Career Progression and Formal versus on the Job Training

J. Adda[*], C. Dustmann[†], C. Meghir[‡], J.-M. Robin[§]

December 8, 2004

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. 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.

[*]University College London and IFS

[†]University College London, IFS and CEPR

[‡]University College London and IFS

[§]Université Paris-I 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.

1 Introduction

Human capital can be accumulated in many different ways and at different points in the life cycle. For example on the job training can be to some extent a substitute for formal education.

There has been a proliferation of active labour market and welfare to work policies in both the US and Europe. Key examples are the programs in Sweden and the UK (New Deal) as well as the Working Families Tax credit in the UK and the EITC in the US. [1] However the existence of these programs can have a potentially profound effect on the entire life cycle accumulation of Human Capital both in terms of early education choices as well as in terms of labour market careers as pointed out by Cossa et al. (1999).

Welfare to work programs may well encourage on the job training at the expense of formal education. In addition wage floors offered by these kinds of programs may discourage job mobility and change the nature of matching in the labour market.

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.

[1] Such programs have been evaluated usually ex post and many such examples can be given, such as Heckman et al. (1997) on the US job training Partnership Act (JTPA), Sianesi (2002) for the Swedish programs, Blundell et al. (2003) for the New Deal and Eissa and Liebman (1996) for the EITC are but a few examples.

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.

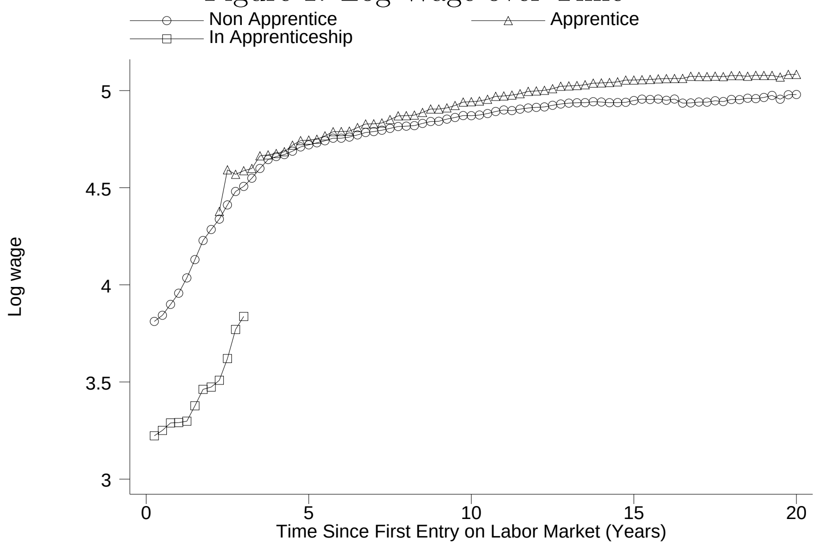
In total, we follow 27525 individuals through time, quarter after quarter up to 1995. In total, we have 996 872 observations on wages, transitions and education choices. The average age at first observation is 16.7. The oldest individual in our data is 35 years old.

2.2 Descriptive Data

2.2.1 Wage Profile

Figure 1 displays the log wage profile as a function of years of labor market experience for apprentices and non apprentices. Individuals with an apprenticeship training have on average higher wages, but a flatter wage profile. In contrast, non apprentices start at a low wage and experience a rapid wage growth, but the wage gap never closes.

Figure 1: Log Wage over Time



2.2.2 Labour Market Transitions

Apprentices are also more likely to work as shown in Figure 4. This is especially true in the first years in the labor market. Figure 5 shows the proportion of workers exiting from the labor market as a function of labor market experience and education. Similarly, apprentices are more likely to

Figure 2: Accumulated Work Experience and Time

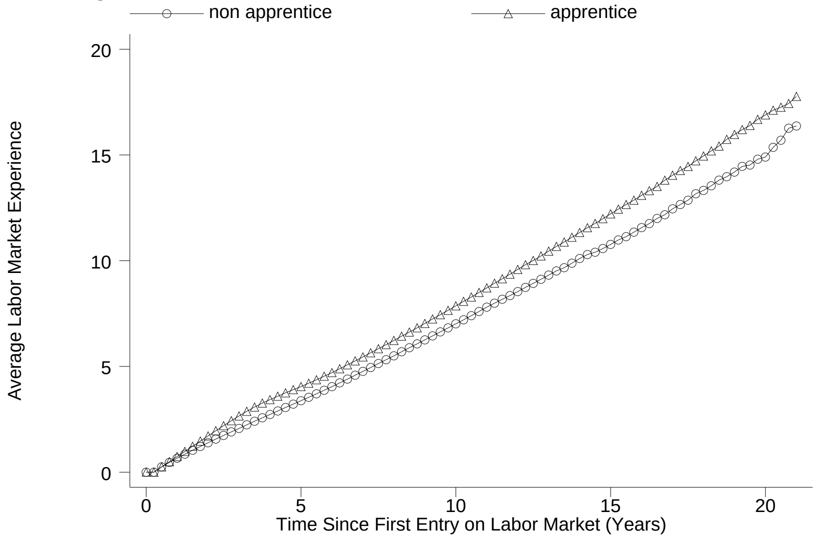
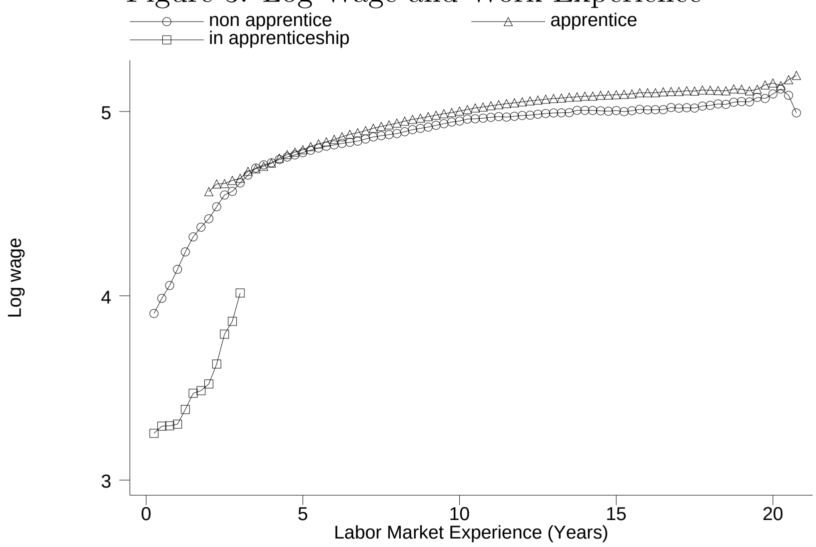
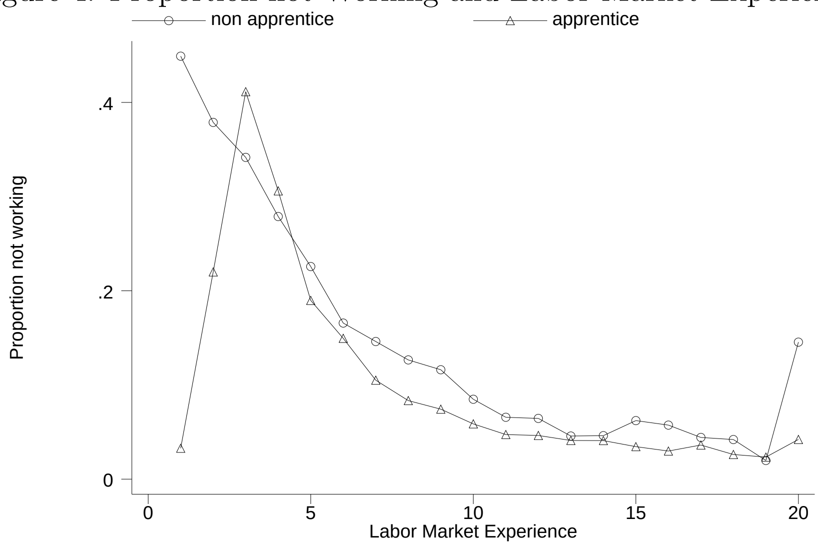


Figure 3: Log Wage and Work Experience



re-enter the labor market, conditional on not working. After ten years of experience, both education groups appear to have the same re-entry behavior.

Figure 4: Proportion not Working and Labor Market Experience



The two groups also differs by the number of jobs hold over time. Non apprentices are much more mobile, going from firm to firm, especially in the first years (Figure 7).

2.2.3 Decomposing Wage Growth

Next, we try to decompose the wage growth into different components. Figure 9 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 8 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 years. However, the results in both Figures are potentially biased, because

Figure 5: Probability of Exit from Labor Market

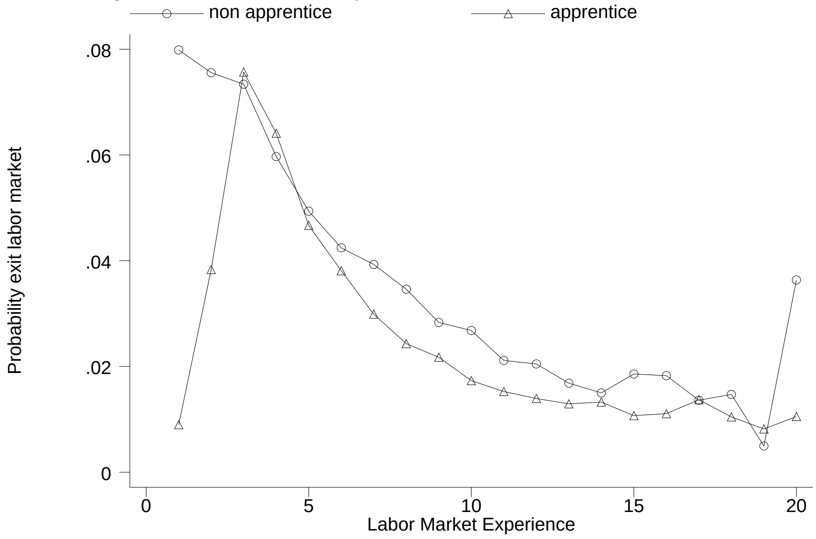


Figure 6: Probability of Entry in Labor Market

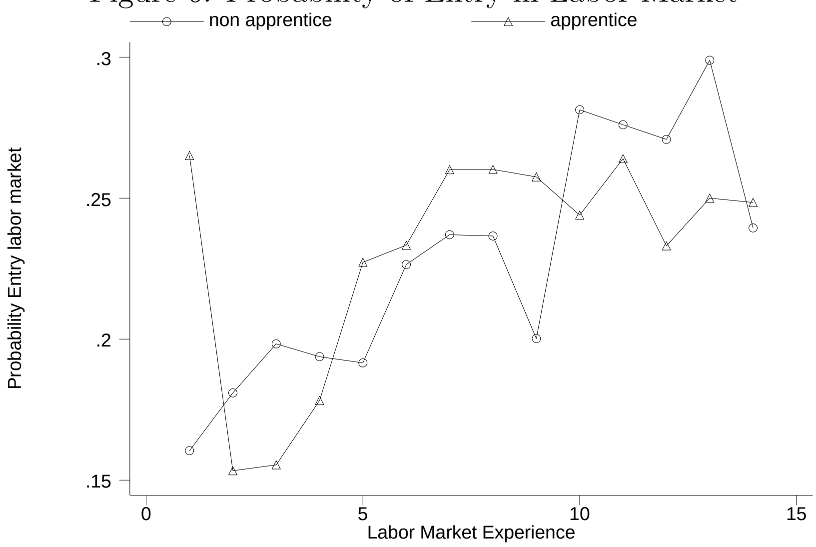


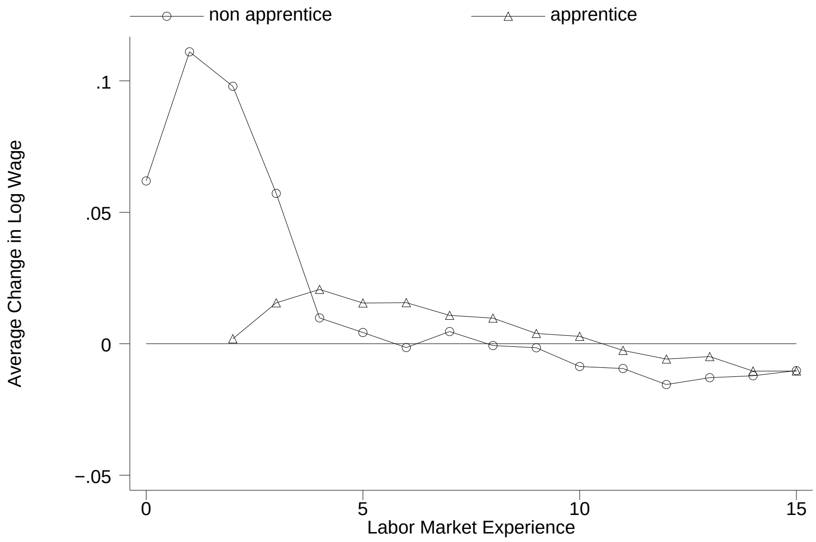
Figure 7: Cumulative Number of Jobs and Labor Market Experience

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

mobility may be endogenous. Our model will be able to disentangle the selection effect from the determinant of wage growth.

Figure 8: Annual Changes in Log Wage (Within) and Labor Market Experience



2.2.4 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. 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.

Figure 9: Changes in Log Wage (Between) and Labor Market Experience

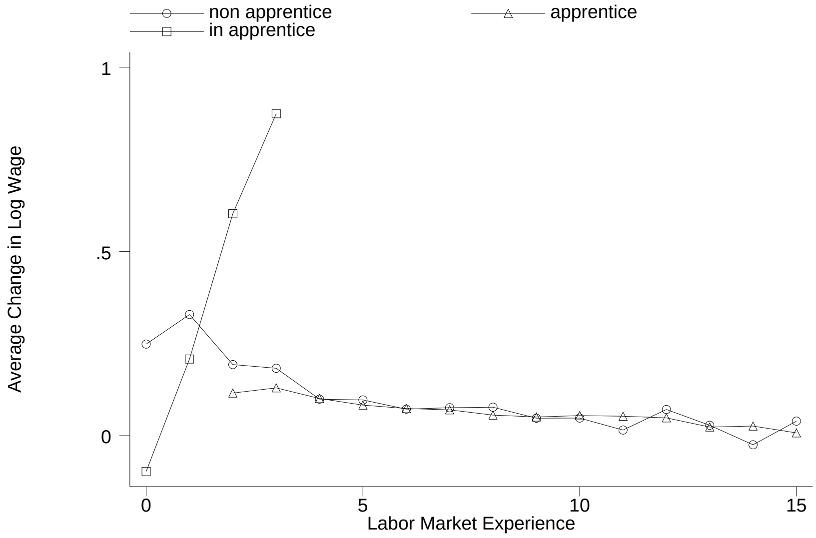


Table 2: Local Wage Effects on Apprenticeship Decision. Marginal Effects

Variable	Marginal Effect	s.e.	t-stat
Local wage Apprentice	.128393	.051	2.49
Local wage Non Apprentice	-.0365127	.025	-1.41
Region 2	.0225809	.023	0.92
Region 3	-.0423446	.029	-1.53
Region 4	.0170435	.021	0.79
Region 5	.0275428	.027	0.94
Region 6	-.0454237	.022	-2.11
Region 7	.0165816	.021	0.76
Region 8	.0243119	.021	1.09
Region 9	-.0119084	.021	-0.55
Region 10	.0569074	.018	2.79
Region 11	-.1495905	.034	-4.97

Note: The regression also controls for time effects.

3 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.

3.1 Instantaneous Rewards

When a worker and a firm meet, they draw a match specific effect constituted by a monetary part, κ_t, and a non monetary part μ. We allow the monetary part of the match to be time varying and persistent through time, as it follows a random walk:

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

The match (κ_t, μ) is dissolved when the firm and the worker separate, and a new match (κ_0', μ') is drawn anew for the next match. We assume that $\kappa_0 \sim \mathcal{N}(0, \sigma_0^2)$ and $\mu \sim \mathcal{N}(0, \sigma_\mu^2)$.

The instantaneous utility of a worker is defined as his wage plus the non pecuniary match:

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

Wages are skill-specific functions of the macroeconomic environment G_t, experience X_t, tenure T_t, of the current value of a match-specific effect $\kappa(T_t)$ and of unobserved heterogeneity, denoted by ε.

While unemployed, the individual derives a utility from unemployment benefits calculated as a fraction of the last wage. In addition, there is a utility of leisure γ_0^ε(E,X), which varies across individuals, education choices, time on the labor market and with the unobserved heterogeneity ε and an i.i.d shock *eta*:

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

3.2 Employment transitions

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

The value of unemployment is defined recursively as:

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

The variable with a prime denote next period values. β is the discount factor. With a probability $\pi_U(E,X)$, the individual gets a job offer $(\tilde{\kappa}_0',\tilde{\mu}')$ next period. He then decides whether to accept the offer or to decline it and wait until the next period. If the offer is accepted, the worker starts with a zero tenure in the new firm. If the individual does not receive a job offer, then he stays for one more period in unemployment.

The value of working is defined as:

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

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 $\pi_W(E)$. The outside offer is a pair $(\tilde{\kappa}_0,\tilde{\mu})$, to be compared with the next period match with the current employer $(\kappa+u',\mu)$. 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 go into unemployment.

The experience evolve as:

$$\begin{aligned}&X' = X+1 &&\text{if the individual is working}\\ &X' = X &&\text{if not}\end{aligned}$$

Tenure evolve as:

$$\begin{array}{ll} T' = T + 1 & \text{if the individual is working in the same firm} \\ T' = 0 & \text{if not} \end{array}$$

We consider an economy subject to exogenous macro 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.

3.3 Educational choice

After finishing high school, an individual can work in a regular job or start as an apprentice. Apprenticeship lasts τ^A periods. The initial educational decision is conditioned by the the macro environment G. Its effect is identified if we follow several cohorts of individuals.

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 of its productivity. In addition, we do not allow individuals to experience unemployment during apprenticeship:

- for $X < \tau^A$,

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

  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 stay on for one more period and accumulate experience and tenure within the firm.

- for $X = \tau^A$, apprenticeship is finished at the end of the period:

$$W_A^\varepsilon(G, X, T, \kappa, \mu) = \lambda_A \cdot w^\varepsilon\left(0, G, \tau^A, T, \kappa\right) + \mu$$

$$+\beta\pi_A \mathrm{E}_{(G',\eta',u',\widetilde{\kappa}'_0,\widetilde{\mu}')} \max \begin{pmatrix} U^\varepsilon(1,G',X',0,\eta') \\ W^\varepsilon(1,G',X',0,\widetilde{\kappa}_0,\widetilde{\mu}) \\ W^\varepsilon(1,G',X',T',\kappa+u',\mu) \end{pmatrix}$$
$$+\beta(1-\pi_A)\mathrm{E}_{(G',\eta',u')} \max \begin{pmatrix} U^\varepsilon(1,G',X',0,\eta') \\ W^\varepsilon(1,G',X',T',\kappa+u',\mu) \end{pmatrix}$$

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

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

where λ_0 is a cost of going into apprenticeship. We allow that parameter to differ with unobserved heterogeneity and with 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_μ^2.

3.4 Estimation Method

The model is solved numerically using a value function iteration technique. The model is estimated by maximum likelihood. We refer the reader to the appendix for the likelihood function. The estimation was done at a quarterly frequency, using a random sample of 1635 individuals, totaling about 77137 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.

4 Results

4.1 Fit of the Model

4.2 Results [TO BE UPDATED. OLD STUFF]

In total we have 49 parameters. The results are presented in Table 3 in appendix B. The results are best described by a series of graphs.

Figure 12 plots the proportion of each type in the sample together with the wage fixed effect ($\alpha_0 + \alpha^\varepsilon$). The estimation distinguish between a large

Figure 10: Goodness of Fit, Wages

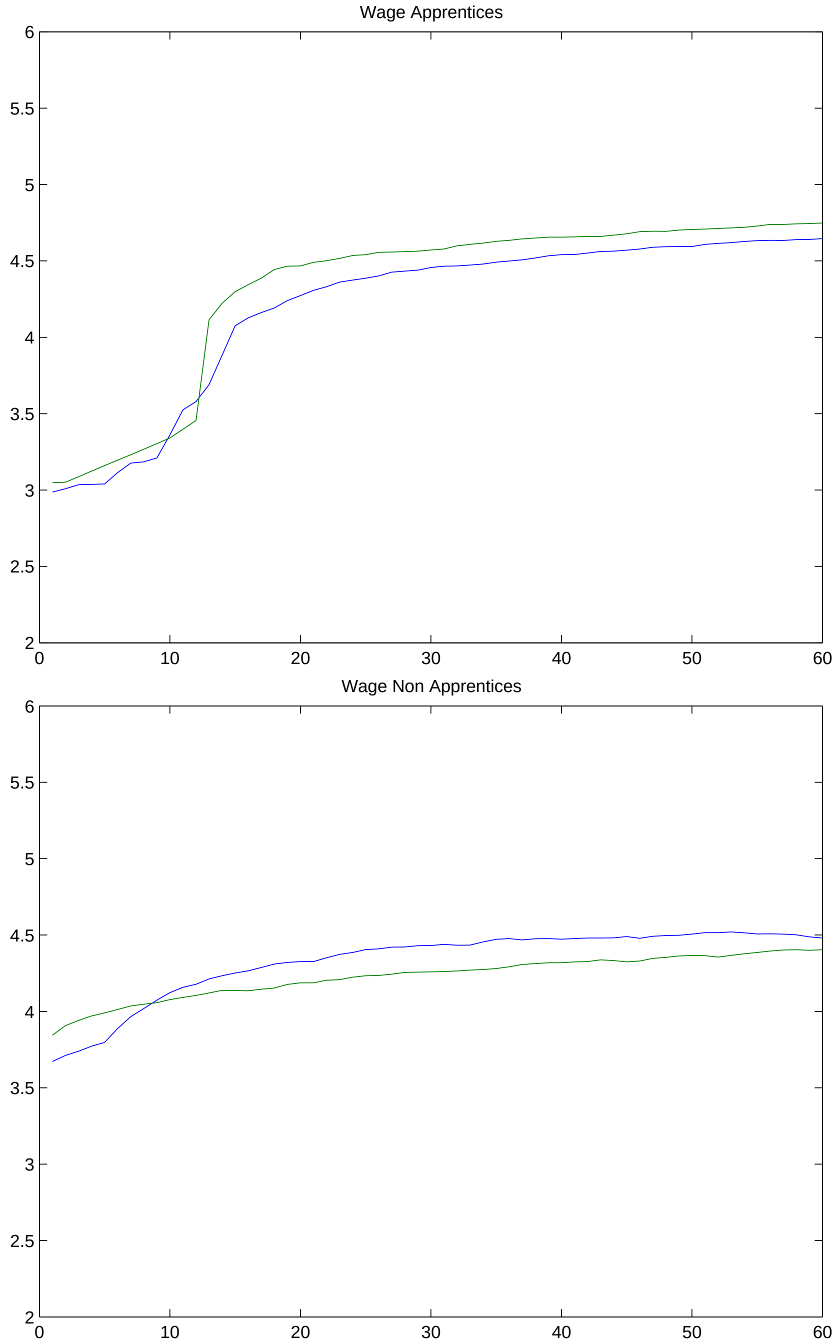
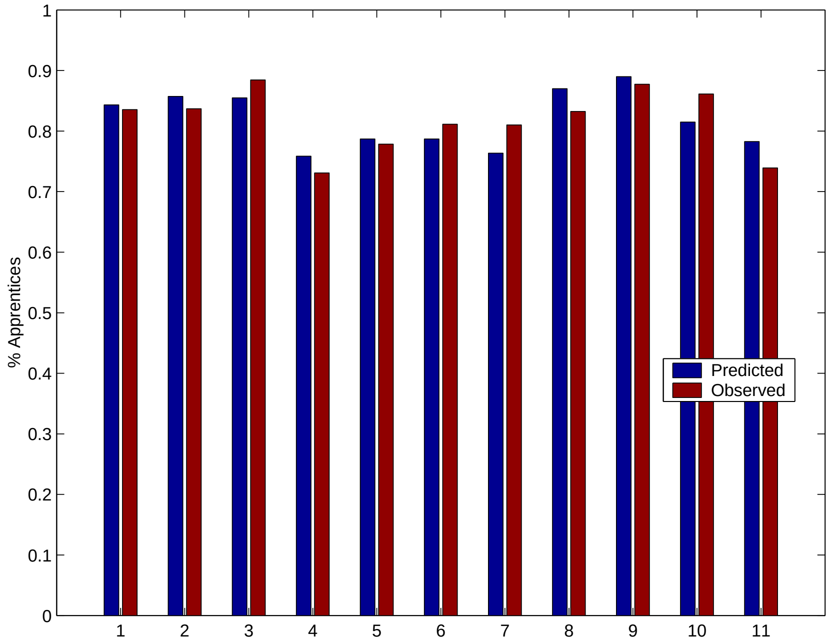


Figure 11: Goodness of Fit, Education Choices by Region



group (56%) with a medium level of wages and two smaller groups (14% and 36%) with respectively lower and higher wages (conditional on observables).

Figure 13 displays the average log wage for apprentices and non apprentices, conditional on unobserved heterogeneity and no mobility. It cumulates both the return to experience and to tenure.

Both apprentices and non apprentices see a sizable wage growth within the firm, at about 8% per year for apprentices and 7% per year for non apprentices. Note that most of the wage growth occurs within 4 years after finishing basic schooling. There is very little wage growth after that period. However, even after many years, there is still a gap between education groups suggesting that the value of training is positive.

We now decompose the return to experience and to tenure. Figure 14 displays the return to experience, conditional on ability and on staying forever in the same firm (with a zero match specific effect), for apprentices and non apprentices. In both cases, the return to experience is steep in the first 3-4 years and then flat. The return to experience is larger for apprentices than non apprentices (respectively about 6% and 4% per year).

Figure 15 displays the return to tenure for apprentices and non appren-

Figure 12: Unobserved Heterogeneity

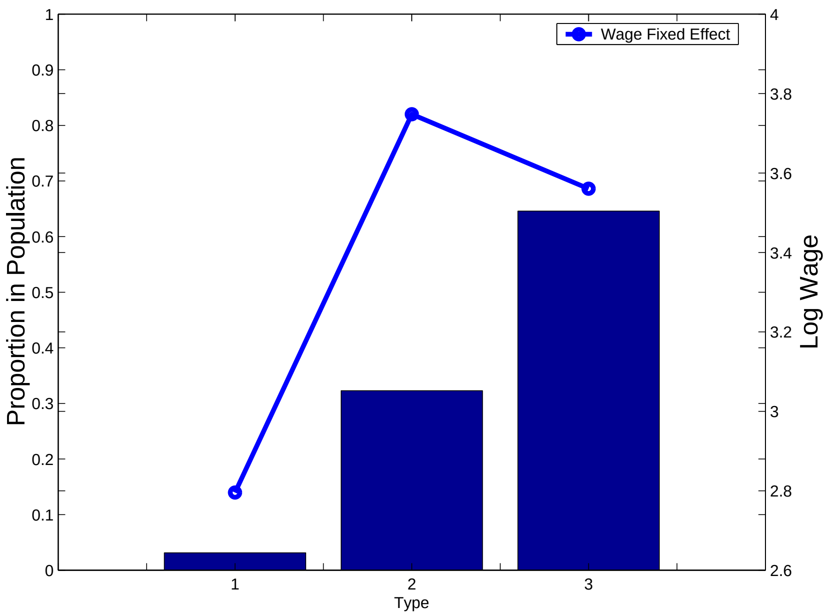


Figure 13: Wage Growth Within Firm

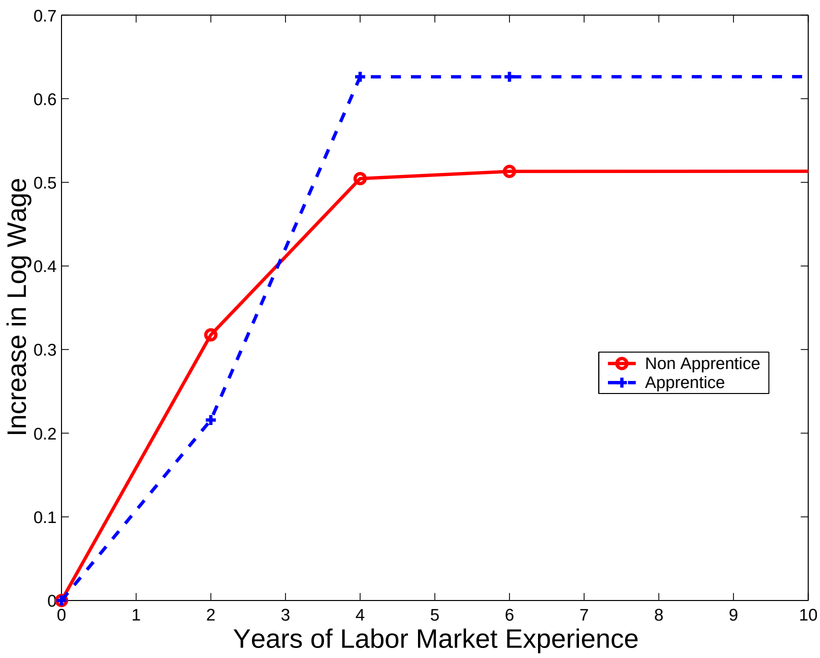
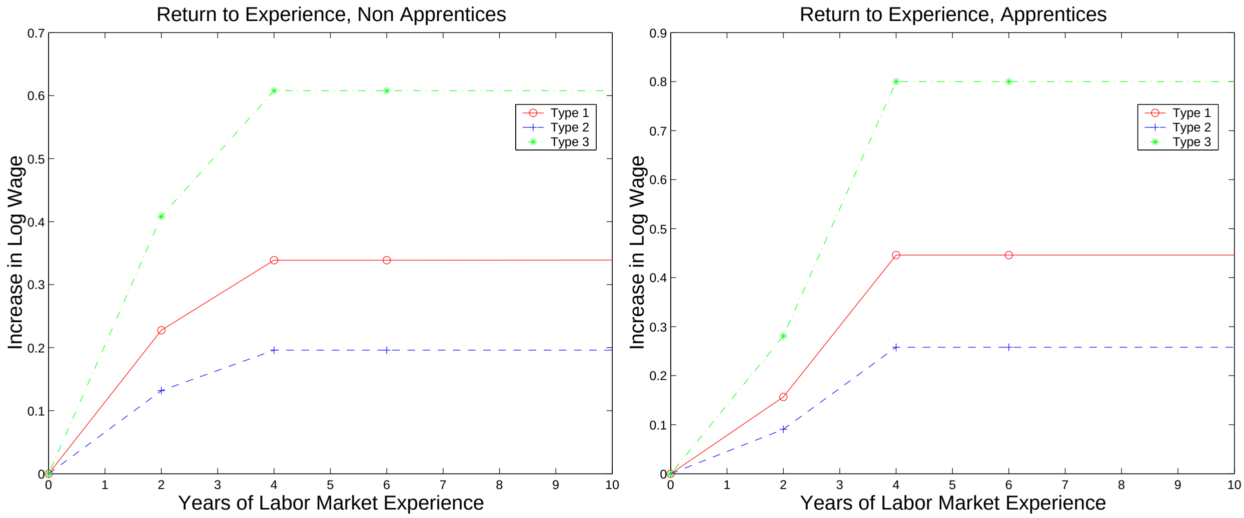


Figure 14: Effect of Labor Market Experience on Wages, Conditional on Ability



tices. These are lower than the return to experience, at about 3% per year for non apprentices and 2% for apprentices. Note also that the return to tenure is almost zero in the first two years.

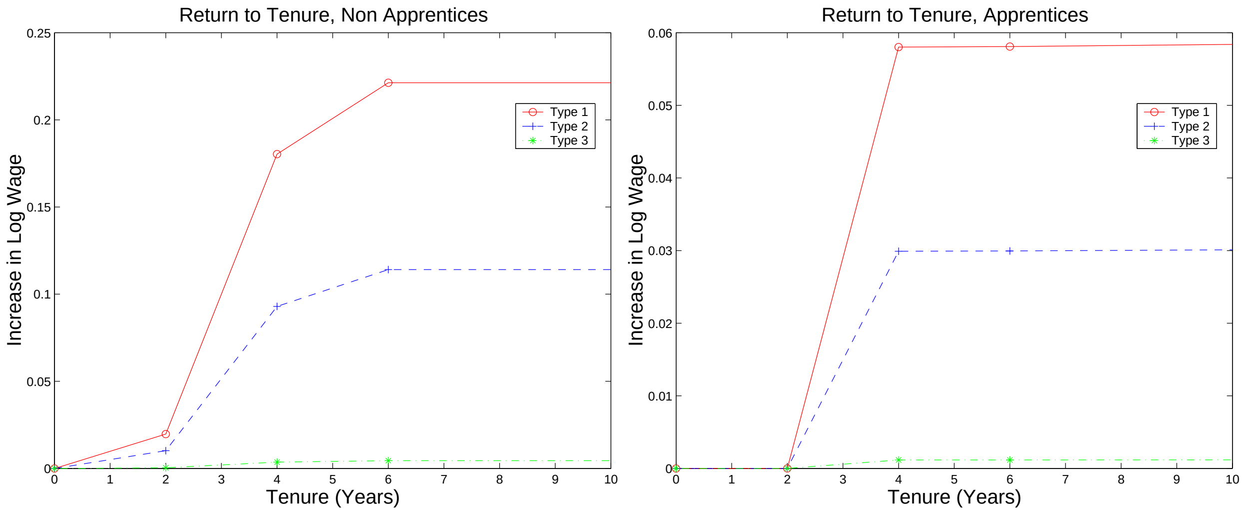
Figure **??** displays the probability of apprenticeship training by region and by ability. High ability individuals are more likely to go through apprenticeship. This explains partly the differences in wages, but not entirely as shown in Figure 14. The probabilities also vary by region, reflecting different costs.

5 Policy Evaluations [TO BE EXPANDED AND UPDATED]

- Separate simulation of EITC by education
- Impact of reform on education
- how high should the return to education be be to offset the effect of EITC?

In this section, we evaluate the effect of labor market policies on career

Figure 15: Effect of Tenure on Wages, Conditional on Ability



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.

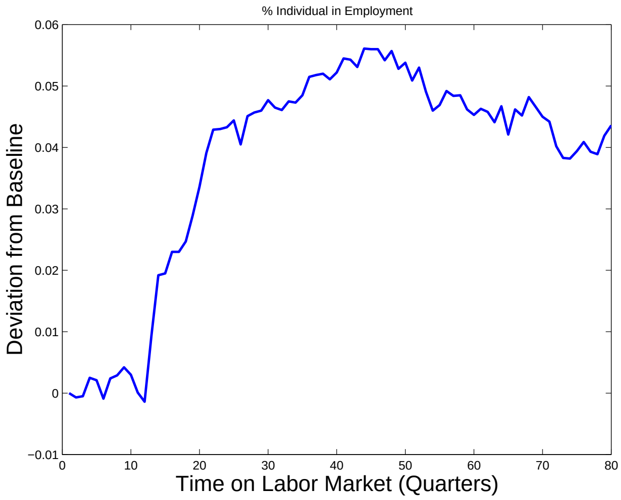
We simulate a reform similar to the EITC, where low wage individuals get a subsidy. This subsidy starts at 0 for a zero wage, increases with the wage up to a first limit, stays constant over a range of income and finally declines to zero. Hence, two categories of individuals do not receive a subsidy: 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 16 to 18 show the results of the policies simulated from the estimated model. The policy has a positive impact on labor participation. It raises participation by about 3 to 4%. However, the policy has a negative impact on the match between firm and workers. The match is about 2% lower. Finally, the in-work benefit decreases the incentive to undertake further education. The proportion of individuals going into apprenticeship decreases by about 4%.

but a negative one on overall productivity.

Figure 16: Effect of In-Work Benefits on Labor Market Participation



6 Conclusion [to be written]

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.

Figure 17: Effect of In-Work Benefits on Firm-Worker Match

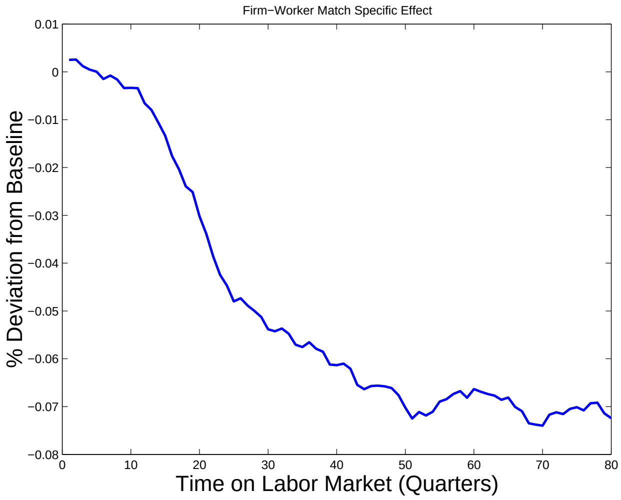


Figure 18: Effect of In-Work Benefits on Education Choices

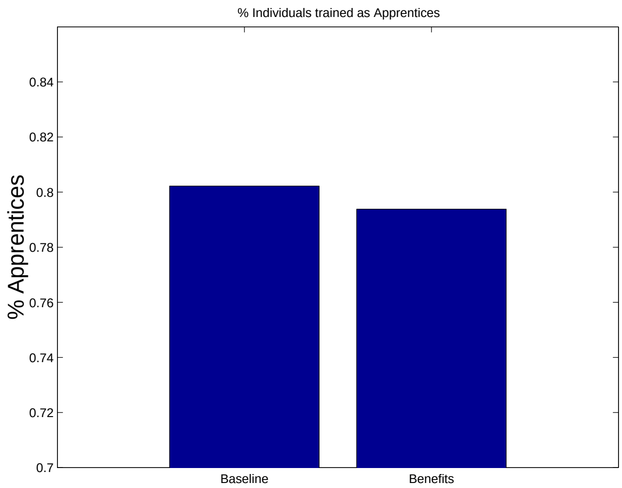
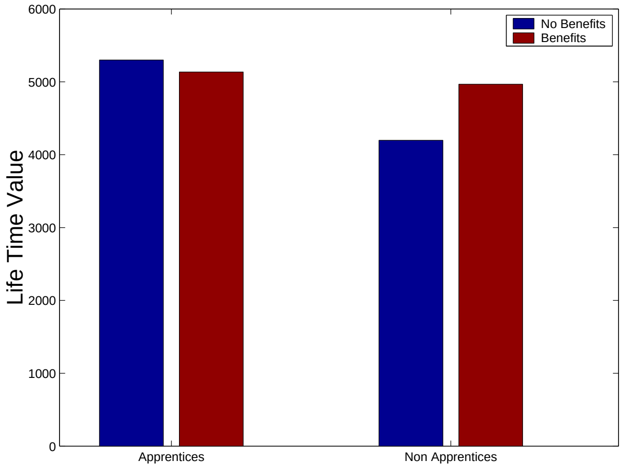


Figure 19: Effect of In-Work Benefits on Life Time Value



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

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

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

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

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

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

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:

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

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 :

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

and

$$
\ell\left(w_t^A|G_t\right) = \frac{1}{w_t^A}\frac{1}{\sigma_{\kappa_0}}\varphi\left(\frac{\kappa\left(0,G_t,t,0,\frac{w_t^A}{\lambda_w}\right)}{\sigma_{\kappa_0}}\right) \qquad (2)
$$

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:

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

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

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

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

The density of the accepted wage is:

$$
\ell\left(w_t|E, G_t, X_t, T_t, w_{t-1}, d_{t-1} = 0\right) = \frac{1}{w_t}\frac{1}{\sigma_{\kappa 0}}\varphi\left(\frac{\kappa\left(E, G_t, X_t, 0, w_t\right)}{\sigma_{\kappa 0}}\right). \tag{4}
$$

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

$$
\Pr\left\{d_t = 0|E, G_t, X_t, w_{-1,t}, d_{t-1} = 0\right\}
$$

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

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

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

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

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

$$\times\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\left\{\begin{array}{c} W_t\left(E,G_t,X_t,T_t,\kappa\left(E,G_t,X_t-1,T_t-1,w_{t-1}\right)+\sigma_u\Phi^{-1}\left(\frac{i}{n}\right)\right) \\ < W_t\left(E,G_t,X_t,0,\kappa\left(E,G_t,X_t,0,w_t\right)\right)\end{array}\right\}$$

and

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

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

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

and the density of that new wage is:

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

B Results

Table 3: Estimation Results

Parameter	Coeff	s.e.
σ_U	0.14	0.00012
σ_0	0.6	0.0075
σ_η	34	0.53
σ_ω	2.4e+02	41
α_0	2.8	0.041
α_G, non apprentice	0.0034	0.0066
α_G, apprentice	0.012	0.002
α_{Ed}	0.00013	0.026
π_A	0.21	0.0042
λ_W	0.8	0.23
$\alpha_X, X = 2$, non apprentice	0.32	0.035
$\alpha_X, X = 4$, non apprentice	0.06	0.028
$\alpha_X, X = 6$, non apprentice	6.6e-06	0.031
$\alpha_X, X = 30$, non apprentice	4.4e-05	0.21
$\alpha_X, X = 2$, apprentice	0.16	0.016
$\alpha_X, X = 4$, apprentice	0.45	0.028
$\alpha_X, X = 6$, apprentice	2.2e-06	0.009
$\alpha_X, X = 30$, apprentice	0.0004	0.052
$\alpha_T, T = 2$, non apprentice	0.0012	0.031
$\alpha_T, T = 4$, non apprentice	0.26	0.054
$\alpha_T, T = 6$, non apprentice	0.13	0.082
$\alpha_T, T = 30$, non apprentice	0.00063	0.53
$\alpha_T, T = 2$, apprentice	0.00023	0.0067
$\alpha_T, T = 4$, apprentice	0.19	0.016
$\alpha_T, T = 6$, apprentice	0.0049	0.02
$\alpha_T, T = 30$, apprentice	0.16	0.22
$\alpha_{X,\varepsilon}$, Type 2	0.33	0.067
$\alpha_{X,\varepsilon}$, Type 3	-0.55	0.086
$\alpha_{T,\varepsilon}$, Type 2	-0.23	0.11
$\alpha_{T,\varepsilon}$, Type 3	-0.67	0.16
γ_0	6.9	0.45
α_ε, Type 2	0.15	0.037
α_ε, Type 3	0.72	0.037
λ, Type 2	-2.2	0.58
λ, Type 3	0.49	0.21
$\lambda_{0,\varepsilon}$	-1.5	59
$\lambda_{0,\varepsilon}$	-2.3	41