Career Progression and Formal versus on the Job Training

J. Adda*, C. Dustmann*, C. Meghir*, J.-M. Robin*

July 12, 2005

PRELIMINARY

Abstract

This paper evaluates the return to formal education over the life-cycle and compare it to informal, on the job training. More specifically, we assess the apprenticeship system in Germany by comparing the long run value of education choices and subsequent labor market outcomes for apprentices and non-apprentices. We develop a structural model of career progression and educational choice, allowing for unobserved ability, endogenous job to job transition, specific firm-worker matches, specific returns to tenure and to general experience. We estimate this model on a large panel data set which describes the career progression of young Germans. We find that formal education is more important than informal training, even when taking into account for the possible selection into education. We use the estimated model to evaluate the long-run impact of labor market policies on educational choices and career progression.

1 Introduction

This paper evaluates the return to formal education over the life-cycle and compare it to informal, on the job training in Germany. We assess the

*University College London and IFS. We are grateful for funding from the DFES through the CEE. We are grateful to seminar participants at NYU, Minneapolis Fed, LBS and ESEM for comments.

apprenticeship system in Germany by comparing the long run value of education choices and subsequent labor market outcomes for apprentices and non-apprentices.

To address these issues it is necessary to link education choices and labour market careers within a complete life cycle setting and to study the way that incentives at different parts of the life cycle affect education choices. This paper specifies and estimates a life cycle model of education choice and labour market careers for men who complete standard schooling at 16. Individuals face the choice of formal apprenticeship or the standard labour market. Once in the labour market they can search so as to improve the quality of the match. While working they face wage growth by experience and job specific learning. Estimation of such a model requires data on complete work and earnings histories which is available to us. We observe individuals from the moment they enter the labour market, whether as candidate apprentices or as workers. Their complete history is thus available from the age of 16 onwards with all transitions and corresponding wages observed. Moreover the fact that we observe many cohorts allows us to estimate the model over different macroeconomic conditions and hence different opportunity costs of education. In fact in descriptive regressions we show that wages in the two sectors (apprentices and non-apprentices) are important determinants of education choice.

The model we estimate combines many features of education choice models (e.g Taber (2001) and wage growth models (e.g. Topel (1991), Topel and Ward (1992), Dustmann and Meghir (2001), Altonji and Williams (1997) Altonji and Shakotko (1987)) and bears some similarities to the Keane and Wolpin (1997) model. In addition it allows for heterogeneous returns to education, experience and tenure and similarly to the Willis and Rosen (1979) model allows for comparative advantage in education choice. Finally we also model the basic elements of the welfare system to help explain the observed welfare spells.

Estimation of the model provides us with measures of the returns to experience and tenure (and their distribution) as well as the return to apprenticeship training and its distribution. It also provides a way of accounting for the sources of wage growth (learning by doing, search and selection).

Having estimated the model we have a tool that allows us to carry out policy analysis. We thus impose an EITC type program and assess its impact on education choice career progression and wage growth.

Section 2 presents the data set and descriptive statistics. Section 3 presents the model. In Section 4 we display the estimation results. Section 5 we evaluate the effect of in-work benefits.

2 The Data Set

2.1 The Data Set

We use a 1% extract of the German social security records. The data set follows a large number of young individual from 1975 to 1995. For each individual in the sample, we get the exact employment date (starting date, end date) for each job. The data set also reports the daily wage each year if the individual stays an entire year, or for the part of the year the individual works for the firm. We aggregate the data to obtain information on a quarterly basis.

The data set also reports the periods of apprenticeship training. For the purpose of this study, we select our sample to consist only of West-German males, with only post-secondary education and who start either work or an apprenticeship after school. This is a rather homogenous group of young individuals. We drop all individuals who continue onto higher education, a rather small fraction in Germany.

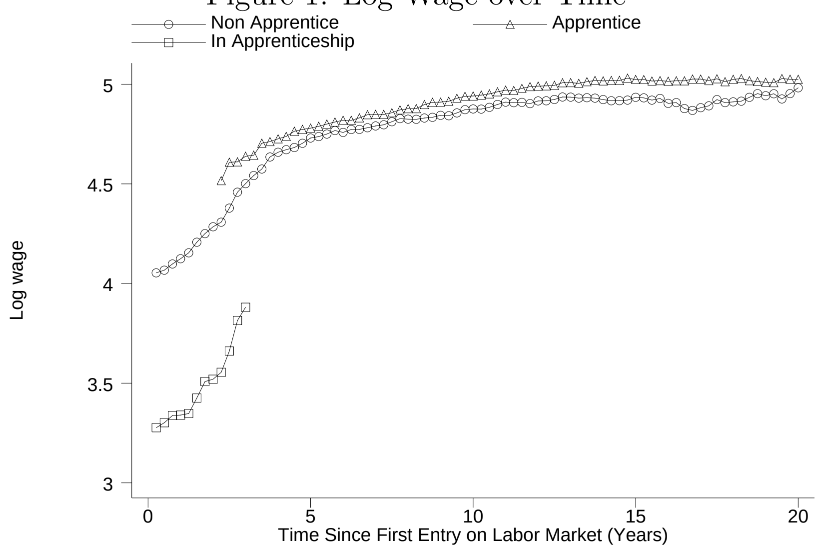
In total, we follow 27525 individuals through time, quarter after quarter up to 1995. In total, we have 996 872 observations on wages, transitions and education choices. The average age at first observation is 16.7. The oldest individual in our data is 35 years old.

2.2 Descriptive Data

2.2.1 Wage Profile and Labor Market Transitions

Figure 1 displays the log wage profile as a function of years of labor market experience for apprentices and non apprentices. Unskilled workers (non-apprentices) have a rapid increase in their wage during the first five years on the labor market. Over the next fifteen years, the wage growth is only about twenty percent, resulting in a 1.5 percentage growth rate per year (wages have been corrected for inflation). Apprentices in apprenticeship starts at a very low wage, as they are working only part-time. At the end of the apprenticeship training, wages increase up to the level of unskilled wages. From there on, the wages of apprentices increases slightly faster than those of non apprentices at rate of 1.6% per year. After fifteen to twenty years, the difference in wages between skilled and unskilled is about eight to ten percent.

Figure 1: Log Wage over Time



Wages are only one dimension in which education groups may differ. An important dimension is labor market attachment. Table 1 displays the transition probabilities by education groups and time. We distinguish the transition from work to work within and between firms. Unskilled workers have a higher probability of dropping out of the labor force. During the first five years on the labor market, each quarter, about four percent of employed skilled workers exit, while this figure is about eight percent for unskilled. The proportion decreases when we look at more senior workers, but the education difference still persists. The probability of job to job transitions are the same for both education groups, at about two to three percent. This probability decreases with the time since first entry on the labor market.

Unskilled workers have a higher probability of exit from the out-of-labor-force state, about four to five percentage points higher. This only compensates in part, the higher probability of unemployment. In total, unskilled spend less time working; over 20 years they work a total of 15 years, compared with a total of 16.5 years for skilled workers.

The education differences in exit and entry probabilities implies that non apprentices are more mobile and have more job experiences with more firms than apprentices. Figure 2 displays the number of firms in which an individual has worked in as a function of time since entry on the labor market. The difference comes from the early years, where apprentices (in apprenticeship) are much less mobile.

Table 1: Labor Market Transitions, Quarterly Frequency

	Work (Same Firm)	Work (New Firm)	Out of Labor Force
Apprentices, First 5 years			
Work	92.8	2.6	4.6
Out of Labor Force	29.6	-	70.4
Non Apprentices, First 5 Years			
Work	88.7	3.0	8.3
Out of Labor Force	25.7	-	74.3
Apprentices, After 5 years			
Work	96.2	1.9	1.9
Out of Labor Force	18.1	-	81.9
Non Apprentices, After 5 Years			
Work	94.4	1.9	3.6
Out of Labor Force	13.1	-	86.8

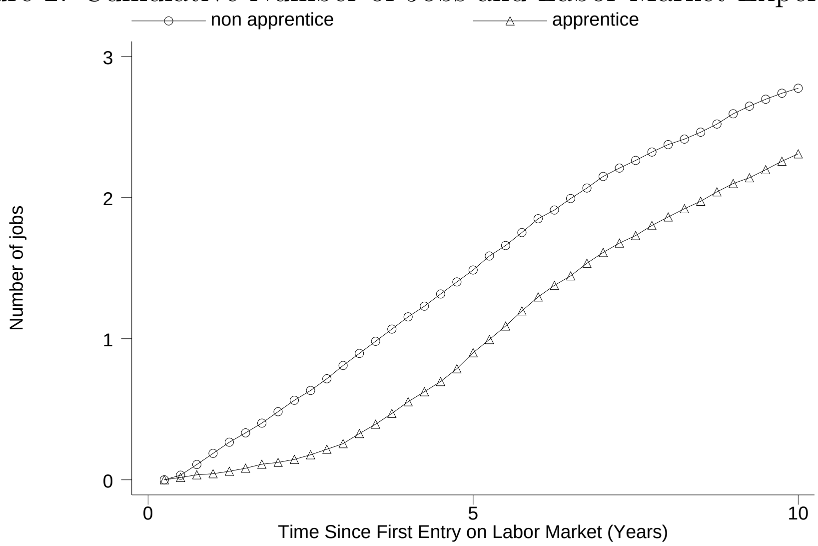
2.2.2 Decomposing Wage Growth

Next, we try to decompose the wage growth into different components. Figure 4 displays the changes in the log wage for individuals who change jobs. In the first years in the labor market, the wage growth can be substantial, at about 30% for non apprentices and 10% to 20% for apprentices. The gain in wages reduces over time, decreasing towards zero.

Figure 3 displays the wage growth conditional on staying with the same firm for two consecutive periods. The wage growth is of an order of 1 to 2% and is higher in the first 4 years for non apprentices.

Hence, most of the wage growth is due to job to job transition and very little to gains in experience or tenure. It appears that the rapid wage growth of non apprentices is mostly due to better matches and job search in the early

Figure 2: Cumulative Number of Jobs and Labor Market Experience



years. However, the results in both Figures are potentially biased, because mobility may be endogenous. Our model will be able to disentangle the selection effect from the determinant of wage growth.

2.2.3 Education Choices

Table 2 presents the marginal effect of the determinants of going into apprenticeship. In particular, we regress an indicator of apprenticeship on local wages, both skilled and unskilled, *at the time of the decision*. We also include regional indicators as well as time dummies. As apparent, educational choices are influenced by local labor market variables. Young Germans are more likely to become apprentices in region and times where the apprentice wage is higher and the unskilled wage is lower. This provides us with an exogenous variation that shifts the decision of apprenticeship and will help us, in the structural model, to identify both unobserved heterogeneity and the effect of wages on education decisions.

Table 3 presents the return to apprenticeship estimated from a regression of log wages on an education dummy, indicators for region of living, year and time since entry into the labor market. The first row presents the results using OLS, for all individuals (aged fifteen and above) and for those eighteen and above, at an age when apprenticeship is finished. The overall return at

Figure 3: Annual Changes in Log Wage (Within) and Labor Market Experience

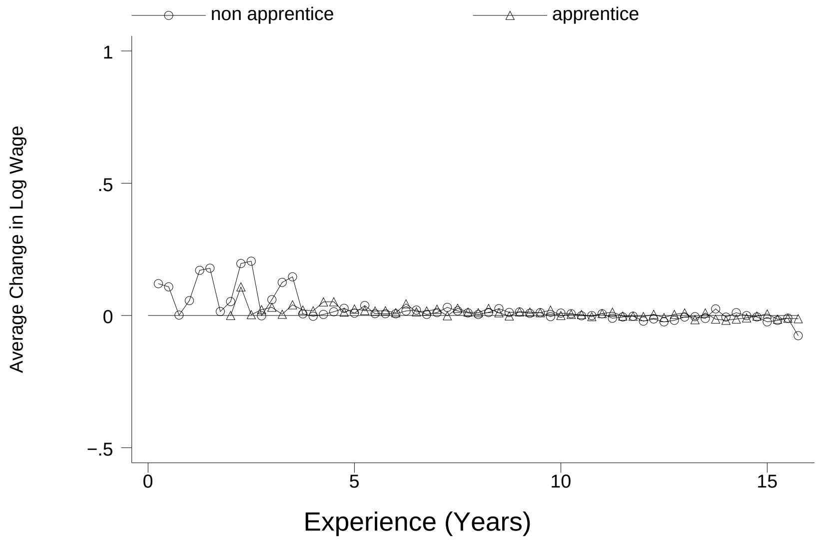


Figure 4: Changes in Log Wage (Between) and Labor Market Experience

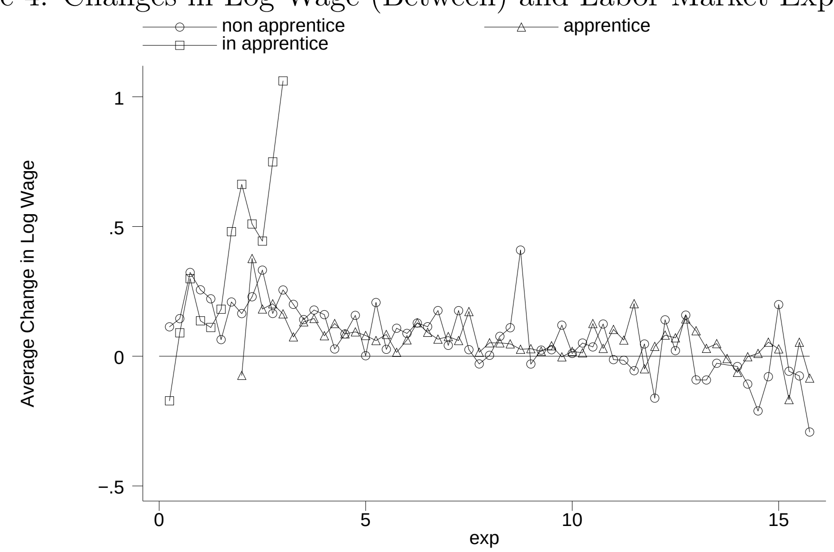


Table 2: Local Wage Effects on Apprenticeship Decision. Marginal Effects

Variable	Marginal Effect	s.e.	t-stat
Local wage Apprentice	.128393	.051	2.49
Local wage Non Apprentice	-.0365127	.025	-1.41

Note: The regression also controls for time and regional effects.

Table 3: Return to Apprenticeship: OLS and IV. Dependent variable: log wage

	Age 15 +		Age 18 + (Excluding Apprenticeship Period)	
Method	Marginal Effect	s.e.	Marginal Effect	s.e.
OLS	-0.15	(0.002)	0.07	(0.002)
IV	-0.12	(0.01)	0.09	(0.009)

Note: The regression controls for time since entry on labor market, year and region effects. IV instruments: interaction between time and region at time of first entry. Regression on 119625 observations. Apprentices in apprenticeship not included in regression. F stat for first stage instruments: 51.3.

age 15 is estimated at -15%. This negative number is due to the fact that there is a large opportunity cost of going into apprenticeship and to the fact that we do not observe any individual after age 35. It does not mean that the average return at that age is necessarily negative, but that it may take more than 20 years to make the investment profitable. After apprenticeship is finished, the return is equal to 7%.

The second row presents instrumental variable estimates. The instruments are the interaction between time and region *at the time* of the decision of whether to start or not the apprenticeship. The first stage is highly significant, with an F statistic of about 51. The estimated returns are respectively -12% and 9%. The return to apprenticeship calculated using instrumental variable is significantly (but only slightly) higher than the OLS return. This would point to some evidence of negative selection, a point we will return to in section 4.

3 Model

3.1 An Overview of the model

The model we describe takes individuals from the first point at which they make a choice and follows them to mid career. This exploits the key advantage of the data that provides information from the first point where there is a choice to be made implying no initial conditions problem and all education and career choices are observed. We focus on the population that drops out of formal academic schooling at 16 years of age and at that point just has the choice of following an apprenticeship or entering the labour market as an unskilled worker. We allow for a production function where unskilled labour and apprentices are not perfectly substitutable. Hence, the relative price of the two types of human capital will be allowed to vary over time. Utility is linear in earnings so in our model liquidity constraints are not an issue since the timing of consumption is irrelevant. We also allow for a utility of leisure by allowing the weight of income to be different when employed from when one is unemployed as well as by an intercept shift in the utility.

At the start individuals choose whether they will take the apprenticeship route, which offers formal on the job and classroom training at a reduced wage, or no formal training. In taking this decision they trade-off current earnings of an unskilled worker with working at a lower wage and possibly obtaining an improved career path through the formal training. The information they possess at that point is the distribution of idiosyncratic match specific shocks as well as the distribution of aggregate shocks that affect the evolution of relative prices in the two skill categories. They also know their type/ability which affects a number of aspects of their career, namely the costs of education, the wage level as well as the returns to experience and tenure, which are heterogeneous in our model. From an econometric point of view it is worth noting that the time variation of unskilled wages relative to skilled ones will help identify the model.

Once the education choice has been made the individual starts up on his career. First we allow for the possibility that during his apprentisheship he may move to a new employer. Once apprenticeship is completed or from the point of initial "entry" in the labour market for the non-apprentices job offers arrive at some rate, which may differ depending on whether the worker is employed or not. Associated with an offer is a draw of a match specific effect which defines the initial wage level given the person's type and experience. This then evolves as a random walk while the worker remains on the job. In addition the job offer is associated with "fringe benefits" for the job.

The model is set in discrete time. In choosing the time period we needed to address the facs that a) within a firm pay inreases are only visible to us at the end of the callendar year; bowever b) workers may change employer or stop working at any point in time. If they move to a new employer we observe a pay change. To be able to capture the richness of the data without making the model intractable we chose the time period to be a quarter. However, we restrict the arrival of the shocks to the match specific effects to occur only once a year on average.

Apprenticeship lasts more than one period, typically two years in the manufacturing industry, three years in banking services. However both the sectoral choice and the apprenticeship period are both assumed exogenous in this study. Individuals who complete an apprenticeship are hereafter designated as skilled workers ($E = s$), while those who choose the direct labour market route are labelled unskilled workers ($E = u$).

The dynamics in the model are due to the effects of apprenticeship education on future outcomes, the effects of experience and tenure, the different arrival rates of job offers between the employed and the unemployed and the effects of earnings on future unemnployment benefits. We now describe the model formally and then discuss estimation.

3.2 A formal presentation of the model

Time is discrete and individuals live H periods. At $t = 0$ an individual can either leave school and take a regular job or become an apprentice. Apprenticeship lasts τ^A units of time. This training duration is exogenously determined and depends on the particular sector of activity the individual applies to (typically two years in the manufacturing industry, three years in banking services). This sectoral choice may be endogenous but we neglect that possibility. Former apprentices are hereafter designated as skilled workers ($E =$ "S"), school dropouts are unskilled workers ($E =$ "U"). The skill indicator E takes on value "A" while the individual is in apprenticeship.

3.2.1 Business cycle

We consider a stationary economy subject to exogenous macroeconomic shocks. We assume that the economy fluctuates in a stationary way around a deterministic trend. After detrending, the macro shock is an AR(1) process

$$G' = \rho G + v, \qquad v \sim IID, \mathcal{N}(0, \sigma_v^2). \tag{1}$$

In practice, we discretize this AR(1) process into a Markov process of order one. 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.

3.2.2 Instantaneous Rewards

Wages are match specific and there are non-wage benefits to working that vary across firms. Thus, when a worker and a firm meet, they draw a match specific effect comprising a monetary part, κ_0 directly affecting the wage received, and a non monetary part μ affecting the utility of the job. Both of these components are normally distributed ($\kappa_0 \sim \mathcal{N}(0, \sigma_0^2)$ and $\mu \sim \mathcal{N}(0, \sigma_\mu^2)$). We allow the monetary value of the match to follow a random walk:

$$\kappa_t = \kappa_{t-1} + u_t, \qquad u_t \sim \mathcal{N}(0, \sigma_u^2)$$

making it time-varying but persistent. A job offer always consists of a new draw from the distribution of κ_0 and μ. These draws are not correlated across job offers, although individual choice will induce a correlation of realized offers.

Wages or earnings, denoted by $w(E, G_t, X_t, T_t, \kappa_t, \varepsilon)$, are skill-specific ($E$) functions of the macroeconomic environment G_t, experience X_t, tenure T_t, of the current value of the match-specific effect κ_t and of unobserved heterogeneity, denoted by ε.

Workers are assumed risk neutral which also implies that liquidity constraints are not an issue of concern for this model. Thus the instantaneous utility of a worker is defined as his wage plus the non wage value of the match (expressed in monetary terms):

$$R_W = w(E, G_t, X_t, T_t, \kappa_t, \varepsilon) + \mu$$

While unemployed, the individual derives a utility from unemployment benefits calculated as a fraction of the last wage when employed, as in the German UI system. In addition, there is a utility of leisure $\gamma_0^{\varepsilon}(E, X, \varepsilon)$, which varies across individuals on the basis of education, experience X, the unobserved heterogeneity component ε and an i.i.d shock η:

$$R_U(E, X, w_{-1}, \varepsilon, \eta) = \gamma_U w_{-1} + \gamma_0^{\varepsilon}(E, X, \varepsilon) + \eta$$

3.2.3 Employment transitions

Denote $W^{\varepsilon}(E, G, X, T, \kappa, \mu)$ the intertemporal utility flow of an individual who is working and by $U^{\varepsilon}(E, G, X, w_{-1}, \eta)$ the flow of utility for an unemployed person. These allow for optimal actions in the future.

These values are defined recurcively. Thus the value of unemployment is defined by

$$\begin{aligned} U^{\varepsilon}(E, G, X, w_{-1}, \eta) \quad = \quad & R_U(E, X, w_{-1}, \varepsilon, \eta) \\ & + \beta \pi_U(E, X) \mathrm{E}_{(G', \eta', \tilde{\kappa}_0', \tilde{\mu}')} \max \left(\begin{array}{c} U^{\varepsilon}\left(E, G', X, w_{-1}, \eta'\right) \\ W^{\varepsilon}\left(E, G', X, 0, \tilde{\kappa}_0', \tilde{\mu}'\right) \end{array} \right) \\ & + \beta (1 - \pi_U(E, X)) \mathrm{E}_{(G', \eta')} U^{\varepsilon}\left(E, G', X, w_{-1}, \eta'\right) \end{aligned}$$

The variable with a prime denotes next period values and β is the discount factor. With a probability $\pi_U(E, X)$, the individual receives a job offer in the next period defined by the match specific characteristics $(\tilde{\kappa}_0', \tilde{\mu}')$. He then decides whether to accept the offer or to decline it and wait until the next period and resample, possibly obtaining a better offer. The potential differencein the arrival rate of offers creates an option value to waiting. If the offer is accepted however, the worker starts with zero tenure T in the new firm. If the individual does not receive a job offer, then he stays for one more period in unemployment. If he does not receive a job offer, which happens with probability $(1 - \pi_U(E, X))$ he has no choice and receives the expected flow utility of unemployment from the next period on.

We define the value of working by:

$$
\begin{aligned}
W^{\varepsilon}(E, G, X, T, \kappa, \mu) &= w(E, G_t, X_t, T_t, \kappa_t, \varepsilon) + \mu \\
&+ \beta\delta(E, X)\mathrm{E}_{(G', \eta')} U^{\varepsilon}(E, G', X', w(E, G, X', T', \kappa), \eta') \\
&+ \beta\left(1 - \delta(E, X)\right)\pi_W(E, X)\mathrm{E}_{(G', \eta', u', \widetilde{\kappa}_0', \widetilde{\mu}')} \max \begin{pmatrix} U^{\varepsilon}(E, G', X', w(E, G, X', T', \kappa), \eta') \\ W^{\varepsilon}(E, G', X', T', \kappa + u', \mu) \\ W^{\varepsilon}(E, G', X', 0, \widetilde{\kappa}_0', \widetilde{\mu}') \end{pmatrix} \\
&+ \beta\left(1 - \delta\left(E, X\right)\right)\left(1 - \pi_W(E, X)\right)\mathrm{E}_{(G', \eta', u')} \max \begin{pmatrix} U^{\varepsilon}(E, G', X', w(E, G, X', T', \kappa), \eta') \\ W^{\varepsilon}(E, G', X', T', \kappa + u', \mu) \end{pmatrix}
\end{aligned}
$$

With a probability $\delta(E, X)$, the worker looses his job and has no option but to go next period into unemployment. If the job is not destroyed, the individual gets an outside offer with a probability $\pi_W(E, X)$, which depends on the education level E and experience. The outside offer is a pair $(\tilde{\kappa}_0, \tilde{\mu})$, to be compared with the value of the current match as it will be in the next period, i.e. $(\kappa + u', \mu)$. The individual then decides whether to stay in the same job, to accept the outside offer or to go into unemployment. If no outside offer is received, the worker only decides on whether to stay on the same job or to quit into unemployment.

A key point is that the effect of the business cycle on wages is allowed to differ by education group. This allows for the relative prices of the two education levels to change with the cycle reflecting the possibility that the two labour factors may not be perfectly substitutable.

Experience X and tenure T grow by one in each period of work. Tenure is set to zero at the start of a new job; we do not allow for depreciation of skills while unemployed.

3.2.4 Educational choice

After completing high school at 16, an individual can work in a regular job or start as an apprentice, always with experience X being zero at that point.[1] Apprenticeship lasts τ^A periods. We take the actual length of the apprenticeship as exogenous and we condition on it. The decision of whether to follow formal training will depend on the opportunity cost of doing so and this in turn depends on the current economic environment reflected in the value of the business cycle indicator G. This drives the relative wage of the apprentices and non-apprentices and is observable. It will also depend on expectations about future returns as well as on the unobseved ability ε characterising the individual. It is a key feature of our model that it uses the business cycle fluctuations as identifying information for who decides to move into formal on-the-job training and who decides to obtain a regular job.

The value of apprenticeship W_A^ε is similar to the value of employment W^ε except that the training firm pays the worker only a fraction λ_A of his productivity as an unskilled worker ($w^\varepsilon(0, G, X, T, \kappa)$), the rest presumably serving as payment for the general training received. In addition, we do not allow individuals to experience unemployment during apprenticeship, although he can decide to change firm if the opportunity arises, which reflects the facts in the data. Thus during the apprenticeship training period ($X < \tau^A$) the value of work is:

$$
\begin{aligned}
W_A^\varepsilon(G, X, T, \kappa, \mu) \quad = \quad & \lambda_A \cdot w^\varepsilon(0, G, X, T, \kappa) + \mu \\
& + \beta \pi_A \mathrm{E}_{(G', u', \widetilde{\kappa}_0', \widetilde{\mu}')} \max \begin{pmatrix} W_A^\varepsilon(G', X', T', \kappa + u', \mu) \\ W_A^\varepsilon(G', X', 0, \widetilde{\kappa}_0', \widetilde{\mu}') \end{pmatrix} \\
& + \beta(1 - \pi_A) \mathrm{E}_{(G', u')} W_A^\varepsilon(G', X', T', \kappa + u', \mu)
\end{aligned}
$$

With a probability π_A, the apprentice gets an outside offer $(\tilde{\kappa}_0', \tilde{\mu}')$ and choose optimally. If no offer is received, the apprentice stays on for one more period and accumulates experience and tenure within the firm.

At the end of the training period ($X = \tau^A$), he may stay in the same form or not, and he may also choose to stop working. This the value of work

[1] He can also continue with an academic stream of education. However we condition on stopping full time academic education at 16.

now becomes

$$
\begin{aligned}
W_A^{\varepsilon}(G,X,T,\kappa,\mu) &= \lambda_A \cdot w^{\varepsilon}\left(0,G,\tau^A,T,\kappa\right)+\mu \\
&\quad +\beta\pi_A \mathrm{E}_{(G',\eta',u',\widetilde{\kappa}'_0,\widetilde{\mu}')} \max\begin{pmatrix} U^{\varepsilon}(1,G',X',0,\eta') \\ W^{\varepsilon}(1,G',X',0,\widetilde{\kappa}_0,\widetilde{\mu}) \\ W^{\varepsilon}(1,G',X',T',\kappa+u',\mu)\end{pmatrix} \\
&\quad +\beta(1-\pi_A)\mathrm{E}_{(G',\eta',u')} \max\begin{pmatrix} U^{\varepsilon}(1,G',X',0,\eta') \\ W^{\varepsilon}(1,G',X',T',\kappa+u',\mu)\end{pmatrix}
\end{aligned}
$$

Apprenticeship is the chosen decision at $X = 0$ if the value of apprenticeship at that point is larger than the value of working without vocational training:

$$W_A^{\varepsilon}(G,0,0,\kappa_0,\mu) - \lambda_0(Z,\varepsilon) - \omega > W^{\varepsilon}(E=0,G,0,0,\tilde{\kappa}_0,\tilde{\mu})$$

where $\lambda_0(Z,\varepsilon)$ is a utility cost of going into apprenticeship. We allow that parameter to differ with unobserved heterogeneity and with a set of (exogenous) variables Z. The vector Z region of living, reflecting differences in access into the apprenticeship system. ω is a random cost of going into apprenticeship, normally distributed with mean zero and variance σ_ω^2.

3.3 Estimation Method and Empirical Specification

The model is solved numerically using a value function iteration technique. The expectation operators are computed using Gaussian quadrature. The model is estimated by maximum likelihood. We refer the reader to appendix A for the likelihood function.

The estimation was done at a quarterly frequency, using a random sample of 1403 individuals, totaling about 68932 observations on wages, employment choices and education choices.

We imposed three different "types" of individuals in the likelihood. Each type differ in several ways. First, we allow for different wage levels (fixed effect). Second, we also allow for heterogeneity in the return to experience and tenure, as well as a heterogenous cost of education.

The log wage equation for individual i in firm f and in period t is specified as:

$$
\begin{aligned}
\ln w_{ift} &= \alpha_0(\varepsilon_i) + \alpha_{Ed}(\varepsilon_i)Ed_i + \alpha_X(X_i,Ed_i,\varepsilon_i) + \alpha_T(T_i,Ed_i,\varepsilon_i) \\
&\quad + \alpha_G(Ed_i)G_t + \kappa_{ift}
\end{aligned} \tag{2}
$$

where $\alpha_X()$ and α_T are two smooth functions of experience and tenure, which are education specific. We use a piecewise linear function, with nodes at 0,2,4,6 and 30 years of experience and tenure. Unobserved heterogeneity enters these functions multiplicatively.

The probability of job destruction $\delta(E, X)$ varies with experience and we allow for two different periods, between 0 and 4 years of experience and more than four years of experience. Our experience is that after that period, the probability is fairly stable. The probability of receiving a job offer on the job, $\pi_W(E, X)$ is education specific and takes two distinct values for apprentices, during and after apprenticeship. The probability of receiving a job offer while not working, $\pi_U(E, X)$ is modelled as an education specific linear function in experience.

4 Results

4.1 Fit of the Model

We evaluate the fit of the model by simulating the education decisions and the labor market transitions for a cohort of individuals over time. Table 4 displays the labor market transitions by education groups at a quarterly frequency. We distinguish five possible transitions, from and to unemployment, between same job and job to job. Overall, the model match the transition probabilities closely.

Table 4: Goodness of Fit: Labor Market Transitions

	Apprentices		Non Apprentices	
	Obs	Pred	Obs	Pred
U to U	0.82	0.83	0.82	0.84
U to E	0.18	0.17	0.18	0.16
E to U	0.04	0.03	0.08	0.04
E to new E	0.03	0.02	0.03	0.02
E to same E	0.93	0.89	0.89	0.94

Figure 5 plots the average log wage as a function of time since first entry on the labor market (including the apprenticeship period). Both observed and predicted wages are close. Figure 6 plots the average experience and tenure over time for the two education groups. The model does a good job in both dimension and even picks up the non linearity in the evolution of tenure for apprentices.

Figure 5: Goodness of Fit, Wages

4.2 Results

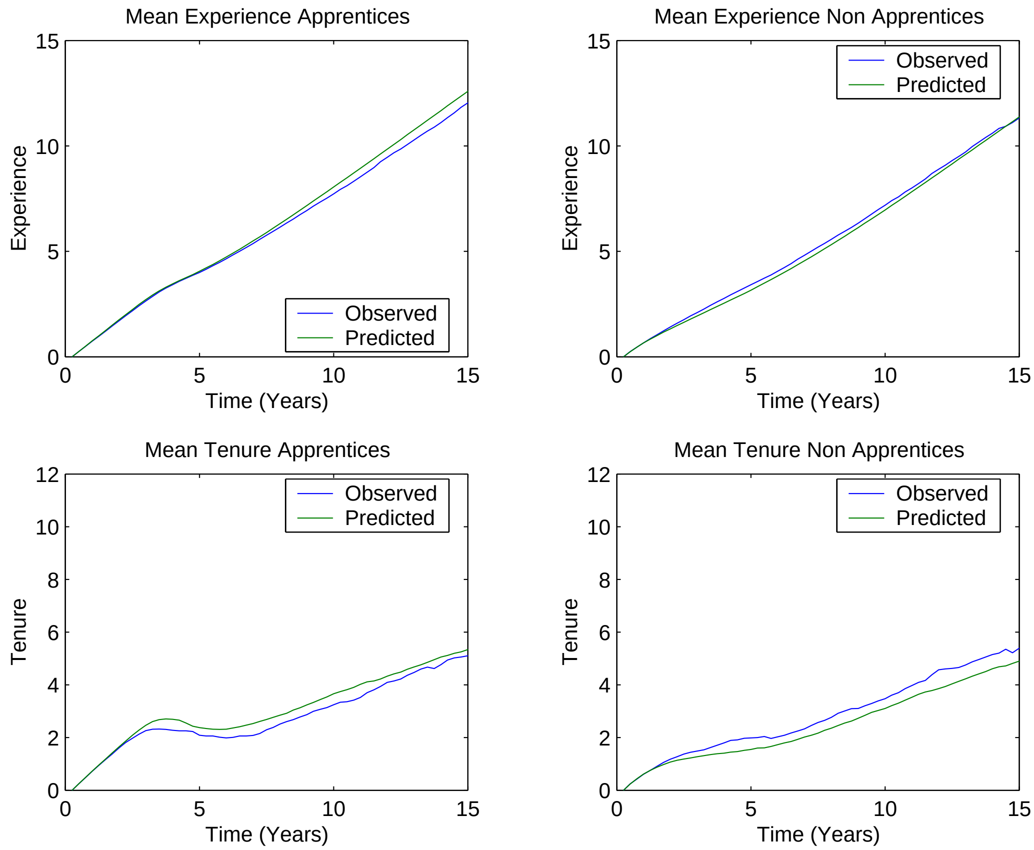
4.2.1 Estimated Parameters

In total we have 68 parameters. The results are presented in Table 8 in appendix B. Given the wealth of data, most of the parameters are estimated with a high precision.

A direct interpretation of some of the coefficients in Table 8 is difficult as these parameters are only interesting in combination with others. We also use simulations to illustrate our results below.

We concentrate the discussion around some interesting parameters. First, the price of human capital for both education group is slightly different.

Figure 6: Goodness of Fit, Average Experience and Tenure over Time by Education



When the business cycle goes from bad to good, wages of unskilled increases by 0.3% and by 0.6% for skilled labor. The probability of a job offer while on the job is estimated at 11% per quarter for apprentices (only 5% if they are in apprenticeship) and at 7.5% for non apprentices. Given that skilled and unskilled have the same probability to move from job to job, it follows that non-apprentices decline job offers more often than apprentices.

In unemployment, skilled individuals have a 21% probability of receiving a job offer. For unskilled, this number is slightly lower at 19%.

4.2.2 Unobserved Heterogeneity

The estimation allows unobserved heterogeneity in the form of three types of individuals. The proportion of these types is respectively 6, 47 and 47%. Table 5 displays summary characteristics for different groups. Type 1 individuals are low ability individuals with the lowest wage at start. Given their low proportion in our sample, we concentrate on the two other types. Type 2 is a relatively low ability type compared with type 3. The unskilled wage for this group, at zero experience and tenure, is about 30% lower. However, they experiencing a slightly higher return to experience than Type 3 individuals. In all cases, the return to tenure is very small, and not statistically different from zero.

In terms of education choice, Type 2 individuals are more likely than Type 3 ones to opt for an apprenticeship. Given that they are also of lower ability, we find again evidence of negative selection.

4.2.3 Return to Apprenticeship

The model allows us to compute the value of apprenticeship, $E_{G,\kappa,\mu} W_A^{\varepsilon}(G, X = 0, T = 0, \kappa, \mu)$ and the value of starting as an unskilled on the labor market, $E_{G,\kappa,\mu} W^{\varepsilon}(E = u, G, X = 0, T = 0, \kappa, \mu)$. These value functions take into account all possible future outcomes, the wage profile over time as well as the future labor market transitions. Based on these estimated values, we can compute the return to apprenticeship computed as:

$$r^{\varepsilon} = \frac{E_{G,\kappa,\mu} W_A^{\varepsilon}(G, X = 0, T = 0, \kappa, \mu)}{E_{G,\kappa,\mu} W^{\varepsilon}(E = u, G, X = 0, T = 0, \kappa, \mu)} - 1$$

We condition on the type of the agent and we compute the average return per year, evaluated over forty years. The results are displayed in Table 6.

Table 5: Unobserved Heterogeneity

	Type 1	Type 2	Type 3
Proportion in Sample	6%	47%	47%
Proportion Apprentices	78%	74%	68%
Log wage constant	3.38	3.76	4.07
Instantaneous Return to Apprenticeship	34%	45%	45%
Average Return to Experience (per year) if App.	1.2%	1.3%	1.2%
Average Return to Experience (per year) if Non App.	2%	2.1%	1.9%
Average Return to Tenure (per year) if App.	0%	0%	0%
Average Return to Tenure (per year) if Non App.	0%	0%	0%
Utility Cost of Apprenticeship (% of total lifetime value)	-7%	-9%	22%

Note: Average return to experience calculated after a period of 20 years of experience. Average return to tenure calculated after a period of 4 years in the job.

On average, the return to apprenticeship is equal to about 15%. However, there is some heterogeneity across types as it is lower for Type 3 individuals, hence the relatively smaller proportion of apprentices in that group. The second row of Table 6 displays the return to apprenticeship, when the utility cost of education is set to zero. The change in the returns reflects the differential costs by type. Type 3 has the lowest cost (actually a benefit) of apprenticeship.

The third row displays the return to apprenticeship calculated at age 18- an age at which all individuals have completed their training in the model. The average return is about 15%, but is very low for Type 2 individuals (even slightly negative), and high for Type 3. High ability individuals are financially better off in apprenticeship.

We next decompose the overall return to apprenticeship at age 18. We first consider an equal variance for the distribution of the firm-worker match specific effect. We set the variance of apprentices equal to that of non apprentices (σ_{NA}=σ_{0NA}). As the variance of the distribution of initial matches is higher for non apprentices, skilled occupation becomes more attractive, as they are more gains in moving to new firms.

Table 6: Return to Apprenticeship

	Average	Type 1	Type 2	Type 3
Return to Apprenticeship at age 15	15.3 %	27.2 %	15.8 %	13.2 %
Net of utility of education at age 15	21.6 %	18.1 %	5.23 %	38.4 %
Return to Apprenticeship at age 18	15.4 %	8.61 %	-2.82 %	34.5 %
Decomposing the Return to Apprenticeship at age 18				
Baseline	15.4 %	8.61 %	-2.82 %	34.5 %
Equal distribution of firm-worker match (σ_0)	20.4 %	12.4 %	1.2 %	40.7 %
Equal job to job mobility (π_W)	15.8 %	8.84 %	-2.54 %	35 %
Equal job destruction rate (δ)	2.11 %	-8.46 %	-14.7 %	20.4 %
Equal growth return to tenure	15.4 %	8.63 %	-2.85 %	34.5 %
Equal growth return to experience	15.4 %	8.63 %	-2.85 %	34.5 %

Note: Average returns to apprenticeship calculated over a period of 40 years.

We next impose an equal job-to-job mobility, by setting $\pi_W(E)$ to the non apprentice value. This has little consequences on the return to apprenticeship.

Imposing the rate of job destruction to be equal across education groups decreases dramatically the return to apprenticeship, because job destruction occurs more frequently in unskilled occupation.

The last two rows impose that the growth rate in the return to tenure or experience are equal across education groups. As very little growth takes place, both in the return to tenure and experience, the effect on the return to apprenticeship is negligible.

5 Policy Evaluations

In this section, we evaluate the effect of labor market policies on career progression and education choices. In particular, we evaluate the effect of in-work benefits on human capital accumulation and acquisition of skills. These policies offer subsidies to employed individuals with a low wage. Examples of such policies are the Earn Income Tax Credit (EITC) in the US and the Working Family Tax Credit (WFTC) in the UK. These policies are in place to encourage labor market participation.

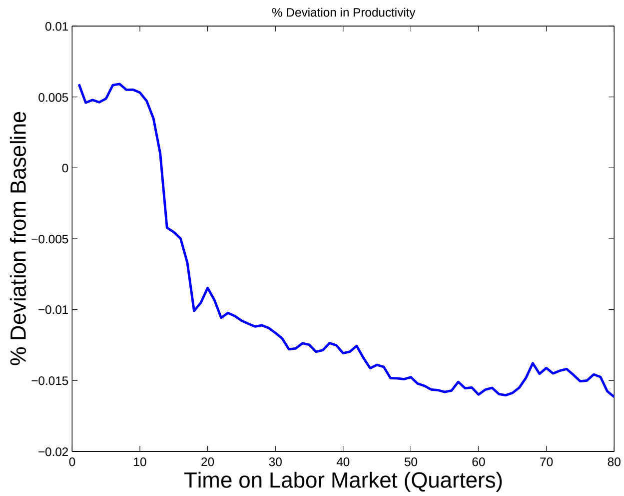
We simulate a reform similar to the EITC, where low wage individuals get a subsidy. This subsidy starts at 0 for a zero wage, increases with the wage up to a first limit, stays constant over a range of income and finally declines to zero. Hence, two categories of individuals do not receive a subsidy: individuals not working and individuals with a high enough wage.

In general, these in-work benefit policies have an effect on labor market participation. However, these policies could also have detrimental long-term effects on education choices and skill acquisition. As lower wages are subsidized, individuals are less likely to obtain higher education levels as the wage gap between education groups might decrease. Second, due to the non linearity of the benefits, the policy might discourage job-to-job mobility. This would reduce the mobility of workers across jobs and slow down or prevent the best matches between firms and workers to form, decreasing over-all productivity.

Figures 7 and 8 show the results of the policies simulated from the estimated model. The policy has a negative impact on the firm-worker specific match. The match is about 1% lower. The policy has a negative effect on overall productivity of about the same magnitude.

Figure 7: Effect of In-Work Benefits on Firm-Worker Match

Figure 8: Effect of In-Work Benefits on Total Productivity



6 Conclusion

This paper evaluates the return to formal education over the life-cycle and compare it to informal, on the job training. More specifically, we assess the apprenticeship system in Germany by comparing the long run value of education choices and subsequent labor market outcomes for apprentices and non-apprentices. We develop a structural model of career progression and educational choice, allowing for unobserved ability, endogenous job to job transition, specific firm-worker matches, specific returns to tenure and to general experience. We estimate this model on a large panel data set which describes the career progression of young Germans. We find that the return to apprenticeship is positive. A large part of the return to apprenticeship comes from non financial returns as the utility of education.

We use the estimated model to evaluate the long-run impact of labor market policies on educational choices and career progression. We find that policies such as the Earned Income Tax Credit which subsidize low wage have a detrimental effect on the probability of further education and on job mobility.

References

Altonji, Joseph and Shakotko, Robert. “Do wages rise with job seniority?” In Orley

Ashenfelter and Kevin Hallock, eds., *Unemployment, trade unions, and dispute resolution*, volume 47, pp. 219–241, International Library of Critical Writings in Economics 1987.

Altonji, Joseph and Williams. "Do wages rise with job security?" *Review of Economic Studies* 1997, *54*(179), 437–460.

Cossa, Ricardo, Heckman, James and Lochner, Lance. "Wage subsidies and skill formation: A study of the earned income tax credit." mimeo University of Chicago 1999.

Dustmann, Christian and Meghir, Costas. "Wages, experience and seniority." IFS working paper W01/01 2001.

Eissa, Nada and Liebman, Jeffrey B. "Labor supply response to the earned income tax credit." *Quarterly Journal of Economics*, May 1996, *111*(2), 605–637.

Heckman, James, Ichimura, Hidehiko and Todd, Petra. "Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme." *Review of Economic Studies*, Oct 1997, *64*(4), 605–654.

Keane, Micheal P. and Wolpin, Kenneth I. "The career decisions of young men." *Journal of Political Economy*, June 1997, *105*(3), 473–522.

Sianesi, Barbara. "An evaluation of the swedish active labour market programmes." IFS working paper W02/01, forthcoming Review of Economics and Statistics 2002.

Taber, Christopher. "The rising college premium in the eighties: Return to college or return to unobserved ability?" *Review of Economic Studies*, July 2001, *68*(3), 665–691.

Topel, Robert. "Specific capital, mobility, and wages: Wages rise with job seniority." *Journal of Political Economy*, Feb 1991, *99*(1), 145–176.

Topel, Robert and Ward, Michael. "Job mobility and the careers of young men." *Quarterly Journal of Economics*, May 1992, *107*(2), 439–479.

Willis, Robert and Rosen, Sherwin. "Education and self-selection." *Journal of Political Economy*, Oct 1979, *87*(5), S7–S36.

Appendix

A Likelihood

$$U^{\varepsilon}(E,G,X,w_{-1},\eta)=\gamma_w w_{-1}^{\varepsilon}+\gamma_0+\eta+\widehat{U}^{\varepsilon}(E,G,X,w_{-1}) \quad \text{(say)},$$

$$W^{\varepsilon}(E,G,X,T,\kappa,\mu)=w^{\varepsilon}\left(E,G,X,T,\kappa\right)+\mu+\widehat{W}^{\varepsilon}(E,G,X,T,\kappa,\mu).$$

$$W^{A}(G,X,T,\kappa,\mu)=\lambda_w\cdot w\left(0,G,X,T,\kappa\right)+\mu+\widehat{W}^{A\varepsilon}(G,X,T,\kappa,\mu)$$

$$W^{A}(G,\tau,T,\kappa,\mu)=\lambda_w\cdot w^{\varepsilon}\left(0,G,\tau,T,\kappa\right)+\mu-\lambda_0-\omega+\widehat{W}_t^{A}(G,T,\kappa,\mu).$$

$$\begin{aligned}W_A^{\varepsilon}(G,0,0,\kappa_0,\mu)-\lambda_0^{\varepsilon}-\omega &> W^{\varepsilon}(E=0,G,0,0,\tilde{\kappa}_0,\tilde{\mu})\\ &\Leftrightarrow \omega<\lambda_w\cdot w^{\varepsilon}\left(0,G,0,T,\kappa_0\right)-w^{\varepsilon}\left(0,G,0,T,\kappa_0'\right)-\lambda_0\end{aligned}$$
$$+\widehat{W}_t^{A,\varepsilon}(G,0,\kappa_0)-\widehat{W}_t^{\varepsilon}(0,G,0,0,\kappa_0'),$$

An individual occupational trajectory is denoted as $y=(\cdots,w_t,d_t,\cdots)$ for $t\geq 0$ or τ depending on education.The variable d_t^A indicates whether an individual in the course of apprenticeship is employed in a new job with tenure zero ($d_t=1$) or employed in the same job as in period $t-1$ with positive tenure ($d_t=2$). The variable d_t indicates whether an individual who has left school or apprenticeship is unemployed in period t ($d_t=0$), employed in a new job with tenure zero ($d_t=1$) or employed in the same job as in period $t-1$ with positive tenure ($d_t=2$). We let $w_t=0$ if $d_t=0$. Employment trajectories are conditioned by the initial educational choice: $E=1$ for apprencices and $E=0$ for non apprentices. Knowledge of y suffices to construct the experience and tenure variables X_t and T_t. Also, one must keep track of the last paid wage for currently unemployed workers (call it $w_{-1,t}$).

Conditional on observed and unobserved heterogeneity, the likelihood of one individual observation (E,y) is constructed as follows.

Educational choice: The apprenticeship probability, conditionally on a business cycle G and an accepted wage as an apprentice w^A is:

$$\Pr\left\{E=1|G,w^A\right\}=\Pr\left\{\begin{array}{c}\omega<w^A-w\left(0,G,0,0,\tilde{\kappa}_0\right)-\lambda_0^{\varepsilon}+\mu-\tilde{\mu}\\+\widehat{W}_A^{\varepsilon}\left(G,0,0,\kappa,\mu\right)-\widehat{W}^{\varepsilon}(0,G,0,0,\tilde{\kappa}_0,\tilde{\mu})|G,w^A\end{array}\right\}$$

$$=\iiint\Phi\left(\frac{\begin{array}{c}w^A-w\left(0,G,0,0,\tilde{\kappa}_0\right)-\lambda_0^{\varepsilon}+\mu-\tilde{\mu}\\+\widehat{W}_A^{\varepsilon}(G,0,0,\kappa,\mu)-\widehat{W}^{\varepsilon}\left(0,G,0,0,\tilde{\kappa}_0,\tilde{\mu}\right)\end{array}}{\sigma_\omega}\right)dF(\tilde{\kappa}_0)dF(\mu)dF(\tilde{\mu})$$

and the likelihood of observing wage w^A is:

$$\ell\left(w^A|G\right)=\frac{1}{w^A}\frac{1}{\sigma_{\kappa_0}}\varphi\left(\frac{\kappa}{\sigma_{\kappa_0}}\right).$$

where $\kappa=w^A-w(0)$

Apprentices changing employers in the course of apprenticeship: The probability of accepting a new apprentice job paid $w_t^A$for in-apprenticeship workers in period $t-1$is such that :

$$\begin{aligned}&\Pr\left\{d_t^A=1|G_t,T_t,w_t^A,w_{t-1}^A\right\}\\=\;&\Pr\left\{W_t^A\left(G_t,0,\underbrace{\kappa\left(0,G_t,t,0,\frac{w_t^A}{\lambda_w}\right)}_{\kappa_t}\right)>W_t^A\left(G_t,T_t,\underbrace{\kappa\left(0,G_t,t-1,T_t-1,\frac{w_{t-1}^A}{\lambda_w}\right)+\widetilde{u}}_{\kappa_{t-1}}\right)\right\}\\=\;&\int_0^1\mathbf{1}\left\{\begin{array}{c}W_t^A\left(G_t,0,\kappa\left(0,G_t,t,0,\frac{w_t^A}{\lambda_w}\right)\right)\\>W_t^A\left(G_t,T_t,\kappa\left(0,G_t,t-1,T_t-1,\frac{w_{t-1}^A}{\lambda_w}\right)+\sigma_u\Phi^{-1}\left(u\right)\right)\end{array}\right\}du\\\approx\;&\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\left\{\begin{array}{c}W_t^A\left(G_t,0,\kappa\left(0,G_t,t,0,\frac{w_t^A}{\lambda_w}\right)\right)\\>W_t^A\left(G_t,T_t,\kappa\left(0,G_t,t-1,T_t-1,\frac{w_{t-1}^A}{\lambda_w}\right)+\sigma_u\Phi^{-1}\left(u\right)\right)\end{array}\right\}\end{aligned}$$

and

$$\ell\left(w_t^A|G_t\right)=\frac{1}{w_t^A}\frac{1}{\sigma_{\kappa_0}}\varphi\left(\frac{\kappa\left(0,G_t,t,0,\frac{w_t^A}{\lambda_w}\right)}{\sigma_{\kappa_0}}\right)\tag{3}$$

Apprentices keeping the same employer in the course of apprenticeship: The probability of keeping the same job given a new wage $w_t^A$is:

$$\begin{aligned}
& \Pr\left\{d_t^A = 2|G_t, T_t, w_t^A, w_{t-1}^A\right\} \\
= & \Pr\left\{W_t^A\left(G_t, 0, \kappa_0\right) > W_t^A\left(G_t, T_t, \kappa\left(0, G_t, t, T_t, \frac{w_t^A}{\lambda_w}\right)\right)\right\} \\
= & \int_0^1 \mathbf{1}\left\{W_t^A\left(G_t, T_t, \kappa\left(0, G_t, t, T_t, \frac{w_t^A}{\lambda_w}\right)\right) > W_t^A\left(G_t, 0, \sigma_{\kappa_0}\Phi^{-1}\left(u\right)\right)\right\} du \\
\approx & \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\left\{W_t^A\left(G_t, T_t, \kappa\left(0, G_t, t, T_t, \frac{w_t^A}{\lambda_w}\right)\right) > W_t^A\left(G_t, 0, \sigma_{\kappa_0}\Phi^{-1}\left(u\right)\right)\right\}
\end{aligned}$$

and the density of that new wage is:

$$\ell\left(w_t^A|G_t, T_t, w_{t-1}^A\right) = \frac{1}{w_t^A}\frac{1}{\sigma_u}\varphi\left(\frac{\kappa\left(0, G_t, t, T_t, \frac{w_t^A}{\lambda_w}\right) - \kappa\left(0, G_t, t-1, T_t-1, \frac{w_{t-1}^A}{\lambda_w}\right)}{\sigma_u}\right) \tag{4}$$

Transition from unemployment to employment: The probability of accepting a job paid $w_t$for unemployed workers in period $t-1$is such that :

$$\begin{aligned}
& \Pr\left\{d_t = 1|E, G_t, X_t, w_{-1,t}, w_t, d_{t-1} = 0\right\} \\
= & \Pr\left\{U_t(E, G_t, X_t, w_{-1,t}, \widetilde{\eta}) \le W_t\left(E, G_t, X_t, 0, \kappa\left(E, G_t, X_t, 0, w_t\right)\right)\right\} \\
= & \Phi\left(\frac{W_t\left(E, G_t, X_t, 0, \kappa\left(E, G_t, X_t, 0, w_t\right)\right) - \gamma_w w_{-1,t} - \gamma_0 - \widehat{U}_t(E, G_t, X_t, w_{-1,t})}{\sigma_\eta}\right).
\end{aligned}$$

If $X_t = 0$then $w_{-1,t} = 0$.

The density of the accepted wage is:

$$\ell\left(w_t|E, G_t, X_t, T_t, w_{t-1}, d_{t-1} = 0\right) = \frac{1}{w_t}\frac{1}{\sigma_{\kappa_0}}\varphi\left(\frac{\kappa\left(E, G_t, X_t, 0, w_t\right)}{\sigma_{\kappa_0}}\right). \tag{5}$$

Long term unemployed: The probability of remaining unemployed in period t given unemployment in period $t-1$ is:

$$
\begin{aligned}
& \Pr\left\{d_t = 0 | E, G_t, X_t, w_{-1,t}, d_{t-1} = 0\right\} \\
= \; & \Pr\left\{U_t(E, G_t, X_t, w_{-1,t}, \widetilde{\eta}) > W_t(E, G_t, X_t, 0, \widetilde{\kappa}_0)\right\} \\
= \; & \int_0^1 \overline{\Phi}\left(\frac{W_t\left(E, G_t, X_t, 0, \sigma_{\kappa_0}\Phi^{-1}(u)\right) - \gamma_w w_{-1,t} - \gamma_0 - \widehat{U}_t(E, G_t, X_t, w_{-1,t})}{\sigma_\eta}\right) du \\
\approx \; & \frac{1}{n}\sum_{i=1}^{n} \overline{\Phi}\left(\frac{W_t\left(E, G_t, X_t, 0, \sigma_{\kappa_0}\Phi^{-1}\left(\frac{i}{n}\right)\right) - \gamma_w w_{-1,t} - \gamma_0 - \widehat{U}_t(E, G_t, X_t, w_{-1,t})}{\sigma_\eta}\right).
\end{aligned}
$$

Transition from employment to unemployment: The probability of losing one's job in period t is:

$$
\begin{aligned}
& \Pr\left\{d_t = 0 | E, G_t, X_t, T_t, w_{t-1}, d_{t-1} > 0\right\} \\
= \; & \Pr\left\{ U_t(E, G_t, X_t, w_{t-1}, \widetilde{\eta}) > \underbrace{\max\begin{pmatrix} W_t(E, G_t, X_t, 0, \widetilde{\kappa}_0) \\ W_t\left(E, G_t, X_t, T_t, \kappa\left(E, G_t, X_t - 1, T_t - 1, w_{t-1}\right) + \widetilde{u}\right) \end{pmatrix}}_{\text{call that}\Lambda(\widetilde{\kappa}_0, \widetilde{u})} \right\} \\
= \; & \int_0^1 \int_0^1 \overline{\Phi}\left(\frac{\Lambda\left(\sigma_{\kappa_0}\Phi^{-1}(u), \sigma_u\Phi^{-1}(v)\right) - \gamma_w w_{t-1} - \gamma_0 - \widehat{U}_t(E, G_t, X_t, w_{t-1})}{\sigma_\eta}\right) dudv \\
\approx \; & \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \overline{\Phi}\left(\frac{\Lambda\left(\sigma_{\kappa_0}\Phi^{-1}(\frac{i}{n}), \sigma_u\Phi^{-1}(\frac{j}{n})\right) - \gamma_w w_{,t-1} - \gamma_0 - \widehat{U}_t(E, G_t, X_t, w_{t-1})}{\sigma_\eta}\right).
\end{aligned}
$$

Job movers: The probability of accepting a new job paid $w_t$for employed workers in period $t-1$is such that :

$$\begin{aligned}
& \Pr\left\{d_t = 1 | E, G_t, X_t, T_t, w_{t-1}, d_{t-1} > 0, w_t\right\} \\
= \; & \Pr\left\{W_t\left(E, G_t, X_t, 0, \kappa\left(E, G_t, X_t, 0, w_t\right)\right) > \max\left(U_t(E, G_t, X_t, w_{,t-1}, \widetilde{\eta}),\right.\right. \\
& \qquad \left.\left. W_t\left(E, G_t, X_t, T_t, \kappa\left(E, G_t, X_t - 1, T_t - 1, w_{t-1}\right) + \widetilde{u}\right)\right)\right\} \\
= \; & \Phi\left(\frac{W_t\left(E, G_t, X_t, 0, \kappa\left(E, G_t, X_t, 0, w_t\right)\right) - \gamma_w w_{t-1} - \gamma_0 - \widehat{U}_t(E, G_t, X_t, w_{t-1})}{\sigma_\eta}\right) \\
& \times \int_0^1 \mathbf{1}\left\{\begin{array}{c} W_t\left(E, G_t, X_t, T_t, \kappa\left(E, G_t, X_t - 1, T_t - 1, w_{t-1}\right) + \sigma_u \Phi^{-1}(u)\right) \\ < W_t\left(E, G_t, X_t, 0, \kappa\left(E, G_t, X_t, 0, w_t\right)\right) \end{array}\right\} du \\
\approx \; & \Phi\left(\frac{W_t\left(E, G_t, X_t, 0, \kappa\left(E, G_t, X_t, 0, w_t\right)\right) - \gamma_w w_{t-1} - \gamma_0 - \widehat{U}_t(E, G_t, X_t, w_{t-1})}{\sigma_\eta}\right) \\
& \times \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left\{\begin{array}{c} W_t\left(E, G_t, X_t, T_t, \kappa\left(E, G_t, X_t - 1, T_t - 1, w_{t-1}\right) + \sigma_u \Phi^{-1}\left(\frac{i}{n}\right)\right) \\ < W_t\left(E, G_t, X_t, 0, \kappa\left(E, G_t, X_t, 0, w_t\right)\right) \end{array}\right\}
\end{aligned}$$

and

$$\ell\left(w_t | E, G_t, X_t, T_t, w_{t-1}, d_{t-1} > 0\right) = \frac{1}{w_t} \frac{1}{\sigma_{\kappa_0}} \varphi\left(\frac{\kappa\left(E, G_t, X_t, 0, w_t\right)}{\sigma_{\kappa_0}}\right) \qquad (6)$$

Job stayers: The probability of keeping the same job given a new wage $w_t$is:

$$\begin{aligned}
& \Pr\left\{d_t = 2 | E, G_t, X_t, T_t, w_{t-1}, d_{t-1} > 0, w_t\right\} \\
= \; & \Pr\left\{W_t\left(E, G_t, X_t, T_t, \kappa\left(E, G_t, X_t, T_t, w_t\right)\right) > \max\left(\begin{array}{c} U_t(E, G_t, X_t, w_{,t-1}, \widetilde{\eta}) \\ W_t(E, G_t, X_t, 0, \widetilde{\kappa}_0) \end{array}\right)\right\} \\
= \; & \Phi\left(\frac{W_t\left(E, G_t, X_t, T_t, \kappa\left(E, G_t, X_t, T_t, w_t\right)\right) - \gamma_w w_{t-1} - \gamma_0 - \widehat{U}_t(E, G_t, X_t, w_{t-1})}{\sigma_\eta}\right) \\
& \times \int_0^1 \mathbf{1}\left\{W_t\left(E, G_t, X_t, T_t, \kappa\left(E, G_t, X_t, T_t, w_t\right)\right) \geq W_t\left(E, G_t, X_t, 0, \sigma_{\kappa_0} \Phi^{-1}(u)\right)\right\} du \\
\approx \; & \Phi\left(\frac{W_t\left(E, G_t, X_t, T_t, \kappa\left(E, G_t, X_t, T_t, w_t\right)\right) - \gamma_w w_{t-1} - \gamma_0 - \widehat{U}_t(E, G_t, X_t, w_{t-1})}{\sigma_\eta}\right) \\
& \times \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left\{W_t\left(E, G_t, X_t, T_t, \kappa\left(E, G_t, X_t, T_t, w_t\right)\right) \geq W_t\left(E, G_t, X_t, 0, \sigma_{\kappa_0} \Phi^{-1}\left(\frac{i}{n}\right)\right)\right\}\right)
\end{aligned}$$

and the density of that new wage is:

$$\ell\left(w_t|E, G_t, X_t, T_t, w_{t-1}, d_{t-1} > 0\right) = \frac{1}{w_t}\frac{1}{\sigma_u}\varphi\left(\frac{\kappa\left(E, G_t, X_t, T_t, w_t\right) - \kappa\left(E, G_t, X_t - 1, T_t - 1, w_{t-1}\right)}{\sigma_u}\right) \quad (7)$$

B Estimation Results

Table 7: Estimation Results

Parameter	Coeff	s.e.
σ_{0A}	0.24	0.0018
σ_{0AA}	0.28	0.0052
σ_{0NA}	0.35	0.0046
σ_{η}	8.3	0.6
σ_{μ}	1e+02	4.3
$\alpha_X, X = 2$, non apprentice	0.18	0.012
$\alpha_X, X = 4$, non apprentice	0.34	0.014
$\alpha_X, X = 6$, non apprentice	0.34	0.03
$\alpha_X, X = 30$, non apprentice	0.35	0.04
$\alpha_X, X = 2$, apprentice	0.17	0.01
$\alpha_X, X = 4$, apprentice	0.28	0.01
$\alpha_X, X = 6$, apprentice	0.28	0.02
$\alpha_X, X = 30$, apprentice	0.37	0.03
$\alpha_T, T = 2$, non apprentice	6.7e-05	0.01
$\alpha_T, T = 4$, non apprentice	0.0012	0.03
$\alpha_T, T = 6$, non apprentice	0.0027	0.3
$\alpha_T, T = 30$, non apprentice	0.025	0.3
$\alpha_T, T = 2$, apprentice	0.00067	0.01
$\alpha_T, T = 4$, apprentice	0.0012	0.02
$\alpha_T, T = 6$, apprentice	0.0017	0.4
$\alpha_T, T = 30$, apprentice	0.032	0.4
α_0	3.26	0.2
α_G, non apprentice	0.0028	0.002
α_G, apprentice	0.0064	0.0016
λ_W	0.52	0.075
$\alpha_E d$	0.37	0.018
α_{ε}, Type 2	0.42	0.024
α_{ε}, Type 3	0.7	0.022
$\alpha_{X,\varepsilon}$, Type 2	1.06	0.005
$\alpha_{X,\varepsilon}$, Type 3	0.93	0.004
$\alpha_{T,\varepsilon}$, Type 2	0.94	0.7
$\alpha_{T,\varepsilon}$, Type 3	0.87	1.5

Table 8: Estimation Results, Cont.

Parameter	Coeff	s.e.
π_A	0.11	0.0027
π_{NA}	0.075	0.0024
π_{AA}	0.051	0.0031
π_{AU}	0.21	0.0033
π_{NAU}	0.19	0.0034
π_{AUs}	0.0001	0.00013
π_{NAUs}	0.0012	0.00019
γ_{0A}	-1.5e+02	8.8
γ_{0NA}	-2.1e+02	10
γ_{0Ab}	96	4.8
γ_{0NAb}	5	1.3
δ_A	0.17	0.0067
δ_{NA}	0.12	0.0033
δ_{AA}	-0.012	0.00081
δ_{NNA}	0.015	0.00085
λ_0, region 1	1.7e+03	14
λ_0, region 2	2.2e+03	4.1
λ_0, region 3	1.7e+03	1.3e+02
λ_0, region 4	2.1e+03	1.7
λ_0, region 5	2.2e+03	1.2e+02
λ_0, region 6	1.4e+03	16
λ_0, region 7	1.6e+03	20
λ_0, region 8	1.5e+03	1.4e+02
λ_0, region 9	1.2e+03	1.3e+02
λ_0, region 10	1.6e+03	3.2
λ_0, region 11	1.3e+03	3.9
$\lambda_{0,\varepsilon}$, Type 1	0	-
$\lambda_{0,\varepsilon}$, Type 2	-6.2e+02	1.1e+02
$\lambda_{0,\varepsilon}$, Type 3	2.1e+02	1e+02