed. by Blank, Danziger, and Shoeni.

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

Gladden, T., and C. Taber (2000): "Wage Progression Among Low Skilled Workers," in *Finding Jobs: Work and Welfare Reform*, ed. by Card, and Blank.

——— (2009a): "The relationship between wage growth and wage levels," *Journal of Applied Econometrics*, 24(6), 914–932.

——— (2009b): "Turnover and Wage Growth for Low Skilled Young Men," .

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

Griliches, Z. (1971): "Estimating the Returns to Schooling: Some Econometric Problems," 45(1), 1–22.

Heckman, J. (1993): "Assessing Clinton's program on job training, workfare, and education in the workplace," Discussion paper, NBER Working Paper 4428.

Heckman, J., and G. Sedlacec (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., and B. Singer (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52(2), 271–320.

Heckman, J., and E. Vytlacil (2005): "'Structural equations, treatment effects and econometric policy evaluation', Fisher-Schultz Lecture," 73(3), 669–738.

Kennan, J., and J. R. Walker (2010): "The Effect of Expected Income on Individual Migration Decisions," forthcoming Econometrica.

Lemieux, T. (2006): "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?," *The American Economic Review*, 96(3), 461–498.

Lise, J., C. Meghir, and J.-M. Robin (2009): "Matching, Sorting and Wages," mimeo, UCL.

Low, H., C. Meghir, and L. Pistaferri (2009): "Wage Risk and Employment Risk over the Life Cycle," Forthcoming American Economic Review.

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

Meghir, C., and L. Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72, 1–32.

Mincer, J., and B. Jovanovic (1981): "Labor Mobility and Wages," in *Studies in Labor Markets*, ed. by S. Rosen. University of Chicago Press.

Mortensen, D., and C. Pissarides (1994): "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, 61(3), 397–415.

Nagypal, E. (2005): "Worker Reallocation over the Business Cycle: The Importance of Job-to-Job Transitions," mimeo, Northwestern University.

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

Oddbjorn, R., and K. Roed (2006): “Do Business Cycle Conditions at the Time of Labour Market Entry Affect Future Employment Prospects?,” *The Review of Economics and Statistics*, 88(2), 193–210.

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.

Parent, D. (1999): “Wages and Mobility: The Impact of EmployerProvided Training,” *Journal of Labor Economics*, 17(2), 298–317.

Parey, M. (2009): “Vocational Schooling versus Apprenticeship Training. Evidence from Vacancy Data,” mimeo, University of Essex.

Pavan, R. (2011): “Career Choice and Wage Growth,” *Journal of Labor Economics*, 29(3), 549–587.

Petrongolo, B., and C. Pissarides (2008): “The ins and outs of European unemployment,” *American Economic Review*, 98, 256–262.

Rosen, S. (1972): “Learning and Experience in the Labor Market,” *Journal of Human Resources*, 7, 326–342.

Rosen, S. (1976): “A Theory of Life Earnin,” *Journal of Political Economy*, 84, S45–.

Ryan, P. (2001): “The School-to-Work Transition: A Cross-National Perspective,” *Journal of Economic Literature*, 39, 34–92.

Schmieder, J. F., and T. von Wachter (2010): “Does Wage Persistence Matter for Employment Fluctuations? Evidence from Displaced Workers,” *American Economic Journal: Applied Economics*, 2(3), 1–21.

Shimer, R. (2012): “Reassessing the ins and outs of unemployment,” *Review of Economic Dynamics*, 15(2), 127–148.

Solon, G., R. Barsky, and J. A. Parker (1994): “Measuring the Cyclicality of Real Wages: How Important is Composition Bias,” *The Quarterly Journal of Economics*, 109(1), 1–25.

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

Stock, J., and M. Watson (1999): “Business cycle fluctuations in US macroeconomic time series,” in *Handbook of macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1, pp. 3–64. Elsevier.

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

Taber, C. (2001): “The Rising College Premium in the Eighties: Return to College or Return to Unobserved Ability?,” *Review of Economic Studies*, 68(3), 665–691.

Tauchen, G., and R. Hussey (1991): “Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models,” *Econometrica*, 59, 371–396.

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.

von Wachter, T., and S. Bender (2006): “In the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers’ Careers,” *American Economic Review*, 96, 1679–1705.

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.

Wolpin, K. I. (1992): "The Determinants of Black-White Differences in Early Employment Careers: Search, Layoffs,Quits, and Endogenous Wage Growth," *Journal of Political Economy*, 100(3), 835–560.

Table 1: Observed Quarterly Labor Market Transitions

	Potential Experience (Years)					
Labor Market Transitions	Unskilled			Skilled		
	0-5	5-10	10-20	0-5	5-10	10-20
Out of work to Out of work	.88	.92	.95	.73	.81	.92
Out of Work to Work	.12	.07	.05	.27	.19	.08
Work to out of Work	.09	.05	.03	.03	.04	.02
Work to new Work	.04	.03	.02	.02	.04	.03
Work to same Work	.87	.92	.94	.95	.92	.95

Notes: results derived from IAB data, 1975-2004, aggregated at a quarterly frequency.

Table 2: Proportion in different education tracks by Year of birth

	Birth Cohorts		
	1960	1965	1970
Academic Track	20%	21%	24%
Apprentices	64%	67%	65%
Non Apprentices	16%	12%	11%

Notes: results derived from IAB data, 1975-2004.

Table 3: Estimated parameters: variance of shocks, job destruction and job arrival rates and mobility costs

Parameter	In Apprenticeship	Qualified Apprentices	Non-Apprentices
Std dev initial match specific effect (σ_0)	0.264	0.249	0.228
	(0.013)	(0.0066)	(0.052)
Std dev innovation to match specific effect (σ_u)	0.023	0.0131	0.0251
	(0.041)	(0.00662)	(0.019)
Job offers and job destruction rates			
Quarterly job destruction rate (δ)			
if experience $\leq$ 4 years	-	0.0252	0.0792
		(8.4e-06)	(0.00073)
if experience $\in$ [4,6] years	-	0.04	0.053
		(1.2e-05)	(0.033)
if experience $>$ 6 years	-	0.019	0.03
		(1.1e-06)	(6.4e-06)
additional effect if business cycle high	-	-0.00249	-0.00353
		(0.00065)	(3.6e-05)
Quarterly offer arrival rate when employed (π_W)			
if business cycle low, experience <6	0.0448	0.485	0.738
	(0.0014)	(0.0021)	(8.3e-07)
if business cycle high, experience<6	0.471	0.912	1
	(0.28)	(0.28)	-
if business cycle low, experience $\geq$6	-	0.498	0.324
	-	(0.023)	(0.055)
if business cycle high, experience$\geq$6	-	0.924	0.636
	-	(0.28)	(0.055)
Quarterly offer arrival rate when unemployed (π_U)			
if business cycle low, experience=0	-	0.137	0.182
		(5.2e-05)	(0.00025)
if business cycle high, experience=0	-	0.16	0.192
		(7.8e-05)	(0.00073)
if business cycle low, experience=10	-	0.208	0.5
		(5.8e-05)	(0.0016)
if business cycle high, experience=10	-	0.231	0.51
		(6.4e-05)	(0.0016)

Note: ^{a}: as a percentage of lifetime value at age 16. Asymptotic standard errors in parenthesis. Utility of leisure and the standard deviation of mobility costs have been restricted to be common across all skill groups.

Table 4: Estimated parameters: wage equations

Parameter	Apprentices		Non-Apprentices	
Log Wage Constant	3.83	(0.016)	3.68	(0.052)
In apprenticeship indicator	-0.98	(0.02)	-	
Experience=0 yrs	0		0	
Experience=2 yrs	0.0063	(0.015)	0.31	(0.03)
Experience=4 yrs	0.25	(0.017)	0.46	(0.028)
Experience=6 yrs	0.28	(0.018)	0.46	(0.063)
Experience=10 yrs	0.31	(0.021)	0.46	(0.052)
Experience=30 yrs	0.32	(0.037)	0.46	(0.095)
Tenure=0 yrs	0	-	0	-
Tenure=2 yrs	0.00011	(0.012)	0.02	(0.029)
Tenure=4 yrs	0.0099	(0.012)	0.026	(0.033)
Tenure=6 yrs	0.02	(0.011)	0.044	(0.055)
Tenure=20 yrs	0.042	(0.045)	0.067	(0.16)
Effect of high business cycle	0.0169	(0.0043)	0.0528	(0.02)

Log wage is the dependent variable. The wage equation for apprentices during and following training differ only in the indicator for apprenticeship training (and the variance of the shocks). Asymptotic standard errors in parenthesis.

Table 5: The Life-cycle returns to apprenticeship

	Ratio of net present values	
	Age 16	Age 19
OLS	7.2%	15.7%
Model	3.8%	10.9%
Model, excl. UI benefits	5.3%	11.6%
Model, conditional on working	1.9%	10.2%

Note: Returns calculated over a horizon of 40 years, and with a discount factor set at 0.95 annually. The returns displayed are average treatment effects.

Table 6: Annual wage growth, by skill levels

	Non Apprentices			Apprentices
Potential Experience	0-20	0-5	5-20	5-20
Baseline	3.8%	11%	0.71%	1.4%
No return to tenure	3.5%	10%	0.64%	1.4%
No evolution of firm-worker match	3.6%	11%	0.58%	1.4%
No return to experience	0.68%	0.83%	0.55%	1%
No job-to-job mobility	2.9%	9.6%	0.14%	0.5%

Note: Annual wage growth, conditional on working, calculated by simulating the model over a horizon of 20 years.

Figure 1: Log Wage by skill and the wage gain for qualified apprentices

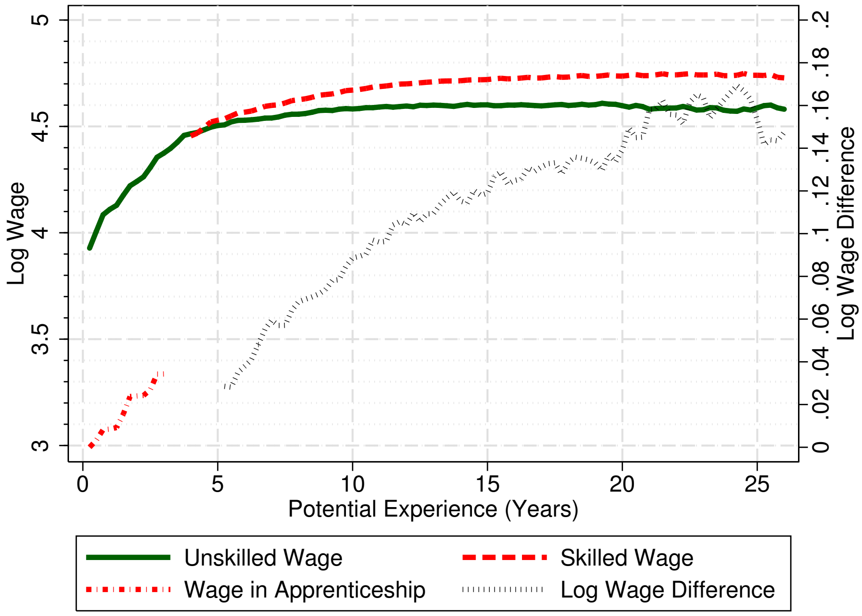


Figure 2: Proportion Working by skill

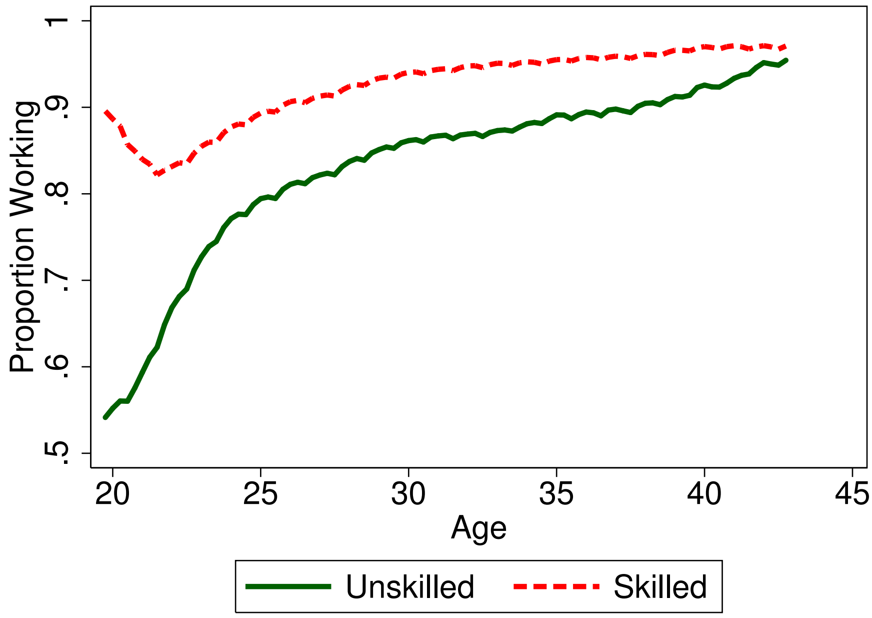


Figure 3: Annual Change in Log Wage

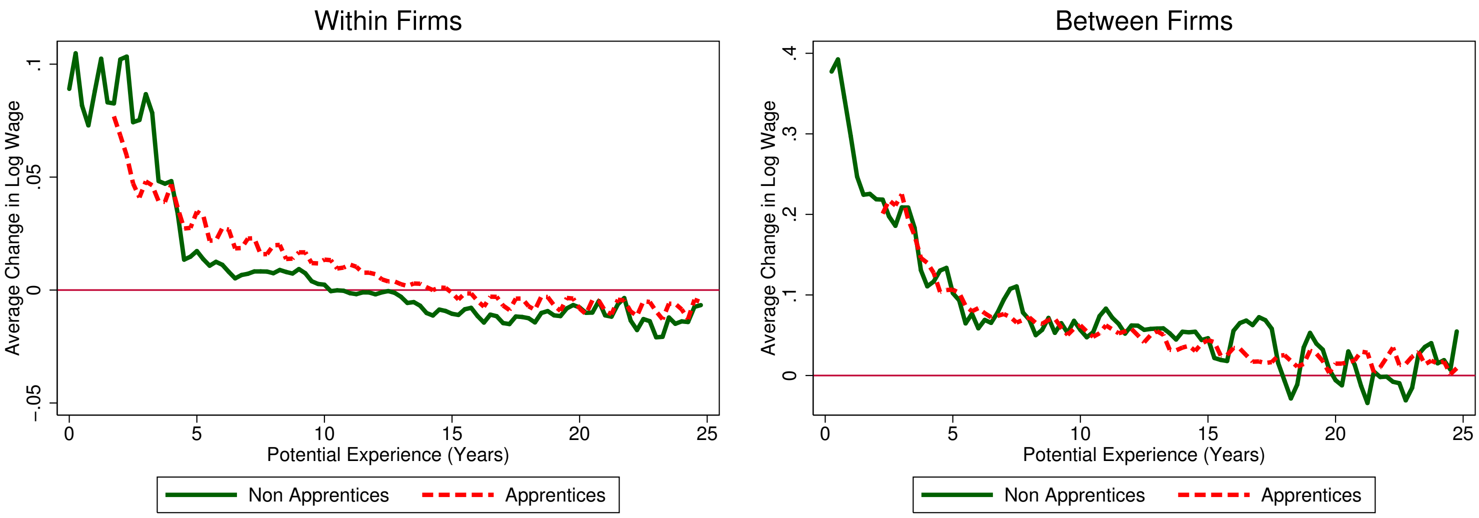


Figure 4: Mobility: Number of Jobs, by Education

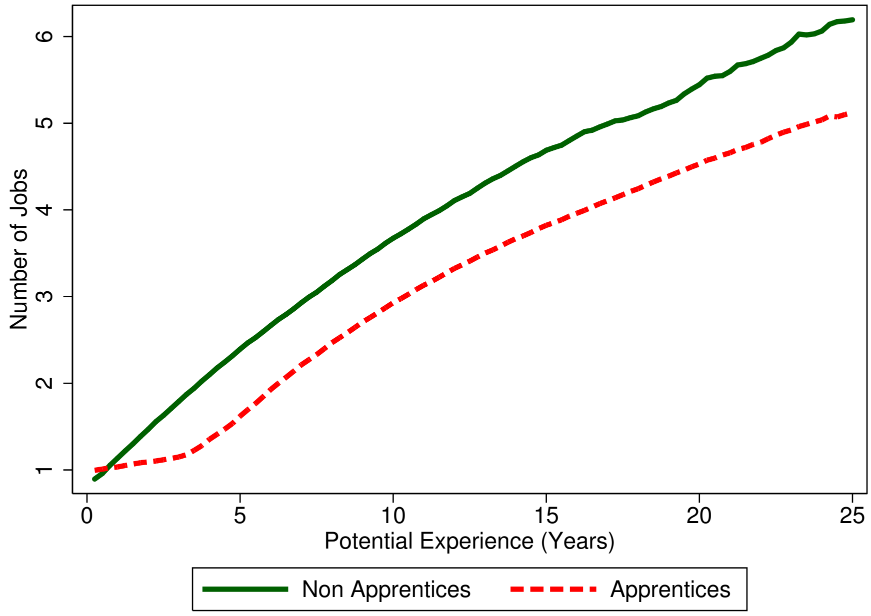
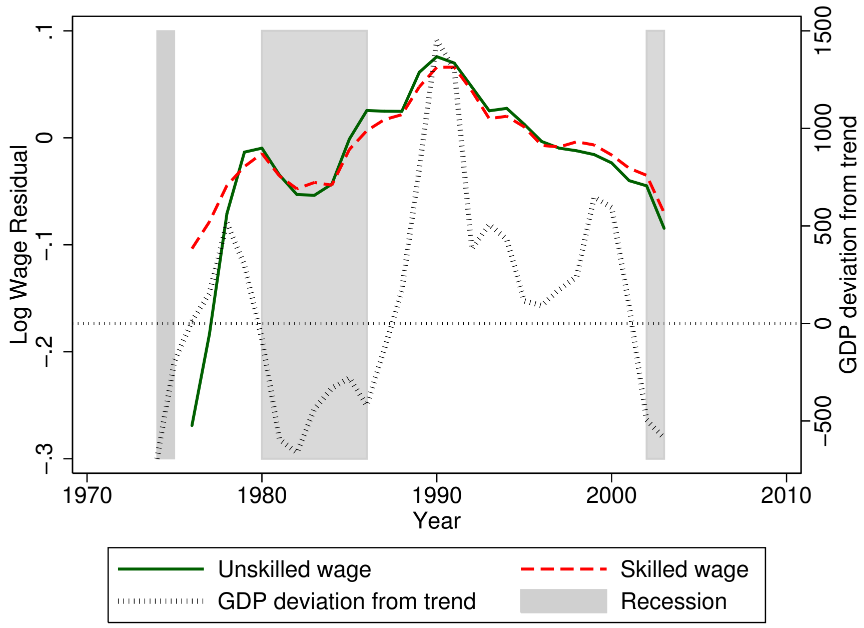


Figure 5: Wages and Business Cycles



Note: GDP per capita US $, constant prices obtained from OECD. Wage data derived from IAB sample. Residual log wages are obtained by projecting log wages on potential experience and time trend.

Figure 6: Observed and Predicted Employment, Wage and Standard deviation of Wage Profiles

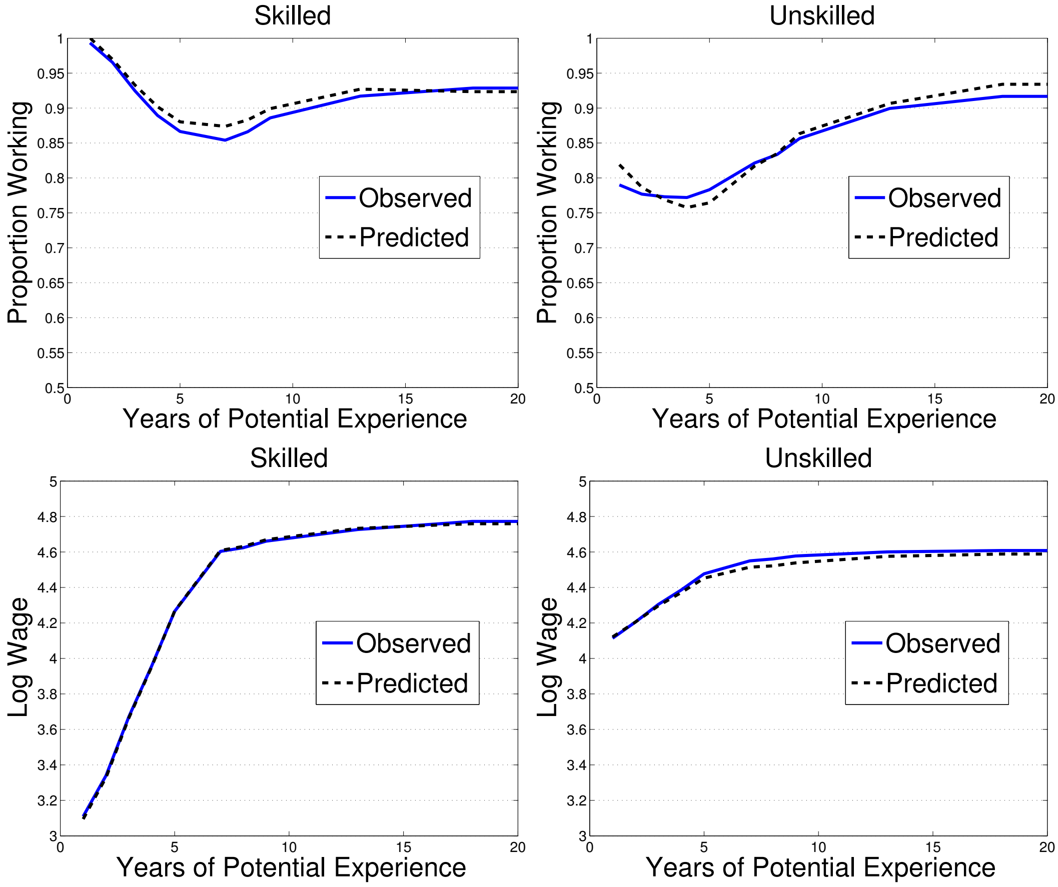


Figure 7: Change in Employment Following a Recession

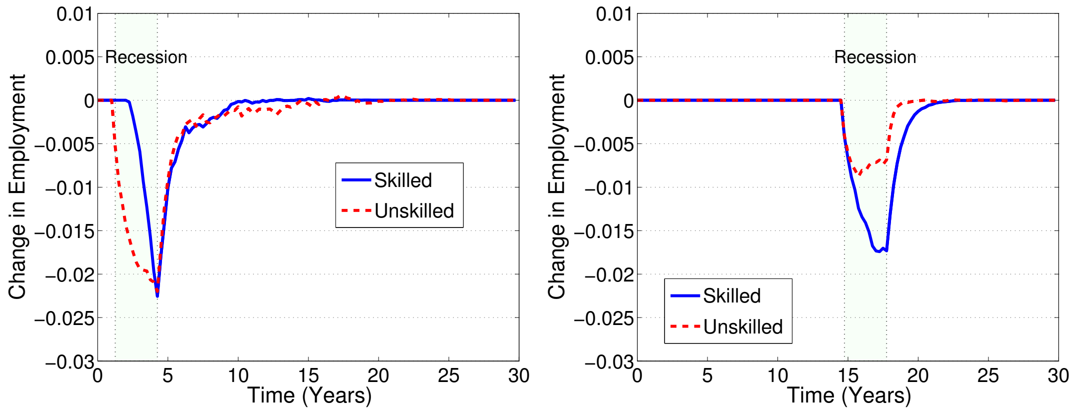


Figure 8: Change in Experience Following a Recession

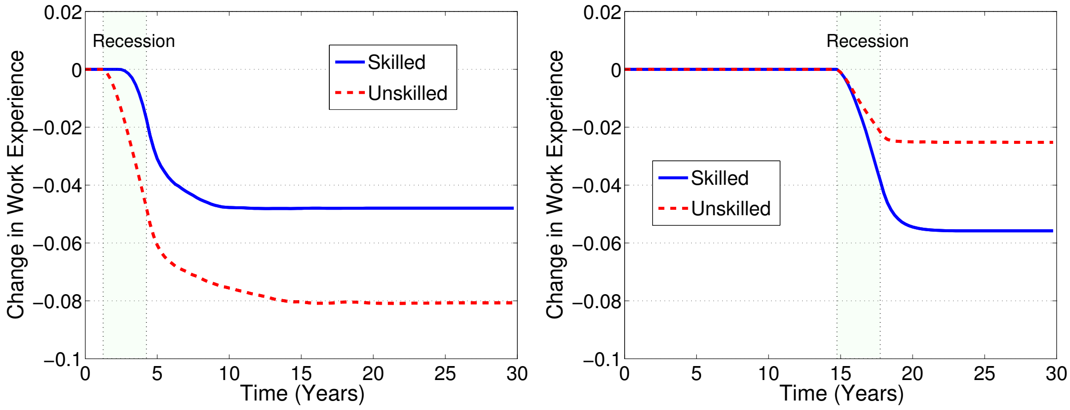


Figure 9: Change in Firm Tenure Following a Recession

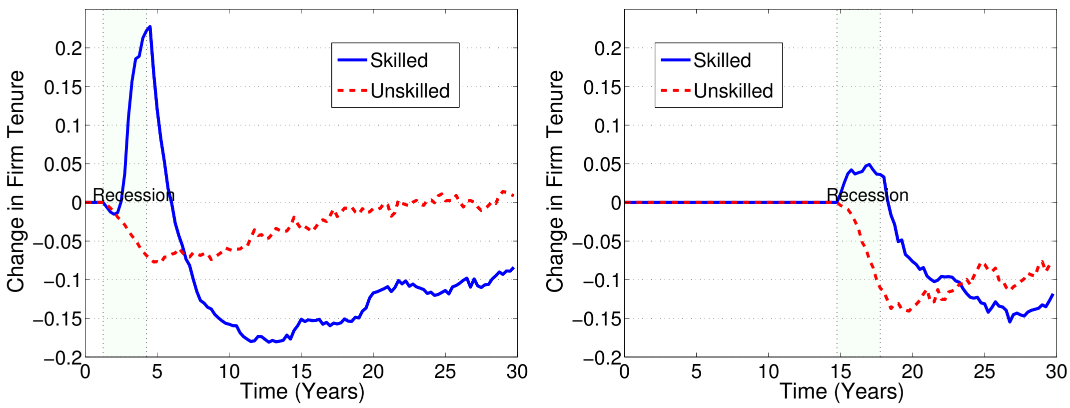
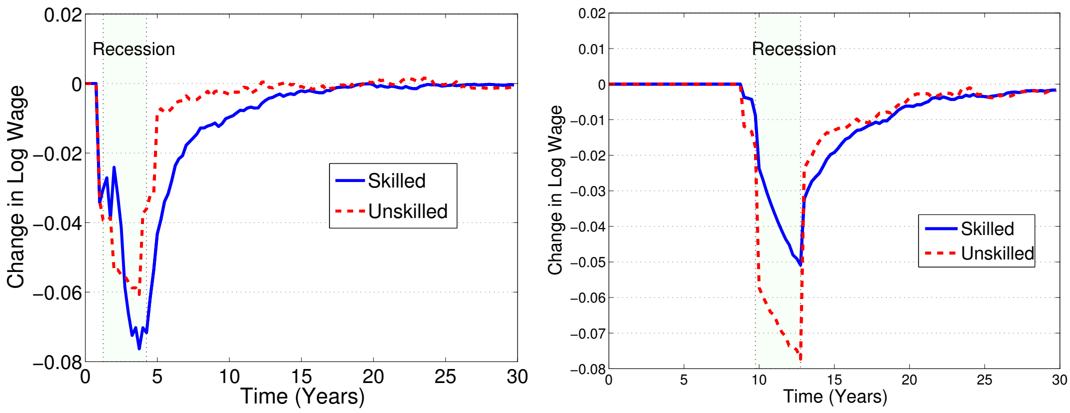
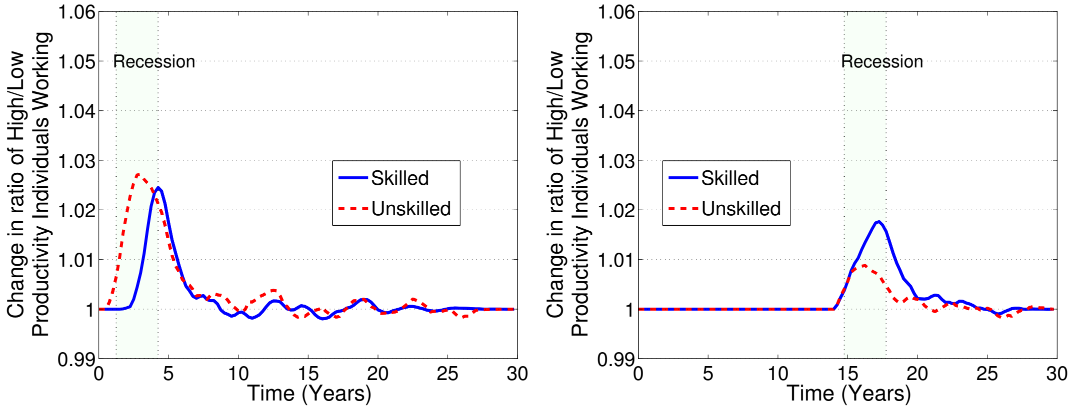


Figure 10: Change in Earnings Following a Recession[a]



[a] *Note:* The figures display earnings, including zero income when not working.

Figure 11: Change in Composition of Workers, along Unobserved Productivity Following a Recession[a]



[a] *Note:* The figures display the difference in the ratio of high versus low productivity individuals who are working. The comparison is between two paths, without and with a recession.

Appendix

A Data: Sample Used for Estimation

We select all male individuals who are born between 1960 and 1972. Thus, we make sure that no individual is older than 15 in 1975 (the minimum age at which post-secondary labor market entry is possible), which is the first year of our data. We consider all years between 1975 and 2004. We exclude all individuals who live in East-Germany. We drop individuals who work in the agricultural industry, and individuals who work in the family businesses. We restrict our sample to those who are not older than 23 when they enter the labor market the first time, and who enter the labor market with only a lower secondary school education, who either enrol into apprenticeship training directly, or who enter the labor market without further training.[30] [31] We further 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 (which we set to 1600 days). We restrict our analysis to individuals with German citizenship, as individuals with non-German citizenship may 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 the period for which the individual has been in employment. Thus, every wage we observe belongs to one particular worker-firm spell. We compute real wages in 1995 prices.

The precise distinction between individuals who enrol in a traditional apprenticeship

[30]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 whether to enter one of three secondary school branches: lower secondary school (which takes another 5-6 years), intermediate secondary school (which takes another 6 years), and higher secondary school (which takes another 9 years). For our analysis, we concentrate on individuals who choose lower or intermediate secondary school. These two options do not allow for direct access to university, and individuals typically enrol into apprenticeship training, or enter the labor market directly.

[31]As the comparison group of individuals who choose upper track secondary school, which we use to implement our selection correction, we define 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.

scheme, and individuals who enter the labor market without that training, is as follows. We define as "apprentices" all those individuals who entered the labor market with a lower or intermediate secondary school degree, and who can be observed after entry on an apprenticeship training scheme for at least 24 months, and who transit to a "skilled" status afterwards.[32] We define as "non-apprentices" all those individuals who enter the labor market without further training, or who have been on an apprenticeship training schemes for less than 7 months, without obtaining a degree (i.e. dropouts). This group may include individuals who enrolled in one-year vocational courses before entering the labor market – preparatory courses that do not lead to vocational degrees. Thus, among our non-apprentices may be individuals who did receive some preparatory training.

Another mode of entry, as discussed in Parey (2009), is attendance of 2-3 year vocational schools, which provide vocational training with unpaid work experience in specialised schools for a limited number of occupations.[33] These occupations are mainly in female-dominated occupation groups, like caring and health-related occupations. In our sample, these individuals constitute about 6% of individuals.[34] In line with Parey (2009), we find that the wage paths of this group are very similar to those of individuals undergoing firm-based training, and higher than those of individuals entering the labor market without further training. We also find that they are experiencing lower employment probabilities than apprentices. The way we deal with these individuals is to include them among our apprentices, assuming that the choice to undergo training at a full time school rather than within the firm is equivalent to choosing apprenticeship training in a firm.

B Model and Numerical Solution

B.1 The value of unemployment.

The value of unemployment consists of a predetermined part and a stochastic shock η_{it} reflecting changes in the utility of being out of work. Denoting the predetermined part

[32] For apprentices who finish their training within a calendar without changing firms, we do not observe the date of graduation, neither can we distinguish the apprenticeship wage during that year from the skilled worker wage. To compute the number of apprenticeship training months, we assign to these individual 6 months of training. Further, when we compute wages after the apprenticeship period, we discard these observations.

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

[34] The size of this group is smaller than in Parey (2009). One reason for this is that we consider only the years up to 1996, where these school based vocational schemes were less frequent than in later years.

by $U_a\left(Ed_i, G_t, X_{it}, w_{i(-1)}, \varepsilon_i\right)$, where the subscript a denotes the age of the individual, we can write

$$
\begin{aligned}
&U_a\left(Ed_i, G_t, X_{it}, w_{i(-1)}, \varepsilon_i\right) = \log(\gamma_U w_{i(-1)}) + \gamma_0(\varepsilon_{iP}) + \gamma_X(X_{it}, Ed_i) && A \\
&+\beta\pi_{it}^U\ \mathbb{E}\max\begin{pmatrix} \underline{\mu_{if}} + W_{a+1}\left(Ed_i, \underline{G_{t+1}}, X_{it}, T_{ift+1}=0, \underline{\kappa_{if}^0}, \varepsilon_i\right) \\ U_{a+1}\left(Ed_i, \underline{G_{t+1}}, X_{it}, w_{i(-1)}, \varepsilon_i\right) + \underline{\eta_{it+1}} \end{pmatrix} && B \\
&+\beta(1-\pi_{it}^U)\ \mathbb{E}U_{a+1}\left(Ed_i, \underline{G_{t+1}}, X_{it}, w_{i(-1)}, \varepsilon_i\right) && C
\end{aligned}
\tag{B1}
$$

where we underline the variables over which we are taking expectations (because they are unknown to the individual in period t) and where β is the discount factor.

In (B1) the first line of the right hand side (A) represents the within period value of being out of work (up to the stochastic shock η_{it}). This consists of the unemployment insurance income plus a value for leisure. The lines denoted by (B) represent the expected future value for the case where the worker gets a job offer, which happens with probability π_{it}^U. In that case the worker will choose the best of taking the job offer or continuing as an unemployed worker. The value of taking the job offer is equal to the sum of the present value of the future flow of earnings defined below, $W_{a+1}(\cdot)$, plus a (stochastic) amenity μ_{if}. The final line (C) represents the case where the individual obtains no offer and thus just has to continue out of work.

B.2 The value of employment.

Their value of employment is then given by

$$
\begin{aligned}
&W_a\left(Ed_i, G_t, X_{it}, T_{ift}, \kappa_{ift}, \varepsilon_i\right) = \log(w_{it}) && A \\
&+\beta\delta_{it}\ \mathbb{E}\left[U_{a+1}\left(Ed_i, \underline{G_{t+1}}, X_{it}+1, w_{it}\right) + \underline{\eta_{it+1}}\right] && B \\
&+\beta\left(1-\delta_{it}\right)\pi_{it}^W\ \mathbb{E}\max\begin{pmatrix} U_{a+1}\left(Ed_i, \underline{G_{t+1}}, X_{it}+1, w_{it}, \varepsilon_i\right) + \underline{\eta_{it+1}} \\ W_{a+1}\left(Ed_i, \underline{G_{t+1}}, X_{it}+1, T_{ift}+1, \kappa_{ift} + \underline{u_{ift+1}}, \varepsilon_i\right) \\ \underline{\mu_{i\tilde{f}}} + W_{a+1}\left(Ed_i, \underline{G_{t+1}}, X_{it}+1, T_{i\tilde{f}t+1}=0, \underline{\kappa_{i\tilde{f}}^0}, \varepsilon_i\right) \end{pmatrix} && C \\
&+\beta(1-\delta_{it})(1-\pi_{it}^W)\ \mathbb{E}\max\begin{pmatrix} U_{a+1}\left(Ed_i, \underline{G_{t+1}}, X_{it}+1, w_{it}, \varepsilon_i\right) + \underline{\eta_{it+1}} \\ W_{a+1}\left(Ed_i, \underline{G_{t+1}}, X_{it}+1, T_{ift}+1, \kappa_{ift} + \underline{u_{ift+1}}, \varepsilon_i\right) \end{pmatrix} && D
\end{aligned}
\tag{B2}
$$

The current value of work is just the wages w_{it}. Following job destruction, which occurs with probability δ_{it} the individual will receive the value of unemployment as shown in line B. The group of lines marked C represent the events when the job is not destroyed and the individual obtains an alternative job offer. In this case they have to choose between becoming unemployed; remaining with the firm; or taking the alternative offer, which is associated with the one off random switching cost $\underline{\mu_{i\tilde{f}}}$ of joining a new firm $\tilde{f}$. The following group of lines marked by D represent the expected value of a worker not being laid off and not having access to an alternative offer. Given that a shock can occur to the match specific effect, the worker may decide it is best to quit, in which case they receive the value of unemployment. Otherwise they receive the value of working with the same firm, at the updated wage.

B.3 The value of employment while in training.

Going back, earlier into the individual's history, we consider choices available when training. During apprenticeship (which lasts τ_A periods[35]) we assume that the training firm pays the worker only a fraction λ_A of his productivity as a non-apprentice $(w\,(Ed_i = 0, G_t, X_{it}, T_{it}, \kappa_{it}, \varepsilon_i))$, the rest presumably serving as payment for the general training received.[36] Reflecting the facts in the data, we do not allow the individual to experience unemployment during apprenticeship, although they can decide to change firm if the opportunity arises. Thus, during the apprenticeship training period $(X_{it} < \tau^A)$ the value of work is:

$$
\begin{aligned}
&W_a^A\left(G_t, X_{it}, T_{ift}, \kappa_{ift}, \varepsilon_i\right) = \log(\lambda_A \cdot w\left(Ed_i = 0, G_t, X_{it}, T_{ift}, \kappa_{ift}, \varepsilon_{iP}\right)) \quad A \\
&+\beta\pi_A(G_t)\ \mathbb{E}\max\begin{pmatrix} W_{a+1}^A\left(\underline{G_{t+1}}, X_{it}+1, T_{ift}+1, \kappa_{ift}+\underline{u_{ift+1}}, \varepsilon_i\right) \\ \underline{\mu_{i\tilde{f}}} + W_{a+1}^A\left(G_{t+1}, X_{it+1}, T_{i\tilde{f}t+1}=0, \underline{\kappa^0_{i\tilde{f}}}, \varepsilon_i\right)\end{pmatrix} \quad B \qquad \text{(B3)} \\
&+\beta[1-\pi_A(G_t)]\ \mathbb{E}W_{a+1}^A\left(\underline{G_{t+1}}, X_{it}+1, T_{ift+1}, \kappa_{ift}+\underline{u_{ift+1}}, \varepsilon_i\right) \quad C
\end{aligned}
$$

where as before, the expectation operator $\mathbb{E}$ relates to the underlined variables, which are unknown to the individual in period t.

Similarly to the value of working described above, the first line (A) is earnings while training, (B) represents the part of the value due to the possibility of changing training

[35] Apprenticeship courses last between two and three years. We equate τ_A to whatever is the actual duration in the data.

[36] In actual fact this is only part payment towards the general training: at least the classroom component is funded by the government.

firms if an offer arrives (with probability π_A). As before there is a mobility cost associated with the decision to join the alternative firm $\widetilde{f}$. Finally, line (C) represents the continuation value for the case where no alternative training firm is available. While in the last period of apprenticeship the value function becomes as in equation (B2) with all options available.

B.4 The time horizon and the terminal condition

We solve for the value functions at each age by backwards induction from retirement, which occurs at 65 years of age, to the start of the labor market career when the apprenticeship choice is made at 16. At retirement the value is assigned to zero: in a linear utility framework, such as ours, this is equivalent to assuming that individuals finance retirement through their own savings out of their wages.[37] Having a terminal point beyond our observation window requires assumptions on the returns to experience and tenure. Noting from the data that there is almost no wage growth beyond 11 years of potential experience we imposed that the returns to experience and tenure are constant between 10 and 30 years of actual experience.[38] We then assume that there is no wage growth beyond 30 years of experience and tenure respectively. The gain from this tight specification is that we avoid having to use a separately parameterized terminal value function. Further computational details can be found in Appendix C.

C Computational Details

This section presents the computational details for solving and estimating our model.

C.1 GDP growth and Markov transition matrix

To compute business cycles, we use the per capita West-German GDP expressed in constant prices, obtained from the OECD for the period 1975-2009. We linearly detrend the series and use transitions between above trend (good times) and below trend (bad times). Table A1 presents the transition matrix for this first order Markov process, estimated over our sample period.

[37] Note that the model uses gross wages, before any pension contributions.
[38] Thus, extrapolating from our data which stops at 30 years of experience

C.2 Construction of the Moments

As the model does not include regional variation in wages or employment, nor aggregate time trends, we remove those variations from the moments, by including regional indicator variables and a quadratic trend in all our regressions.

C.3 Computing the Value Functions

The model is solved recursively backward, starting at age 65 and until age 16. We allow the value function to depend on age as well as the other state variables.

We integrated out analytically as many state variables as possible (shocks to the value of leisure (η), shocks to the cost of training ω, and shocks to cost of moving μ) as shown in the subsection below. We approximate the value functions by evaluating them at a number of discrete points in the state space and interpolating linearly in between. For experience and tenure the points where we evaluate are 0, 2, 4, 6, 10 and 30 years of experience and 0, 2, 4, 6 and 30 years of tenure; this level of detail turned out to be sufficient. The other state variable is the firm-worker match specific effect which evolves as a random walk while the worker remains in the same job. We use 10 points on a grid which depends on education and on tenure to take into account the non-stationary nature of the process. More specifically, given the assumptions made, the match effect is a normal variable with mean zero and variance $T\sigma_U(Ed)^2 + \sigma_0(Ed)^2$ for an individual with T years of tenure. We use a quadrature-based method as in the Tauchen and Hussey (1991) procedure to generate a grid and transition matrices. We interpolate between the points.

The code was solved using parallel processing to increase speed. Solving, simulating and computing the moments for a particular set of parameters takes about 10 seconds.

D The Fit of the Model

In this section, we present the fit of the model in detail in Tables A2 to A11. The tables list all the moments used in the estimation, apart from the ones used to identify the educational choices at age 10 and 16, as they involve more than 100 entries each and are too long to display.

E Additional Parameters

In Table A12 we display the parameters of the model which are associated with unobserved heterogeneity. We model these two types of ability as a bivariate mass-point distribution with two points of support, and allow for the possibility that the two dimensions of unobserved heterogeneity are correlated. This results in four groups: individuals with high ability (which we denote "Type 3" and "Type 4") and individuals with high costs of training ("Type 2" and "Type 4"). As shown in Table A12, high ability individuals and those with lower cost of education are more likely to become apprentices. This is because the returns to choosing a skilled career is higher.

Table A1: Quarterly transition matrix for below and above trend GDP

	Below Trend in t+1	Above Trend in t+1
Below Trend in t	0.9302 (0.039)	0.069 (0.039)
Above Trend in t	0.075 (0.042)	0.925 (0.042)

Note: Note: Data source: OECD, GDP per capita, constant prices, constant PPP, period 1975-2009. Asymptotic standard errors in parenthesis.

Table A2: Goodness of Fit: Wage Level and Potential Experience

	Apprentices			Non Apprentices		
	Observed	Std Error	Simulated	Observed	Std Error	Simulated
Potential Exp ∈ [0,2]	3.09	(0.002)	3.07	4.09	(0.01)	4.12
Potential Exp ∈]2,4]	3.78	(0.003)	3.78	4.37	(0.009)	4.37
Potential Exp ∈]4,6]	4.52	(0.002)	4.52	4.5	(0.007)	4.5
Potential Exp ∈]6,8]	4.62	(0.002)	4.64	4.55	(0.008)	4.54
Potential Exp ∈]8,10]	4.71	(0.003)	4.71	4.58	(0.01)	4.59
Potential Exp ∈]10,15]	4.75	(0.004)	4.74	4.59	(0.01)	4.61
Potential Exp ∈]15,30]	4.78	(0.005)	4.75	4.58	(0.02)	4.61
Business Cycle Good	0.0336	(0.002)	0.0315	0.046	(0.009)	0.0459
Business Cycle Good, Pot. Exp>4	0.00819	(0.002)	0.0112	-0.0106	(0.009)	0.0148

Table A3: Goodness of Fit: Proportion Working and Potential Experience

	Apprentices			Non Apprentices		
	Observed	Std Error	Simulated	Observed	Std Error	Simulated
Potential Exp ∈ [0,2]	0.984	(0.001)	0.995	0.76	(0.007)	0.792
Potential Exp ∈]2,4]	0.907	(0.001)	0.903	0.751	(0.008)	0.719
Potential Exp ∈]4,6]	0.815	(0.002)	0.842	0.786	(0.008)	0.761
Potential Exp ∈]6,10]	0.876	(0.002)	0.889	0.847	(0.008)	0.85
Potential Exp ∈]10,15]	0.915	(0.002)	0.923	0.901	(0.01)	0.9
Potential Exp ∈]15,20]	0.926	(0.003)	0.92	0.918	(0.01)	0.934
Potential Exp ∈]20,40]	0.935	(0.003)	0.921	0.952	(0.01)	0.941
Business Cycle Good	0.0188	(0.001)	0.011	0.061	(0.007)	0.0545
Business Cycle Good, Pot. Exp>4	-0.014	(0.001)	-0.00349	-0.0633	(0.007)	-0.0489

Table A4: Goodness of Fit: Experience Levels and Potential Experience

	Apprentices			Non Apprentices		
	Observed	Std Error	Simulated	Observed	Std Error	Simulated
Potential Exp ∈ [0,2]	0.942	(0.01)	0.875	0.765	(0.05)	0.779
Potential Exp ∈]2,4]	2.68	(0.01)	2.82	2.19	(0.05)	2.31
Potential Exp ∈]4,6]	3.98	(0.01)	4.56	3.6	(0.06)	3.85
Potential Exp ∈]6,10]	6.07	(0.01)	7.15	5.88	(0.07)	6.31
Potential Exp ∈]10,15]	9.73	(0.02)	11.3	9.63	(0.09)	10.2
Potential Exp ∈]15,20]	14	(0.03)	15.9	14	(0.1)	14.9
Potential Exp ∈]20,40]	18.8	(0.04)	21.2	19.1	(0.2)	20.3

Table A5: Goodness of Fit: Firm Seniority and Potential Experience

	Apprentices			Non Apprentices		
	Observed	Std Error	Simulated	Observed	Std Error	Simulated
Potential Exp ∈ [0,2]	0.845	(0.03)	0.872	0.866	(0.08)	0.832
Potential Exp ∈]2,4]	2.27	(0.03)	2.35	2	(0.09)	1.97
Potential Exp ∈]4,6]	2.67	(0.03)	2.6	2.8	(0.09)	2.6
Potential Exp ∈]6,10]	3.48	(0.03)	3.18	4.03	(0.1)	3.44
Potential Exp ∈]10,15]	5.09	(0.05)	4.77	5.84	(0.1)	4.77
Potential Exp ∈]15,20]	7	(0.06)	6.22	7.85	(0.2)	5.37
Potential Exp ∈]20,40]	8.92	(0.09)	7.4	9.74	(0.3)	5.74
Business Cycle Good	-0.0111	(0.006)	-0.0518	-0.0813	(0.03)	-0.135
Business Cycle Good, Pot. Exp>4	0.0814	(0.02)	0.0833	0.0905	(0.06)	0.17

Table A6: Goodness of Fit: Number of Firms and Potential Experience

	Apprentices			Non Apprentices		
	Observed	Std Error	Simulated	Observed	Std Error	Simulated
Potential Exp ∈ [0,2]	1	(0.01)	1.04	0.91	(0.06)	1.13
Potential Exp ∈]2,4]	1.13	(0.01)	1.27	1.56	(0.06)	1.46
Potential Exp ∈]4,6]	1.56	(0.01)	1.74	2.14	(0.06)	1.83
Potential Exp ∈]6,10]	2.32	(0.02)	2.41	2.89	(0.08)	2.38
Potential Exp ∈]10,15]	3.2	(0.02)	3.15	3.86	(0.1)	3.02
Potential Exp ∈]15,20]	3.91	(0.03)	3.84	4.67	(0.1)	3.74
Potential Exp ∈]20,40]	4.62	(0.05)	4.62	5.5	(0.2)	4.55
Business Cycle Good	0.00241	(0.004)	-0.00789	0.101	(0.02)	-0.0573
Business Cycle Good, Pot. Exp>4	0.0362	(0.007)	0.0387	-0.0296	(0.03)	0.0472

Table A7: Goodness of Fit: Standard Deviations of Wages and Potential Experience

	Apprentices			Non Apprentices		
	Observed	Std Error	Simulated	Observed	Std Error	Simulated
Potential Exp ∈ [0,2]	0.337	(0.004)	0.339	0.489	(0.01)	0.381
Potential Exp ∈]2,4]	0.485	(0.03)	0.501	0.4	(0.009)	0.395
Potential Exp ∈]4,6]	0.312	(0.007)	0.332	0.353	(0.004)	0.402
Potential Exp ∈]6,10]	0.301	(0.002)	0.288	0.35	(0.002)	0.399
Potential Exp ∈]10,15]	0.334	(0.002)	0.272	0.377	(0.001)	0.383
Potential Exp ∈]15,40]	0.31	(0.002)	0.276	0.323	(0.005)	0.387

Table A8: Goodness of Fit: Wages, Experience and Tenure

	Apprentices			Non Apprentices		
	Observed	Std Error	Simulated	Observed	Std Error	Simulated
Exp ∈]2,4]	0.221	(0.002)	0.163	0.256	(0.006)	0.25
Exp ∈]4,6]	0.437	(0.003)	0.434	0.339	(0.009)	0.346
Exp ∈]6,8]	0.509	(0.003)	0.504	0.365	(0.01)	0.359
Exp ∈]8,10]	0.552	(0.004)	0.542	0.384	(0.01)	0.375
Exp ∈]10,15]	0.594	(0.004)	0.56	0.397	(0.01)	0.407
Exp ∈]15,40]	0.647	(0.005)	0.573	0.414	(0.02)	0.405
Tenure ∈]2,4]	0.00417	(0.0009)	0.043	0.0418	(0.004)	0.0369
Tenure ∈]4,6]	0.0326	(0.001)	0.0855	0.0704	(0.005)	0.047
Tenure ∈]6,8]	0.039	(0.002)	0.118	0.0778	(0.007)	0.057
Tenure ∈]8,10]	0.0473	(0.002)	0.137	0.0841	(0.008)	0.0802
Tenure ∈]10,40]	0.065	(0.003)	0.171	0.0817	(0.01)	0.0752
Business Cycle Good	0.0293	(0.002)	0.0355	0.0432	(0.009)	0.0343
Business Cycle Good, Pot. Exp>4	0.0129	(0.002)	0.00668	-0.0053	(0.009)	0.0264
In Apprenticeship Training	-1.01	(0.003)	-0.994	-	-	-
Constant	4.12	(0.003)	-0.994	4.15	(0.009)	0.0264

Table A9: Goodness of Fit: Standard Deviation of Wages, Experience and Tenure

	Apprentices			Non Apprentices		
	Observed	Std Error	Simulated	Observed	Std Error	Simulated
Exp	-0.0024	(0.0003)	-0.00404	-0.02	(0.001)	-0.000495
Exp squared	0.000123	(1e-05)	0.000116	0.000689	(5e-05)	4.27e-05
Tenure	-0.00332	(0.0002)	-0.00495	-0.00683	(0.001)	-0.00257
Tenure squared	9.49e-05	(1e-05)	0.00016	0.000278	(5e-05)	6.61e-05
Business Cycle Good	0.0156	(0.001)	0.0211	0.0184	(0.007)	0.00732
Business Cycle Good, Pot. Exp>4	-0.0306	(0.001)	-0.0298	-0.0144	(0.007)	-0.00574
In Apprenticeship Training	0.00503	(0.001)	0.0214	-	-	-
Constant	0.0963	(0.001)	0.126	0.215	(0.006)	0.154

Table A10: Goodness of Fit: Wages Changes, Experience and Tenure

	Apprentices			Non Apprentices		
	Observed	Std Error	Simulated	Observed	Std Error	Simulated
Exp	-0.022	(8e-05)	-0.0139	-0.00354	(0.0001)	-0.00294
Exp squared	0.000709	(3e-06)	0.000419	0.000104	(5e-06)	9.17e-05
Tenure	-0.000376	(5e-05)	0.00109	-0.00191	(0.0001)	-0.00062
Tenure squared	1.19e-05	(3e-06)	-9.37e-05	8.73e-05	(6e-06)	1.61e-05
In Apprenticeship Training	-0.0911	(0.0004)	-0.0534	-	-	-
Constant	0.155	(0.0005)	0.113	0.032	(0.0008)	0.03

Table A11: Goodness of Fit: Standard Deviation of Wages Changes, Experience and Tenure

	Apprentices			Non Apprentices		
	Observed	Std Error	Simulated	Observed	Std Error	Simulated
Exp	-0.00157	(6e-05)	0.00141	-0.00299	(0.0002)	-0.00054
Exp squared	6.14e-05	(3e-06)	-9.3e-05	0.000127	(8e-06)	1.79e-05
Tenure	-0.0159	(9e-05)	-0.0113	-0.00155	(0.0002)	-0.000314
Tenure squared	0.000521	(3e-06)	0.000347	5e-05	(7e-06)	7.66e-06
In Apprenticeship Training	-0.0846	(0.0005)	-0.0434	-	-	-
Constant	0.119	(0.0006)	0.0835	0.0257	(0.001)	0.00974

Table A12: Estimated parameters: unobserved heterogeneity

Parameter	Type 1	Type 2	Type 3	Type 4
Proportion in sample (π_j)	0.21	0.2	0.21	0.38
Proportion in Apprenticeship	0.98	0.83	0.99	0.87
Log wage constant Apprentices ($\alpha_0(\epsilon)$)	0	0	0.35	0.35
			(0.023)	(0.023)
Log wage constant Non Apprentices ($\alpha_0(\epsilon)$)	0	0	0.55	0.55
			(0.33)	(0.33)
Utility gain of apprenticeship ($-\lambda_0(\epsilon)$)	483	0	483	0
	(0.0228)		(0.0228)	
Correlation between types		-0.15		

Note: [a]: as a percentage of the value of leisure for apprentices. [b]: as a percentage of lifetime value. Asymptotic standard errors in parenthesis.