MODELING WORK-RELATED TRAINING AND TRAINING EFFECTS USING COUNT DATA TECHNIQUES

Wiji Arulampalam, Alison L. Booth and Peter Elias

No.448

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

MODELING WORK-RELATED TRAINING AND TRAINING EFFECTS USING COUNT DATA TECHNIQUES

Wiji Arulampalam* Alison L. Booth** and Peter Elias**

No.448

January 1996

* Department of Economics
University of Warwick

** University of Essex & CEPR

*** University of Warwick

This paper is circulated for discussion purposes only and its contents should be considered preliminary.

MODELING WORK-RELATED TRAINING AND TRAINING EFFECTS USING COUNT DATA TECHNIQUES*

Wiji Arulampalam**
University of Warwick

Alison L Booth
University of Essex and CEPR

and

Peter Elias
University of Warwick

January 1996

Address for correspondence:
AL Booth
ESRC Research Center on Micro-social Change
University of Essex
Wivenhoe Park CO4 3SQ Essex UK
email: albooth@essex.ac.uk

JEL Classification: C25, I21, J24, J30, J42.
Keywords: Count data models, hurdle, training, skills segmentation, wages growth.

* Financial support from the ESRC (under Research Grant Number ROOO234411 "Measuring the Dynamic Impact of Job-related Training") is gratefully acknowledged. For their helpful comments, we are grateful to Dennis Leech, Neil Shephard, Chris Skeels, Mark Stewart, John Van Reenen, Ken Wallis and seminar participants at the University of Warwick, the Center for Labor Market and Social Research (Aarhus University), and an Institute for Fiscal Studies Conference on "Economic Policy and the Distribution of Earnings". We should also like to thank Cristian Herrera-Salas and Chris Jones for assistance with the data. An earlier version of this paper appeared as "Count Data Models of Work-related Training: A Study of Young Men in Britain", University of Essex Working Paper No. 95-14.

** nee Narendranathan

ABSTRACT

This paper examines the impact of work-related training on expected wages growth, using longitudinal data from the British National Child Development Study. The analysis covers a crucial decade in the working lives of a cohort of young men - the years from the age of 23 to the age of 33. We use hurdle negative binomial models to estimate the number of work-related training events. This approach, which has not been used for training before, allows us to account for the fact that more than 50% of sample members experienced no work-related training over the period 1981 to 1991. We find evidence of strong complementarities between past general education and training, suggesting that reliance on job-related training to increase the skills of the British workforce will result in an increase in the skills of the already-educated, but will not improve the skills of individuals entering the labor market with a low level of education. The results generated from the hurdle count model are subsequently used in estimation of the wages growth model. We find that each additional training event is estimated to increase wages growth by 0.7 per cent, for young men experiencing at least one training occurrence over the decade.

I. INTRODUCTION

Governments in the US and in Britain have, over the past decade, increasingly been emphasizing the importance of employer-led training in providing the skilled workforce necessary for improving competitiveness, adaptability and economic growth into the next millennium.[1] Employers are best placed to provide such skills, it has been argued, since firms are more responsive to market forces than are governments. In this context, we estimate in this paper first the determinants of the number of work-related training courses received by a group of young men over the decade 1981 to 1991, and second, the impact of these training events on wages growth (as a proxy for productivity) over the same period. The data set used is the National Child Development Study, a cohort of individuals born in Britain in the first week of March 1958.

We estimate the determinants of the number of training events using count models, in which the dependent variable takes only non-negative integer values corresponding to the number of work-related training courses occurring in the interval 1981 to 1991. This modeling procedure has not been used for training events before. Over half the sample of young men experienced no training at all over the period 1981-1991, a decade covering the crucial years from age 23 to 33. In view of this bunching of observations at zero counts, we extend the count modeling approach in order to estimate negative binomial hurdle models, in which the process generating training incidence is allowed to differ from the process generating positive training counts. We use the estimates generated from this procedure to control for training endogeneity in the wages growth model.

An additional goal of the paper is to establish whether there is any evidence of the "low-skill, bad-job" trap outlined in recent theoretical work by Snower (1995) and Burdett and Smith (1995). According to this approach, where there is a high proportion

[1] See the UK Government White Paper Employment for the 1990s and the US Department of Labor report Work-based Learning, inter alia.

of uneducated workers, firms may have little incentive to provide good jobs requiring high skills and training, and if there are few good jobs, workers may have little incentive to obtain such skills. As a result, some workers may get caught in a cycle of low productivity, deficient training and insufficient skilled jobs.

An important finding of the paper is that there are strong complementarities between past general education and training, a finding which provides some evidence for the low skill, bad job trap in Britain. An implication of the observed positive correlation between education and subsequent training is that individuals entering the labor market with low educational attainment have limited training opportunities in the work-place. Moreover, since our estimates show that wages growth is significantly increasing in the number of training events, such workers also face lower wage growth prospects. While it is not surprising that firms should offer the most able and better educated workers more training, a clear implication of our results is that reliance on employer-provided training alone will not up-grade the skills of all workers in the labor market.

The remainder of this paper is set out as follows. Section II describes the data, while Section III sets out the theoretical background. In Section IV, the count data models of training courses are presented, and the estimates discussed. Section V describes the econometric methodology used for estimation of the impact of training courses on wages growth over the period 1981 to 1991, and presents the wage growth estimates. The final section concludes.

II THE DATA SOURCE

The paper first estimates models based on count data, in which the dependent variable takes only non-negative integer values corresponding to the number of training courses lasting 3 days or more occurring in the interval 1981 to 1991. Second, the results generated from this model are then used in estimation of the wages growth model. The data set is the National Child Development Study (NCDS), a longitudinal study of individuals living in Britain and born in the week of 3-9 March 1958. Data were collected on each individual at birth, and at five follow-ups at ages 7, 11, 16, 23 and 33.[2] Particular use is made of the information collected at age 23 in 1981 (Wave 4 data) and at age 33 in 1991 (Wave 5 data). A sub-set of these data are used in the analysis in this paper, which is confined to young men in the birth cohort.[3] The advantages of using the NCDS for analysis of the determinants of the occurrence of work-related training courses are as follows. First, earlier waves of the NCDS (in particular Waves 3 and 4) provide data on time-varying and fixed individual characteristics before the individual has received training over the period 1981-1991. The education variable used is the highest educational qualification obtained by the survey date of March 1981. Work-related training courses received between leaving school and 1981 are proxied by a number of dummy variables. The rich data available in Wave 4 of the NCDS allow for the estimation of the impact of predetermined and exogenous variables on human capital acquisition between Waves 4 and 5 of the survey.

A second advantage to using the NCDS data is that problems of unobservable age-related effects (that may be found in surveys of individuals from a variety of age groups) are not present, since the data come from a specific cohort of individuals who were aged 23 in 1981.

A third advantage is that Wave 5 of the NCDS is a remarkably rich source of information about training and education received over the period 1981 to 1991. These

========================

2 Immigrants arriving in Britain in the period 1958-74 and born in the week 3-9 March were added to the survey sample. For further details of the NCDS see Shepherd (1993) and references therein.

3 Women are not analyzed in this present paper, given the complexity of modeling simultaneously their labor force participation decision, family formation plans and access to jobs providing training.

training data were elicited by a question asking respondents "Since March 1981 have you been on any training courses designed to help you develop skills that you might use in a job (apart from any courses you have already told me about)". If the respondent had been on any courses lasting at least 3 days in total, the number of such courses was requested. From this, we construct the variable NUWKTR.[4]

Well over half of the sample experienced no training at all over a crucial decade in their working lives, the 10 year period between the ages of 23 and 33 (Waves 4 and 5 of the NCDS). Some of the characteristics of the raw data for NUWKTR are as follows: 54% of the 3133 young men for whom there is complete information reported no work-related training courses in the period 1981-1991, 14% had one such course, 9% had two courses, 6% had three courses, and the remainder had up to a maximum of 21+. The sample mean is 1.903, while the sample standard deviation is 3.54. Thus there is considerable over-dispersion in raw terms, in the sense that the variance is substantially greater than the mean. The frequency distribution for this variable is given in Table 1, for all men with complete data in the sample, and separately for men with complete data who were employed in 1981. The raw data in Table 1 are characterized by a unimodal skewed distribution,

[Insert Table 1 near here]

The count data in Table 1 show signs of clustering after 9 training courses: there are spikes at 10, 12 and 15 occurrences. We believe these spikes may have arisen because individuals experiencing a lot of training over the period 1981-91 (and who were asked about training occurrences retrospectively) may have recalled them as rounded up or down numbers, that is, as a dozen, or fifteen, or twenty. For this reason, we experimented in our estimation with censoring the raw data at various points (*viz.* 10,

========================

[4] In preliminary estimation, we also experimented with estimating the number of courses leading to qualifications over the period 1981-1991, that is, general education. The explanatory power of these models was very low; it would appear that unobservables are determining individuals' decisions to undertake education over the period.

15 and 20), but found that the various censoring assumptions made little difference to the results.

Table 1: FREQUENCY DISTRIBUTION OF THE NUMBER OF TRAINING COURSES

NUWKTR **No. of training courses of 3 or more days duration 1981-91**	**All men** **Observed frequencies** *Column 1*	**Men employed 1981** **Observed frequencies** *Column 2*
0	1686	1106
1	434	293
2	274	191
3	180	126
4	118	77
5	99	68
6	89	63
7	34	25
8	30	22
9	14	12
10	72	56
11	4	2
12	31	15
13	4	2
14	3	3
15	15	11
16	2	2
17	2	2
18	1	1
19	0	0
20+	41	34
Total number of men	3133	2111

While the raw data indicate that only 46-48% of young men in a crucial decade of their life-cycle received any training at all, we need to control for covariates before making any inferences about what sort of young man was being trained over the period 1981 to 1991 in Britain. The next section sets out the modeling framework for this analysis.

III. THE THEORETICAL BACKGROUND

The purpose of the estimation of the count data models is to provide some stylized facts about the type of young man experiencing work-related training. In particular, we first wish to establish whether there is any evidence of the "low-skill, bad-job" trap outlined in recent theoretical work by Snower (1995) and Burdett and Smith (1995). According to this literature, there may be multiple equilibria in the market for skills, and policy intervention may be required to shift workers stuck in a low-skill trap to the equilibrium characterized by a high skills level. Where there is a high proportion of uneducated workers, firms may have little incentive to provide good jobs requiring high skills and training, and if there are few good jobs, workers may have little incentive to obtain such skills. As a result, certain workers may get caught in a cycle of low productivity, deficient training and insufficient skilled jobs. While our data and estimation do not represent a direct test of this theory, we are able to provide some stylized facts consistent with the theory. For example, we are able to establish that there are strong complementarities between education and training: workers entering the labor market with high levels of general education are more likely to experience work-related training courses as they progress through their working lives.

It is of course a rational response of firms to train individuals most able to benefit from the training and perhaps faster to learn. The cost of work-related training will be lower for higher ability workers, and for better-educated workers, ceteris paribus, since bright workers will learn faster than their less able colleagues. We would therefore expect to observe a positive correlation between ability and work-related training, and between higher levels of educational attainment and training. However, as recent theoretical work makes clear, the upshot may be that reliance on job-related training leads to a skills-segmented labor market and an under-class of uneducated, and perhaps unemployable, workers.

Another important question about work-related training that we address in the paper is the following: is there evidence of discrimination after controlling for other characteristics? Although in Britain there is legislation against discriminatory practices in hiring workers, employer discrimination may take the form of not offering places on courses to nonwhites. Or it may be that nonwhites do not volunteer for such training on the expectation of discrimination.

Thirdly, does experience of unemployment in the past have an adverse effect on the amount of training individuals undertake? According to the orthodox human capital approach, agents will invest in training courses if the present discounted value of training benefits exceeds training costs.[5] Irrespective of whether training is general or specific, the amount of any training investment should be greater the longer is the post-training period over which the investment can be amortized. For this reason, it might be expected that training is more likely to be offered to, or undertaken by, workers with a strong attachment to the labor market. Uncertainty about future incomes and opportunities will affect both individual workers' decisions to train and firms' decisions to offer training. The demand by workers for vocational training is likely to be influenced by the probability of unemployment and the perceived risk of not completing or of failing a training course. To the extent that unemployment is state-dependent (for example, young men who have experienced unemployment may not be confident about retaining a job in the future), past unemployment experience may have a negative effect on future training. Or it may be that workers with low motivation are the first to be laid off in a slump, and the last to be offered or to accept training courses, since the returns are likely to be low.

========================

[5] In the case of general training, the benefits are held to accrue to trainees who can take their embodied human capital with then if they change jobs in the future. It is therefore argued that trainees will bear all the costs of general training. In the case of specific human capital, both parties are held to share in training costs, and therefore both also share in post-training returns.

Fourthly, are workers who have been members of a trade union more likely to experience training? There are several hypotheses about the expected impact of trade unions and training. Trade unions in their monopoly role use their power over labor supply to extract a larger share of the surplus, and thereby induce deadweight losses. In union establishments, employer incentives to provide training are often thought to be low, because of high wages, restrictive work practices and problems with the introduction of new skill-intensive technologies that threaten union jobs. On the other hand, unions are in some circumstances cooperative, and are sometimes associated with improvements in worker morale and organization at the work place, and thereby increase training and productivity. Ultimately it is an empirical question as to whether unions are associated with an increase or decrease in training.[6]

A final question is about the relationship between the number of training courses and firm size and sector. Are larger firms and public sector firms more likely to train workers, perhaps because they are more forward looking or better placed to bear the risk associated with training? Competing hypotheses about the relationship between the number of training courses and firm size abound in the literature; for example large firms may be associated with more work-related training courses because of economies of scale in training provision (Greenhalgh and Mavrotas (1992), or perhaps because they face more regulations, more bureaucracy, and so provide more training of the nature of meeting safety regulations etc. - see Felstead and Green (1995). While we cannot hope to distinguish between these competing hypotheses, we are able to establish some stylized facts about the relationship between firm size and sector and the extent of training experienced by young men in the British labor market.[7]

========================

[6] Most empirical evidence for Britain to date suggests that union workers receive more training than nonunion workers. See for example Booth (1991), Tan et al (1992), and Greenhalgh and Mavrotas (1994). Tan et al (1992) used Wave 4 of the NCDS.

[7] Since we are restricted to using information about employer attributes for 1981, and the training counts refer to the period 1981-1991, we are actually estimating the impact of early employer attributes on training.

Human capital theory predicts that investment in training increases worker productivity. A second goal of our paper is therefore to estimate the impact of training events on expected wages growth, as a proxy for individual productivity. In particular, we wish to estimate the impact on expected earnings of additional training events, for young men who have experienced at least one training event over the period 1981 to 1991. We use the results generated from the hurdle count model to do this.

IV. MODELING THE NUMBER OF TRAINING OCCURRENCES

IV.1 The Econometric Models

In count data models, the dependent variable takes only non-negative integer values corresponding to the number of training events occurring in the interval 1981 to 1991.[8] The experience of a work-related training event is the result of optimizing decisions made by both an individual and an employer. In the case of employer-provided training, the employer decides to offer a course to an employee, who then decides whether or not to accept. Since the data preclude it, we do not model the structural framework for the training decision, but rather estimate reduced form models of the probability of individuals in the sample experiencing training events that occur $n=0,1,2,\ldots$ times in the given time interval 1981 to 1991. Given the nature of our data, the natural starting point is the Poisson model.

Let Y_i denote the number of occurrences of training courses for individual i, $i=1,2,\ldots,N$, in the interval 1981 to 1991. Then the probability density of this variable is given by

========================

[8] For surveys of these models, see Cameron and Trivedi (1986), Winkelmann (1994), Gurmu and Trivedi (1994) and Winkelmann and Zimmermann (1995).

$$\Pr(Y_i=y_i) = \frac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!} \qquad y_i=0,1,2,\ldots \tag{1}$$

where y_i is the realized value of the random variable, and λ_i is the expected number of training events, parameterized as

$$\lambda_i = \exp(\mathbf{X}_i'\beta) \tag{2}$$

where $\mathbf{X}_i$ is a vector of exogenous variables, and β is the associated vector of coefficients. The exponential form ensures non-negativity of λ_i. The Poisson distribution in (1) imposes the restriction that the conditional mean is equal to the conditional variance of y_i, given by λ_i, where the conditioning is on the the observable individual characteristics $\mathbf{X}_i$.[9] But, as shown in Table 1, the raw data indicate over-dispersion. There are at least two possible causes of such over-dispersion. One is unobserved heterogeneity in the mean function λ. Another is when the probability of experiencing an event is increased as a result of past experiences of the event. Panel data are necessary in order to distinguish between these two competing hypotheses, but unfortunately the form of the NCDS data for occurrences of training counts in the interval 1981 to 1991 is a simple cross-section (where the number of training occurrences over the period 1981-91 is measured retrospectively at the 1991 NCDS). Given the cross-section nature of the data, we take a reduced form approach, in the sense that models allowing for over-dispersion are directly specified and estimated, in order to explain the number of training events experienced by our sample members.

A common generalization of the Poisson model that allows for over-dispersion is the negative binomial distribution (see Cameron and Trivedi (1986), Winkelmann (1994)

========================

[9] For ease of exposition, from now on we shall not specifically state that the distributions being considered are conditional on the observed $\mathbf{X}_i$.

and Winkelmann and Zimmermann (1995)). This is given by

$$\Pr(Y_i=y_i) = \frac{\Gamma(\alpha_i + y_i)}{\Gamma(y_i + 1)\Gamma(\alpha_i)} \left(\frac{\alpha_i}{\alpha_i + \lambda_i}\right)^{\alpha_i} \left(\frac{\lambda_i}{\alpha_i + \lambda_i}\right)^{y_i} \quad y_i = 0,1,2,... \tag{3}$$

with $E(Y_i) = \lambda_i$, $var(Y_i) = \lambda_i + \lambda_i^2/\alpha_i$ and $\lambda_i, \alpha_i \in \mathbb{R}^+$.[10]

One model which generates the negative binomial distribution is a model of random mean function for Y_i. Suppose that the mean function of Y_i is $\tilde{\lambda}_i = \lambda_i u_i$, where u_i is an unobservable heterogeneity term and $u_i \sim$ Gamma(α_i,α_i), or equivalently $\tilde{\lambda}_i \sim$ Gamma$(\alpha_i,\alpha_i/\lambda_i)$.[11] Marginalization with respect to the unobservable u_i yields the unconditional distribution for Y_i given in equation (3), which is known as the *compound Poisson* model. Cameron and Trivedi (1986) show how to generate various versions of the negative binomial model by linking the λ_i with the α_i. Setting $\alpha_i = c\lambda_i^k$, for $c>0$ and an arbitrary constant k, produces the models they term Negbin I and Negbin II in the special cases where k=1 and k=0 respectively. The model we estimate is the Negbin II, obtained by imposing the restriction k=0, which is equivalent to the assumption that the variance is a quadratic function of the mean λ_i.[12] Thus the Poisson model is obtained with the restriction $a = 1/\alpha = 1/c = 0$ for all i.

One limitation of the model discussed above is that the zeros, as well as the positive counts, are generated by the same process. As can be seen from Table 1, there are a great many zeros in the sample. Since it is clear that some individuals never experience any training, it is sensible to model the process generating training

=========================

10 $\Gamma(n)$ is the standard gamma function.

11 If $Z \sim$ Gamma(a,b), then the probability density is

$$g(z;a,b) = \frac{a^b}{\Gamma(a)} z^{a-1} e^{-zb}$$

with $E(Z) = a/b$ and $var(Z) = a/b^2$.

12 This assumes a homoskedastic u.

incidence differently from the process generating positive counts. To do this, we estimate a hurdle model, where it is assumed that a binomial process governs the binary outcome of whether or not the individual experiences any training events and, once the hurdle is crossed, the conditional distribution of the positive values is governed by a truncated-at-zero count data model.[13] This model also allows for over-dispersion.

Formally, let f_1 be the probability density function (pdf) of the process governing the hurdle (that is, the incidence of training), and let f_2 be the pdf of the process governing the number of training events once the hurdle has been crossed.[14] Then the probability distribution of the hurdle model variable Y_{ih} for the i-th individual is given by

$$\text{Prob(no training over the period)} = \Pr(Y_{ih}=0) = f_{1i}(0) \tag{4a}$$

and

$$\begin{aligned}\text{Prob}(y_i \text{ training events over the period}) &= \Pr(Y_{ih}=y_i) \\ &= f_{2i}(y_i)[1-f_{1i}(0)]/[1-f_{2i}(0)] \qquad y_i=1,2,\ldots \\ &= f_{2i}(y_i)\ \theta_i \end{aligned} \tag{4b}$$

where $\theta_i=[1-f_{1i}(0)]/[1-f_{2i}(0)]$. Thus the mean $E(Y_{ih})$ and the $Var(Y_{ih})$ are given by:

$$E(Y_{ih}) = \sum_{y_i=1}^{\infty} y_i f_{2i}(y_i)\theta_i \tag{5}$$

and

========================

[13] This was first introduced in economics by Mullahy (1986), who considers a Poisson hurdle model. See Winkelmann (1994) for additional references.

[14] This was called the *parent*-process by Mullahy (1986).

$$\mathrm{Var}(Y_{ih}) = \theta_i \sum_{y_i=1}^{\infty} y_i^2 f_{2i}(y_i) - \theta_i^2 \left[\sum_{y_i=1}^{\infty} y_i f_{2i}(y_i)\right]^2 \tag{6}$$

Hence the over/under-dispersion is now defined at the individual level, and depends on the value of θ_i. It is interesting to note that the expected value of the hurdle model differs from the expected value of the parent model by the factor θ_i.

The likelihood for the sample is given by

$$L = \prod_{(y=0)} f_1(0) \prod_{(y>0)} [1-f_1(0)] \prod_{(y>0)} \{f_2(y)/[1-f_2(0)]\} \tag{7}$$

The first two terms on the right-hand side (RHS) of (7) refer to the likelihood for training incidence, while the third term is the likelihood for positive counts for the number of training events. The log-likelihood is therefore separable, and maximization is simplified by first maximizing a binary model log-likelihood, and then separately maximizing the log-likelihood for a truncated variable. If it is assumed that both distribution functions f_1 and f_2 are identical, but that they may be characterized by different parameter values, then standard tests can be used to test the restriction that the parameter values are the same. Some possible choices for the distribution functions are Poisson, geometric, or negative binomial.[15] We choose the Negbin II model for estimation of the hurdle model, which nests both the Poisson and the previous Negbin II models as special cases.

Let f_{1i} and f_{2i} be Negbin II with parameters (λ_{1i}, α_1) (λ_{2i}, α_2) respectively. This implies a binary model for the hurdle part of the form:

========================

15 The geometric distribution is obtained by restricting $\alpha=1$ in equation (3). The hurdle part of the specification of these models is easily estimated, by setting the censoring threshold at unity, using a software package such as LIMDEP (which allows estimation of censored Negbin II models). All models presented in this paper are estimated using LIMDEP 6.0 (See Greene, 1992).

$$\Pr(Y_{ih}=0) = f_{1i}(0) = \{\alpha_1/[\alpha_1+\exp(\mathbf{X}_{1i}'\beta_1)]\}^{\alpha_1} = [1 + a_1.\exp(\mathbf{X}_{1i}'\beta_1)]^{-1/a_1}$$

where the mean λ_{1i} is parameterized as $\exp(\mathbf{X}_{1i}'\beta_1)$, and $a_1 = 1/\alpha_1$.

In summary, we estimate two types of count data models which allow for the possibility of over-dispersion. These are first, the Negbin II model, and second, the hurdle Negbin II model. The hurdle Negbin II model nests both the simpler Negbin II model and the Poisson model as special cases.

IV.2 Estimating the Count Models for the Entire Sample

Table 2 contains the estimates for the entire sample of the 3133 young men with complete information. (In Section IV.3 we present estimates for a sub-sample of the 2111 young men who were in employment in 1981.) Since some individuals were not in employment in 1981, no characteristics of the firms where individuals might have been employed in 1981 are included as regressors in Table 2. The dependent variable is NUWKTR - the number of training events experienced by sample members over the period 1981 to 1991, and which lasted at least 3 days and were designed to develop skills used in a job. Columns 1 to 4 refer to the non-hurdle model, while Columns 5 to 8 refer to the hurdle model, where both stages have been estimated under the Poisson as well as the Negbin II assumptions. The means of the variables are given in Column 9.

First consider the results presented in the first four columns of Table 2. An interesting issue is whether or not past experience of training increases the probability of receiving training in the future, that is, the issue of state dependence in training incidence. True state dependence can only be distinguished from spurious state dependence through the use of panel data. Given the cross-section nature of our data (with retrospective information for training between 1981 and 1991), we are unable to address this issue properly. Nonetheless, we feel we should try to control for this in the estimation. We therefore report in Table 2 the estimates of two specifications - Specification 1 which omits variables measuring the number of training events experienced by our sample members prior to 1981, and Specification 2 which includes these variables.[16] Any interpretation of the impact of the pre-1981 training variables must be made with caution, since it could simply be proxying unobservable characteristics rather than measuring the true impact of state dependence in training experiences. Specification 2 in Table 2 is chosen as the preferred specification on the basis of a likelihood ratio test. The estimated effects in Specification 2 are slightly smaller in absolute terms than those in Specification 1 (which does not include pre-sample information about training receipt).

[Insert Table 2 near here]

As noted in Section III, a test of the Poisson model (where the mean equals the variance) against the Negbin II model is to test if $a=1/\alpha=0$. Since this parameter restriction is on the boundary of the parameter space, the standard Wald test and the likelihood ratio (LR) test for this restriction do not have the usual distribution. Under the null, the Wald test has a probability mass of 0.5 at zero and a 0.5 N(0,1) distribution for positive values. Similarly, under the null, the LR test statistic has a probability mass of 0.5 at zero and 0.5 $\chi^2(1)$ for positive values.[17] On the basis of these two tests and using the set of results presented in Table 2, we reject the Poisson model; that is, Column 1 is rejected against Column 3, and Column 2 is rejected against Column 4.

As noted in Section III, the non-hurdle model is nested within the hurdle model; we can therefore test this using a simple likelihood ratio test. The null hypothesis,

========================

[16] Specification 1 may be thought of as a reduced form of Specification 2 in which the past training variables have been substituted out.

[17] Thus a one-sided 5% significance level test requires the use of the 10% critical value. See Lawless (1987) for a discussion of this issue.

that the non-hurdle model is appropriate, is easily rejected for both the Poisson and the Negbin II variants. The respective likelihood ratio values are (i) 90.04 with 23 degrees of freedom in the comparison of Columns 7 and 8 with Column 4; and (ii) 1065.6 with 22 degrees of freedom in the comparison of Columns 5 and 6 with Column 2.

We now consider the suitability of the Poisson model for the hurdle specification. To do this, we compare Columns 7 and 5 of Table 2, and Columns 8 and 6. Two widely used tests are the Wald test (a simple "t-test" in this instance)[18] and the LR test. Column 7 of Table 2 shows that the parameter a ($=1/\alpha$) in the hurdle part of the process is estimated to be 0.100, with an associated standard error of 0.634. This implies that, using the Wald test, we cannot reject the null hypothesis that the assumption of a Poisson process for the hurdle part is appropriate. But in contrast, the LR test statistic gives a value rejecting the same null hypothesis. The reason for the conflicting result is as follows. For programming convenience the software package estimates α and not a. The package then returns a value for a that is estimated as the reciprocal of the estimated α (since the parameter of interest is a and not α). The program also calculates the approximate standard error for this re-parameterized value of α. However it is a well-known result that LR tests are invariant to reparameterization, whereas the Wald test is not. Gregory and Veall (1985) show that, depending on how the reparameterization is carried out, a range of different values for the Wald test may be obtained. We therefore use only the LR test for model comparison here.[19]

The LR test rejects the null hypothesis that the Poisson model is appropriate for the positive counts of training events, given in Columns 8 and 6. In summary, on the basis of this testing procedure, the Negbin II process is preferred, both for training

[18] This is an asymptotic test, and thus necessitates use of the standard normal tables instead of the t-tables.

[19] Note that this problem would not have arisen if the software package had implemented the maximization of the likelihood function in terms of the parameter a rather than α.

incidence and for positive counts conditional on incidence.

The estimates in Table 2 of the expected number of training events produce some interesting findings, which are qualitatively robust across different models. Since the preferred specification is the hurdle model, we concentrate on these estimates (given in Columns 7 and 8). First, although workers of white ethnic origin do not have a significantly higher probability of experiencing training events, they do experience significantly more training events conditional on training receipt. The size of this coefficient is large. This suggests there may be some employer discrimination in providing access to training courses, or that non-white workers may not volunteer for training on the expectation of discrimination.[20] Secondly, the estimated coefficient to the variable "reading score below average" is significantly negative. Young men scoring below average in reading tests at age 11 have both a lower probability of training incidence and experience significantly fewer training events conditional on incidence.

Thirdly, consider the effects of the variables "married by 1981" (defined to include both marriage and cohabitation) and "married and kids by 1981" (the interaction of marriage or cohabitation with the number of children). The base category here includes men neither married nor living as married, with and without children.[21] From Column 7, we see that a married or cohabiting man without any children has a significantly higher probability of receiving training, relative to the base group. But this effect is reduced to zero if he has children (the total effect is measured by the sum of 0.212 and -0.207, with a t-ratio of 0.05). Conditional on receipt of training, married or cohabiting men with children are estimated to experience on average fewer

========================

20 Where employers are relied on to provide training, the issue of whether or not there is discrimination in access to work-related training becomes very important. Booth (1993) shows that, even in the graduate labor market in Britain, women were found to receive significantly less training; however this effect was not found for black graduates.

21 We also experimented with the inclusion of the number of children without marriage, but since there were very few cases in this group and it proved to be insignificant we did not include it in the final specification.

training events relative to the base group (see Column 8).[22]

Fourthly, of the variables under the heading "Employment Status 1981", only unemployment in 1981 has a significant impact on training experiences. As expected, it has a negative effect. This may arise because individuals unemployed in 1981 have a lower attachment to the labor market, or perhaps lower motivation. Alternatively, employers may view previous unemployment as a signal of a lower attachment to the labor market, and therefore offer less training on the expectation that the investment will not be amortized.

Finally, the variables measuring human capital acquisition prior to Wave 4 of the NCDS (carried out in 1981 when respondents were aged 23) consistently have a significantly positive effect on the number of training courses over the period 1981-1991, ceteris paribus. This effect is found both for the formal human capital dummy variables measuring highest educational qualifications prior to 1981, and for the variables measuring employer-related training prior to 1981. The pre-1981 education variables having the largest impact on training incidence and the conditional number of training events are "Degree" (the highest qualification in 1981 was a university degree) and "A-level" (one or more advanced-level qualifications representing university entrance-level qualifications usually taken at or around the age of 18). The variables "O-level" (one of more ordinary-level qualifications obtained at or around the age of 16) and "Vocational qualification" (one or more business, technical or industrial vocational qualifications) also have a significant positive effect, although the impact of the latter on training incidence is significant only at the 10% level. "Apprenticeship completed" is a dummy variable indicating completion of a trade

========================

22 In preliminary regressions we also experimented with inclusion of dummy variables for paternal socio-economic class, to test if the sons of men from a higher social class experienced more training. A variable taking the value unity if the father left school at under age 16 was also included. These variables were found without exception to be insignificant, and hence were not included in the reported regressions.

apprenticeship (typically after a 3-5 year indenture period begun at age 16). This has an insignificant impact on incidence, but a positive impact on training experiences (conditional on incidence), although this is significant only at the 10% level.

This evidence of strong complementarities between past general education and training suggests that reliance on employer-provided training to increase the level of skills of the British workforce will result in an increase in the skills of the already-educated, but will not improve the skills of individuals entering the labor market with a low level of education. While it is a rational response of firms to train individuals most able to benefit from the training and perhaps faster to learn, the upshot may be that reliance on employer-provided training leads to a segmented labor market and an under class of uneducated (and possibly unemployable) workers.[23]

IV.3 Estimating the Count Models for the Sub-sample of Young Men Employed in 1981

Table 3 contains the estimates of the models derived from the sub-sample of 2111 young men in employment in 1981, and for whom we therefore have recorded information about the characteristics of the firm where they were employed in 1981. These characteristics are included as regressors in Table 3. The attributes of firms employing young men *at the time the training event was experienced* are endogenous, and therefore could not be included as regressors even if these data were available in NCDS5. The dependent variable is, as before, the NUWKTR (the number of training courses lasting at least 3 days and designed to develop skills used in a job).

A comparison of the estimates in Tables 2 and 3 shows that the inferences made in the previous sub-section for NUWKTR are generally robust across the samples. Moreover

========================

[23] Our results show that workers with low levels of general education receive relatively less work-related training. However we cannot determine if these workers choose not to train on the expectation there will be no jobs, or if instead firms do not offer these workers training in the belief that low educational levels make them untrainable.

the findings are qualitatively robust across different estimation methods. For the smaller sub-sample results given in Table 3, it can be seen that young men employed in the private sector in 1981 receive significantly less training relative to the base of public sector employees.[24] Training incidence is significantly higher for men employed in larger firms in 1981 (100 or more employees).[25] However, the number of training events conditional on training receipt is significantly higher only for men working in firms with 26-99 employees in 1981. The base is small firms (with 25 or fewer employees). It is striking that union membership status in 1981 has no significant effect on the number of training courses, especially so since other studies using British data have found that trade unionism is associated with greater training incidence (see for example Booth (1991), Greenhalgh and Mavrotas (1994) and Tan *et al* (1992)).[26]

V. MODELING THE IMPACT OF TRAINING ON EARNINGS GROWTH

V.1 The Econometric Model of Wages Growth

An implication of the observed positive correlation between education and subsequent training is that individuals entering the labor market with low educational attainment have limited training opportunities in the work place. Since training increases worker productivity, such workers may also face lower wage growth prospects. We test for this by estimating an earnings growth model. Consider the following log-linear earnings equations:

========================

24 This finding raises questions about increasing reliance on the private sector to provide a more skilled workforce. Moreover, it suggests that privatisation may well be associated with a reduction in training, ceteris paribus.

25 This firm size effect is found in many studies of training incidence; see inter alia Lynch (1992) and Tan et al (1992) for the US, and Booth (1991, 1993) and Tan et al (1992) for Britain.

26 Booth (1991) and Greenhalgh and Mavrotas (1992) analyze data sets representative of the entire adult population of Britain, in contrast to the NCDS which looks at one specific cohort of workers.

$$w_{i1} \equiv \ln(W_{i1}) = z_{i1}\gamma_1 + v_i + u_{i1} \tag{9}$$

$$w_{i2} \equiv \ln(W_{i2}) = z_{i2}\gamma_2 + \delta_1 T_i + \delta_2 T_i N_i + v_i + u_{i2} \tag{10}$$

where w_{it} denotes the natural logarithm of earnings of the i-th individual at time t, and t is 1981 in equation (9) and 1991 in equation (10). The usual observable variables representing both individual and firm characteristics are given by the vector z_{it}; these variables can be either time-fixed or time-varying, and the vectors γ_t are parameters associated with the z_{it} variables. (Unlike standard panel data models that assume the γ coefficients are constant over time, we allow the effects to change across time.) T_i is a dummy variable denoting training incidence, where $T_i=1$ if individual i experienced at least one training event over the decade 1981 to 1991, and zero otherwise. The number of training events experienced by individual i (conditional on experiencing at least one event during the interval) is given by N_i, and δ_1 and δ_2 are parameters associated with the training variables. Individual-specific error terms are denoted by v_i, which captures the effects of unobservable characteristics such as "motivation", while the random error terms of each equation are given by u_{i1} and u_{i2}.

Subtraction of (9) from (10) yields the earnings growth equation:

$$\Delta w_i = w_{i2} - w_{i1} = z_{i2}\gamma_2 - z_{i1}\gamma_1 + \delta_1 T_i + \delta_2 T_i N_i + \varepsilon_i \tag{11}$$

where $\varepsilon_i = u_{i2} - u_{i1}$. Since the earnings are in logarithms, the first difference can be interpreted as measuring approximate earnings growth over the period. We make the standard assumption that the random errors have zero means and are distributed independently across individuals. In addition, it is assumed that ε is uncorrelated with the observable characteristics z.[27]

The Endogeneity Issue

The issue of endogeneity arises when participation in a training program is not random. The earnings of untrained workers do not provide a reliable estimate of what trained workers would have received had they not participated in training. For example, suppose that individuals receiving training are more motivated than non-participants, and motivation is unobservable. If highly motivated individuals also receive higher earnings, the error term in equation (11) will be correlated with unobservables in the training determination equation. Hence OLS estimation of (11) will produce inconsistent parameter estimates.

To address this problem, assumptions have to be made about various correlations. If it is assumed that the training variables are only correlated with the unobservable individual-specific error term v_i, then OLS of (11) can be used to estimate the parameters of interest. The results of OLS estimation are given in Table 4, Column 1.[28] Column 2 of Table 4 reports the results from relaxation of this assumption, that is, where it is assumed that there may be correlation between the training variables and the error term ε in equation (11).[29] The standard practice in the literature is to use the Heckman correction to control for endogeneity of the training variable T. This assumes joint normality of ε and the error terms of the training incidence equations. However,

========================

27 In an interesting paper Bartel (1995) also examines training incidence and its impact on wages. Our paper differs from hers in that she uses company-level data to which we do not have access. In addition, unlike Bartel, we allow the effects of covariates on wages to vary across time, and we examine simultaneously the effect of training incidence and the number of training events on wages growth.

28 Arulampalam, Booth and Elias (1995) estimate their earnings growth equations under this assumption. Arulampalam *et al* (1995) and Blundell, Dearden and Meghir (1995) estimate the impact of various types of training and education on earnings growth using the NCDS data.

29 This may arise, for example, where there is a temporary unobserved adverse demand shock, that causes both less training and lower earnings growth.

because our training models are different to the standard probit model for training incidence,[30] we take a different approach. Assume that the <u>number</u> of training events experienced by individual i (conditional on training incidence) is independent of ε. Now the expected value for T_i can be substituted into (11) and the resulting equation estimated by OLS. This means we can concentrate on the endogeneity of T_i without making strong assumptions about the nature of any correlation between the N_i and the ε_i.[31] Since we are using a generated regressor which is the expected value of T_i, we correct the standard errors to take this into account, using the procedure set out in Murphy and Topel (1985).

For expositional ease, rewrite equation (11) as:

$$\Delta w_i = z_i\gamma + \delta_1 T_i + \delta_2 T_i N_i + \varepsilon_i \tag{12}$$

Note that the signs of the coefficients in (12) depend on the relative magnitudes of the corresponding γ_1 and γ_2 in (9) and (10). Conditional on the z_i, we can write

$$E(\Delta w_i) = z_i\gamma + \delta_1 E(T_i) + \delta_2\ E(T_i N_i) \tag{13}$$

From equations (4a) and (5),

$$E(T_i) = 1\text{-}f_{1i}(0) \quad \text{and} \quad E(T_i N_i) = N_i\ [1\text{-}f_{1i}(0)] \tag{14}$$

In summary, using the estimated parameters from the hurdle model, we first calculate the expected values as given in (14), and then use these in place of the two

========================

30 See Heckman and Robb (1985) for a very clear exposition of estimation of this type of probit selectivity model.

31 This assumption is implicit in all models of this form, where there are disaggregated training variables in the equation in addition to the aggregated training incidence variable.

endogenous variables T_i and T_iN_i in (12). We also correct the standard errors in order to take into account the nature of the generated regressor.

Sample Selection Issues

In order to carry out the estimation procedure described above, we require a sub-sample of individuals in employment in both 1981 and 1991, since earnings are available only for individuals employed at both dates. A natural question therefore arises as to the possible endogeneity of employment status. However, our earlier work on earnings growth models (using the same data-set) found no evidence of sample selection bias (see Arulampalam, Booth and Elias, 1995).[32] On this basis, we therefore carry out estimation of equation (12) using OLS.

V.2 Estimating the Impact of Training on Earnings

Table 4 presents estimates of the impact on earnings growth of the number of training events experienced by an individual over the period 1981 to 1991. The earnings data used are usual gross hourly earnings received at the survey dates of 1981 and 1991 deflated using the Consumer Price Index.[33] Mean real hourly earnings were £2.72 in 1981, and £4.24 in 1991, with a growth rate of 40% over the period. The estimating sub-sample for Table 4 is the 949 young men in employment in both 1981 and 1991 for whom we have complete information.

[Insert Table 4 near here]

The estimated effect of a time-invariant variable on wages growth will be significant in our model only if the effects differ in each of the wages equations given in (9) and (10). The relative magnitude of the effects at each separate time period

---

32 This paper examined the earnings impact of a variety of different forms of training, and their interactions with job changes and accreditation.

33 The 1991 earnings data we used were obtained from Heather Joshi of the City University, to whom we are very grateful.

determines the sign of the net effect of a variable on wages growth. Column 1 of Table 4 presents the fixed effects OLS estimates. If the training variables are not endogenous, OLS is a valid estimation procedure that will produce consistent parameter estimates. Column 2 presents the estimates of wages growth in which the training variables are treated as endogenous, using the predictions from the preferred hurdle model. A comparison of Columns 1 and 2 indicates that the parameter estimates are generally very similar, with the exception of the pre-1981 highest educational qualification variables and the training variables. The Hausman test does not reject the null hypothesis that the reults in Column 1 best describe the data. The test statistic is 3.2 $\sim \chi^2(38)$; the critical value at the 10% level is 49.4. We therefore focus on the estimates in Column 1 when interpreting the coefficients.

The estimates reveal some interesting findings. From Column 1, note that the number of training events is seen to have a significant positive impact on wages growth. For young men experiencing at least one course over the period 1981-1991, each additional course is associated with nearly 1% higher wages growth, ceteris paribus A man who has experienced the mean number of training events conditional on experiencing at least one training occurrence (five events), will have wages growth of approximately 4% relative to a man with no training.

While the focus of this paper is on training courses, it is also interesting to note a few of the other results of the earnings growth model. Men who were union members in both 1981 and 1991, and men who were members in 1981 but not 1991, have significantly lower earnings growth than the base - men who were not in a union in either 1981 or 1991. The negative coefficients can be interpreted in several ways: unions may be associated with flatter age-earnings profiles, or alternatively, the anti-union legislation introduced in Britain over the decade may have reduced the power of trade unions, resulting in negative wage growth for men involved in unions relative to men who were nonunion at both dates. Second, men who received secondary education at a private school enjoy a wages growth premium of over 20% relative to the base (men educated at comprehensive government-funded schools). This finding is suggestive of an old-boy network, and is an effect also found in Booth (1993) with a sample of male British graduates. It is also interesting that professional or managerial workers are characterized by high earnings growth, relative to the base of non-managerial or professional. We saw from the previous section that being white has a large significant positive impact on the number of training occurrences. Controlling for training effects in both wages growth models, we find that white ethnic origin is associated with significantly lower earnings growth than the base - men of non-white ethnic origin. This suggests that returns to being white have diminished over the decade 1981-1991, implying that anti-discrimination legislation has had some effect in eliminating earnings discrimination over the period. Finally, men with university degrees by 1981 have 20% higher wages growth than men with no formal qualifications in 1981, while men with a highest qualification in 1981 of one or more A-levels enjoy a wages growth premium of 15%.

VI. CONCLUSIONS

The paper estimates models of training based on count data (in which the dependent variable takes only non-negative integer values corresponding to the number of work-related training courses occurring in the interval 1981 to 1991). The data set is the National Child Development Study. We use hurdle negative binomial models to estimate the number of work-related training events. This approach, which has not been used for training before, allows us to account for the fact that more than 50% of sample members experienced no work-related training over the period 1981 to 1991. The results generated from this model are subsequently used in estimation of the wages growth model.

The principal findings of the paper are as follows. First, workers of white ethnic origin undertake significantly more training courses. Second, young men who scored below average in reading tests at age 16 undertake fewer training courses. Third, young men marrying or cohabiting early (but with no children) experience significantly more training occurrences. Fourth, young men who were unemployed in 1981 were significantly less likely to undertake training courses over the period 1981-91. Fifth, past human capital acquisition has a large significant positive effect on the number of training courses over the period 1981-1991. This effect is found both for the formal human capital dummy variables measuring highest educational qualifications prior to 1981, and for the employer-related training prior to 1981. Finally, we find that each additional training event is estimated to increase wages growth by nearly one per cent, for young men experiencing at least one training occurrence over the decade.

This evidence of strong complementarities between past general education and training provides some evidence for the low skill, bad job trap in Britain. An implication of the observed positive correlation between education and subsequent training is that individuals entering the labor market with low educational attainment have limited training opportunities in the work place. Moreover, since the estimates show that wages growth is increasing in the number of training events, such workers also face lower wage growth prospects. Our analysis suggests that reliance on work-related training to improve the skills of the work force will result in an increase in the skills of the already educated, but will not improve the skills of individuals entering the labor market with relatively low levels of education.

REFERENCES

Arulampalam S, AL Booth and P Elias (1995) "Work-related Training and Earnings Growth for Young Men in Britain", Department of Economics Research Paper No. 440, University of Warwick.

Bartel, AP (1995) "Training, Wage Growth and Job Performance: Evidence from a Company Database", Journal of Labor Economics 13(3), July, 401-425.

Blanchflower D and LM Lynch (1995) "Training at Work: A Comparison of US and British Youth", in Lynch LM Training and the Private Sector: International Comparisons Chicago: The University of Chicago Press.

Blundell R,L Dearden and C Meghir (1994) "The Determinants and Effects of Work-related Training in Britain", mimeo University College London.

Booth AL (1991) "Job-related Formal Training: Who Receives it and What is it Worth?" Oxford Bulletin of Economics and Statistics August, 281-294.

Booth AL (1993) "Private Sector Training and Graduate Earnings" Review of Economics and Statistics 76, 164-170.

Burdett K and E Smith (1995) "The Low Skill Trap", mimeo, University of Essex.

Cameron AC and PK Trivedi (1986) "Econometric Models based on Count Data: Comparisons and Applications of Some Estimators and Some Tests" Journal of Applied Econometrics 1, 29-54.

Cameron AC and PK Trivedi (1990) "Regression- Based Tests for Over dispersion in the Poisson Model" Journal of Econometrics 46, 347-64.

Felstead A and F Green (1995) "Training and the Business Cycle", in AL Booth and DJ Snower Acquiring Skills: Market Failures, their Symptoms and Policy Responses Cambridge University Press.

Greene WH (1992) LIMDEP Version 6.0 User's Manual and Reference Guide, Econometric Software Inc., New York.

Greene WH (1993) Econometric Analysis (second edition), Macmillan Publishing Company.

Greenhalgh C and G Mavrotas (1994) "The Role of Career Aspirations and Financial Constraints in Individual Access to Vocational Training" Oxford Economic Papers, 46(4), October, 579-604.

Gregory AW and MR Veall (1985) "Formulating Wald Tests of Nonlinear Restrictions", Econometrica 53 (6), 1465-1468.

Gurmu S and PK Trivedi (1994) "Recent Developments in Models of Event Counts: A Survey", University of Virginia Discussion Paper No. 261.

Heckman J and G Borjas (1980) "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model for Heterogeneity and State Dependence", Economica 47, 247-283.

Heckman JJ and R Robb (1985) "Alternative Methods for Evaluation of the Impact of Interventions:An Overview",Journal of Econometrics 30, 239-267.

Lawless JF (1987) "Negative Binomial and Mixed Poisson Regression", Canadian Journal of Statistics, 15(3), 209-225.

Lynch LM (1992) "Private Sector Training and the Earnings of Young Workers", American Economic Review 82 (March), 299-312.

Mullahy J (1986) "Specification and Testing of Some Modified Count Data Models", Journal of Econometrics, 33, 341-365.

Murphy KM and RH Topel (1985) "Estimation and Inference in Two-step Econometric Models" Journal of Business and Economic Statistics, 3, 370-379.

Shepherd P (1993) "Analysis of Response Bias", in E Ferrie (ed) Life at 33: the Fifth Follow-up of the National Child Development Study, London: National Children's Bureau.

Snower DJ (1995) "The Low Skill, Bad Job Trap" in AL Booth and DJ Snower Acquiring Skills: Market Failures, their Symptoms and Policy Responses Cambridge University Press.

Tan H, B Chapman, C Peterson and AL Booth (1992) "Youth Training in the US, Britain and Australia" Research in Labor Economics 13, 63-99.

UK Department of Employment (1988) Employment for the 1990s London: Her Majesty's Stationery Office (HMSO) Books.

US Department of Labor (1989) Work-based Training: Training America's Workers, Washington: US Government Publishing Office.

Winkelmann R (1994) Count Data Models - Economic Theory and an Application to Labor Mobility, Lecture Notes in Economics and Mathematical Systems Vol. 410, Berlin: Springer-Verlag.

Winkelmann R and KF Zimmermann (1995) "Recent Developments in Count Data Modeling: Theory and Application", Journal of Economic Surveys 9(1), 1-24.

Table 2: Determinants of the number of Training Courses for all young men. - uncensored model

Variable	Poisson specification 1 (1)	Poisson specification 2 (2)	Negbin II specification 1 (3)	Negbin II specification 2 (4)	Poisson Hurdle Training Incidence Spec. 2 (5)	Poisson Hurdle Positive Counts Spec. 2 (6)	Negbin II Hurdle Training Incidence Spec. 2 (7)	Negbin II Hurdle Positive Counts Spec. 2 (8)	Mean (9)
Constant	-1.299 (0.187)	-1.261 (0.187)	-1.184 (0.353)	-1.024 (0.343)	-0.733 (0.205)	-0.023 (0.233)	-1.396 (0.292)	-0.950 (0.425)	
Individual attributes									
Reading score below average	-0.476 (0.036)**	-0.429 (0.036)**	-0.433 (0.074)**	-0.407 (0.079)**	-0.297 (0.052)**	-0.224 (0.038)**	-0.332 (0.099)**	-0.265 (0.095)**	0.432
Maths score below average	-0.097 (0.035)**	-0.053 (0.035)	-0.152 (0.081)*	-0.124 (0.084)	-0.103 (0.052)**	0.015 (0.036)	-0.117 (0.082)	-0.046 (0.100)	0.491
Educated at grammar school	-0.083 (0.039)**	-0.066 (0.039)*	-0.017 (0.111)	-0.051 (0.116)	-0.116 (0.064)*	0.0002(0.040)	-0.134 (0.013)	0.035 (0.123)	0.099
Educated at private school	-0.031 (0.065)	-0.070 (0.065)	-0.032 (0.246)	-0.012 (0.234)	-0.016 (0.104)	-0.037 (0.067)	-0.021 (0.149)	-0.024 (0.233)	0.031
White	0.953 (0.170)**	0.952 (0.170)**	0.906 (0.322)**	0.885 (0.312)**	0.359 (0.176)**	0.773 (0.217)**	0.410 (0.262)	1.025 (0.367)**	0.985
Registered disabled 1981	-0.103 (0.079)	-0.090 (0.079)	-0.063 (0.167)	-0.024 (0.160)	-0.147 (0.113)	0.030 (0.084)	-0.164 (0.165)	0.130 (0.225)	0.038
married by 1981	0.382 (0.027)**	0.324 (0.028)**	0.409 (0.072)**	0.365 (0.071)**	0.186 (0.044)**	0.205 (0.029)**	0.212 (0.082)**	0.310 (0.087)**	0.401
married and kids by 1981	-0.335 (0.050)**	-0.356 (0.050)**	-0.349 (0.110)**	-0.408 (0.120)**	-0.182 (0.072)**	-0.236 (0.053)**	-0.207 (0.115)*	-0.364 (0.158)**	0.116
full-time experiece (months)/100	0.545 (0.001)**	0.250 (0.074)**	0.426 (0.157)**	0.228 (0.159)	0.149 (0.103)	0.166 (0.082)**	0.162 (0.157)	0.172 (0.216)	0.663
Employment Status 1981									
In full-time education	-0.275 (0.085)**	-0.183 (0.086)**	-0.205 (0.268)	-0.099 (0.248)	0.064 (0.119)	-0.242 (0.092)**	0.068 (0.172)	-0.289 (0.263)	0.023
Unemployed	-0.467 (0.066)**	-0.409 (0.066)**	-0.403 (0.125)**	-0.349 (0.123)**	-0.225 (0.081)**	-0.228 (0.074)**	-0.250 (0.121)**	-0.304 (0.175)*	0.089
Out-of-the-labour force	-0.531 (0.175)**	-0.432 (0.178)**	-0.492 (0.477)	-0.369 (0.454)	-0.021 (0.205)	-0.479 (0.204)**	-0.027 (0.291)	-0.669 (0.550)	0.009
Highest Educational Qualifications 1981									
Degree	1.430 (0.074)**	1.256 (0.073)**	1.355 (0.165)**	1.238 (0.163)**	0.819 (0.102)**	0.729 (0.081)**	0.912 (0.237)**	0.922 (0.201)**	0.109
A-levels	1.202 (0.063)**	1.003 (0.063)**	1.160 (0.160)**	1.004 (0.154)**	0.757 (0.087)**	0.519 (0.070)**	0.842 (0.215)**	0.647 (0.179)**	0.117
O-levels	0.822 (0.053)**	0.619 (0.054)**	0.854 (0.099)**	0.653 (0.099)**	0.455 (0.069)**	0.305 (0.060)**	0.493 (0.121)**	0.437 (0.130)**	0.383
Vocational qualification	0.577 (0.062)**	0.379 (0.062)**	0.640 (0.107)**	0.418 (0.114)**	0.194 (0.082)**	0.249 (0.068)**	0.202 (0.121)*	0.405 (0.150)**	0.161
Apprenticeship completed	-0.034 (0.035)	0.209 (0.036)	-0.082 (0.081)	0.147 (0.090)	0.031 (0.055)	0.207 (0.038)**	0.037 (0.082)	0.200 (0.105)*	0.275
Training prior to 1981									
1 Training course - excl. apprenticeship		0.397 (0.033)**		0.438 (0.075)**	0.248 (0.050)**	0.240 (0.034)**	0.280 (0.097)**	0.327 (0.089)**	0.198
2 Training courses		0.424 (0.043)**		0.508 (0.121)**	0.356 (0.066)**	0.194 (0.044)**	0.404 (0.142)**	0.270 (0.129)**	0.086
3-4 Training courses		0.885 (0.045)**		0.934 (0.147)**	0.611 (0.085)**	0.509 (0.046)**	0.715 (0.245)**	0.692 (0.150)**	0.047
More than 4 courses		1.430 (0.058)**		1.429 (0.344)**	0.964 (0.139)**	0.900 (0.059)**	1.144 (0.474)**	1.094 (0.291)**	0.016
Variance Parameter a=1/α	0.00 fixed	0.00 fixed	2.624 (0.100)**	2.453 (0.097)**	0.00 fixed	0.00 fixed	0.100 (0.634)	1.896 (0.247)**	
Model log-likelihood	-8663.57	-8301.94	-5319.62	-5268.75	-3285.79	-4483.35	-1980.31	-3243.42	
Number of cases	3133	3133	3133	3133	3133	1427	3133	1427	3133

Notes: (i) Standard errors are given in parentheses.
(ii) ** Coefficient significant at 5%.
* Coefficient significant at 10%.

Table 3: Determinants of the Number of Training Courses for Employed Young Men. - Uncensored Model

Variable	Poisson specification 2 (1)	Negbin II specification 2 (2)	Poisson Hurdle Training Incidence Spec. 2 (3)	Poisson Hurdle Positive Counts Spec. 2 (4)	Negbin II Hurdle Training Incidence Spec. 2 (5)	Negbin II Hurdle Positive Counts Spec. 2 (6)	Mean (7)
Constant	-1.431 (0.226)	-1.492 (0.424)	-1.023 (0.256)	-0.053 (0.279)	-1.715 (0.363)	-0.867 (0.548)	
Individual attributes							
Reading score below average	-0.309 (0.041)**	-0.276 (0.093)**	-0.249 (0.061)**	-0.158 (0.043)**	-0.280 (0.106)**	-0.172 (0.111)	0.439
Maths score below average	-0.080 (0.041)**	-0.100 (0.100)	-0.102 (0.062)	-0.010 (0.042)	-0.117 (0.096)	-0.011 (0.120)	0.504
Educated at grammar school	-0.027 (0.047)	-0.000 (0.150)	-0.079 (0.079)	-0.004 (0.048)	-0.093 (0.119)	0.039 (0.164)	0.098
Educated at private school	-0.129 (0.091)	-0.067 (0.311)	0.171 (0.138)	-0.217 (0.096)**	0.195 (0.211)	-0.352 (0.283)	0.025
White	1.211 (0.202)**	1.170 (0.348)**	0.359 (0.176)**	0.877 (0.256)**	0.620 (0.317)**	1.195 (0.428)**	0.983
Registered disabled 1981	-0.038 (0.092)	0.001 (0.209)	-0.002 (0.132)	-0.075 (0.098)	-0.004 (0.189)	-0.001 (0.255)	0.035
Married by 1981	0.331 (0.032)**	0.385 (0.090)**	0.205 (0.052)**	0.191 (0.034)**	0.236 (0.099)**	0.348 (0.179)**	0.426
Married and kids by 1981	-0.459 (0.058)**	-0.513 (0.143)**	-0.178 (0.083)**	-0.336 (0.063)**	-0.204 (0.131)	-0.511 (0.174)**	0.122
Full-time experiece (months)/100	0.120 (0.087)	0.262 (0.204)	0.075 (0.124)	0.096 (0.097)	0.080 (0.186)	0.207 (0.278)	0.696
Union Member	0.103 (0.034)**	0.063 (0.085)	0.102 (0.055)*	0.029 (0.036)	0.118 (0.085)	0.036 (0.101)	0.482
Local unemployment rate - TTWA	-0.013 (0.004)**	-0.012 (0.012)	0.004 (0.006)	-0.162 (0.005)**	0.004 (0.010)	-0.029 (0.014)**	10.81
Job is Professional/Managerial	0.073 (0.037)**	0.078 (0.105)	0.050 (0.060)	0.042 (0.037)	0.052 (0.088)	0.004 (0.114)	0.247
Highest Educational Qualifications 1981							
Degree	1.113 (0.089)**	1.236 (0.230)**	0.714 (0.132)**	0.675 (0.096)**	0.802 (0.261)**	1.018 (0.278)**	0.080
A-levels	0.982 (0.072)**	1.079 (0.192)**	0.772 (0.103)**	0.484 (0.079)**	0.868 (0.248)**	0.749 (0.226)**	0.116
O-levels	0.594 (0.061)**	0.655 (0.116)**	0.499 (0.080)**	0.264 (0.068)**	0.547 (0.145)**	0.395 (0.155)**	0.405
Vocational qualification	0.361 (0.071)**	0.421 (0.137)**	0.241 (0.097)**	0.209 (0.076)**	0.258 (0.144)*	0.347 (0.179)*	0.164
Apprenticeship completed	0.233 (0.042)**	0.136 (0.109)	-0.020 (0.066)	0.245 (0.044)**	-0.021 (0.099)	0.250 (0.126)**	0.263
Employer Characteristics - 1981							
Private Sector	-0.163 (0.035)**	-0.275 (0.095)**	-0.157 (0.055)**	-0.070 (0.035)**	-0.178 (0.092)**	-0.178 (0.108)**	0.685
26-99 employees	0.312 (0.046)**	0.335 (0.104)**	0.130 (0.070)*	0.223 (0.048)**	0.142 (0.105)	0.304 (0.133)**	0.229
100-499 employees	0.127 (0.048)**	0.206 (0.108)*	0.237 (0.070)**	-0.045 (0.050)	0.265 (0.118)**	-0.007 (0.131)	0.227
500 or more employees	0.353 (0.046)**	0.352 (0.100)**	0.322 (0.070)**	0.125 (0.048)**	0.361 (0.133)**	0.106 (0.125)	0.243
Training prior to 1981							
1 Training course - excl. apprenticeship	0.459 (0.038)**	0.460 (0.088)**	0.196 (0.060)**	0.338 (0.040)**	0.221 (0.103)**	0.434 (0.105)**	0.482
2 Training courses	0.278 (0.053)**	0.305 (0.152)**	0.166 (0.081)**	0.157 (0.055)**	0.186 (0.126)	0.159 (0.176)	0.211
3-4 Training courses	0.933 (0.052)**	0.861 (0.189)**	0.542 (0.100)**	0.595 (0.053)**	0.642 (0.247)**	0.736 (0.188)**	0.093
More than 4 courses	1.387 (0.067)**	1.288 (0.431)**	0.809 (0.165)**	0.976 (0.067)**	0.967 (0.431)**	1.049 (0.393)**	0.054
Variance Parameter a=1/α	0.00 fixed	2.335 (0.111)**	0.00 fixed	0.00 fixed	0.100 (0.637)	1.748 (0.255)**	0.018
Model log-likelihood	-5768.28	-3643.40	-2193.00	-3182.96	-1324.84	-2278.21	
Number of cases	2111	2111	2111	1005	2111	1005	2111

Notes: See Table 2 Notes.

Table 4- Approximate Earnings Growth 1981 - 1991 (dependent var=ln(wage$_{91}$)-ln(wage$_{81}$))

Variable	OLS (1)	IV (2)	Mean (3)
Intercept	0.698 (0.15)**	0.590 (0.19)**	
Fixed individual characteristics			
Ethnicity - white	-0.235 (0.11)**	-0.294 (0.14)**	0.988
Maths score below average	-0.020 (0.03)	-0.060 (0.03)	0.491
Reading score below average	-0.005 (0.03)	0.028 (0.04)	0.407
School - private	0.211 (0.08)**	0.215 (0.08)**	0.023
- direct grant/grammar	0.017 (0.04)	0.018 (0.04)	0.132
Individual characteristic in 1981			
Has a disability which affects work	0.200 (0.08)**	0.197 (0.08)**	0.024
Individual characteristic in 1991			
Has a disability which affects work	-0.137 (0.08)*	-0.129 (0.08)*	0.025
Changes across period 1981-1991			
Trade union membership			
Non member to a member	-0.073 (0.05)	-0.076 (0.05)	0.063
Member to a non member	-0.131 (0.03)**	-0.145 (0.04)**	0.210
Member to a member	-0.117 (0.03)**	-0.134 (0.04)**	0.427
Regional changes in residence			
London to outside	0.097 (0.05)**	0.097 (0.05)**	0.069
Outside to London	-0.001 (0.13)	-0.004 (0.13)	0.008
London to London	-0.011 (0.07)	-0.002 (0.07)	0.035
Firm type			
Private to Public	0.016 (0.05)	0.042 (0.05)	0.082
Public to Private	-0.063 (0.04)	-0.069 (0.04)*	0.161
Private to Private	-0.014 (0.03)	-0.002 (0.04)	0.537
Firm size			
to larger	0.060 (0.03)**	0.071 (0.03)**	0.291
to smaller	-0.008 (0.03)	-0.015 (0.03)	0.308
Job type			
Prof./Manag/Admin to other	0.047 (0.06)	0.059 (0.06)	0.047
Other to Prof/Manag/Admin	0.164 (0.03)**	0.173 (0.03)**	0.203
no change at Prof/Manag/Admin	0.189 (0.04)**	0.193 (0.04)**	0.200
Marital status			
not married to married	0.010 (0.04)	0.010 (0.04)	0.434
married to not married	-0.067 (0.07)	-0.092 (0.07)	0.035
married to married	0.016 (0.04)	-0.002 (0.04)	0.390
Regional unemployment rate - %			
1981	-0.003 (0.08)	-0.001 (0.01)	11.278
1991	-0.011 (0.02)	-0.015 (0.02)	7.925
Highest qualification obtained prior to 1981$^{(@)}$			
Degree	0.208 (0.06)**	0.134 (0.09)	0.076
Advanced Level (A.L.)	0.150 (0.05)**	0.064 (0.08)	0.129
Ordinary Level (O.L.)	0.051 0.04)	-0.008 (0.06)	0.408
Vocational	0.035 (0.04)	0.002 (0.06)	0.165
Training received prior to 1981			
Apprenticeship completed	-0.034 (0.03)	-0.019 (0.04)	0.275
Employer provided training in current job in 1981	-0.051 (0.03)**	-0.059 (0.027)**	0.568
Employer provided training in first job if current job # first job	-0.031 (0.03)	-0.041 (0.03)	0.348

Table 4 continued

Education and Training received over the period 1981-1991			
Atleast one educational course followed(dummy)	0.046 (0.03)*	0.038 (0.03)	0.365
Number of educational courses followed+	0.0029(0.01)	0.002 (0.01)	
Atleast one Training course (dummy) - T_i	0.044 (0.03)	0.340 (0.22)	0.533
Number of training courses experienced+ - N_i T_i	0.007(0.003)**	0.0096(0.003)**	
Sample size	949	949	949

Notes: (i) Mean of dep. var=0.419.

(ii) ** Coefficient significant at 5% significance level.

(iii) *Coefficient significant at 10% significance level.

(iv) + Since this variable is only defined for those individuals who have had atleast one such course, the variable entered in the equation is this variable interacted with the incidence variable (see equation 10).

(v) In column 2 the training variables T_i and $T_i N_i$ variables are replaced with their predicted values calculated using the Poisson incidence results from Table 3 column 3 and the resulting equation estimated by OLS. The standard errors are corrected for the use of these generated regressors (see text).