Education and Training over the Lifecycle: The Causal E®ect of Accumulated Human Capital on Training Opportunities[¤]

Christian Belzil
Centre National de Recherche Scienti¯que,
Concordia University, Cirano, and IZA

JÄrgen Hansen
Concordia University, IZA, and CEPR

June 1, 2004

Abstract

We estimate a structural dynamic programming model of education and training choices made over the lifecycle using a panel of young white males taken from the National Longitudinal survey of Youth (79-95). We examine two competing hypothesis for the explanation of a positive correlation between the incidence of training and both schooling and accumulated training experience in the past; namely that the correlation is explained by unobserved persistent individual speci¯c tastes and abiities or that it is explained by a true causal e®ect of accumulated human capital on future training incidence. Our results indicate that individual skill endowments explaining grade attainment are strongly (positively) correlated with skills and tastes for on-the-job training. We ¯nd that both accumulated on-the-job and o®-the-job training increase the probability of receiving training in the future. However, given individual endowments, reaching a higher grade level reduces the probability of receiving on-the-job training but increases the incidence of o®-the-job training. This is consistent with the hypothesis that ¯rms may view on-the-job training as a substitute for education.

Key words: Training, Education, Human Capital, Dynamic Programming,
JEL Classi¯cation: J2-J3

[¤]We would like to thank Peter Arcidiacono and Thierry Magnac for helpful comments. We also thank the Social Sciences and Humanities Research Council of Canada for generous funding. The usual disclaimer applies.

1 Introduction

The positive effect of training on labor market productivity is a central prediction of human capital theory (Becker, 1964 and Mincer, 1974). While various behavioral models have been advanced to explain the prevalence of upward wage profiles, human capital theory remains the most popular. In a standard Mincerian framework, individuals sacrifice present consumption in order to accumulate "skill units". These units, although intrinsically unobservable, are assumed to be correlated with schooling and post-schooling experience, through a production function. These assumptions have lead to the popular mincerian wage regression, in which the effects of education and accumulated experience is separable.

In actual data, the independence between education and wage growth is rarely verified. Indeed, for several decades, economists have observed that the age earnings profiles of those who are more educated tend to be steeper than those of low educated people (Mincer 1974). This is often explained by the conjunction of the existence of post-schooling training opportunities and heteroginity in abilities and costs. The argument may be justified if, for example, tastes for schooling/academic abilities are negatively correlated with the innate cost of receiving training. If so, those who invest in schooling are also more likely to invest in on-the-job training.

An alternative explanation for this correlation may be that accumulated schooling increase the incidence of training opportunities, even after conditionning on innate abilities. This may be justified if the marginal cost of training is decreasing with schooling (after conditiioning on innate abilities) or if education magnifies the increase in productivity (the return to training). In a context where actual post-schooling human capital investments are proxied by measured experience, this suggests that log wages regression may not be separable in education and experience and, in particular, that the return to experience may be affected by schooling.[1]

While a positive correlation between education and training is plausible, it is not the only possibility. In practice, training decisions are jointly decided by workers and firms. In an environment where resources devoted to training are scarce, firms may prefer to train the low educated if the marginal benefit of training the low educated is higher than the highly educated workers. Under such a scenario, training may be viewed as a substitute for schooling. The sign of the correlation between education and training is therefore ambiguous.

In the empirical literature on training, it is customary to report a positive correlation between training and education as well as a certain degree of persistence in the individual incidence of training (see Lynch 1992 and Altonji and Spletzer,

[1] Issues related to non-separability are discussed in Heckman, Lochner and Taber (forthcoming) and Lemieux (2003). The non-separability hypothesis is tested formally in a context where schooling is endogenous in Belzil (2004).

1991). However, one should be reluctant to give a structural interpretation to the correlation between training and education and.between training and past traning incidence. As indicated above, the measured correlation may be explained by the fact that preferences, prices or other constraints a®ecting future training decisions are directly a®ected by the occurrence of training and/or education as well as by unobserved di®erences, correlated over time and improperly treated, which create a spurious correlation between future and past experience (Heckman, 1981). Indeed, the distinction between true and spurious state dependence is central to several empirical issues related to the labor market. As of now, a thorough review of the literature reveals that it is impossible to establish whether the correlation between training and schooling is causal or spurious.[2]

We believe that investigating the existence of a structural (causal) e®ect of education on training (and quantifying this e®ect) is an important issue for two main reasons. First, it may help understand the empirical correlation between wage growth and schooling. Second, structural estimates of the causal e®ect of education on the incidence of post-schooling training may help evaluate the e®ectiveness of various education policies aimed at increasing high school graduation rates or college attendance. Indeed, disregarding the potential increase in training opportunities caused by schooling may lead to an under-statement of the economic bene¯ts of these policies.

This is precisely the issue addressed in this paper. in what follows, we investigate the empirical correlation between the incidence of training and accumulated human capital. We estimate a dynamic model of education and training choices over a ¯nite horizon. For various reasons related to the inherent di±culties of measuring training, we only model the incidence of training.[3] In our model, the intertemporal utility of choosing a particular option is function of initial individual endowments, which have an observable component (proxied by parents background variables and Armed Force Quali¯cation tests) as well as an observed components, and also depends on accumulated human capital (accumulated years of schooling, accumulated years of on-the-job training and accumulated years of o®-the-job training). The dependence of the utility of choosing training on accumulated human capital (say schooling or past training) may be explained by the fact that accumulated human capital reduced the marginal costs of training or that, other things equal, employers who o®er training opportunities, tend to favor those who have accumulated more human capital (conditional on tastes and abilities). The model is therefore able to quantify the portion of the correlation between training and education which is explained by sorting (correlated tastes and abilities) and the portion of the correlation explained by structural dependence. It is also able to o®er a similar decomposition of the persistence in

[2] A similar conclusion would apply with respect to the correlation between wage growth and education (Belzil, 2004).

[3] The problems encountered in measuring training intensity are discussed in Barron, Berger and Black (1977).

lifecycle training decisions.

2 The Data

2.1 Training Data in the NLSY

We use the 1979 cohort of the NLSY and restrict ourselves to white males. As is well known, the NLSY has relatively comprehensive information on education, employment and training. The NLSY is therefore most appropriate for analyzing the causal link between education and training. Respondents are asked about what types of training they had received. The different types of training are separated in three categories: company training (on-the-job training), apprenticeships and training obtained outside the firm (off-the-job training). The off-the-job training category includes business courses, barber and beauty schools, vocational institutes, nursing programs and correspondence courses. Despite its name, the incidence of off-the-job training does not require current employment. Our definitions are quite standard, for instance they are the same as those used in Lynch (1991). Inasmuch as it is natural to associate on-the-job training to firm specific training and off-the-job training to general training, the distinction between general and specific training does not play a central role in our analysis. This is because we are not modeling job mobility.

While the information regarding the type of training is detailed, the measure of training intensity is far from being perfect. Before 1988, the NLSY specifies both starting and ending dates of all training spells that lasted at least one month. After 1988, all spells are reported. very short spells of training are therefore likely to be under-reported before 1988. Furthermore, as the NLSY does not report actual hours of training per week, it is not possible to measure actual training duration (or intensity) in a meaningful fashion. For this reason, we decided to focus on the incidence of training.

2.2 Construction of the Sample

In this section, we document all steps undertaken in order to construct the sample data analyzed in this paper and explain how each state has been defined

- As we need to observe the full realization of the incidence of training for every individuals, we need to focus on individuals who, most likely, could not have received training before 1979. For this reason, we selected white males aged between 14 and 16 years old in 1979. This is a sample very close to the sample analyzed by Eckstein and Wolpin (1999).

- As a second step, we kept the individuals for whom we had non-missing observations for the most important measured characteristics (parents' education, income, # of siblings, presence of both parents at age 14, rural/urban indicator and armed Forces Quali¯cation Test (AFQT) scores. These characteristics are standard in the literature. They are the same as those used in various studies such as Cameron and Heckman (1998 and 2000) and Belzil and Hansen (2002,a and 2002,b). In total, we obtained a sample of 667 individuals.

- In order to control for the fact that individuals might have taken the AFQT at a di®erent ages (and di®erent schooling levels), we use a corrected measure. This corrected measure is based on the residual of a OLS regression of AFQT scores on age and education. This is common in the literature (Cameron and Heckman, 1998).

- The individual histories are described as a sequence of mutually exclusive states. These states correspond to potential combinations of the potential fundamental choices taken by the individuals in our sample. These fundamental choices include schooling, home production, work, o®-the-job training, apprenticeship and on-the-job training. Given the size of the sample and the very large number of combinations, we decided to group Apprenticeship with on-the-job training. We also chose not to distinguish between schooling and o®-the-job training since both activities are closely related to the notion of "general" training.[4] As we also found a very small number of individuals who report both on-the-job and o®-the job training (only 3), we decided to disregard these individuals. To reduce the number of states, we also decided to group those who work while in school with those who are in school without working into a single group. As a result, we obtain seven potential states.[5] These are the following;

1. School and/or O®-the-job training
2. Home
3. Work (no training)
4. Work/on-the job training
5. Work/o®-the job training

[4] The classical distinction between general and speci¯c training has been strongly questioned in recent years. For more discussions, see Acemoglu and Pischke.

[5] It should be noted that the number of combinations is limited by the fact that some actions are mutually exclusive by construction (school and home production).

2.3 Some Features of the Data

The main features of the data may be found upon looking at Table 2 and Table 3. Overall, training is relatively common, especially around the age of 25-26. At age 26, around 22% of the young individuals report having received some training during that year and on-the-job training appears the dominant form of training (13% having received on-the-job training and 9% having received o®-the-job training). Before 20, o®-the-job training is the dominant form of training. After 20, it is on-the-job training which becomes more common. This is essentially explained by the work patterns of young individuals; namely that the majority of young individuals is still in school at age 18.

2.4 The Correlation between Training and Accumulated Human Capital in the NLSY

As reported in the literature (Lynch, 1992 and Altonji and Spletzer 1991), we also ¯nd that the incidence of training is positively correlated with schooling and that there is a certain degree of persistence in training. This may be veri¯ed upon looking at the results obtained from simple OLS regressions of the propensity to obtain on-the-job training and o®-the-job training in a given year on some measures of accumulated human capital. These are found in Table 4A and Table 4B. The dependent variable is equal to 1 if the young individual has received on-the-job training (OJT) during his last year observed in the sample and 0 if not. Accumulated education, accumulated on-the-job training (OJT), accumulated o®-the-job training (OFT) and accumulated experience are measured at the beginning of the last year of observation and re°ect all past human capital decisions from the age of 14 until the second last year of observation.

Regarding the determinants of on-the-job training (Table 4A), the results indicate that, while there is a positive correlation between schooling and on-the-job training, the positive correlation between receiving training and the amount of training accumulated in the past is much stronger. When accumulated on-the-job training is controlled for, the e®ect of accumulated schooling drops from 0.0109 to 0.0080 and its level of signi¯cance also drops substantially. On the other hand, accumulated on-the-job training has a positive e®ect on the incidence of training and this e®ect remains quite robust (around 0.05) whether or not schooling is accounted for or not. The correlation between the incidence of on-the-job training and accumulated o®-the-job training is very weak.

The main di®erence between the incidence of on-the-job training and o®-the-job training (Table 4B), is the signi¯cant positive e®ect that accumulated education has on o®-the-job training. Accumulated o®-the-job training is also strongly and positively correlated with the incidence of o®-the-job training but the e®ect of accumulated on-the-job training appears insigni¯cant.

To summarize, our data indicate that the positive correlation between education and training is much stronger for o®-the-job training than on-the job training, and that there is a high degree of persistence in both types of training. These are the features of the data that we will now try to explain with our structural model.

3 The Model

The individuals maximize expected lifetime utility by choosing the optimal state over a ¯nite horizon T. Lifetime utility is time additive and there are K mutually exclusive states. The objective function is therefore

$$Max_{\{d_{kt}\}}\ E(\sum_{t=0}^{T}\beta^t \cdot (\sum_{k=1}^{K} U_{kt} \cdot d_{kt}) \mid \Omega_t) \tag{1}$$

where the control variables, d_{kt}, are equal to one when option k is chosen and 0 if not, where U_{kt} denotes the contemporaneous (per-period) utility of choosing option k at age t and β is the yearly discount factor. The information set, at date t, is denoted Ω_t.

The maximum expected value achieved at date t, denoted $V(\Omega_t)$, is given as follows

$$Max_{\{d_{kt}\}}E(\sum_{t=0}^{T}\beta^t \cdot (\sum_{k=1}^{K} U_{kt} \cdot d_{kt}) \mid \Omega_t) = Max_{k\in K}V_k(\Omega_t) = V(\Omega_t) \tag{2}$$

where the alternative speci¯c value functions, $V_{kt}(\Omega_t)$, are given by the following expression,

$$V_{kt}(\Omega_t) = U_{kt} + \beta EV_{t+1}(\Omega_{t+1} \mid d_{kt} = 1) \tag{3}$$

and where $EV_{t+1}(\Omega_{t+1} \mid d_{kt} = 1)$ denotes the value of following the optimal policy in period t+1 .

We follow an approach similar to Cameron and Heckman (2000) and approximate the alternative speci¯c value functions,$V_{kt}(.)$, using a °exible (quadratic) functional form. That is the intertemporal utility of choosing a given state k at age t is assumed to be of the following form

$$\begin{aligned} V_{kt} \quad = \quad & X'\beta_{kt} + \psi_{kt}(S_t) + \varphi_{1kt}(EX_t) \\ & +\varphi_{2kt}(OJT_t) + \varphi_{3kt}(OFT_t) + \varphi_{4kt} \cdot H_t + \eta_k + \varepsilon_{kt} \end{aligned} \tag{4}$$

for k=1,2...K, and where the dependence of all the regression parameters $(\beta_{kt}, \varphi_{1kt}, \varphi_{2kt}, \varphi_{1kt})$ and the function(ψ_{kt}) on k and t allows for a maximum degree of °exibility at the estimation level. The variables and parameters are de¯ned as follows,

- S_t is accumulated schooling at age t.

- ² EX_t is accumulated years of experience at age t.
- ² OJT_t is accumulated years in which on-the-Job Training took place.
- ² OFT_t is accumulated years in which o®-the-Job Training took place.
- ² H_t is accumulated years of home time.
- ² The vector X contains household human capital variables which act as proxies for the initial ability/taste endowments. These include mother's education, father's education, family income (as measured in thousands of 1978 dollars), number of siblings, an indicator equal to 1 (Nuclear) for the presence of both parents at age 14 (and 0 if not) and Armed Forces Quali¯cation Test (AFQT) scores.
- ² The function ψ_{kt} captures the structural e®ect of accumulated schooling on the utility of choosing state k (including training).
- ² The functions $\varphi_{1k}(.), \varphi_{2k}(.)$ and $\varphi_{3k}(.)$ capture the structural e®ects of accumulated experience, accumulated on-the-job training and accumulated o®-the-job training on the utility of choosing state k.
- ² The term η_k represents a state speci¯c unobserved heterogeneity term representing individual di®erences in tastes for all relevant combinations of schooling, work, home production and training.

4 Estimation Strategy

In order to estimate the model, some restrictions need to be imposed. These restrictions will re°ect the necessity to keep the number of parameters at a manageable level as well as the necessity to hold the model to a certain level of coherency.

- ² To reduced the number of parameters, we assume that the vector of parameters β_{kt} remains constant over some age intervals. We actually experienced with 2 possibilities. In a ¯rst case, the intervals chosen are 14-19, 20-25 and 26 or more. The second option considered is to have 2 intervals; 14 to 21 and 22 to 30.
- ² Because most individuals are in school in the early phase of the lifecycle, it is practically impossible to allow the e®ects of parents background to vary with age. For this reason, the e®ects of parents background are assumed to be constant for two options; School and School/o®-the-job training. for a similar reason, it is also practically impossible to allow the utility of attending school to depend on accumulated experience and training. The corresponding parameters are therefore set to 0.

- The function $\psi_{kt}(.)$ is estimated flexibly so to mimic a non-parametric regression. With respect to the utility of school (as well as school/off-the-job training), the $\psi_{kt}(.)$ function is estimated using a specific intercept term for each potential grade level. As most people reach their maximum schooling attainment without any interruption, we do not allow for age/grade specific effects.[6] For other choices (training and work), the $\psi_{kt}(.)$ is specified as a spline function with 4 segments; high school dropouts ($S_t < 12$), high school graduates ($S_t = 12$), some college ($12 < S_t < 16$) and college graduate ($S_t > 16$).

- The functions $\varphi_{1k}(.), \varphi_{2k}(.)$ and $\varphi_{3k}(.)$ are assumed to be quadratic.

- As the seventh option, off-the-job training/no work, is only rarely chosen, we disregard the estimation of age specific affects of parents background variables for this option as well.

- We assume that

$$\eta_k = \alpha_{0k} + \eta.\alpha_k \tag{5}$$

where the distribution of η is approximated by a discrete distribution with both marginal distributions having 2 points of support point (η_{11} and η_{12}). The type probabilities are estimated using logistic transforms. In order to obtain identification, we f normalize α_{11} to 1 and α_{01} to 0 (the unobserved taste for schooling)

- The term ε_{kt} represents a pure stochastic i.i.d. shock observed (by the agent) at the beginning of period t. We assume that the cumulative distribution of the $\varepsilon_{kt}^{0}s$ is an Extreme Value of type 1 (i.e.: Prob $(\varepsilon < e)$ = $F(e) = \exp(-\exp(-e)))$

- The final date,T, is set at age 31.

The distributional assumption, coupled with the model structure already laid out, will imply that,

$$\Pr(d_{kt} = 1) = \frac{\exp(\tilde{V}_{kt})}{\sum_{j=1}^{K} \exp(\tilde{V}_{jt})} \tag{6}$$

where

$$\tilde{V}_{kt} = V_{kt} - \varepsilon_{kt} \tag{7}$$

The model is estimated by maximum likelihood techniques. Altogether, the implementation of the model requires estimation of 197 parameters. The type specific likelihood function, L(.;η) is given by

[6] If we did, this would increase substantially the number of parameters to be estimated.

$$L(.\eta) = \prod_{t=1}^{T} \Pr(d_{kt} = 1 \mid \eta)$$

and the unconditional likelihood function is just a weighted average of $L(.\eta)$, that is

$$L(.) = \sum_{i=1}^{I} L(. \mid \eta_i) \cdot p_i$$

5 Predicted Frequencies and Goodness of Fit

Despite the relatively high degree of asymmetry in the actual frequencies between all the possible options (some options are only rarely chosen), our predicted frequencies (found in table 5) indicate that our model is able to fit the data quite well. In particular, we capture the increase in the incidence of on-the-job training from age 22 to age 26 (the peak age for on-the-job training). The incidence of off-the-job training/work and off-the-job training alone are also predicted quite accurately. The incidence of household activities (home) are also quite accurate. Finally, we note that our model predicts the high proportions of young individuals in school until age 17 and the rapid decline in school attendance, although it seems to over-predict slightly school attendance beyond age 23.

6 Some Parameter Estimates

Incomplete

7 Some Preliminary Conclusions

- There is a weak correlation (positive) between Schooling and on-the-job Training Incidence in the data. This correlation is decomposed into
  - a weak positive causal effect of schooling on Training (on-the-job)
  - a weak negative correlation between unobserved abilities/tastes explaining schooling and on-the-job training
- Training, as formally measured, may be irrelevant over the life cycle.
- Open Question: Training and wage growth

References

[1] Altonji, Joseph and James Spletzer (1991) \Worker Characteristics, Job Characteristics, and the Receipt of On-The-Job Training", Industrial and Labor Relations Review, vol 45, no 1, 58-79.

[2] Barron, John, Mark Berger and Dan Black (1997) \How Well do we measure Training?" Journal of Labor Economics, vol 15 (3), 507-528.

[3] Bartel, Ann (1995) \Training, Wage growth and Job Performance: Evidence from a Company Database". Journal of Labor Economics, vol 13 (3), 401-425.

[4] Becker, Gary (1964), Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, New-York: National Bureau of Economic Research

[5] Bellmann, Richard (1959) Dynamic Programming, Princeton, New-Jersey.

[6] Belzil, Christian and Hansen, JÄrgen (2002) \Unobserved Ability and the Return to Schooling", Econometrica, vol 70, no 6, 575-591.

[7] Belzil, Christian and Hansen, JÄrgen (2003) \Structural Estimates of the intergenerational Education Correlation" forthcoming in Journal of Applied Econometrics

[8] Ben-Porath, Yoram (1967) \The Production of Human Capital and the Life Cycle of Earnings" Journal of Political Economy, 75 (4): 352-65.

[9] Blundell, Richard and James Powell (2000) \Endogeneity in Non-Parametric and Semi-Parametric Regression Models" Paper presented at the 2000 World Congress of the Econometric Society"

[10] Card, David (2000) \The Causal E®ect of Education on Earnings" Handbook of Labor Economics, edited by David Card and Orley Ashenfelter, North-Holland Publishers.

[11] Eckstein, Zvi and Kenneth Wolpin (1999) \Youth Employment and Academic Performance in High School", Econometrica 67 (6), 1295-1339.

[12] Heckman, James (1981) \Statistical Models for Discrete Panel Data" in Structural Analysis of Discrete Data with economic Application, edited by Charles Manski and Daniel McFadden, MIT Press.

[13] Heckman, James and E. Vytlacil (1998) \Instrumental Variables Methods for the Correlated Random Coe±cient Model", Journal of Human Resources, 33, (4), 974-987.

[14] Keane, Michael P. and Wolpin, Kenneth (1997) "The Career Decisions of Young Men" Journal of Political Economy, 105 (3), 473-522.

[15] Lynch, Lisa (1992) \Private Sector Training and the Earnings of Young Workers" American Economic Review" vol. 82 (March): 299-312.

[16] Mincer, Jacob (1974) Schooling, Experience and Earnings, New-York: Columbia University Press,

[17] Rosenzweig Mark and K.Wolpin (2000) \Natural Natural Experiments in Economics" Journal of Economic Literature, December, 827-74.

[18] Rust, John (1987) "Optimal Replacement of GMC Bus Engines: An Empirical Analysis of Harold Zurcher" Econometrica, 55 (5), 999-1033.

[19] Sauer, Robert (2001) \Education Financing and Lifetime Earnings" Working Paper, Brown University.

[20] Staiger, Douglas and James H. Stock (1997),\Instrumental Variables Regression with Weak Instruments" Econometrica, 65: 557-586.

[21] Topel, Robert (2000) \Labor Markets and Economic Growth", Handbook of Labor Economics, edited by David Card and Orley Ashenfelter, North-Holland Publishers.

Table 1 - Descriptive Statistics

	Mean	St dev.	# of individuals
Family Income (in $)	28877	15086	667
father's education	12.5	3.2	667
mother's education	12.1	2.3	667
# of siblings	2.7	1.7	667
proportion raised in urban areas	0.74	0.44	667
AFQT scores	49.4	26.8	667
proportion raised in nuclear family	0.82	0.39	667
Schooling completed (1994)	12.7	2.4	667
# of years with OJT (1994)	1.0	1.5	667
# of years with OFT (1994)	0.8	1.2	667
average 3 time periods	15.1	3.5	667

Note: Family income is an average of two values taken as of May 1978 and May 80 respectively.

Table 2
Occupation by Age in the NLSY

Age	Empirical Frequencies 1 School only	 3 Work only	 4 Work & OJT	 5 Work & OFT	 7 Home
14	0.999	0.000	0.000	0.000	0.002
15	0.977	0.000	0.005	0.000	0.009
16	0.943	0.006	0.021	0.000	0.030
17	0.858	0.076	0.030	0.006	0.011
18	0.624	0.267	0.032	0.028	0.048
19	0.387	0.442	0.020	0.030	0.120
20	0.328	0.496	0.044	0.017	0.115
21	0.262	0.550	0.055	0.011	0.123
22	0.189	0.635	0.057	0.020	0.100
23	0.112	0.673	0.090	0.041	0.085
24	0.081	0.657	0.137	0.052	0.071
25	0.052	0.689	0.128	0.064	0.068
26	0.049	0.667	0.147	0.078	0.060
27	0.040	0.694	0.140	0.066	0.061
28	0.024	0.740	0.132	0.050	0.054
29	0.029	0.780	0.107	0.047	0.038
30	0.008	0.777	0.106	0.044	0.065

Table 3
The Incidence of Training

Empirical Frequencies

Age	OJT	OFT	Total
14	0.000	0.002	0.002
15	0.005	0.044	0.049
16	0.021	0.080	0.101
17	0.030	0.085	0.115
18	0.032	0.084	0.116
19	0.020	0.059	0.079
20	0.044	0.027	0.071
21	0.055	0.015	0.070
22	0.057	0.024	0.081
23	0.090	0.055	0.145
24	0.137	0.067	0.204
25	0.128	0.073	0.201
26	0.147	0.091	0.238
27	0.140	0.076	0.216
28	0.132	0.054	0.186
29	0.107	0.055	0.162
30	0.106	0.044	0.150

Note:

Table 4A
OLS Regressions of the incidence of on-the-job Training on accumulated human capital
(T-ratios in parentheses)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(7)
constant	-0.0639	0.0396	0.0626	0.0687	-0.0621	-0.0771	-0.0830	
	(1.13)	(3.29)	(4.98)	(3.12)	(1.12)	(1.38)	(1.27)	
acc. educ	0.0113	-	-	-	0.0082	0.0085	0.0087	
	(2.57)				(1.89)	(1.97)	(1.94)	
acc OJT	-	0.0461	-	-	0.0444	0.0426	0.0424	
		(6.22)			(5.96)	(5.67)	(5.60)	
acc OFT	-		0.0223	-		0.0161	0.0158	
			(2.41)			(1.77)	(1.71)	
acc Exper	-		-	0.0017	-	-	0.0005	
				(0.56)			(0.18)	

Note: The dependent variable is equal to 1 if the young individual has received on-the-job training (OJT) during his last year in the sample. Accumulated education, OJT and Experience are measured at the beginning of the last year of observation and re°ect all past human capital decisions from the age of 14 until the second last year of observation.

Table 4B

OLS Regressions of the incidence of o®-the-job Training on accumulated human capital

(T-ratios in parentheses)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(7)
constant	-0.0751	0.0323	0.0241	0.0319	-0.0882	-0.0884	-0.1037	
	(1.97)	(3.87)	(2.85)	(2.15)	(2.29)	(2.29)	(2.30)	
acc. educ	0.0087	-	-	-	0.0089	0.0089	0.0095	
	(2.92)				(2.99)	(2.99)	(3.05)	
acc OJT	-	0.0025	-	-	-	-0.0009	-0.0014	
		(0.48)				(0.18)	(0.26)	
acc OFT	-		0.0137	-	0.0142	0.0143	0.0137	
			(2.19)		(2.29)	(2.29)	(2.15)	
acc Exper	-		-	0.0004	-	-	0.0014	
				(0.20)			(0.65)	

Note: The dependent variable is equal to 1 if the young individual has received o®-the-job training (OJT) during his last year in the sample. Accumulated education, OJT and Experience are measured at the beginning of the last year of observation and re°ect all past human capital decisions from the age of 14 until the second last year of observation.

Table 5
Occupation by Age: Goodness of ¯t

	Predicted Frequencies				
	1	3	4	5	7
Age	School only	Work only	Work & OJT	Work & OFT	Home
14	0.986	0.000	0.000	0.000	0.013
15	0.986	0.000	0.004	0.000	0.010
16	0.923	0.006	0.021	0.000	0.050
17	0.820	0.080	0.028	0.007	0.055
18	0.583	0.271	0.030	0.029	0.087
19	0.384	0.446	0.020	0.035	0.115
20	0.327	0.500	0.043	0.024	0.106
21	0.274	0.556	0.056	0.019	0.095
22	0.202	0.628	0.056	0.031	0.084
23	0.122	0.652	0.081	0.063	0.082
24	0.095	0.631	0.124	0.078	0.072
25	0.082	0.656	0.111	0.093	0.059
26	0.083	0.628	0.126	0.109	0.055
27	0.074	0.646	0.115	0.109	0.056
28	0.063	0.686	0.115	0.089	0.047
29	0.073	0.711	0.091	0.086	0.039
30	0.064	0.717	0.088	0.077	0.053

Table 6
The causal e®ects of Accumulated Education on the Intertemporal Utility of choosing various Options:
Spline estimates (T-ratios)

	accumulated human capital				
	educ	experience	OJT	OFT	Home
curent choices					
Work (no training)	0.3426	0.5374	-0.2114	1.27423	-1.4467
	(4.91)	(28.38)	(3.32)	(12.13)	(15.83)
Work/OJT	0.0369	0.2664	0.5733	0.9450	-1.1630
	(1.12)	(9.09)	(7.31)	(4.70)	6.55
Work/OFT	0.2770	0.3343	0.1294	3.7956	-4.0907
	(3.49)	(5.49)	(1.05)	(15.77)	(15.90)
School					-1.1949
					(19.05)

Table 7A
The E®ects of family background

	Parameters (standard errors)					
	Family background variables					
	father's education	mother's education	fath.*moth education	family income	# of siblings	Nuclear family
Choices						
Work	-0.0243	0.0025	-0.0014	-0.0003	0.0210	0.0081
	(0.58)	(0.06)	(0.45)	(0.07)	(0.81)	(0.10)
Work/OJT	0.3002	0.3675	-0.0242	-0.0028	-0.0732	0.1482
	(4.86)	(5.31)	(4.23)	(0.49)	1.84	(1.17)
Work/OFT	0.1934	0.2936	-0.0213	-0.0066	0.0295	-0.0116
	(1.62)	(3.16)	2.30	(0.83)	(0.56)	(0.10)
School	0.1569	0.1330	-0.0035	0.0078	-0.0954	0.3273
	(3.12)	(2.73)	(0.92)	(1.86)	(3.08)	(2.82)

note: Corrected AFQT scores are measured as the residual of the OLS regression of original scores (out of 100) on age and education. the residuals are then rescaled.

Table 8B
The Distribution of Unobserved Ability:
Unobserved tastes for Work and Training

	Parameter t-ratio
η_2	1.9784 (7.44)
q	-0.2987 (1.99)
School/OFT	
α_{02}	0.0 (¯xed)
α_{12}	1.0 (¯xed)
Work	
α_{03}	-2.5892 (4.38)
α_{13}	-0.6193 (3.87)
Work/OJT	
α_{04}	-5.7676 (7.92)
α_{14}	-0.1255 (1.28)
work/OFT	
α_{05}	-7.3524 (-4.51)
α_{05}	-0.5707 (3.22)

Table 9A

Some Marginal E®ects of Accumulated Human capital
on the incidence of on-the-job and o®-the-job training

		Potential Choices				
unobs. het		yes			no	
AFQT scores		no			yes	
	work/ OJT	work/ OFT	Work	work OJT	work OFT	work
acc. education	0.0014	0.0013	0.0714	-0.0002	0.0003	0.0276
acc. experience	0.0098	0.0015	0.1120	0.0106	0.0016	0.1550
acc. OJT	0.0211	0.0006	-0.0441	0.0181	0.0004	-0.0582
Acc. OFT	0.0348	0.0176	0.2655	0.0293	0.0156	0.3169
Acc. Home	-0.0428	-0.0189	-0.3015	-0.0368	-0.0152	-0.3773
AFQT scores	-	-	-	-	-	-
Father's educ	0.0110	0.0009	-0.0051	0.0090	0.0007	-0.0017
Mother's educ	0.0135	0.0014	0.0005	0.0111	0.0010	0.0025
Family income	-0.0001	-0.0001	-0.0004	-0.0001	0.0000	0.0001