DISCUSSION PAPER SERIES

No. 1582

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

Wiji Arulampalam, Alison L Booth and Peter Elias

HUMAN RESOURCES

Centre for Economic Policy Research

CEPR Discussion Paper No. 1582

February 1997

ABSTRACT

Modelling Work-Related Training and Training Effects Using Count Data Techniques*

This paper estimates the determinants of the number of work-related training courses, and their impact 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 – from the age of 23 to 33. We use hurdle negative binomial models to estimate the number of work-related training events. This approach allows us to account for the fact that half of all sample members experienced no work-related training over the period from 1981 to 1991. We then estimate a wages growth model where the returns from the first training experience are allowed to differ from subsequent training experiences. The results generated from the hurdle count model are used to control for training endogeneity in the wages growth equation. This has not been done before in the training literature. The sensitivity of the wage growth estimates to alternative modelling strategies is also examined. Since we find a strong correlation between education and subsequent training experiences, we experiment with estimating the joint impact of previous education and subsequent training on wages growth, in order to tease out the combined effect of these variables. We find that the biggest returns to training are to highly educated men.

JEL Classification: C25, I21, J24, J30, J42
Keywords: count data models, hurdle, training endogeneity, selectivity bias, wages growth

Wiji Arulampalam
Department of Economics
Warwick University
Coventry CV4 7AL
UK
Tel (44 1203) 523 471
Fax: (44 1203) 523 032
Email: wiji.arulampalam@warwick.ac.uk

Alison L Booth
ESRC Research Centre for
Micro-social Change
University of Essex
Wivenhoe Park
Colchester CO4 3SQ
UK
Tel: (44 1206) 873 789
Fax: (44 1206) 873 151
Email: albooth@essex.ac.uk

Peter Elias
Institute for Employment Research
University of Warwick
Coventry CV4 7AL
UK
Tel: (44 1203) 523 286
Fax: (44 1203) 524 241
Email: p.elias@warwick.ac.uk

*This paper is produced as part of a CEPR research programme on *The UK Labour Market: Microeconomic Imperfections and Institutional Features,* which is co-sponsored by the UK Department for Education and Employment and the Department of Trade and Industry. The views expressed in this paper are those of the author(s) and do not necessarily reflect those of the DfEE, DTI or CEPR. The authors should like to thank an anonymous referee for many useful suggestions and William Greene for software advice. For helpful comments they are grateful to Jeff Borland, Costas Meghir, Mark Stewart, John Van Reenen, Ken Wallis and seminar participants at the University of Warwick, the Centre for Labour Market and Social Research (Aarhus University), an Institute for Fiscal Studies Conference on 'Economic Policy and the Distribution of Earnings', and the 1996 Australian Labour Market Research Conference at the ANU. Thanks are also due to Heather Joshi and Peter Shepherd of City University, who provided us with an algorithm for cleaning the 1991 earnings data, and Cristian Herrera-Salas and Chris Jones for assistance with the data. Financial support from the ESRC (under Research Grant Number ROOO234411 'Measuring the Dynamic Impact of Job-related Training') is gratefully acknowledged. An early 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.

Submitted 29 December 1996

NON-TECHNICAL SUMMARY

Governments in OECD countries have increasingly been emphasizing the importance of work-related training in providing the skilled workforce necessary for improving competitiveness, adaptability and economic growth into the next millennium (OECD (1995)). It has also been suggested that skills acquisition will reduce the growing earnings inequality observed in some countries since the 1980s. Employers are best placed to provide such skills, it has been argued, since firms are more responsive to market forces than governments. In this context, we estimate first the determinants of the number of work-related training courses received by a group of young men over a decade from 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 a cohort of individuals born in Britain in the first week of March 1958; the National Child Development Study, which contains unique data for Britain about the number of training courses of at least three days duration.

The experience of work-related training is the result of optimizing decisions made by both an individual worker and an employer. For employer-provided training, the employer decides to offer a course to a worker, who then decides whether or not to accept. Since the data preclude it, we do not model the structural framework for the training decision. Instead, using count models, we estimate reduced-form equations of the determinants of the number of training courses. We then use the estimates generated from this procedure to control for training endogeneity in the wages growth model. In the count models, 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. Over half the sample of young men experienced no training at all over the period 1981–91, 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 modelling approach to estimate negative binomial hurdle models, in which the process generating training incidence is allowed to differ from the process generating positive training counts.

Since human capital theory predicts that investment in training increases worker productivity, we then estimate the impact of training events on expected wages growth as a proxy for individual productivity. In particular, we estimate the impact on expected wages growth of additional training events for young men who have experienced at least one training event over the period from 1981 to 1991. We use the results generated from the hurdle count model to do this. To our knowledge, this has not been done before in the training

literature. The issue of training endogeneity in wages models arises when participation in a training course is not random. The wages 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 individuals receiving training are more motivated than non-participants, and motivation is unobservable. If highly motivated individuals also have higher wages, the error term in the wages equation will be correlated with unobservables in the training determination equation. Hence ordinary least squares estimation will produce inconsistent parameter estimates.

Earlier studies using surveys with data on the amount of training to estimate the impact of training on wages or wages growth have either not controlled for self-selection into training, or have controlled for self-selection by treating only training incidence as endogenous. The methodology adopted in this paper offers a relatively straightforward way of controlling for self-selection into multiple training occurrences. The determinants of the number of training courses are modelled, and then used in a wages growth equation to control for potential self-selection into training, allowing for both self-selection into training incidence and for training counts conditional on incidence. The sensitivity of the estimates to alternative modelling strategies is also examined.

The principal findings of the paper are as follows. First, the hurdle Negbin II count model describes the data best. The significant determinants of work-related training are ability (as measured by reading ability at age eleven), occupation, establishment size interacted with sector, and educational qualifications received prior to training. The biggest impact is from the latter. Second, we find that human capital increases wages growth for young men in our sample. While the estimates for non-human capital variables are remarkably robust to different treatments of training endogeneity, the human capital variables are not. After controlling for self-selection into training, we find that young men with a higher educational background are not only more likely to be trained, they are also more likely to receive substantially higher wages growth as a result. The sensitivity of the wage growth estimates to alternative modelling strategies is also examined. Since we find a strong correlation between previous education and subsequent training experiences, we experiment with estimating the joint impact of previous education and subsequent training on wages growth, in order to tease out the combined effect of these variables. We find that the biggest returns to training are to highly educated men. An implication of the observed positive correlation between education and subsequent training is that individuals entering the labour market with low educational attainment have limited training opportunities in the work place. Moreover, since the estimates show that

I. INTRODUCTION

Governments in OECD countries have increasingly been emphasizing the importance of work-related training in providing the skilled workforce necessary for improving competitiveness, adaptability and economic growth into the next millennium (OECD, 1995). It has also been suggested that skills acquisition will reduce the growing earnings inequality observed in some OECD countries since the 1980s.[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 a cohort of individuals born in Britain in the first week of March 1958, the National Child Development Study, which contains unique data for Britain about the number of training courses of at least 3 days duration.[2]

The experience of work-related training is the result of optimizing decisions made by both an individual worker and an employer. For employer-provided training, the employer decides to offer a course to a worker, who then decides whether or not to accept. Since the data preclude it, we do not model the structural framework for the training decision. Instead, using count models, we estimate reduced form equations of the determinants of the number of training courses. We then use the estimates generated from this procedure to control for training endogeneity in the wages growth model. In the count models, 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. 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.[3] In view

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

1 See Freeman and Katz (1995) for an account.

2 This paper represents the first attempt to use the NCDS data to model the number of training courses and their impact on wages growth. Other papers using the NCDS5 data on training have focused on using detailed information on only the most recent training courses; see Arulampalam et al (1995) and Blundell et al (1996).

3 Lillard and Tan (1992:12) note a similar phenomenon in the US, where they find that "as much as one third of the young men in the NLS get no training, even after nine

of this bunching of observations at zero counts, we extend the count modelling approach to estimate negative binomial hurdle models, in which the process generating training incidence is allowed to differ from the process generating positive training counts.

Since human capital theory predicts that investment in training increases worker productivity, we estimate the impact of training events on expected wages growth as a proxy for individual productivity. In particular, we estimate the impact on expected wages growth 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. To our knowledge, this has not been done before in the training literature.[4] Earlier studies using surveys with data on the amount of training to estimate the impact of training on wages or wages growth have either not controlled for self-selection into training, or have controlled for self-selection by treating only training incidence as endogenous.[5] The methodology adopted in this paper offers a relatively straightforward way of controlling for self-selection into multiple training occurrences. The determinants of the number of training courses are modelled, and then used in a wages growth equation to control for potential self-selection into training, allowing for both self-selection into training incidence <u>and</u> to training counts conditional on incidence. The sensitivity of the estimates to alternative modelling strategies is also examined. Since we find a strong correlation between previous education and subsequent training experiences, we also experiment with estimating the joint impact of previous education and subsequent training on wages growth, in order to try to tease out the combined effect of these variables.

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

survey periods, while many report getting multiple training events."

4 While Lillard and Tan (1992) note the importance of multiple training occurrences, they treat these as exogenous when examining the impact of training on economic outcomes. They also note (p.31) that multiple training occurrences within a period are typically not known from US survey data. The NLS data for young men, for example, contain training information for every survey period, but multiple sources of training are not known within each period; data about sources and types of training are available <u>only</u> for the longest event. Thus Lillard and Tan use as their "events" measure of training the accumulated sum of all training events, where there is only one event measured at each wave.

5 See for example papers in Lynch (1994) and Ashenfelter and LaLonde (1996).

The principal findings of the paper are as follows. First, the hurdle Negbin II count model describes the data best. The significant determinants of work-related training are ability (as measured by reading ability at age 11), occupation, establishment size interacted with sector, and educational qualifications received prior to training. The biggest impact is from the latter. Second, we find that human capital increases wages growth for young men in our sample. While the estimates for non-human capital variables are remarkably robust to different treatments of training endogeneity, the human capital variables are not. After controlling for self-selection into training, we find that young men with a higher educational background are not only more likely to be trained, they are also more likely to receive substantially higher wages growth as a result. 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 increased by training, such workers also face lower wage growth prospects.

The remainder of this paper is set out as follows. Section II describes the data. In Section III, the count data models of training courses are presented, and the estimates discussed. Section IV 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 AND EXPLANATORY VARIABLES

II.1 The Data

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. 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

(1985).) Particular use is made of the information collected at age 23 in 1981 (Wave 4) and at age 33 in 1991 (Wave 5). A sub-set of these data is used in the analysis in this paper, which is confined to young men in the birth cohort.[6]

Wave 5 of the NCDS devoted a significant amount of time and interviewer resources to the collection of information about education and training received since March 1981. A basic distinction was made between "Courses for qualification" on the one hand, and "Training courses designed to help you develop skills that you might use in a job (excluding courses already classified as 'courses for qualification'." This procedure meant that respondents first discussed those courses specifically intended to lead to an educational qualification, and then were questioned in more detail about courses helping them develop skills for use in a job. The latter category we designate as "work-related training" (WRT). All WRT courses must have been of at least 3 days duration before they were included in the survey responses. From this information, we construct the variable "Number of training courses lasting at least 3 days".

Since information on WRT was collected via retrospective inquiry about the ten year period prior to Wave 5 of the survey, recall error is bound to be a problem for WRT courses. However, we believe that the problem is lessened by the fact that all very short (one or two day) courses are excluded. Respondents were only asked to record those which are reasonably significant and which constituted a "course" (that is, a planned and purposeful programme of instruction, experience and learning). Moreover, while studies have shown that recall error increases with age (see inter alia Sudman and Bradburn, 1973; Bradburn et al, 1994), all respondents in the NCDS are the same age, and hence our data cannot suffer from selective age-related recall bias. Finally, in our analysis we do not use information on the duration of training courses, but only on their incidence, about which there is likely to be less recall error.[7]

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

6 Women are not analyzed in this present paper, given our focus on training occurrences and their impact on wages growth. To measure the impact of female training on wages, we would need to model simultaneously women's labour force participation decisions, family formation plans and access to jobs providing training. For a comparison of male and female training experiences, see Arulampalam and Booth (1996).

7 Information about the duration of training courses was asked only for the three most

The frequency distribution for the variable "Number of training courses lasting at least 3 days" is given in Table 1, for all men with complete data who were employed in 1981. This sub-sample of young men was chosen in order to condition on 1981 job characteristics in the training models in the next section.[8] The raw data in Table 1 are characterized by a unimodal skewed distribution.

Table 1: FREQUENCY DISTRIBUTION OF THE NUMBER OF TRAINING COURSES

No. of training courses of 3 or more days duration 1981-91	Men employed 1981 Observed frequencies	Percentages
0	876	49.63
1	251	14.22
2	162	9.18
3	112	6.35
4	73	4.14
5	66	3.74
6	58	3.29
7	23	1.30
8	18	1.02
9	11	0.62
10+	115	6.52
Total number of men	1765	

Approximately half 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 the "Number of training courses lasting at least 3 days" are as follows: 50% of the 1765 young men in employment in 1981 and for whom there is complete information reported no work-related

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

recent courses lasting at least one week. We were therefore unable to undertake any within-sample evaluation of recall bias by looking at incidence by year, because only the three most recent courses were dated.

[8] In an earlier version of this paper we estimated the training models both for all young men, and also for the sub-sample including only those employed in 1981. Since the qualitative results were the same, we focus only on the latter in the interests of space.

training courses in the period 1981-1991, 14% had one course, 9% had two courses, 6% had three courses, and the remainder had up to a maximum of 10+.[9] The sample mean is 2.2, while the sample standard deviation is 4.01. Thus there is considerable over-dispersion in raw terms, in the sense that the variance is substantially greater than the mean.

Because of the structure of the NCDS questionnaire, detailed information was not requested about each of the work-related training courses lasting at least 3 days. Only for the (up to three) most recent training courses was detailed information elicited about the duration and type of training. Thus there is a trade-off, in analysis employing the NCDS training data, between using the aggregate information about the number of significant training courses, and the detailed information about only the 3 most recent courses. In this paper, we have chosen to focus on the former, and estimate the determinants of the total number of significant training courses and their impact on earnings growth.[10] Nonetheless, the reader may be interested in some salient features about the most recent training course experienced by our sample of young men. For 87% of men reporting the most recent training course, the course was less than a week long. Only for men reporting a course lasting at least a week was duration data collected: for these men, the mean duration of the most recent training course was 5 weeks, and the standard deviation 12 weeks. Of men reporting the most recent training course, 45% were on courses conducted on the employer's premises, while the remainder were off-the-job. Some 17% ended in a qualification, while 91% were employer-provided. Of those men who said the employer did not provide the most recent course, 53% said the employer paid the fees (in part or in full). Just 2.7% of young men reporting their most recent training event were on a government training scheme.

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

9 Although Table 1 reports 10+ training courses, the training data used in estimation are uncensored.

10 Arulampalam, Booth and Elias (1995) estimate the disaggregated impact of the most recent forms of training on wages growth, and the extent to which training of various forms decays across time. Blundell, Dearden and Meghir (1996) also examine the impact of various detailed forms of training and education on wages growth using the NCDS. Arulampalam and Booth (1996) examine gender differences in training receipt.

While the raw data in Table 1 indicate that only 50% of young men in a crucial decade of their life-cycle received any training at all, we need to control for covariates in order to make inferences about what sort of young man was being trained over the period 1981 to 1991 in Britain. The next sub-section sets out the modelling framework used.

III. MODELLING THE NUMBER OF TRAINING OCCURRENCES

III.1 The Econometric Models

In count data models, the dependent variable takes only non-negative integer values corresponding to the number of events occurring in an interval.[11] We estimate reduced form models of the probability of individuals in the sample experiencing training courses occurring y=0,1,2,... times in the given time interval 1981 to 1991. Given the nature of our data, the natural starting point is the Poisson model.[12]

Let Y_i denote the number of training courses for individual i, i=1,2,...,n, in the interval 1981 to 1991. Then the probability distribution of this variable is given by

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

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

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

[12] Dagsvik (1994) provides an individual choice-theoretic model in the presence of multiple discrete or continuous alternatives.

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 observable individual characteristics $\mathbf{X}_i$.[13] But, as shown in Table 1, the raw data indicate over-dispersion. Over-dispersion in the raw data at the marginal level need not imply over-dispersion at the conditional level after controlling for covariates. As we shall see later, it is straightforward to test for over-dispersion at the conditional level.

There are at least two possible causes of 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 in the interval 1981 to 1991 is a simple cross-section (where the number of training courses over the period 1981-91 is measured retrospectively at the 1991 NCDS). Given the cross-sectional 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 courses experienced by 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) and Winkelmann and Zimmermann (1995)). This is given by

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

with $E(Y_i) = \lambda_i$, $var(Y_i) = \lambda_i + \lambda_i^2/\alpha_i$ and $\lambda_i, \alpha_i \in \mathbb{R}^+$.[14]

One model which generates the negative binomial distribution is a model with a random mean function for Y_i. Suppose that the mean function of Y_i is $\tilde{\lambda}_i = \lambda_i u_i$, where

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

[14] $\Gamma(n)$ is the standard gamma function.

u_i is an unobservable heterogeneity term and $u_i \sim$ Gamma(α_i,α_i), or equivalently $\tilde{\lambda}_i \sim$ Gamma(α_i,α_i/λ_i).[15] 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.[16] Thus the Poisson model is obtained with the restriction $a = 1/\alpha = 1/c = 0$ for all i. A test of this restriction is a test of no over-dispersion at the conditional level.

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 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 courses and, once the hurdle is crossed, the conditional distribution of the positive values is governed by a truncated-at-zero count data model.[17] This model also allows for over-dispersion.

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

15 If Z ~ 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 $var(Z) = a/b^2$.

16 This assumes a homoskedastic u.

17 This was first introduced in economics by Mullahy (1986), who considers a Poisson hurdle model. See Winkelmann (1994) for additional references. Although we assume that the hurdle occurs at the first event, the model could be easily generalized to allow for the hurdle to occur at a different point, or to have several hurdles. We do not do that here, first because it makes economic sense to consider whether training incidence has a different effect to the number of training courses, and second, allowing the model to have more than one hurdle imposes severe demands on the data in terms of identifying restrictions.

Formally, let f_1 be the probability distribution function (pdf) of the process governing the hurdle (that is, the incidence of training). Let f_2 be the pdf of the process governing the number of training events once the hurdle has been crossed, called the parent-process by Mullahy (1986). 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 $\mathrm{Var}(Y_{ih})$ are given by:

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

and

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

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 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.[18] 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.[19] 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.[20]

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:

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

where the mean λ_{1i} is parameterized as $\exp(\mathbf{X}_{1i}'\beta_1)$, $a_1=1/\alpha_1$, λ_{2i} is parameterized as $\exp(\mathbf{X}_{2i}'\beta_2)$, and $a_2=1/\alpha_2$. The distributions commonly used in the literature modelling binary outcomes are the normal and logistic, which are symmetric with respect to their means and which give rise to the probit and logit models respectively. In contrast, equation (8) gives a non-symmetric distribution for the binary outcome, and nests the logit as a special case when $a_1=1$ and the extreme value distribution as a special case when $a_1=0$ (the Poisson model). In the table of results (Table 3), we also provide estimates of the probit model of training incidence.[21]

In summary, we estimate two types of count data models which allow for the

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

18 The dependence between the two processes is assumed to act via the observed covariates.

19 The geometric distribution is obtained by restricting $\alpha=1$ in equation (3).

20 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 7.0 (See Greene, 1995).

21 The probit (rather than the logit) is presented in Table 3 because it is typically used in wages or wages growth models that correct for self-selection.

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. On the basis of tests described below, the preferred model is the Negbin II hurdle specification, the estimates of which are presented in Table 3. The estimating sample comprises the 1765 young men with complete information, and who were in employment in 1981. The dependent variable is the number of training courses 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.

III.2 Discriminating Between the Models

Before discussing the estimates of the Negbin II hurdle model, we first consider the testing procedure used to discriminate between the various models. The model log-likelihoods are presented in Table 2a, and the likelihood ratio (LR) tests in Table 2b. In Tables 2a and 2b, Rows 1 and 2 refer to the Poisson and Negbin II non-hurdle models respectively. Rows 3 and 4 refer to the hurdle Poisson model, while Rows 5 and 6 refer to the hurdle Negbin II model. 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 (a t-test in this case) 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.[22] On the basis of these two tests, we reject the Poisson model; that is, Row 1 is rejected against Row 2.

[22] 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.

Table 2a: Model Log-likelihoods

Model	Log-lklhd	$a=1/\alpha$ (SE)
Non-hurdle Models		
1. Poisson	-4621.508	fixed
2. Negbin II	-3133.806	1.949 (0.107)
Hurdle Models		
3. Poisson incidence	-1787.198	fixed
4. Poisson positive counts	-2594.701	fixed
5. Negbin II incidence	-1084.579	1.304 (0.987)
6. Negbin II positive counts		1.2 (0.1604)

Table 2b: Specification Tests
(Row number refers to that given in Table 2b)

Test: χ^2(d.f.)	Null Hypothesis	LR Statistic
Row 1 against 2: $\chi^2(1)$	$a_1=0$	2975.404
Row 1 against 3 and 4: $\chi^2(40)$	$\beta_1=\beta_2$	479.218
Row 2 against 5 and 6: $\chi^2(41)$	$a_1=a_2,\ \beta_1=\beta_2$	99.656
Row 3 against 5: $\chi^2(1)$	$a_1=0$	1405.238
Row 4 against 6: $\chi^2(1)$	$a_2=0$	1190.604

Because the non-hurdle model is nested within the hurdle model, the non-hurdle model can be tested using a simple likelihood ratio test. The null hypothesis (that the non-hurdle model is appropriate) is easily rejected for both the Poisson and the Negbin II variants. This is shown in Table 2b as a test of Row 1 against Rows 3 and 4 for the Poisson model, and for the Negbin II model Row 2 against Rows 5 and 6.

Since the hurdle model is preferred, we now test the null hypothesis of the Poisson model for the hurdle specification. To do this, we test Row 3 against Row 5, and Row 4 against Row 6, as shown in Table 2b. The appropriate test in this instance is the LR test.[23] For both the hurdle incidence and for the positive counts conditional on incidence, the LR test rejects the null hypothesis that the Poisson model is appropriate. On the basis of this testing procedure, the Negbin II hurdle process is preferred, both for training incidence and for positive counts conditional on incidence.

From Columns 2 and 3 of Table 4, it can be seen that the maximized log-likelihood values for the hurdle Negbin II incidence and the probit models are very similar for this data set. We also estimated a logit model for training incidence; not surprisingly, it produced a model log-likelihood value of -1084.62, and thus did not reject the null hypothesis that $a_1=0$. From inspection of the estimated coefficients in Columns 2 and 4 of Table 4, it can be seen that the hurdle Negbin II incidence model coefficients are between 1.6 and 1.9 times larger than the probit coefficients. We choose to work with the hurdle Negbin II model, since it provides a natural model from which can be generated the necessary instruments for the two endogenous variables in the wages growth equation - training incidence and the number of training courses.

III.3 The Negbin II Hurdle Estimates

The coefficient estimates of the preferred specification, the Negbin II hurdle model, are presented in Columns 2 and 3 of Table 3. As a comparison, Columns 1 and 4

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

23 Another widely used test is the Wald test. Row 5 of Table 2a shows that the parameter a_1 is estimated to be 1.304 with an associated standard error of 0.987. 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 ($a_1=0$). (We also cannot reject the null hypothesis that the logit model is appropriate for incidence ($a_1=1$)). But in contrast, the LR test statistic gives a value rejecting the null hypothesis that $a_1=0$. A reason for this conflicting result may be the following. 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.

report respectively the estimates of the Negbin II non-hurdle model and a probit model of training incidence. Full definitions of the variables are given in Table A1 in the Appendix. The explanatory variables for the training counts models fall under five broad headings, given below.

Individual Attributes

According to human capital theory, agents will invest in training if the present discounted value of training benefits exceeds training costs (see for example Becker (1962), Mincer (1962), Oi, (1962), Hashimoto (1981), and Parsons (1990)). 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. Thus training is more likely to be offered to, or undertaken by, workers with a strong attachment to the labour market. Past work experience may be used by employers to make inferences about individual's future work commitment (Duncan and Hoffman, 1979). Full-time work experience was therefore included as a regressor; this may also proxy unmeasured on-the-job training. However, from Column 2 in Table 3 it can be seen that experience has an insignificant impact on training incidence, and (from Column 3) an insignificant effect on the expected number of training courses.

Family responsibilities may proxy the attachment of young men to the labour market and their motivation to invest in human capital. The variables married and kids by 1981 (where marriage is broadly defined to include both marriage and cohabitation), married and no kids by 1981, and unmarried but with kids by 1981, were therefore included to see if family responsibilities were associated with more training. The base category comprises childless single men. From Columns 2 and 3 it can be seen that a married or cohabiting man without any children has a significantly higher probability of receiving training and a significantly higher expected number of training courses, relative to the base group. The other variables proxying family responsibilities are insignificant.

A worker's commitment to the labour market may be reduced by disability, hence a control was included indicating whether or not the individual was registered as disabled in 1981. This had no significant impact on training incidence or the expected number of training courses. We also control for ethnic origin, since men of non-white ethnic origin may experience poorer quality schooling, which increases the costs of acquiring subsequent training (Duncan and Hoffman, 1979). However, the results indicate that men of white ethnic origin have neither a significantly higher probability of experiencing training events, nor a higher expected number of courses. However, there are only 42 non-white men in the estimating sample.

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 in the future. For this reason, the travel-to-work area (TTWA) unemployment rate in 1981 was included in the estimation, since expectations of future unemployment may be extrapolated from past unemployment rates. This variable has a significantly positive impact on training incidence, but a significantly negative effect on the expected number of training courses.

Altonji and Spletzer (1991) note that the positive correlation between education and training observed in many empirical studies may reflect complementarity in production or the presence of factors that influence investment in both forms of human capital. For example, firms will prefer to train individuals most able to benefit from 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 well-educated workers will learn faster than their less able colleagues.[24] We proxy ability by two dummy variables - reading score at age 11 below average, and mathematics score at age 11 below average. The estimated coefficient to the variable

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

24 Complementarity between worker ability and training in production also means that higher ability workers will be matched to positions requiring more training (Barron et al, 1989).

"reading score below average" is significantly negative. Young men scoring below average in nationally administered reading tests at age 11 have both a lower probability of training incidence and a lower expected number of training courses. This finding is consistent with the hypothesis that the costs of training are higher for lower ability workers.

The quality of schooling is proxied by variables indicating whether or not secondary education was received at a private school (fee paying) or a grammar or direct grant school (admission to which was typically determined by entrance examination at age 11). Neither variable has a significant effect on training.[25]

Highest educational qualification by 1981

Education is measured by five dummy variables, indicating the highest qualification attained by the respondent by 1981. These variables are Degree (the highest qualification in 1981 was a university degree); A-level (one or more advanced-level qualifications representing university entrance-level qualifications usually taken at or around the age of 18); O-level (one or more ordinary-level qualifications obtained at or around the minimum school-leaving age of 16); Vocational qualification (one or more business, technical or industrial vocational qualifications); and Apprenticeship (a trade apprenticeship typically achieved after a 3-5 year indenture period begun at age 16).

Most of the variables measuring highest educational qualifications reached by 1981 have a large significantly positive effect on the number of training courses over the period 1981-1991, ceteris paribus. The education variables having the largest impact on training incidence and the expected number of training courses are Degree and A-level. Men whose highest educational qualification was one or more O-level qualifications also experience more training. However, the impact of Vocational qualification has a significant positive effect only on training incidence (and only at

[25] Private schooling may also proxy family income, and there is no reason why this should directly affect training receipt.

the 10% level). Apprenticeship completed has an insignificant impact on incidence and on the expected number of training courses. These results are consistent with the human capital prediction that the costs of training are lower for more highly educated workers, or that training and education are complements in production. Workers with less general training may also have higher discount rates, and hence be less willing to invest in training through lower earnings.[26]

Characteristics of the 1981 Job

Wave 5 of the NCDS does not provide information on the attributes of the job or firm in which the individual received training over the period 1981-91. Even if these data were available, they would arguably be endogenous: individuals may choose to work in particular occupations or large firms, for example, because these are perceived to offer more training. The only available data on job and firm attributes is for 1981. Since this information is pre-determined, we include it in our estimation.

Union status may affect training experiences, although it is not clear a priori in which direction.[27] Unions in their monopoly role use their power over labour supply to extract a larger share of the surplus, and thereby induce deadweight losses. Higher union wages, restrictive work practices, and any union resistance to the introduction of new skill-intensive technologies, may reduce employer incentives to provide training. On the other hand, unions are in some circumstances cooperative and instrumental in improving worker morale and organization at the work place, and may thereby increase training and productivity. However, our estimates show that union membership status in 1981 has no significant effect on training, ceteris paribus.

Individuals in particular occupational groups in 1981 may be more likely to

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

[26] We are grateful to the referee for making this point.

[27] Studies using British data have found that members of a trade union are more likely to experience training (see for example Booth (1991), Tan et al (1992), Greenhalgh and Mavrotas (1994), and Green, Machin and Wilkinson (1996)). The evidence is more mixed for the US (see Brown, 1990); however, Lynch (1992) and Lillard and Tan (1992) find a significant positive correlation between union status or coverage and many forms of training.

experience training, through complementarity in production of particular occupational groups and training, or because occupational classification may proxy individual ability and motivation. We find that training incidence is higher for workers in professional or managerial occupations, or in skilled manual occupations, ceteris paribus; however, the expected number of training courses is unaffected.

Employer characteristics 1981

There are a number of hypotheses about the relationship between the incidence of work-related training courses and firm size, sector or industry. Larger firms and public sector firms may be more likely to train workers because they are more forward looking or better placed to bear any risk associated with training. Large firms may also benefit from economies of scale in training provision, or they may face more regulations and bureaucracy and so provide more training in the nature of meeting safety regulations etc. (Felstead and Green (1996)). Moreover, particular industries may by their nature require more training, or may have a past legacy of training provision through the old Industry Training Boards.[28] For this reason, 1981 industry controls were included.

While we are unable to distinguish between the various hypotheses suggesting a positive correlation between establishment size, sector and work-related training, our results in Table 3 do show that more training was received by young men employed in 1981 in private sector establishments with more than 25 employees, or in public sector establishments of all sizes, relative to the base group of private sector establishments with fewer than 25 employees.[29]

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

28 The 1964 Industrial Training Act established 27 Industrial Training Boards, which determined and exacted at the industry level training levies from firms, and then reallocated the funds to firms in the industry providing their workforces with training. This was an attempt to overcome the "poaching" problem, whereby non-training firms poached trained workers from training firms (see Stevens (1996)).

29 This firm size effect is found in many studies of training incidence; see inter alia Brown (1990), Lynch (1992) and Tan et al (1992) for the US, and Booth (1991, 1993) and Tan et al (1992) for Britain.

Training prior to 1981

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-sectional nature of our data (with retrospective information for training between 1981 and 1991), we are unable to address this issue properly. Nonetheless, to try to control for this in the estimation, we include a set of variables measuring pre-1981 training, under the heading training prior to 1981. These variables generally have a significantly positive effect on the incidence and expected number of training courses over the period 1981 to 1991, ceteris paribus. The appropriate LR tests for the specifications with and without this set of controls rejected the null hypothesis that the pre-training variables had no effect.[30]

Summary

Perhaps the most striking result to emerge from the estimates is the strong positive correlation between past general education or past training on the one hand, and subsequent training experiences on the other, a result also found in many earlier studies of training incidence. This finding is consistent with less educated workers having greater discount factors, training and education being complements in production, or with training costs being lower for more highly educated workers. It also suggests that young men entering the labour market with low levels of education have limited training opportunities in the workplace.

Recent theoretical work by Snower (1996) and Burdett and Smith (1995) has considered a "low-skill, bad-job" trap, in which there may be multiple equilibria in the market for skills, and policy intervention may be required to shift workers stuck

---

30 However, interpretation of their impact should be made with caution, since they could simply be proxying unobservable characteristics rather than measuring the true impact of state dependence in training experiences.

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, our results are consistent with it.

In the next section we investigate the extent to which workers with work-related training receive higher wages growth. In particular, we estimate the impact on expected wages growth of additional training courses, for young men who have experienced at least one training course over the period 1981 to 1991. The results generated from the Negbin II hurdle count model are used to control for training endogeneity in the wages growth equation.

IV. MODELLING THE IMPACT OF TRAINING ON WAGES GROWTH

IV.1 The Econometric Model of Wages Growth

We have already noted that 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 a wages growth model. Consider the following log-linear wage equations:

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

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

where w_{i1} and w_{i2} denote the natural logarithm of real hourly wages of the i-th individual in 1981 and 1991 respectively. 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}. The usual observable variables representing both individual and firm characteristics are given by the vectors z_{i1} and z_{i2}; these variables can be either time-fixed or time-varying, and the vectors γ_1 and γ_2 are parameters associated with these 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 course 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.

It is uncommon in the literature to have both T_i and T_iN_i in the wage equation. A model which includes only T_i assumes that what matters for wages growth is whether or not an individual ever experiences a training event, and not the actual number of experiences. Another way of looking at this is that expected wages growth is determined solely by the first training incident and not subsequent experiences. In contrast, a wages growth model including only T_iN_i (the number of training experiences) assumes that all training experiences matter <u>and</u> that they all have the same effect, ceteris paribus. Here we include both variables in the model.

Subtraction of (9) from (10) yields the wages 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 wages are in logarithms, the first difference can be interpreted as measuring approximate wages 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.[31]

The Endogeneity Issue

The issue of endogeneity arises when participation in a training program is not random. The wages 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 individuals receiving training are more motivated than non-participants, and motivation is unobservable. If highly motivated individuals also have higher wages growth, 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. However, there may be correlation between the training variables and the error term ε in equation (11). This may arise, for example, where there is a temporary unobserved adverse demand shock, that causes both less training and lower wages growth. 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, because our training models differ from the standard probit model for training incidence, we take a different approach.[32] For expositional ease, rewrite equation (11) as:

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

---

31 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.

32 See Heckman and Robb (1985) for a clear exposition of selectivity models.

Note that the signs of the coefficients in the vector γ in (12) depend on the relative magnitudes of the corresponding γ_1 and γ_2 in (9) and (10). The percentage contribution of each coefficient estimate of the impact of a dummy variable on wages growth is given by $[\exp(\gamma)-1].100$. Conditional on the z_i, we can write

$$E(\Delta w_i) = z_i\gamma + \delta_1 E(T_i) + \delta_2 E(T_iN_i) \qquad (13)$$

From equations (4a) and (5),

$$E(T_i) = 1-f_{1i}(0) \quad \text{and} \quad E(T_iN_i) = E(Y_{ih}) = \theta_i\lambda_{2i} \qquad (14)$$

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 endogenous variables T_i and T_iN_i in (12). These will be valid instruments if $E(T_i|z_i,\varepsilon_i) = E(T_i|z_i)$ and $E(T_iN_i|z_i,\varepsilon_i) = E(T_iN_i|z_i)$. Using the procedure set out in Murphy and Topel (1985), we also correct the standard errors to take into account the nature of the generated regressors; details of the derivation are given in the Appendix.

Identification

Identification is achieved by functional form, and through the omission from the wages growth equation of some of the variables included in the training model, and vice versa. Wages growth is determined not only by exogenous variables expressed as levels in 1981, but also by changes occurring in these variables over the period 1981-91. The inclusion of changes in the wages growth model is equivalent to including z_t, z_{t-1} and z_tz_{t-1} as regressors, where z represents an exogenous characteristic from the vector of exogenous variables z and the subscripts t and $t-1$ refer to 1991 and 1981 respectively. In the training models, only variables pre-determined in 1981 were included as regressors. The TTWA unemployment rate in 1981 was included in the training counts models, but since this variable was unavailable for 1991, regional unemployment rates and shifts into and out of London were included in the wages growth models. In the training counts models, children were interacted with marital status, while in the wages growth model, only changes to marital status were included. In the training counts models, the number of pre-1981 training courses was included as a set of dummy variables, whereas in the wages growth equation a different specification was adopted.[33] Disaggregated establishment size interacted with sector was found to have a significant effect on training counts, but an insignificant effect on wages growth. Hence only the more aggregated establishment size variables of moving from a smaller to a larger establishment, or from a larger to a smaller establishment, were included in the wages growth equation. Moreover, changes in sector were included as separate regressors in the wages growth equation.

Sample Selection Issues

In order to carry out the estimation procedure described above, we require a sub-sample of individuals who were employed in both 1981 and 1991, since wages are available only for individuals employed at both dates. A natural question therefore arises as to the possible endogeneity of employment status. However, our preliminary work (using bivariate sample-selection corrections) found no evidence of sample selection bias; we therefore report only the results from estimation of models that do not make this correction.

V.2 Estimating the Impact of Training on Wages Growth

Tables 4 and 5 present estimates of the impact on wages growth of the number of training events experienced by an individual over the period 1981 to 1991. The wages data used are usual gross hourly wages received at the survey dates of 1981 and 1991, deflated to 1981 values using the Consumer Price Index. Mean <u>real</u> hourly earnings were

[33] Although some of these variables were insignificant in some specifications, out of interest they have been retained in reported specifications.

£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 1301 young men in employment in both 1981 and 1991 for whom we have complete information.[34]

Table 4 summarizes the estimated effects on wages growth obtained from a number of specifications with different treatment of training endogeneity. Column 1 in Table 4 presents the training effects from a model in which T_i and T_iN_i are included as regressors, Column 2 includes only T_i, while Column 3 includes only T_iN_i. The rows in Table 4 indicate which (if any) of the training variables are treated as endogenous, and what instruments are used. Thus, for example, Row 1 contains estimates from specifications in which both predicted T_i and T_iN_i from the Negbin II hurdle model are used as instruments.

The most interesting results summarized in Table 4 are the following. First, in the models where the effects of all training courses on expected wages growth is assumed to be the same (Column 3), significant returns to training are found only when training is treated as exogenous. Second, in the models where only the effect of the first training course is assumed to affect wages growth (Column 2), significant returns to the first training event are found only when training is treated as endogenous. When the standard probit model is used instead of the Negbin II hurdle model to instrument training incidence, the former estimates that training incidence increases wages growth by 52.2%, about 4.5 percentage points higher than the latter.

Third, consider models allowing the first training experience to have a different effect on wages growth to subsequent training experiences (Column 1). These models produce a significant training incidence effect only if training is treated as endogenous. Moreover, when the relevant instruments are generated from the Negbin II hurdle model, the first training experience is found to increase expected wages growth

[34] One-digit industry dummy variables were included in both the wages equations and the training counts model. To maximize the sample size, we also included a dummy variable taking the value unity for "missing industry". In 1981 2% of the sample had missing industry data, compared to 26% in 1991. (The larger number of missing industry cases in 1991 may have been due to the fact that there were changes over the period in the Standard Industrial Classification).

by 55%. But when a standard probit is used to instrument training incidence, expected wages growth is increased by only 45.7%. If incidence and the number of courses is not treated as endogenous, there is an under-prediction of incidence.

[Insert Table 4 near here]

We now turn to a more detailed comparison of the OLS and IV wages growth estimates, given in Table 5. The estimated effect on wages growth of a time-invariant variable, such as the highest educational qualification received by 1981, will be significant in our model only if the effects differ in each of the wage equations (9) and (10). The relative magnitude of the effects at each separate time period determines the sign of the net effect of a variable on wages growth. Columns 1 and 3 of Table 5 present the OLS estimates. If the training variables are weakly exogenous, OLS is a valid estimation procedure that will produce consistent parameter estimates. Columns 2 and 4 present 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.

[Insert Table 5 near here]

From Column 1 in Table 5, it can be seen that the block of variables indicating highest educational qualification by 1981 has a significantly positive effect on wages growth, while training over the period 1981-1991 has an insignificant effect. Now consider the estimates in Column 2, which reports the specification in which training incidence and the number of courses are instrumented using the Negbin II hurdle estimates. While most of the coefficient estimates remain broadly unchanged across the specifications in Columns 1 and 2, there is a dramatic change in the estimated impact of the human capital variables. Education received prior to 1981 now has no significant impact on wages growth, while the estimated effect of training incidence becomes large and statistically significant. It appears that, once control has been taken of training endogeneity, educational qualifications by 1981 do not directly affect wages growth. However, they do affect wages growth indirectly, through significantly increasing individual selection into work-related training courses.

But do men with different educational backgrounds have different returns to training? Columns 3 and 4 in Table 5 present the OLS and IV estimates of an augmented specification of wages growth, in which training courses are interacted with highest education qualification by 1981, in order to see if men with different educational backgrounds have different returns to training.[35] Only the interaction of Degree and A-level with training incidence has a significant impact, and the impact is large: wages growth for men whose highest qualification in 1981 was a degree or an A-level, and who experience a WRT course, is 127 percent higher than for men with no qualifications in 1981 (obtained from $\{(e^{0.820}-1)100\}$). An O-level qualification by 1981 and no WRT is associated with wages growth of 28 percent.

Given that most of the coefficient estimates (with the exception of the human capital variables) are broadly unchanged across OLS and IV specifications, it is perhaps not surprising that the Wu-Hausman-Durbin test comparing all the coefficients in each specification was unable to reject the null hypothesis of exogeneity. When a similar test is carried out on the sub-set of human capital variables, however, the null hypothesis is strongly rejected. On this basis, our preferred specification is Column 4 in Table 5.

While the focus of this paper is on training courses, it is interesting to note the other significant results of the wages growth models. The magnitude of the other variables' coefficients is robust across all four specifications. Of the fixed individual characteristics, only <u>secondary education at a direct grant or grammar school</u> significantly affects wages growth. This variable proxies both ability and school quality, since access to such schools was typically determined on the basis of exams results at the age of 11, and the curriculum at these schools was geared to the

---

35 Because of small cell sizes, it was not possible to estimate wages growth equations separately for each pre-1981 highest qualification.

academically able. Young men who had in 1981 a disability which affects work also experienced significantly higher wages growth, but with the exception of Column 4, this is significant only at the 10% level. This may reflect the protected nature of employment and wages for disabled workers in Britain.

Men who were union members in 1981 but not 1991 have significantly lower wages growth than the base - men who were not in a union in either 1981 or 1991. This finding is consistent with a positive union-nonunion wage gap in 1981, and a positive age-wage profile for nonunion workers over the decade 1981-91.[36]

Some changes in firm characteristics also have a significant impact on wages growth. A shift from the public sector in 1981 to the private sector in 1991 is associated with significantly lower wages growth, relative to the base of public sector employment at both dates. This is as expected in the British context, where public sector wages have traditionally been determined predominantly by union bargaining, whereas in the private sector wages are more responsive to market forces. Men who have moved from a smaller to a larger size establishment over the period 1981 to 1991 experience significantly higher wages growth, relative to men who work in the same size establishments at both periods. This reflects the fact that in Britain larger firms pay more than smaller firms ceteris paribus, perhaps due to compensating wage differentials and the higher likelihood of wages being determined by union-bargaining in larger establishments.[37]

Finally, note that men whose occupation was professional, managerial or administrative workers in both 1981 and 1991 are characterized by significantly higher wages growth, relative to the base of non-managerial or professional at both periods. Men who shift into the professional, managerial or administrative category over the

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

36 The negative coefficient to the member-to-member variable, although insignificant, is consistent with unions being associated with flatter age-wage profiles, or alternatively with the anti-union legislation in Britain in the 1980s reducing the power of trade unions. See Booth and Frank (1996) for a study of union-nonunion age-wage profiles in Britain.

37 We found that job tenure has no significant effect on affect wages growth; other studies using British data also find that tenure is insignificant (see for example Booth (1993) and Tan et al (1992).

period also experience significantly higher wages growth, although the impact is not so large as for men who have been professional, managerial or administrative over the entire period.

The variables that have an insignificant impact on wages growth are also of interest. Marital status and changes in marital status, changes in region of residence and regional unemployment rates, ability proxies (reading and math scores below average at age 11) and ethnic origin all have an insignificant impact on wages growth over the decade.

Summary

Our analysis is novel, in that we are first to use data on the number of work-related training courses to control for training endogeneity in a wages growth equation. But what has been learnt from this exercise? First, we have found that hurdle model best describes the data, and therefore we used predictions from the hurdle model to control for self-selection into training, which has not been done before. Second, we found that the direct impact of the pre-training highest educational qualification (apart from O-levels) on wages growth becomes insignificant when training endogeneity is controlled for by any method. While pre-training highest educational qualifications are the major significant determinant of training incidence and the expected number of WRT courses, they affect wages growth only indirectly through increased access to productivity-augmenting WRT courses, rather than directly. An exception is provided by the two highest pre-training educational qualifications - Degrees and A-levels: workers with these educational backgrounds have higher returns to training ceteris paribus. Third, we find that only training incidence matters for wages growth; the expected number of courses has no significant impact when incidence and/or the number of courses are treated as endogenous.

What implications do our results have for other researchers in this area? In our comparison of various methods of controlling for training endogeneity, we found a large significant impact on wages growth of training incidence modelled in any way (probit, Negbin II etc). The biggest impact is through the Negbin II hurdle approach in which both T_i and T_iN_i are instrumented. However, all methods used show that the number of events has little effect when incidence and/or the number of courses is endogenized. We would emphasize that these results cannot be interpreted as a license to continue endogenizing only training incidence, for two reasons. First, our results show that other methods produce a smaller estimated training effect on wages growth, so if people do endogenize only incidence using a probit/logit they may be under-estimating the training impact on wages growth. Second, our results may be an artifact of our particular data set; we would want to see if the same results hold using data from other surveys, before drawing the conclusion that only incidence matters.

V. CONCLUSIONS

This paper estimates the determinants of the number of work-related training courses, and their impact 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 allows us to account for the fact that half of all sample members experienced no work-related training over the period 1981 to 1991. The results generated from the hurdle count model are then used to control for training endogeneity in the wages growth equation. This has not been done before in the training literature. The sensitivity of the estimates to alternative modelling strategies is also examined. Since we find a strong correlation between previous education and subsequent training experiences, we also experiment with estimating the joint impact of previous education and subsequent training on wages growth, in order to try to tease out the combined effect of these variables. We find that the biggest returns to training are to highly educated men.

Table 3: Determinants of the Number of Training Courses for Employed Young Men

Variable	Negbin II (1)	Negbin II Hurdle Training Incidence (2)	Negbin II Hurdle Positive Counts (3)	Probit Training Incidence (4)	Mean (5)
Constant	-1.044 (0.413)**	-2.237 (0.542)**	0.384 (0.581)	-1.327 (0.297)**	
Individual attributes					
Full-time experience (months)/100	0.112 (0.245)	0.138 (0.309)	-0.017 (0.313)	0.067 (0.170)	0.704
Married and kids by 1981	0.067 (0.151)	0.212 (0.194)	-0.071 (0.184)	0.125 (0.104)	0.122
Married and NO kids by 1981	0.426 (0.097)**	0.394 (0.174)**	0.389 (0.104)**	0.228 (0.071)**	0.318
Not married with kids by 1981	-0.270 (0.821)	-0.161 (1.238)	-0.287 (1.251)	-0.097 (0.610)	0.003
Registered disabled 1981	0.049 (0.244)	-0.052 (0.332)	0.053 (0.294)	-0.025 (0.179)	0.033
White	0.374 (0.302)	0.411 (0.351)	0.271 (0.447)	0.241 (0.208)	0.976
Local unemployment rate - TTWA	-0.009 (0.013)	0.033 (0.019)*	-0.033 (0.013)**	0.018 (0.009)**	11.040
Reading score below average	-0.342 (0.111)**	-0.377 (0.177)**	-0.233 (0.127)*	-0.213 (0.080)**	0.435
Maths score below average	-0.033 (0.111)	-0.035 (0.153)	-0.076 (0.124)	-0.026 (0.084)	0.509
Educated at grammar or direct grant school	-0.113 (0.157)	-0.216 (0.227)	-0.061 (0.159)	-0.111 (0.111)	0.109
Educated at private school	0.003 (0.283)	0.225 (0.392)	-0.162 (0.258)	0.082 (0.207)	0.026
Highest Educational Qualifications 1981					
Degree	1.058 (0.260)**	1.270 (0.507)**	0.656 (0.304)**	0.707 (0.188)**	0.077
A-levels	1.002 (0.223)**	1.346 (0.505)**	0.586 (0.250)**	0.729 (0.144)**	0.128
O-levels	0.651 (0.134)**	0.656 (0.232)**	0.402 (0.175)**	0.378 (0.102)**	0.399
Vocational qualification	0.385 (0.154)**	0.378 (0.220)*	0.214 (0.196)	0.228 (0.118)**	0.161
Apprenticeship completed	0.135 (0.122)	0.219 (0.178)	0.028 (0.136)	0.117 (0.094)	0.255
Job Characteristics - 1981					
Union Member	0.095 (0.108)	0.149 (0.156)	0.072 (0.112)	0.081 (0.080)	0.613
Job is Professional/Managerial	0.136 (0.121)	0.400 (0.200)**	-0.090 (0.125)	0.221 (0.092)**	0.234
Job is Skilled Manual	0.213 (0.156)	0.647 (0.278)**	-0.069 (0.170)	0.358 (0.104)**	0.156
Employer Characteristics - 1981					
Private Sector and 26-99 employees	0.441 (0.144)**	0.263 (0.211)	0.402 (0.170)**	0.146 (0.111)	0.148
Private Sector and 100-499 employees	0.491 (0.151)**	0.602 (0.267)**	0.207 (0.175)	0.334 (0.116)**	0.143
Private Sector and 500 or more employees	0.583 (0.168)**	0.840 (0.330)**	0.243 (0.182)	0.464 (0.127)**	0.139
Public Sector and less than 25 employees	0.542 (0.221)**	0.585 (0.346)*	0.171 (0.239)	0.332 (0.172)*	0.053
Public Sector and 26-99 employees	0.587 (0.215)**	0.478 (0.313)	0.335 (0.238)	0.275 (0.158)*	0.081
Public Sector and 100-499 employees	0.456 (0.212)**	0.637 (0.332)*	0.101 (0.219)	0.357 (0.151)*	0.096
Public Sector and 500 or more employees	0.595 (0.182)**	0.817 (0.342)**	0.249 (0.196)	0.449 (0.137)**	0.143
Training prior to 1981					
1 Training course - excl. apprenticeship	0.381 (0.100)**	0.272 (0.166)	0.375 (0.107)**	0.152 (0.082)*	0.218
2 Training courses	0.387 (0.154)**	0.488 (0.264)*	0.226 (0.167)	0.256 (0.113)**	0.099
3-4 Training courses	0.662 (0.215)**	0.893 (0.428)**	0.505 (0.200)**	0.486 (0.145)**	0.061
More than 4 courses	1.122 (0.437)**	1.858 (0.961)*	0.820 (0.330)**	0.973 (0.264)**	0.021
Variance Parameter a=1/α	1.949 (0.107)**	1.304 (0.987)	1.204 (0.168)**		
Model log-likelihood	-3133.81	-1084.58	-1999.40	-1084.72	
Log-likelihood at a=0 (i.e Poisson)	-4621.51	-1787.20	-2594.70		
Number of cases	1765	1765	889	1765	1765

Notes: (i) All models include 1981 industry dummies. (ii) Standard errors are given in parentheses.

Table 4 - Estimated effects of training experiences on wages-growth - Summary [instruments used]

	Wages-growth equation has both variables T_i and T_iN_i [1]	Wages-growth equation only has T_i [2]	Wages-growth equation only has T_iN_i [3]
[1] Model used for training experiences is the NegbinII hurdle model. Endogenous variables are T_i and/or T_iN_i	55.0% on expected wages-growth from the first training experience and no significant additional effects from subsequent experiences. $[E(T_i)=1-f_{1i}(0);\ E(T_i N_i) = \theta_i \lambda_{2i}]$	47.7% on expected wages-growth from the first training experience. $[E(T_i)=1-f_{1i}(0)]$	No significant effects from training experiences. $[E(T_i N_i) = \theta_i \lambda_{2i}]$
[2] Model used for training experiences is the NegbinII hurdle model. Endogenous variable is T_i	41.9% on expected wages-growth from the first training experience and no significant additional effects from subsequent experiences. $[E(T_i)=1-f_{1i}(0);\ E(T_i N_i) = (1-f_{1i}(0)) N_i]$	47.7% on expected wages-growth from the first training experience. $[E(T_i)=1-f_{1i}(0)]$	No significant effects from training experiences. $[E(T_i N_i) = (1-f_{1i}(0)) N_i]$
[3] Probit model is used for training incidence. Endogenous variable is T_i.	45.7% on expected wages-growth from the first training experience and no significant additional effects from the subsequent experiences. $[E(T_i)=1-\Phi_i(.);\ E(T_i N_i) = (1-\Phi_i(.)) N_i]$	52.2% on expected wages-growth from the first training experience. $[E(T_i) = 1-\Phi_i(.)]$	No significant effects from training experiences. $[E(T_i N_i) = (1-\Phi_i(.)) N_i]$
[4] Both incidence and the number of training experiences are treated as exogenous in the wages-growth equation.	No significant effects from training experiences.	No significant effects from training experiences.	0.6% per training event on expected wages-growth.

Notes:

(i) See text for definitions of T, N, θ, f_1, λ_2. Φ is the distribution function of a Normal variate.

(ii) All equations had other control variables as reported in Table 5.

Table 5- Approximate Earnings Growth 1981 - 1991 (dependent var=$\ln(wage_{91})-\ln(wage_{81})$)

Variable	OLS[@] (1)	IV (2)	OLS[@] (3)	IV (4)	Mean (5)
Intercept	0.558 (0.153)**	0.486 (0.163)**	0.554 (0.154)**	0.411 (0.179)**	
Fixed individual characteristics					
Reading score below average	-0.001 (0.034)	0.029 (0.039)	-0.001 (0.034)	0.033 (0.039)	0.417
Maths score below average	-0.040 (0.038)	-0.033 (0.038)	-0.041 (0.038)	-0.033 (0.037)	0.489
School - direct grant/grammar	0.097 (0.044)**	0.106 (0.050)**	0.093 (0.044)**	0.108 (0.050)**	0.113
- private	0.115 (0.110)	0.103 (0.090)	0.114 (0.110)	0.090 (0.095)	0.027
Ethnicity - white	-0.100 (0.085)	-0.131 (0.098)	-0.108 (0.084)	-0.130 (0.098)	0.979
Individual characteristic in 1981					
Has a disability which affects work	0.179 (0.097)*	0.198 (0.104)*	0.180 (0.097)*	0.208 (0.105)**	0.020
Individual characteristic in 1991					
Has a disability which affects work	-0.031 (0.079)	-0.024 (0.090)	-0.032 (0.079)	-0.024 (0.090)	0.025
Changes across period 1981-1991					
Trade union membership					
Non member to a member	-0.049 (0.058)	-0.068 (0.061)	-0.042 (0.058)	-0.061 (0.061)	0.067
Member to a non member	-0.092 (0.040)**	-0.116 (0.043)**	-0.088 (0.040)**	-0.115 (0.044)**	0.201
Member to a member	-0.007 (0.039)	-0.031 (0.042)	-0.004 (0.039)	-0.027 (0.042)	0.416
Regional changes in residence					
Outside to London	-0.008 (0.131)	-0.005 (0.143)	0.003 (0.130)	-0.005 (0.143)	0.009
London to outside	-0.016 (0.069)	-0.011 (0.059)	-0.013 (0.068)	-0.012 (0.059)	0.068
London to London	0.021 (0.070)	0.030 (0.078)	0.027 (0.069)	0.043 (0.078)	0.035
Firm type					
Public to Private	-0.131 (0.045)**	-0.126 (0.051)**	-0.127 (0.045)**	-0.122 (0.051)**	0.151
Private to Public	-0.016 (0.067)	0.027 (0.066)	-0.020 (0.067)	0.029 (0.067)	0.085
Private to Private	-0.052 (0.049)	-0.024 (0.056)	-0.047 (0.049)	-0.018 (0.056)	0.543
Firm size					
to larger	0.113 (0.033)**	0.127 (0.034)**	0.113 (0.033)**	0.128 (0.034)**	0.278
to smaller	-0.009 (0.033)	-0.019 (0.033)	-0.012 (0.033)	-0.014 (0.033)	0.304
Job type					
Other to Prof/Manag/Admin	0.131 (0.040)**	0.128 (0.038)**	0.131 (0.040)**	0.127 (0.038)**	0.203
Prof./Manag/Admin to other	0.011 (0.071)	-0.006 (0.071)	0.004 (0.071)	-0.008 (0.070)	0.044
No change at Prof/Manag/Admin	0.186 (0.044)**	0.168 (0.045)**	0.183 (0.044)**	0.173 (0.045)**	0.203
Marital status					
not married to married	-0.027 (0.046)	-0.026 (0.042)	-0.024 (0.046)	-0.025 (0.042)	0.422
married to not married	-0.088 (0.072)	-0.110 (0.089)	-0.092 (0.071)	-0.126 (0.090)	0.029
married to married	-0.005 (0.047)	-0.022 (0.045)	-0.000 (0.047)	-0.029 (0.046)	0.404
Regional unemployment rate - %					
1981	-0.006 (0.010)	-0.006 (0.010)	-0.006 (0.010)	-0.006 (0.010)	11.400
1991	-0.008 (0.020)	-0.012 (0.021)	-0.008 (0.020)	-0.012 (0.021)	7.964

Table 5 continued

Variable	OLS@ (1)	IV (2)	OLS@ (3)	IV (4)	Mean (5)
Highest qualification obtained prior to 1981 (base is no educational qual)					
Degree	0 .205 (0.068)**	0.123 (0.086)	0.249 (0.086)**	0.020 (0.244)	0.076
Advanced Level (A.L.)	0.115 (0.056)**	0.009 (0.078)	0.152 (0.074)**	-0.107 (0.255)	0.139
Ordinary Level (O.L.)	0.079 (0.041)*	0.015 (0.055)	0.073 (0.049)	0.248 (0.133)*	0.404
Vocational	0.094 (0.044)**	0.047 (0.057)	0.071 (0.051)	-0.005 (0.168)	0.161
Training received prior to 1981					
Apprenticeship completed	-0.082 (0.036)**	-0.075 (0.041)*	-0.083 (0.036)**	-0.080 (0.040)**	0.266
Employer provided training in current job in 1981	-0.027 (0.030)	-0.033 (0.031)	-0.026 (0.030)	-0.031 (0.031)	0.553
Employer provided training in first job if current job # first job	-0.009 (0.030)	-0.011 (0.030)	-0.007 (0.030)	-0.014 (0.031)	0.339
Education and Training received over the period 1981-1991					
Atleast one educational course followed(dummy)	0.016 (0.033)	0.009 (0.033)	0.013 (0.033)	0.008 (0.033)	0.355
Number of educational courses followed+	0.004 (0.003)	0.003 (0.006)	0.004 (0.003)	0.003 (0.006)	0.790
Atleast one Training course (dummy) - Ti	0.015 (0.033)	0.438 (0.206)**			
and (Degree+AL) pre 81			-0.002 (0.077)	0.820 (0.414)**	0.149
and OL pre 81			0.018 (0.052)	-0.019 (0.270)	0.228
and vocational qual pre 81			0.015 (0.078)	0.676 (0.502)	0.074
and no qualification pre 81			-0.029 (0.057)	0.674 (0.496)	0.070
Number of training courses experienced+ - Ni Ti	0.005 (0.004)	-0.007 (0.018)			
and (Degree+AL) pre 81			-0.002 (0.006)	-0.032 (0.027)	
and OL pre 81			0.007 (0.006)	0.027 (0.025)	
and vocational qual pre 81			0.018 (0.013)	-0.001 (0.066)	
and no qualification pre 81			0.019 (0.008)**	-0.016 (0.087)	
Standard error of the regression	0.475	0.475	0.476	0.475	
R-bar squared	0.120	0.122	0.119	0.123	
Sample size	1301	1301	1301	1301	

Notes: (i) Mean of dep. var=0.432.

(ii) All equations have, in addition to the above covariates, the industry dummies for the two years.

(iii)@ Standard errors are the heteroskedastic consistent standard errors.

(iv) **, *Coefficient significant at 5% and 10% significance levels respectively.

(v) In columns 2 and 4, the training variables T_i and T_iN_i variables are replaced with their predicted values and the resulting equation estimated by OLS. The standard errors are corrected for the use of these generated regressors (see text).

Table A1: Definitions of Variables

Variable Name	Definition
TRAINING COUNT MODELS	
Individual attributes	
Full-time experience (months)/100	Full-time experience in months since labour market entry/100
Married and kids by 1981	Married/cohabiting interacted with any children, by 1981
Married and NO kids by 1981	Married/cohabiting but without any children, by 1981
Not married with kids by 1981	Neither married/cohabiting, but with children, by 1981
Registered disabled 1981	Registered as disabled in 1981
White	White ethnic origin
Local unemployment rate - TTWA	Travel-to-work area unemployment rate in 1981
Reading score below average	Scored below average in national test of reading ability at age 11
Maths score below average	Scored below average in national test of math reasoning at age 11
Educated at grammar or direct grant school	Secondary education at state-funded school, selection to which was typically based on ability as measured at age 11
Educated at private school	Secondary education at fee-paying private school
Highest Educational Qualifications 1981	
Degree	University degree
A-levels	One or more Advanced level qualifications representing university entrance-level qualification, typically taken at about age 18
O-levels	One or more Ordinary level qualifications taken at about age 16, at end of compulsory schooling; selection mechanism into A-level courses
Vocational qualification	One or more business/technical/industrial vocational qualiifications
Apprenticeship completed	Trade apprenticeship, with 3-5 year indenture period, begun at age 16
Job Characteristics - 1981	
Union Member	Member of a union or staff association
Job is Professional/Managerial	Professional, managerial or administrative occupation
Job is Skilled Manual	Skilled manual occupation
Employer Characteristics - 1981	
Private Sector and 26-99 employees	Private sector establishment , with 26-99 employees at workplace
Private Sector and 100-499 employees	Private sector establishment, with 100-499 employees at workplace
Private Sector and 500 or more employees	Private sector establishment, with 500 or more employees at workplace
Public Sector and less than 25 employees	Public sector establishment, with fewer than 25 employees at workplace
Public Sector and 26-99 employees	Public sector establishment, with 26-99 employees at workplace
Public Sector and 100-499 employees	Public sector establishment, with 100-499 employees at workplace
Public Sector and 500 or more employees	Public sector establishment, with 500 or more mployees at workplace
Training prior to 1981	
1 Training course - excl. apprenticeship	One work-related training course (excluding apprenticeship), up to 1981
2 Training courses	Two work-related training courses up to 1981
3-4 Training courses	Three or four work-related training courses up to 1981
More than 4 courses	More than four work-related training courses up to 1981
WAGES GROWTH MODELS*	
Changes across period 1981-1991	
Trade union membership	
Non member to a member	Non-member of union in 1981, but member in 1991
Member to a non member	Member of union in 1981, but non-member by 1991
Member to a member	Member of union or staff association in both 1981 and 1991
Regional changes in residence	
Outside to London	Resident outside London in 1981, but in London in 1991
London to outside	London resident in 1981, but outside London by 1991
London to London	London resident in both 1981 and 1991
Firm type	
Public to Private	Worked in public sector in 1981, but private sector by 1991
Private to Public	Worked in private sector in 1981, but public sector by 1991
Private to Private	Worked in private sector in both 1981 and 1991
Firm size	
to larger	Over 1981-91, shifted to working in a larger establishment
to smaller	Over 1981-91, shifted to working in a smaller establishment
Job type	
Other to Prof/Manag/Admin	Moved into prof/managerial/admin occupation 1981-91
Prof./Manag/Admin to other	Moved out of prof/managerial/admin occupation 1981-91
No change at Prof/Manag/Admin	Professional/managerial/admin occupation 1981 and 1991
Marital status	
not married to married	Single in 1981, but married/cohabiting in 1991
married to not married	Married/cohabiting 1981,but neither married/cohabiting in 1991
married to married	Married/cohabiting in 1981 and 1991

Variable Name	Definition
Regional unemployment rate - %	
1981	Percentage regional unemployment rate in 1981
1991	Percentage regional unemployment rate in1991
Training received prior to 1981	
Apprenticeship completed	Completed an apprenticeship by 1981
Employer provided training in current job in 1981	Received employer-provided training course in current job in 1981
Employer provided training in first job if current job # first job	Employer provided training received in first job, conditional on 1981 job not being first job
Education and Training received over the period 1981-1991	
At least one educational course followed (dummy)	One or more courses intended to lead to an educational qualification followed over the period 1981-91
Number of educational courses followed	Number of courses intended to lead to an educational qualification followed over the period 1981-91

* Notes: Some of the variables included in the TrainingCounts Models were also included in the Wages Growth Models; these are not listed under the Wages Growth heading here to avoid repetition.

APPENDIX: STANDARD ERROR CORRECTION

Consider equation (12) in the text, which we rewrite below as equation (A1):

$$\Delta w_i = z_i \gamma + \delta_1 T_i + \delta_2 (T_i N_i) + u_i \tag{A1}$$

In the most general model we estimate, both the variables T as well as (TN) are treated as endogenous. We therefore replace these variables by their estimated conditional expectations obtained from the NegbinII hurdle model estimates, and estimate the resulting equation by Ordinary Least Squares (OLS). The conditional expections for the above two variables in this model are given by (where the notation is as given in the main text):

$$E(T_i) = 1\text{-}f_{1i}(0); \quad E(T_i N_i) = \theta_i \lambda_{2i} \tag{A2}$$

If we assume that, conditional on training receipt, the number of training experiences is exogenous in the wages growth model, we only need to replace T and not N by its estimated conditional expectation. Equation (A2) becomes

$$E(T_i) = 1\text{-}f_{1i}(0); \quad E(T_i N_i) = [1\text{-}f_{1i}(0)]N_i \tag{A3}$$

In the case where we assume that the process governing the training incidence is Normal, f_1 will be the cumulative distribution function of a Normal variate.

The OLS standard errors were adjusted in order to take into account the nature of the generated regressors, using Murphy and Topel's equation (15) (see Murphy and Topel (1985)). Let $E(T_i) = \omega_{1i}$ and $E(T_i N_i) = \omega_{2i}$.

The calculation of the standard errors requires calculation of the derivatives of ω_1 and ω_2 with respect to the parameters of the models for the determinants of training. Each NegbinII model in the hurdle specification has two sets of parameters: the parameters in the β vector and the variance parameter a. The relevant derivatives for the models estimated in the paper are given below.

(i) Hurdle Negbin II model where both T_i and T_iN_i are treated as endogeneous:

$$\frac{\partial \omega_{1i}}{\partial \beta_{1j}} = \frac{f_{1i}(0)\, \lambda_{1i}\, x_{1ij}}{[1 + a_1 \lambda_{1i}]} \qquad \text{(for the jth parameter in } \beta_1\text{)} \tag{A4}$$

$$\frac{\partial \omega_{1i}}{\partial a_1} = f_{1i}(0) \left[\frac{\lambda_{1i}}{a_1(1 + a_1 \lambda_{1i})} - \frac{\ln(1 + a_1 \lambda_{1i})}{a_2} \right] \tag{A5}$$

$$\frac{\partial \omega_{1i}}{\partial \beta_{2j}} = \frac{\partial \omega_{1i}}{\partial a_2} = 0 \qquad \text{(A6) and (A7)}$$

$$\frac{\partial \omega_{2i}}{\partial \beta_{1j}} = \frac{\lambda_{2i}\, \lambda_{1i}\, f_{1i}(0)\, x_{2ij}}{(1 - f_{2i}(0))\, [1 + a_1 \lambda_{1i}]} \qquad \text{(for the jth parameter in } \beta_1\text{)} \tag{A8}$$

$$\frac{\partial \omega_{2i}}{\partial a_1} = \frac{\lambda_{2i}}{(1 - f_{2i}(0))} \frac{f_{1i}(0)}{a_1} \left[\frac{\lambda_{1i}}{[1 + a_1 \lambda_{1i}]} + \ln(f_{1i}(0)) \right] \tag{A9}$$

$$\frac{\partial \omega_{2i}}{\partial \beta_{2j}} = (1 - f_{1i}(0)) \frac{\lambda_{2i}}{(1 - f_{2i}(0))} \left[1 - \frac{\lambda_{2i}\, f_{2i}(0)}{(1 - f_{2i}(0))\,(1 + a_2 \lambda_{2i})} \right] x_{ij} \tag{A10}$$

$$\frac{\partial \omega_{2i}}{\partial a_2} = - \frac{(1 - f_{1i}(0))\, f_{2i}(0)}{\lambda_{2i}\, a_2} \left[\frac{\lambda_{2i}}{[1 + a_2 \lambda_{2i}]} + \ln(f_{2i}(0)) \right] \tag{A11}$$

Note that the above derivatives are equal to the product of two terms: one which only varies with i and the other is the x_{ij} (equal to 1 in the case of the a parameters). This is important in the construction of the Q_1 matrix in equation (15) of Murphy and Topel. For our models, Q_1 matrix is the cross product matrix of variables entering the training equation with that of the wages growth equation where each individual's observation is weighted by the first terms appearing in equations (A4) to (A11).

(ii) Hurdle Negbin II model where T_i is treated as endogeneous:

$$\frac{\partial \omega_{1i}}{\partial \beta_{1j}} = \text{(A4)}$$

$$\frac{\partial \omega_{1i}}{\partial a_1} = \text{(A5)}$$

$$\frac{\partial \omega_{2i}}{\partial \beta_{1j}} = N_i \text{ (A4)}$$

$$\frac{\partial \omega_{1i}}{\partial a_1} = N_i \text{ (A5)}$$

(iii) <u>Probit model for the incidence of training:</u>

$E(T_i) = prob(T_i{=}0) = \omega_{1i} = \Phi(x_{1i}'\beta_1) = \Phi_{1i}.$

$E(T_iN_i) = E(T_i)\ N_i = \omega_{2i}.$

$$\frac{\partial \omega_{1i}}{\partial \beta_{1j}} = \phi_{1i}\, x_{1ij} \qquad \text{(A12)}$$

where ϕ is the pdf of the Normal.

$$\frac{\partial \omega_{2i}}{\partial \beta_{1j}} = N_i \text{ (A12)}$$

REFERENCES

Altonji JG and JR Spletzer (1991) "Worker Characteristics, Job Characteristics, and Receipt of On-the-job Training" Industrial and Labor Relations Review 45(1), October, 58-79.

Arulampalam W, 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, forthcoming Research in Labor Economics, 1997.

Arulampalam W and AL Booth (1996) "Who Gets over the Training Hurdle? A Study of the Training Experiences of Young Men and Women in Britain", CEPR Discussion Paper Number 1470, forthcoming Journal of Population Economics, 1997.

Ashenfelter OC and RJ LaLonde (1996) The Economics of Training, Edward Elgar Publishing Company.

Barron JM, DA Black and MA Loewenstein (1989) "Job Matching and On-the-job Training", Journal of Labor Economics 7(1), January, 1-19.

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

Becker GS (1962) "Investment in Human Capital: A Theoretical Analysis", Journal of Political Economy 70, Supplement, 9-49.

Blanchflower D and LM Lynch (1994) "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 (1996) The Determinants and Effects of Work-related Training in Britain, Institute for Fiscal Studies, 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" The Review of Economics and Statistics 76, 164-170.

Booth AL and J Frank (1996) "Seniority, Earnings and Unions", forthcoming Economica.

Bradburn NM, J Huttenlocher and L Hedges (1994) "Telescoping and Temporal Memory", in N Schwarz and S Sudman (eds) Autobiographical Memory and the Validity of Retrospective Reports, New York: Springer-Verlag.

Brown C (1990) "Empirical Evidence on Private Training" Research in Labor Economics, 11, JAI Press, 97-113.

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.

Dagsvik JK (1994) "Discrete and Continuous Choice, Max-stable Processes, and Independence from Irrelevant Attributes" Econometrica 62(5) 1179-1205.

Duncan GJ and S Hoffman (1979) "On-the job Training and Earnings Differences by Race and Sex" The Review of Economics and Statistics LXI 594-603.

Felstead A and F Green (1996) "Training and the Business Cycle", in AL Booth and DJ Snower (eds) Acquiring Skills: Market Failures, their Symptoms and Policy Responses Cambridge University Press.

Freeman RB and LF Katz (eds) (1995) Differences and Changes in Wage Structures, NBER Comparative Labor Markets Series, Chicago: University of Chicago Press.

Green F, S Machin and D Wilkinson (1996) "Trade Unions and Training pracices in British Workplaces", mimeo Centre for Economic Performance, London School of Economics.

Greene WH (1995) LIMDEP Version 7.0 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.

Hashimoto M (1981) "Specific Human Capital as a Shared Investment" American Economic Review 71, 475-482.

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.

Lillard LA and HW Tan (1992) "Private Sector Training: Who Gets It and What Are Its Effects?" Research in Labor Economics 13, 1-62.

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

Lynch LM (1994) (ed) Training and the Private Sector, NBER Comparative Labor Markets Series, Chicago: University of Chicago Press.

Mincer J (1962) "On-the-job Training: Costs, Returns and Some Implications" Journal of Political Economy 70, Supplement, 50-79.

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.

OECD (1995) The Jobs Study, Paris.

Oi W (1962) "Labor as a Quasi-fixed Factor" Journal of Political Economy 70, 538-555.

Parsons DO (1990) "The Firm's Decision to Train" Research in Labor Economics, 11, 53-75.

Shepherd P (1985) "The National Child Development Study: An Introduction to the Origins of the Study and the Methods of Data Collection", NCDS User Support Group Working Paper No. 1, City University, London.

Snower DJ (1996) "The Low Skill, Bad Job Trap" in AL Booth and DJ Snower (eds) Acquiring Skills: Market Failures, their Symptoms and Policy Responses Cambridge University Press.

Stevens, Margaret (1996) "Transferable Training and Poaching Externalities", Chapter 2 in AL Booth and DJ Snower (eds) Acquiring Skills: Market Failures, their Symptoms, and Policy Responses, Cambridge University Press.

Sudman S and NM Bradburn (1973) "Effects of Time and memory Factors on Response in Surveys", Journal of the American Statistical Association 63, 805-15.

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.

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.