J Popul Econ (1997) 10:197–217

Journal of
Population
Economics

Who gets over the training hurdle? A study of the training experiences of young men and women in Britain

Wiji Arulampalam[1], Alison L. Booth[2]

[1] Department of Economics, University of Warwick, Coventry CV4 7AL, UK
(Fax: +44-1203 523 032, e-mail: wiji.arulampalam@warwick.ac.uk)
[2] ESRC Centre on Micro-social Change, University of Essex, Wivenhoe Park CO4 3SQ, UK
(Fax: +44-1206 873 151, e-mail: albooth@essex.ac.uk)

Received February 9, 1996/Accepted August 14, 1996

Abstract. Using longitudinal data from the British National Child Development Study, this paper examines gender differences in the determinants of work-related training. The analysis covers a crucial decade in the working lives of this 1958 birth cohort of young men and women – the years spanning the ages of 23 to 33. Hurdle negative binomial models are used to estimate the number of work-related training events lasting at least three days. This approach takes into account the fact that more than half the men and two thirds of the women in the sample experienced no work-related training lasting three or more days over the period 1981 to 1991. Our analysis suggests that reliance on work-related training to improve the skills of the work force will result in an increase in the skills of the already educated, but will not improve the skills of individuals entering the labor market with relatively low levels of education.

JEL classification: C25, I21, J24.

Key words: Hurdle count data models, training, skills segmentation.

---

1. Introduction

According to the British National Child Development Study, over half of young men and two thirds of young women experienced no training at all over the period 1981–1991, a decade covering the crucial years of their

---

Wiji Arulampalam: née Narendranathan
All correspondence to Alison L. Booth. We should like to thank Klaus F. Zimmermann and two anonymous referees for their helpful comments, and Cristian Herrera-Salas and Chris Jones for assistance with the data. *Responsible editor:* Klaus F. Zimmermann.

working lives from age 23 to 33. Moreover, young men experienced on average twice as many training events as young women. Why do some individuals experience multiple training events while others never experience any training? To what extent do ability and education contribute to repeated occurrences of work-related training? Are there significant gender differences in the determinants of training experiences? This paper provides some answers to these questions by estimating count data models (in which the dependent variable takes only non-negative integer values) of the determinants of the number of work-related training courses lasting at least three days experienced by a cohort of British men and women over the period 1981 to 1991.

Governments in industrialized countries have, over the past decade, increasingly been emphasizing the importance of employer-led training in providing the skilled work force necessary for improving competitiveness, adaptability and economic growth into the next millennium (see for example the UK Government White Paper *Employment for the 1990s* and the US Department of Labor report *Work-based Learning*). 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 investigate the extent to which there are gender differences in the receipt of work-related training. If women are consistently less likely to receive training ceteris paribus, this raise issues of equity in reliance on employer-provided training to improve the skills of the British work force.

A second and related goal is to investigate whether or not there is any evidence of a "low-skill, bad-job" in Britain (Snower 1996; Burdett and Smith 1995). An important finding of our paper is that there are strong complementarities between past general education and training, a finding which provides some evidence for the low skill, bad job trap. 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. While it is not surprising that firms should offer the most able and better educated workers more training, a clear implication of our results is that reliance on employer-provided training alone will not up-grade the skills of all workers in the labor market.

We estimate the determinants of the number of training events using count models, in which the dependent variable takes only non-negative integer values corresponding to the number of work-related training courses occurring in the interval 1981 to 1991. This modeling procedure has not been used for training events before. In view of the bunching of observations at zero counts of training, we extend the count modeling approach in order to estimate negative binomial hurdle models, in which the process generating training incidence is allowed to differ from the process generating positive training counts.

The remainder of this paper is set out as follows. Section 2 outlines the theory, while Sect. 3 describes the data set. The count data models of training courses are outlined in Sect. 4, while in Sect. 5 the estimates for men and women are presented and discussed. The final section concludes.

2. The theoretical framework

The experience of a work-related training event is the result of optimizing decisions made by both an individual worker and an employer. In the case of employer-provided training, the employer decides to offer a course to an employee, who then decides whether or not to accept. Since the data preclude it, we do not model the structural framework for the training decision. Instead we estimate reduced form models of the probability of individuals in the sample experiencing training events that occur $n = 0, 1, 2, \ldots$ times in the given time interval 1981 to 1991. Nonetheless, it is helpful to consider the determinants of training as a result of optimizing decisions made by both parties, since this suggests what variables should be included in the reduced form training count models.

Firms will want to train individuals most able to benefit from the training and perhaps faster to learn. The cost of work-related training will be lower for higher ability workers, and for better-educated workers, ceteris paribus, since bright workers and workers with a sound educational background will learn faster than their less able colleagues. We would therefore expect to observe a positive correlation between ability and work-related training, and between higher levels of educational attainment and training.

However, if workers enter the labor market with poor general education, then it may be the case that reliance on job-related training leads to a skills-segmented labor market and an under-class of uneducated, and perhaps unemployable, workers. Recent theoretical work by Snower (1996) and Burdett and Smith (1995) shows that, where there is a high proportion of uneducated workers, firms may have little incentive to provide good jobs requiring high skills and training, and if there are few good jobs, workers may have little incentive to obtain such skills. As a result, certain workers may get caught in a cycle of low productivity, deficient training and insufficient skilled jobs. While our data and estimation do not represent a direct test of this theory, we are able to provide some stylized facts consistent with the theory. For example, we are able to establish whether or not workers entering the labor market with high levels of general education are more likely to experience work-related training courses as they progress through their working lives.

Firms may also discriminate against particular types of workers when they select workers for training programmes. Although in Britain there is legislation against discriminatory practices in hiring workers, employer discrimination may take the form of not offering places on training courses to women or to nonwhites. Or it may be that women or nonwhites do not volunteer for such training on the expectation of discrimination.

According to the orthodox human capital approach, agents will invest in training courses if the present discounted value of training benefits exceeds training costs.[1] Irrespective of whether training is general or specific, the amount of any training investment should be greater the longer is the post-training period over which the investment can be amortized. For this reason, it might be expected that training is more likely to be offered to, or undertaken by, workers with a strong attachment to the labor market. Gender and the number of children may be used as an indicator of low attachment to the labour market by some employers, in spite of the fact that female labor force participation in Britain is high. To the extent that women do not fit this stereotype, this may represent statistical discrimination.

It is possible that experience of unemployment in the past may have an adverse effect on the amount of training individuals undertake. Uncertainty about future incomes and opportunities will affect both individual workers' decisions to train and firms' decisions to offer training. The demand by workers for vocational training is likely to be influenced by the probability of unemployment and the perceived risk of not completing or of failing a training course. To the extent that unemployment is state-dependent (for example, young men or women who have experienced unemployment may not be confident about retaining a job in the future), past unemployment experience may have a negative effect on future training. Or it may be that workers with low motivation are the first to be laid off in a slump, and the last to be offered or to accept training courses, since the returns are likely to be low.

Firms' attributes are also likely to affect work-related training. For example, in the British context it has been found that members of a trade union are more likely to experience training. There are several hypotheses about the expected impact of trade unions and training. Trade unions in their monopoly role use their power over labor supply to extract a larger share of the surplus, and thereby induce deadweight losses. It is sometimes thought that, in union establishments, employer incentives to provide training are low, because of high wages, restrictive work practices and problems with the introduction of new skill-intensive technologies that threaten union jobs. On the other hand, unions are in some circumstances cooperative, and are sometimes associated with improvements in worker morale and organization at the work place, and thereby increase training and productivity. Ultimately it is an empirical question as to whether unions are associated with an increase or decrease in training.

There are a number of competing hypotheses about the relationship between the incidence of training courses and firm size or sector. For example, larger firms and public sector firms may be more likely to train workers because they are more forward looking or better placed to bear the risk associated with training. Large firms may also be associated with more work-related training courses because of economies of scale in training provision (Greenhalgh and Mavrotas 1994), or perhaps because they face more regulations, more bureaucracy, and so provide more training to meet safety regulations etc. (Felstead and Green 1996).

However, there are compelling econometric reasons for not including current firm characteristics in any model of the determinants of training, on the grounds that these characteristics are likely to be endogenous. Individuals may choose to work in large firms or unionised firms because they are perceived to offer more training. Moreover, information on the attributes of the firm in which an individual received training are not available in the NCDS data. Although we are able to date the three most recent training courses (lasting at least 3 days) experienced by an individual over the decade 1981 to 1991, the characteristics of the firm in which the respondent was employed at the training date were not requested. Thus the only available information on firms attributes is 1981 data (conditional on the individual being in employment). Whether or not the use of 1981 firm characteristics tells us anything is an empirical issue.

The theoretical arguments advanced in this section relate to training incidence at any point in time. These arguments can easily be generalized to repeated occurrences of training incidence. For example high ability or well-educated workers are more likely to be trained at any time. Women with few family commitments are more likely to be continuously in employment, and therefore more likely to experience training opportunities.

3. The data source

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. An advantage to using the NCDS data is that problems of unobservable age-related effects (that may be found in surveys of individuals from a variety of age groups) are not present, since the data come from a specific cohort of individuals. 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 1958–1974 and born in the week 3–9 March were added to the survey sample (Shepherd 1985). Particular use is made of the information collected at age 23 in 1981 (Wave 4 data) and at age 33 in 1991 (Wave 5 data).

Earlier waves of the NCDS (in particular Waves 3 and 4) provide data on time-varying and fixed individual characteristics before the individual received training over the period 1981–1991. The education variable used in the count data models is the highest educational qualification obtained by the survey date of March 1981. Work-related training courses received between leaving school and 1981 are proxied by a number of dummy variables. The rich data available in Wave 4 of the NCDS allow for the estimation of the impact of predetermined and exogenous variables on human capital acquisition between Waves 4 and 5 of the survey.

Wave 5 of the NCDS is a remarkably rich source of information about training and education received over the period 1981 to 1991. These training data were elicited by a question asking respondents "Since March 1981 have you been on any training courses designed to help you develop skills that you might use in a job (*apart* from any courses you have already told me about)". Some 60% of men in our sample and 43% of women had experienced at least one such course over the period 1981–1991. If the respondent had been on any courses lasting at least 3 days in total, the number of such courses was requested. From this, we construct the variable NUWKTR, the dependent variable in the count data models.[2] The count data models could not be estimated for the short courses, since the number of short courses was not requested.

Well over half of young men and over two thirds of young women in 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). The frequency distribution for NUWKTR is given in Table 1, for men and women with complete data in the sample and who were in the labour force for either Wave 4 or Wave 5 or both. We assume that if individuals were in the labour force at either 1981, 1991, or both,

Table 1. Frequency distribution of the number of training courses lasting 3 or more days

Number of training courses; 3 or more days duration 1981–1991 (NUWKTR)	All women Observed frequencies (proportions) Column 1	All men Observed frequencies (proportions) Column 2
0	1511 (0.682)	1040 (0.509)
1	290 (0.131)	290 (0.142)
2	142 (0.064)	187 (0.092)
3	84 (0.038)	127 (0.062)
4	52 (0.024)	79 (0.039)
5	41 (0.019)	69 (0.034)
6	24 (0.011)	67 (0.033)
7	8 (0.004)	24 (0.012)
8	12 (0.005)	26 (0.013)
9	3 (0.001)	13 (0.006)
10	20 (0.009)	51 (0.025)
11	2 (0.001)	2 (0.001)
12	7 (0.003)	15 (0.007)
13	0 (0.000)	2 (0.001)
14	1 (0.001)	3 (0.002)
15	2 (0.001)	10 (0.005)
16+	16 (0.007)	37 (0.018)
Total number	2215	2042

then they are likely to be in the labour force sometime in the intervening period and therefore eligible for work-related training courses.[3]

The raw data in Table 1 are characterized by a unimodal skewed distribution. The sample mean is 2.2 for men and 1.0 for women, while the sample standard deviation is 4.26 for men and 2.84 for women. Thus there is considerable over-dispersion in raw terms, in the sense that the variance is substantially greater than the mean. Some of the characteristics of the raw data for NUWKTR are as follows: 51% of the 2042 young men and 68% of the 2215 young women for whom there is complete information reported no work-related training courses in the period 1981–1991, 14% men and 13% women had one such course, 9% men and 6% women had two courses, 6% men and 4% women had three courses, and the remaining had up to a maximum of 16+.

The count data in Table 1 show signs of clustering after 9 training courses: there are spikes at 10, 12 and 15 occurrences. These spikes may have arisen because individuals experiencing a lot of training over the period 1981–1991 (and who were asked about training occurrences retrospectively) may have recalled them as rounded up or down numbers, that is, as a dozen, or fifteen, or twenty. For this reason, we experimented in our estimation with censoring the raw data at various points (viz. 10 and 15), but found that the various censoring assumptions made little difference to the results.

The focus of the present paper is on estimating the determinants of the number of training courses (lasting at least three days) received by men *and* women over an important decade in their working lives. The mean *duration* of the most recent training course is 3.5 weeks (standard deviation 13.1) for men and 4.3 weeks (standard deviation 16.0) for women. (The NCDS asks for detailed information of up to three of the most recent training courses, but we do not present the means of the earlier training courses since these are subject to measurement error and problems of missing data.)

While we are taking a broad approach in this paper and combining different forms of training lasting at least 3 days, it is nonetheless interesting to note that, for our sample, 43% of all male training courses (and 46% of all female training courses) lasting at least 3 days were carried out on the employer's premises, while the remainder were off-the-job. For men, 17% of training courses lasting at least 3 days ended in a qualification, while for women the comparable figure was 14%. Some 89% of male and 87% of female training courses lasting at least 3 days were employer-provided.

The respondent, a relative or a friend paid (in part or in full) for 5% of male courses and 7% of female courses lasting at least 3 days. In Arulampalam et al. (1995) the disaggregated impact of these various types of training on *male* wages growth is estimated, and the extent to which the impact of training decays across time is also investigated. Blundell et al. (1996) also examine the impact of various forms of training and education on wages growth using the NCDS data. Arulampalam et al. (1996) explore the impact of the aggregated number of training courses on *male* wages growth, controlling for endogeneity.

It is interesting to note that our figures for the incidence of employer-provided training using NCDS data are closer to those characterising the United States than Germany (see Lynch 1992; Tan et al. 1992; Pischke 1994; Veum 1996; Winkelmann 1996). However, our finding that working women receive less work-related training than men is also found in studies using US and German data (see inter alia Lynch 1992; Lillard and Tan 1992; Winkelmann 1996).

In order to make inferences about what sort of person was being trained over the period 1981 to 1991 in Britain, we need to control for covariates. Before doing this, the next section sets out the modeling framework for the analysis.

4. Modeling the number of training occurrences

In our models, the dependent variable takes only non-negative integer values corresponding to the number of training events occurring in the interval 1981 to 1991. (For surveys of these count models, see Cameron and Trivedi 1986; Winkelmann 1994; Gurmu and Trivedi 1994; Winkelmann and Zimmermann 1995.) We estimate reduced form models of the probability of individuals in the sample experiencing training events that occur $n = 0, 1, 2, \ldots$ times in the given time interval 1981 to 1991. Given the nature of our data, the natural starting point is the Poisson model.

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

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

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

$$\lambda_i = e\mathrm{xp}(\boldsymbol{X}_i'\beta) \ . \tag{2}$$

The vector of exogenous variables is denoted by $\boldsymbol{X}_i$, while β is the associated vector of coefficients. The exponential form ensures non-negativity of λ_i. The Poisson distribution 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 $\boldsymbol{X}_i$. (For expositional ease, from now on we shall not specifically state that the distributions being considered are conditional on the observed $\boldsymbol{X}_i$.) But, as shown in Table 1, the raw data indicate over-dispersion – the variance exceeds the mean.

There are at least two possible causes of such over-dispersion. One is unobserved heterogeneity in the mean function λ. Another is when the probability of experiencing an event is increased as a result of past experiences of the event. Panel data are necessary in order to distinguish between these two competing hypotheses, but unfortunately the form of the NCDS data for occurrences of training counts in the interval 1981 to 1991 is a simple cross-section (where the numbr of training occurrences over the period 1981–1991 is measured retrospectively at the 1991 NCDS). Given the cross-section nature of the data, we take a reduced form approach, in the sense that models allowing for over-dispersion are directly specified and estimated, in order to explain the number of training events experienced by our sample members.

A common generalization of the Poisson model that allows for over-dispersion is the negative binomial distribution (Cameron and Trivedi 1986; Winkelmann 1994; Winkelmann and Zimmermann 1995). This is given by

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

with $E(Y_i) = \lambda_i$, $\mathrm{var}(Y_i) = \lambda_i + \lambda_i^2/a_i$ and λ_i, $a_i \in \mathbb{R}^+$, and $\Gamma(n)$ is the standard gamma function.

One model which generates the negative binomial distribution is a model of random mean function for Y_i. Suppose that the mean function of Y_i is $\tilde{\lambda}_i = \lambda_i \mathrm{u}_i$, where u_i is an unobservable heterogeneity term and u_i ~ Gamma (a_i, a_i), or equivalently $\tilde{\lambda}_i$ ~ Gamma$(a_i, a_i/\lambda_i)$.[4] Marginalization with respect to the unobservable u_i yields the unconditional distribution for Y_i given in (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 a_i. Setting $a_i = c\,\lambda_i^k$, 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. (This assumes a homoskedastic u.) Thus the Poisson model is obtained with the restriction $a=1/\alpha=1/c=0$ for all i.

One limitation of the model discussed above is that the zeros, as well as the positive counts, are generated by the same process. As can be seen from Table 1, there are a great many zeros in the sample. Since it is clear that some individuals never experience any training lasting at least 3 days, 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 events and, once the hurdle is crossed, the conditional distribution of the positive values is governed by a truncated-at-zero count data model. (This was first introduced in economics by Mullahy (1986), who considers a Poisson hurdle model. See Winkelmann (1994) for additional references.) This model also allows for over-dispersion.

Formally, let f_1 be the probability density function (pdf) of the process governing the hurdle (that is, the incidence of training), and let f_2 be the pdf of the process governing the number of training events once the hurdle has been crossed. Note that f_2 is the pdf of a distribution for non-negative integers (and not truncated at zero itself). Thus the probability distribution of the hurdle model variable Y_{ih} for the i-th individual is given by

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

and

$$\begin{aligned}\text{Prob}(y_i \text{ training events over the period}) &= \Pr(Y_{i\text{h}} = y_i) \\ &= \text{f}_{2i}(y_i)[1 - \text{f}_{1i}(0)]/[1 - \text{f}_{2i}(0)] \quad y_i = 1, 2, ... \\ &= \text{f}_{2i}(y_i)\,\theta_i\,, \end{aligned} \tag{4b}$$

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

$$\text{E}(Y_{i\text{h}}) = \sum_{y_i=1}^{\infty} y_i\,\text{f}_{2i}(y)\,\theta_i \tag{5}$$

and

$$\text{Var}(Y_{i\text{h}}) = \theta_i \sum_{y_i=1}^{\infty} y_i^2\,\text{f}_{2i}(y_i) - \theta_i^2\left[\sum_{y_i=1}^{\infty} y_i\,\text{f}_{2i}(y_i)\right]^2. \tag{6}$$

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

The likelihood for the sample is given by

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

The first two terms on the right-hand side (RHS) of (7) refer to the likelihood for training incidence, while the third term is the likelihood for positive counts for the number of training events. The log-likelihood is therefore separable, and maximization is simplified by first maximizing a binary model log-likelihood, and then separately maximizing the log-likelihood for a truncated variable. If it is assumed that both distribution functions f_1 and f_2 are identical, but that they may be characterized by different parameter values, then standard tests can be used to test the restriction that the parameter values are the same. Some possible choices for the distribution functions are Poisson, geometric, or negative binomial.[5] We choose the Negbin II model for estimation of the hurdle model, which nests both the Poisson and the previous Negbin II models as special cases.

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

$$\begin{aligned} \Pr(Y_{ih} = 0) &= f_{1i}(0) = \{a_1/[a_1 + e\mathrm{xp}\,(\boldsymbol{X}'_{1i}\beta_1)]\}^{a_1} \\ &= [1 + \mathrm{a}_1 \cdot e\mathrm{xp}\,(\boldsymbol{X}'_{1i}\beta_1)]^{-1/\mathrm{a}_1} \,, \end{aligned} \tag{8}$$

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

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

5. The estimates for men and women

We estimated both hurdle and non-hurdle models of work-related training, both for the Poisson and for the Negbin II assumptions. On the basis of a number of tests which are described below (and reported in Tables 2 and 3), the preferred models for both men and women are the Negbin II hurdle specifications. These Negbin II estimates are presented (separately for men and women) in Table 4.[6] The dependent variable is NUWKTR – 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.

5.1 Discriminating between models

Before discussing the estimates of the Negbin II hurdle models, we first consider the testing procedure used to discriminate between the various

Table 2. Model log-likelihoods

Model	Women log-likelihood	$a = 1/\alpha$	Men log-likelihood	$a = 1/\alpha$
Non-hurdle models				
1. Poisson	–4083.101	fixed	–5607.720	fixed
2. Negbin II	–2622.570	3.351 (0.19)	–3589.489	2.199 (0.11)
Hurdle models				
3. Poisson incidence	–2126.379	fixed	–2116.724	fixed
4. Poisson positive counts	–1872.434	fixed	–3135.009	fixed
5. Negbin II incidence	–1246.565	5.201 (3.22)	–1282.138	0.100 (0.70)
6. Negbin II positive counts	–1335.693	5.701 (2.78)	–2264.074	1.616 (0.23)

Note: Standard errors are given in parentheses.

Table 3. Specification tests for the models of Table 2

Null hypothesis	LR statistic women	LR statistic men
1. against 2.; χ^2 (1)	2921.106	4036.462
1. against 3. and 4.; χ^2 (29)	168.576	711.974
2. against 5. and 6.; χ^2 (30)	80.624	86.554
3. against 5.; χ^2 (1)	1759.628	1669.172
4. against 6.; χ^2 (1)	1073.482	1741.870

models. The model log-likelihoods are presented in Table 2, and the likelihood ratio (LR) tests in Table 3.

In Tables 2 and 3, 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 Sect. 3, 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. Thus a one-sided 5% significance level test requires the use of the 10% critical value (see Lawless 1987). On the basis of these two tests, we reject the Poisson model; that is, Row 1 is rejected against Row 2.

Because the non-hurdle model is nested within the hurdle model (as noted in Sect. 3), we can test the non-hurdle model 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 3 as a test of Row 1 against Rows 3 and 4 for the Poisson model, and for the Negin 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 3. The appropriate test in this instance is the LR test.[7] 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. In summary, 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.

5.2 Predicted frequency comparisons

We now consider the predicted frequencies of the models reported in Tables 2 and 3. The predicted frequencies are calculated by first calculating the predicted probabilities for the various outcomes for each individual, and then summing over all individuals to obtain the predicted frequencies (see Winkelmann (1994:183–194) for further details). For both the male and female samples, the Poisson model under-predicts the zero outcome and over-predicts the positive counts. For women, the Negbin II non-hurdle and hurdle models perform equally well in terms of predicted frequencies. For men, in contrast, the hurdle Negbin II model performs slightly better in terms of predicted frequencies than does the simple non-hurdle Negbin II

Table 4. Actual and predicted frequencies

Women					Men				
Count	Actual	Poisson	Negbin II	Hurdle Negbin II	Count	Actual	Poisson	Negbin II	Hurdle Negbin II
0	1511	987	1512	1511	0	1040	456	1021	1040
1	290	640	298	298	1	290	512	341	296
2	142	315	136	135	2	187	393	188	183
3	84	148	77	77	3	127	263	119	124
4	52	69	48	50	4	79	164	82	88
5	41	32	33	34	5	69	99	59	65
6	24	14	23	24	6	67	59	44	49
7	8	6	17	18	7	24	36	33	37
8	12	3	13	13	8	26	22	26	29
9	3	1	10	10	9	13	13	21	23
10+	48		48	45	10+	120	25	108	108
Total	2215	2215	2215	2215	Total	2042	2042	2042	2042
χ^2 (r) Goodness of fit test			13.26 (8)	13.29 (6)				28.60 (8)	18.62 (6)

Notes: (i) The χ^2 Goodness of fit test is calculated as $\left\{\sum_{j=0}^{10} \frac{(p_j - o_j)^2}{p_j}\right\}$ where the p_j and the o_j are the predicted and the observed frequency for class j. The degrees of freedom = number of classes under consideration –1 – number of parameters estimated. Since the mean and the variance are estimated in the Negbin II model, the number of parameters estimated is 2 for the simple Negbin II model and 4 for the hurdle Negbin II model.

(ii) The 5% critical values are: $\chi^2(6) = 12.59$ and $\chi^2(8) = 15.51$; the 1% critical values are: $\chi^2(6) = 16.81$ and $\chi^2(8) = 20.09$.

model. The corresponding χ^2 goodness-of-fit statistics are given in Table 4. For women, the null hypotheses that the non-hurdle and hurdle Negbin II models are adequate are not rejected at the 1% significance level. However for men the χ^2 tests produce very large values simply for the reason that the observed frequencies exhibit some spikiness at certain frequencies (as was noted in Sect. 3 of this paper). This problem is more pronounced for men reporting in excess of 9 training events during the 10-year period. Since the number of cases who might be affected is proportionately very small (see Table 1), we believe that our basic estimated results are unaffected.

5.3 The Negbin II hurdle estimates for women and men

We now consider the coefficient estimates of the preferred specification, the hurdle Negbin II model, which are presented in Table 5. The results for women are given in Columns 1 and 2, and for men in Columns 3 and 4. The variables used to explain work-related training occurrences fall under five headings, chosen in the light of the theoretical background outlined in Sect. 2. The five groups of variables are: individual attributes; employment status in 1981 (when the individual was 23); highest educational qualification by 1981; number of training events prior to 1981; and lastly, employer characteristics conditional on the individual being in employment in 1981.

First, consider the significant variables under the heading *individual attributes.* The estimated coefficient to the variable "reading score below average" is significantly negative for training incidence for both men and women. Individuals scoring below average in reading tests at age 11 have a lower probability of training incidence; however the number of training events conditional on incidence is unaffected by reading score. While women scoring below average in mathematics tests at age 11 have a significantly lower probability of training incidence (albeit at the 10% level), this variable has no effect for men.

Ethnic origin has no impact for women. In contrast, men of white ethnic origin have a significantly higher probability of experiencing training and experience more training events (see Columns 3 and 4). The size of this last effect is large. This finding suggests there may be some employer race-discrimination in providing access to training courses for men, or that non-white males may not volunteer for training on the expectation of discrimination. Where employers are relied on to provide training, the issue of whether or not there is discrimination in access to work-related training becomes very important.[8]

The effects of "married by 1981" (defined to include living as married), "kids by 1981", and "married and kids by 1981" (the interaction of marriage or cohabitation with the number of children) differ for men and women. (We also experimented with the inclusion of the number of children without marriage, but since there were very few cases in this group and it proved to be insignificant we did not include it in the final specification.) For women, training incidence (Column 1) is significantly increased if the woman was married (or cohabiting) with children by 1981. However, if a woman was a lone parent in 1981, training incidence over the period 1981–1991 was significantly reduced, perhaps because employers view

Table 5. Determinants of the number of training courses for men and women

Variable	Women Negbin II hurdle-training incidence (1)	Women Negbin II hurdle-positive counts (2)	Men Negbin II hurdle-training incidence (3)	Men Negbin II hurdle-positive counts (4)	Means	
					Women	Men
Constant	0.688 (1.811)	–0.331 (1.126)	–1.955 (0.573)[b]	–0.889 (0.984)	1.00	1.00
Individual attributes						
Reading score below average	–0.696 (0.328)[b]	0.064 (0.212)	–0.266 (0.105)[b]	–0.150 (0.144)	0.424	0.444
Maths score below average	–0.642 (0.343)[a]	–0.040 (0.211)	–0.080 (0.095)	–0.064 (0.123)	0.523	0.512
Educated at grammar school	0.231 (0.389)	0.377 (0.251)	–0.154 (0.131)	0.027 (0.164)	0.107	0.094
Educated at private school	–0.909 (0.760)	0.193 (0.606)	0.249 (0.230)	–0.256 (0.304)	0.023	0.024
Educated at direct grant school	0.361 (0.849)	–0.137 (0.650)	0.150 (0.239)	0.315 (0.335)	0.017	0.021
White	–0.617 (1.328)	0.086 (0.094)	0.546 (0.332)[a]	1.170 (0.534)[b]	0.992	0.985
Registered disabled 1981	0.175 (0.513)	0.727 (0.368)[b]	0.058 (0.189)	0.109 (0.247)	0.032	0.038
Married by 1981	–0.391 (0.270)	–0.312 (0.188)[a]	0.243 (0.105)[b]	0.413 (0.103)[b]	0.629	0.407
Married and kids by 1981	1.443 (0.831)[a]	0.702 (1.094)	0.164 (0.630)	0.052 (0.964)	0.246	0.121
Kids by 1981	–1.671 (0.870)[b]	–1.151 (1.089)	–0.377 (0.622)	–0.582 (0.950)	0.273	0.126
Experience (months)	0.005 (0.005)	–0.001 (0.004)	0.002 (0.002)	–0.002 (0.002)	60.19	66.74
Union member	0.637 (0.364)[a]	0.265 (0.208)	0.065 (0.083)	0.052 (0.100)	0.333	0.457
Local unemployment rate – TTWA	–0.027 (0.029)	0.007 (0.023)	0.009 (0.009)	–0.030 (0.014)[b]	10.986	11.039
Job is professional/managerial	0.423 (0.353)	–0.035 (0.197)	0.067 (0.090)	0.028 (0.115)	0.198	0.233
Employment status in 1981						
In full-time education	0.791 (1.153)	–0.877 (0.776)	0.538 (0.492)	0.038 (0.827)	0.013	0.019
Unemployed	0.364 (0.423)	–1.718 (0.506)[b]	–0.049 (0.436)	0.230 (0.779)	0.068	0.096
Employed	–0.132 (0.434)	0.036 (0.367)	0.265 (0.434)	0.350 (0.762)	0.732	0.878
Highest educational qualification 1981						
Degree	2.124 (1.156)[a]	0.450 (0.450)	0.886 (0.298)[b]	0.750 (0.273)[b]	0.057	0.076
A-levels	1.069 (0.601)[a]	0.283 (0.341)	0.935 (0.284)[b]	0.520 (0.217)[b]	0.117	0.117
O-levels	0.606 (0.294)[b]	–0.021 (0.228)	0.536 (0.149)[b]	0.341 (0.150)[b]	0.505	0.397
Vocational qualification	0.628 (0.414)	0.034 (0.402)	0.192 (0.142)	0.367 (0.173)[b]	0.065	0.158
Apprenticeship completed	0.289 (0.054)	0.484 (0.572)	0.034 (0.099)	0.350 (0.121)[b]	0.028	0.250

Table 5 (continued)

Variable	Women Negbin II hurdle-training incidence (1)	Women Negbin II hurdle-positive counts (2)	Men Negbin II hurdle-training incidence (3)	Men Negbin II hurdle-positive counts (4)	Means	
					Women	Men
Training prior to 1981						
1 training course-excl apprenticeship	0.605 (0.354)[a]	0.481 (0.189)[b]	0.175 (0.100)[a]	0.396 (0.105)[b]	0.215	0.210
2 training courses	1.763 (0.964)[a]	0.664 (0.263)[b]	0.166 (0.128)	0.209 (0.171)	0.073	0.094
3 +training courses	1.557 (1.033)	0.907 (0.468)[b]	0.664 (0.272)[b]	0.813 (0.168)[b]	0.023	0.073
Employer characteristics – 1981						
Private sector	–0.568 (0.369)	–0.190 (0.217)	–0.187 (0.098)[a]	–0.138 (0.109)	0.441	0.584
26–99 employees	1.073 (0.522)[b]	0.028 (0.255)	0.114 (0.112)	0.332 (0.136)[b]	0.169	0.204
100 +employees	0.318 (0.289)	–0.402 (0.203)[b]	0.297 (0.127)[b]	0.074 (0.116)	0.312	0.434
Variance parameter= $1/\alpha$	5.201 (3.217)[a]	5.701 (2.778)[b]	0.100 (0.695)	1.616 (0.226)[b]		
Model log-likelihood	–1246.57	–1335.69	–1282.14	–2264.07		
Number of cases	2215	2215	2042	2042	2215	2042

Notes: Standard errors are given in parentheses. [a] Coefficient significant at 10%; [b] Coefficient significant at 5%.

such women as having a lower commitment to the labour market. From Column 2, we see that the expected number is significantly reduced if the woman was living as married but childless in 1981. From Columns 3 and 4, it can be seen that childless men who were living as married by 1981 have a significantly greater training incidence, and they also experience more training events.[9]

The "travel-to-work-area (TTWA) unemployment rate" was found to have a significant negative effect only for the number of training events for men. The impact of "union status in 1981" has no effect on the training experiences of young men over the period 1981–1991, while for women union status in 1981 has a significant positive impact on training incidence only at the 10% level.[10]

Now consider the second set of variables, under the heading *employment status in 1981*. The only significant effect here is female "unemployment in 1981", which is associated with a large reduction in the number of training events. This effect may arise because women unemployed in 1981 have a lower attachment to the labor market, or perhaps lower motivation. Alternatively, employers may view previous unemployment for women as a signal of a lower attachment to the labor market, and therefore offer less training on the expectation that the investment will not be amortized. The base is women who were out of the labor force in 1981.

Next consider the impact of *highest educational qualification in 1981*. It is striking that while there are strong complementarities between education and training for both men and women, these effects are particularly pronounced for men as we shall see. Consider first the results for men, given in Columns 3 and 4 of Table 5. The base group is individuals with no educational qualifications by 1981. The pre-1981 education variables having the largest impact on male training incidence and the expected number of training events are "Degree" (the highest qualification in 1981 was a university degree) and "A-level" (one or more advanced-level secondary school qualifications representing university entrance-level qualifications usually taken at or around the age of 18).

The variable "O-level" (one of more ordinary-level secondary school qualifications obtained at or around the age of 16) also has a significant positive effect on male training incidence and the expected number of training events. "Vocational qualification" (one or more business, technical or industrial vocational qualifications) and "Apprenticeship completed" (completion of a trade apprenticeship, typically after a 3–5 year indenture period begun at the minimum school leaving age of 16) have a significant positive effect only on the expected number of training events, and not on male training incidence.

We now consider the impact of *highest educational qualification in 1981* on female training experiences, given in Columns 1 and 2 of Table 5. For women, only training incidence is affected by education. While the sizes of the coefficients for "Degree and A-level" are large, these variables are significant only at the 10% level. "O-level" significantly increases the female training probability at the 5% level.

This evidence of strong complementarities between past general education and training suggests that reliance on employer-provided training to increase the level of skills of the British work force will result in an increase in the skills of the already-educated, and particularly men, but will not improve the skills of individuals entering the labor market with a low level of education.[11] While it is a rational response of firms to train individuals most able to benefit from the training and perhaps faster to learn, the upshot may be that reliance on employer-provided training leads to a segmented labor market and an under class of uneducated (and possibly unemployable) workers.[12] It is interesting that empirical studies of private sector training for young workers in countries other than Britain also find the same complementarity between education and work-related training (see for example Lynch 1992; Lillard and Tan 1992; Tan et al. 1992; Pischke 1994; Veum 1996; Winkelmann 1996).

The other set of variables measuring human capital acquisition *prior to* Wave 4 of the NCDS (carried out in 1981 when respondents were aged 23) falls under the heading *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-section nature of our data (with retrospective information for training between 1981 and 1991 obtained at Wave 5 in 1991), we are unable to address this issue properly. Nonetheless, we wanted to try to control for this in the estimation, and hence include pre-1981 training variables.

However, interpretation of the impact of the pre-1981 training variables must be made with caution, since they could simply be proxying unobservable characteristics rather than measuring the true impact of state dependence in training experiences. 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. The variables under the heading *training prior to 1981* generally have a significantly positive effect on the incidence and the number of training courses over the period 1981–1991 for both men and women, ceteris paribus.

We now consider the impact of *1981 job characteristics* on training occurrences. These variables are conditional on being in employment in 1981 (that is, they represent the interaction of employment in 1981 with job characteristics in 1981). For women, it makes no difference to subsequent training experiences whether they were employed in the public or the private sector in 1981. For men, working in the private sector in 1981 is associated with a significantly lower incidence of training subsequently, albeit only at the 10% level. The 1981 workplace size effects are of some interest. The base is employment in 1981 in workplaces with 25 or fewer employees. For women, middle-sized workplaces are associated with a greater training incidence, but larger workplaces are associated with more training events. In contrast, the male probability is significantly higher in larger workplaces, but the expected number of training events is significantly higher in middle-size workplaces.

6. Conclusions

The paper estimates models of training based on count data (in which the dependent variable takes only non-negative integer values corresponding to the number of work-related training courses occurring in the interval 1981 to 1991). The data set is the National Child Development Study. We use hurdle negative binomial models to estimate the number of work-related training events. This approach, which has not been used for training before, allows us to account for the fact that more than half of sample men and two thirds of sample women experienced no work-related training over the period 1981 to 1991.

The principal findings of the paper are as follows. First, women undertake significantly fewer training courses. Secondly, male workers of white ethnic origin undertake significantly more training courses. Thirdly, young men and women who scored below average in reading tests at age 11 have a lower incidence of training. Fourthly, young men marrying or cohabiting early (but with no children) experience significantly more training occurrences, while young women marrying or cohabiting earlier experience significantly fewer training events. Fifthly, young women who were unemployed in 1981 were significantly less likely to undertake training courses over the period 1981–1991. Finally, past human capital acquisition has a large significant positive effect on the number of training courses over the period 1981–1991 for both men and women. This effect is found both for the formal human capital dummy variables measuring highest educational qualifications prior to 1981, and for the employer-related training prior to 1981.

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. This suggests that reliance on work-related training to improve the skills of the work force will result in an increase in the skills of the already educated, but will not improve the skills of individuals entering the labor market with relatively low levels of education. Moreover, women and non-white men will be adversely affected by such a policy, since ceteris paribus they receive significantly less work-related training.

Endnotes

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

[2] In preliminary estimation, we also experimented with estimating the number of courses leading to qualifications over the period 1981–1991, that is, general education. The explanatory power of these models was very low; it would appear that unobservables are determining individuals' decisions to undertake education over the period. For a study using the NCDS to estimate the determinants of training and education incidence and their impact on earnings see Blundell et al. (1996).

[3] If individuals were out of the workforce at both survey dates, then the zero coding for training courses might reflect either their non-participation or their lack of training conditional on participation. Since we are focusing on the latter, we exclude cases who were not in the workforce at both 1981 and 1991.

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

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

with $E(Z) = a/b$ and $\mathrm{var}(Z) = a/b^2$.

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

[6] The full set of estimates is available from the authors on request.

[7] Two widely used tests are the Wald test and the LR test. Row 5 of Table 2 shows that the parameter $a (= 1/\alpha)$ in the hurdle part of the process is estimated to be 5.20 with an associated standard error of 3.22. This implies that, using the Wald test, we cannot reject the null hypothesis that the assumption of a Poisson process for the hurdle part is appropriate. But in contrast, the LR test statistic gives a value rejecting the same null hypothesis. A reason for the conflicting result may be as follows. For programming convenience the software package estimates α and not a. The package then returns a value for a that is estimated as the reciprocal of the estimated α (since the parameter of interest is a and not α). The program also calculates the approximate standard error for this re-parameterized value of α. However it is a well-known result that LR tests are invariant to reparameterization, whereas the Wald test is not. Gregory and Veall (1985) show that, depending on how the reparameterization is carried out, a range of different values for the Wald test may be obtained. We therefore use only the LR test for model comparison here.

[8] Booth (1993) shows that, even in the graduate labor market in Britain, women were found to receive significantly less training; however this effect was not found for black graduates.

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

[10] Most empirical evidence for Britain to date suggests that union workers receive more training than nonunion workers. See for example Booth (1991), Tan et al. (1992), Greenhalgh and Mavrotas (1994), Blanchflower and Lynch (1995), and Green et al. (1995). While Arulampalam and Booth (1996) show there is some persistence in union status for NCDS men, our estimates here of the impact of union status in 1981 on training should be interpreted with caution. This is because the individual may have changed job and/or union status between 1981 and the time training was received.

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

[12] Prais (1995), inter alia, argues convincingly that the British vocational educational system requires reform, and that middle to low attainers are neglected by the schooling system.

References

Arulampalam W, Booth AL (1996) Persistence in Union Status; A Longitudinal Study of Young Men in Britain. Mimeo, University of Warwick (January)

Arulampalam W, Booth AL, Elias P (1995) Work-related Training and Earnings Growth for Young Men in Britain. Department of Economics Research Paper no. 440, University of Warwick. *Research in Labor Economics* (1997 forthcoming)

Arulampalam W, Booth AL, Elias P (1996) Modeling Work-related Training and Training Effects Using Count Data Techniques. Department of Economics Research Paper no. 448, University of Warwick (January)

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

Blundell R, Dearden L, Meghir C (1996) *The Determinants and Effects of Work-related Training in Britain.* IFS Report. 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. *Review of Economics and Statistics* 76:164–170

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

Cameron AC, Trivedi PK (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, Trivedi PK (1990) Regression-based Tests for Over dispersion in the Poisson Model. *Journal of Econometrics* 46:347–364

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

Green F, Machin S, Wilkinson D (1995) Trade Unions and Training Practices in British Work places. Mimeo, University of Leeds

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

Greene WH (1993) *Econometric Analysis*, 2nd edn. Macmillan, London

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

Gregory AW, Veall MR (1985) Formulating Wald Tests of Nonlinear Restrictions. *Econometrica* 53(6):1465–1468

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

Heckman J, Borjas G (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, Robb R (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, Tan HW (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(1):299–312

Mullahy J (1986) Specification and Testing of Some Modified Count Data Models. *Journal of Econometrics* 33:341–365

Pischke J-S (1994) Continuous Training in Germany. Mimeo, Department of Economics, MIT

Prais SJ (1995) *Productivity, Education and Training: An International Perspective.* Occasional Paper Series, NIESR, Cambridge University Press

Shepherd P (1985) *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: Booth AL, Snower DJ (eds) *Acquiring Skills: Market Failures, their Symptoms and Policy Responses.* Cambridge University Press, Cambridge

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

UK Department of Employment (1988) *Employment for the 1990s*. Her Majesty's Stationary Office (HMSO) Books, London

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

Veum JR (1995) Training and Job Mobility among Young Workers. Mimeo, US Bureau of Labor Statistics

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

Winkelmann R (1997) How Young Workers get their Training: A Survey of Germany versus the United States. *Journal of Population Economics* 10:159–170

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