

Persistence in the determination of work-related training participation: evidence from the BHPS, 1991-1997.

Panos Sousounis
University of the West of England

Robin Bladen-Hovell
Keele University

September 2009

Abstract

In this paper we investigate the role of workers' training history in determining current training incidence. The analysis is conducted on an unbalanced sample comprising information on approximately 5000 employees from the first seven waves of the BHPS. Training participation is modelled as a dynamic random effects probit model where the effects of unobserved heterogeneity and initial conditions are accounted for in a fashion consistent with methods proposed by Chamberlain (1984) and Wooldridge (2002) respectively. The results suggest that prior training experience is a significant determinant of a worker's participation in a current training episode comparable with other formal educational qualifications.

Keywords: Training; state dependence; dynamic probit
JEL codes: J24; C23.

I Introduction

In recent years a broad consensus has emerged concerning the role of human capital for the determination of productivity and other economic outcomes for the individual. There is a substantial literature demonstrating the relationship between education and wage or income determination (see, for example, Blundell *et al* (1996) and almost as strong a basis for viewing productivity at the firm-level and human capital as equally well-established (see Black and Lynch ,1996, Dearden *et al* 2005).

Although much of the discussion that has taken place with respect to human capital has focussed upon the role of education and educational attainment, it is generally recognized that training – on-the-job or otherwise – has a substantive part to play in the process. Indeed it may be argued that the role of training becomes increasingly important as the pace of technological or organisational change in the workplace increases. Knowledge and skills acquired in formal education or from previous training episodes rapidly depreciate and become outmoded in this type of environment, requiring workers to engage continually in an ongoing process of skills acquisition.

In the existing literature these issues are invariably addressed in the context of a static framework, albeit a static framework that may extend across several years in calendar time and which may involve multiple training episodes. In the current paper we adopt an alternative approach and investigate whether previous experience of training is itself an important determinant of the incidence of training in the subsequent period. In other words, the focus of the current paper is on whether state dependence is an important channel for the determination of training alongside the usual explanatory variables that reflect individual and workplace characteristics. In order to address this issue we examine the determinants of work-related training-incidence within a panel dataset comprising the first seven waves of the British Household Panel Survey, 1991 – 1997.

Naturally, once we allow an individual's current training-incidence to depend upon previous training experience we introduce a range of additional problems into the analysis. These issues most notably relate to the potentially non-trivial problem associated with the treatment of initial conditions in the data and of how to deal with unobserved heterogeneity. In the current paper we address the problem of initial conditions in dynamic, nonlinear unobserved effects models using the approach suggested by Wooldridge (2005) and follow Chamberlain (1984) in dealing with the unobserved heterogeneity issue..

The structure of the paper is as follows. In section II we describe the background to the study and provide a brief review of the background literature. The review is partial and not intended to be exhaustive but rather indicates the conventional wisdom insofar as the determinants of training are concerned. This is followed in section III by an outline of the econometric model and a discussion of the unobserved heterogeneity and initial conditions problem. The data used in the study are described in section IV followed in section V by a discussion of the main results. The final section summarises and concludes.

II Background

Although an extensive literature has developed in relation to training, most studies have been concerned with the evaluation of government sponsored training schemes and the role of formal educational attainment for employment outcomes (for a non-technical review of this literature see, for example, de la Fuente and Ciccone, 2003). Work-related training, by comparison, has attracted considerably less attention with the evidence adduced by the early literature for the determinants of training being well-documented by Blundell et al (1996) and the OECD (2003). The conventional wisdom for the UK suggests that the probability of participating in work-related training is higher among individuals with a record of prior educational attainment, men rather than women, non-minority groups and younger workers; see, for example, Blundell *et al* (1996) and references therein. Work-related training is also influenced by job/employer characteristics such as employment status (full- and part-time/contract work) firm size and industrial classification (workers in industries that are growing or experiencing rapid

technological change receive more work-related training, on average), the sector of employment (public/private) and the presence (coverage) of trade unions in the workplace (Boheim and Booth, 2004, for example, report a positive impact for unions on training in Great Britain). In terms of outcome, training has been associated with a positive wage effect with consistently higher returns for those who train (Blundell *et al* 1999), and improvements in firm-level productivity and competitiveness; see *inter alia* Black and Lynch (1996), Blundell *et al* (1999) and Dearden *et al* (2005).

Although presented as the conventional wisdom it must be acknowledged that doubt remains around many of these effects. Green and Zanchi (1997), for example, report that the probability of participating in training was equalised between men and women during the 1990s in Britain. Additionally, recent evidence seems to suggest that there has been a change in work-related participation rates in a number of OECD member countries, with women exhibiting higher propensities to train (see Jones et al, 2008 and the references therein). Jones et al (2008) also identifies a number of reasons suggested in the literature for this empirical observation, which vary from wider societal and institutional changes improving the relative role of women in the labour market, an increased training demand and supply to accommodate returning mothers and/or carers, technological changes promoting more desktop-based tasks, which appear to attract more female workers, to women possessing ‘better’ educational profiles and (young) age advantages hence both attracting and investing in more training. In a similar manner, when examining the determinants of training among Australian workers, Almeida-Santos and Mumford (2004) do not find strong evidence of a positive link between unions and training in their sample. Research in the area is ongoing.

However, despite the extensive nature of the literature there appears to have been little discussion of the possibility of state dependence in the determination of training¹. In many respects this is surprising. The link between training and prior formal educational attainment identified elsewhere in the literature suggests that training builds upon previously acquired skills

¹ Indeed the OECD (2003) review of the subject contains no mention of persistence effects whatsoever.

and knowledge; if formal education and training are both part of the skills acquisition process the natural presumption would be that previous training experience is also a determinant of subsequent training. Equally, work-related training is an investment – for the individual and for the firm, if there are costs of adjustment associated with training we might expect to see training and skill acquisition spread over time. Finally, all of the available evidence suggests that training is a process whereby workers’ depreciating skills are updated and enhanced. But with “training opportunities” unevenly distributed across the workforce prior experience of training provides an additional factor that effectively discriminates between workers further.

Establishing a definitive link between past experience and future participation in work-related training has important policy implications pertaining to the potentially long-lasting effects of interventions especially for the individuals in the left tail of the skills distribution. Engaging individuals in training episodes would have two desirable outcomes. First, there would be an immediate effect (increase) on the individual’s human capital as is expected after any training and/or education spell. Secondly, there could be a fundamental preference shift in individuals in favour of training and learning. Policy interventions in the latter case assume a ‘preventive’ character as well as the rather standard ‘remedial’ one. The corollary of this is, of course, that serial persistence in work-related training suggests that members of the workforce who abstain from any skills’ upgrading activities would find themselves more prone to (short- and long-term) unemployment and with escalating costs of recovery. For policy makers this is equally crucial. As almost all Western economies come to rely mainly on the knowledge and skills-intensive tertiary sector of production, and in light of the increasing competition from developing countries, an increasing number of low-skilled individuals could put a considerable strain on the fiscal position of the state.

Cuhna and Heckman (2007) try to provide theoretical justification for the empirical finding of human capital investments motivated by past educational investments in the child development literature. Central to their analysis is the notion that the skill formation process is dynamic in nature and as such the product of a multi-stage technology where skills developed at one stage

are affected by skills acquired earlier in the life cycle. This somewhat contradicts the conventional beliefs of earlier scholars of human capital who viewed the outcome of the skill-investment decision as time invariant (see inter alia Becker and Tomes, 1986). This belief was a consequence of assuming that inputs into the skills production function are perfect substitutes (Becker and Tomes, 1986, Benabou, 2002) and that the productivity of skills investment is time invariant. In a multi-stage model of skills formation, however, inputs at different stages (of the life cycle if we accept that the skill formation process extends over the life of the individual) amplify skills acquired at later stages. Moreover, since skills acquired at different stages complement each other in a multi-stage framework: early investments in skills formation should encourage later investments because the subsequent investment build upon the earlier skills and become more productive. The perceived benefit from training in any period will therefore be influenced by whether or not the individual has trained previously. . Cuhna and Heckman (2007) term these notions ‘self-productivity’ and ‘dynamic complementarity’. Together the concepts provide theoretical justification for how skills acquisition encourage further skills formation, which is amongst the main findings of the empirical literature on child development.

In their formulation, Cuhna and Heckman (2007) allow for *in utero* and genetic experiences influencing the child’s initial skills (Keating and Hertzman, 1999, argue that these initial skills could be influenced from environmental as well as hereditary factors). Similarly, in work-related training, existing human capital as well as non-cognitive abilities such as motivation, self-control and persistence, could be the result of parental and/or environmental influences as well as individual choices; initial conditions are accounted for in either case.

The plausible assumption that such a skill formation technology is applicable to work-related training i.e. that there exists a multiple stage process through which employees build-up human capital throughout their (working) life cycle, means that when modelling the incidence of work-related training, a dynamic term of past training behaviour would capture such propensity of skill creation between reference periods. Early training begets later training and thus past choices

exert a positive influence on future decisions as would be empirically documented through a positive and significant coefficient on the lagged past experience term.

III The model

In the current paper, the training incident is modelled as a function of past training experience and a set of socio-economic, personal and job/employer characteristics. A categorical variable, taking the value of 1 if the individual undertakes any work-related training in the past 12 months and 0 otherwise, is observed when the continuous latent variable y^* crosses a threshold of zero. Hence the model takes the form:

$$y_{it}^* = x_{it}'\beta + \xi y_{it-1} + \varepsilon_{it}, y = 1[y^* > 0], i = 1, \dots, N, t = 1, \dots, T_i. \quad (1)$$

The latent variable y^* may be interpreted as the individual's propensity to train, x is a vector of observable characteristics affecting y^* , β is a vector of coefficients associated with those characteristics, and ε is the idiosyncratic error term. By including y_{it-1} , the latent variable y_{it}^* is modelled as a function of the training experience of the individual in the previous period. The coefficient, ξ , associated with the lagged dependent variable captures the effect of past experience on the current participation event. This formulation enables us to test for true state dependence in work-related training. Since individuals are allowed to exit the sample, the total number of observations per individual is $T_i - 1$.

Caution must, of course, be exercised when modelling state dependence because a significantly positive value for ξ may arise as a result of spurious correlation. As argued by Heckman (1981a, 1981b) this problem can arise from an inability to account for unobservable individual characteristics (unobserved heterogeneity) that influence an individual's propensity to participate in work-related training.

To model the unobserved heterogeneity we assume that the subject-specific heterogeneity is time-invariant and thus that the error term in (1) follows a one-way error components structure that can be written as:

$$\varepsilon_{it} = \alpha_i + u_{it}, \quad (2)$$

where α_i is the subject-specific unobservable effect and u_{it} is a random error. If we further assume that $u_{it} \sim IID(0, \sigma_u^2)$ and the u_{it} are independent of the elements of x_{it} for all i and t , then a standard random effects probit model can be used for estimation.

To marginalize the likelihood function and obtain consistent estimates of β we also need to assume that α_i is independent of the u_{it} and x_{it} and that $\alpha_i \sim IID(0, \sigma_\alpha^2)$. The latter is a strong assumption and, if it does not hold, β is subject to an incidental parameters problem and will reflect some of the effect of the unobservable individual heterogeneity.

Chamberlain (1984) proposed a procedure for relaxing the assumption of independence between α_i and the time-varying elements of x_{it} . He suggests specifying a distribution for α_i conditional on $x' = x_1, \dots, x_T$ which can be written as:²

$$\alpha_i = \gamma_0 + \gamma_1' \dot{x}_i + \gamma_i, \quad (3)$$

where it is further assumed that $\gamma_i \sim IN(0, \sigma_\gamma^2)$ and γ_i is independent of x_{it} and u_{it} for all i and t . The vector \dot{x}_i comprises the means (or lags and/or leads) of the time-varying covariates in (1). The coefficients of γ_i associated with the time-invariant elements of x_{it} are set equal to zero and with the constant, γ_0 absorbed into β , (1) becomes:

$$y_{it}^* = x_{it}' \beta + \xi y_{it-1} + \gamma_1' \dot{x}_i + \gamma_i + u_{it}, \quad (4)$$

which can be computed by estimating T cross-sectional probit specifications by maximum likelihood where x_1, \dots, x_T are included for each T . Moreover, if we assume that the distribution of the unobserved effect, conditional on $x' = x_1, \dots, x_T$, is linear in the mean of x_1, \dots, x_T , equation (4) is simply the standard random effects probit model with an added regressor

² In the linear case, the regression function for α_i given $x' = x_1, \dots, x_T$ would most likely be some nonlinear function but a minimum mean-square error linear predictor could be specified and hence the unobserved effect could be decomposed to that linear projection and an orthogonal residual term. In the nonlinear case the rather restrictive assumption that the regression function $E(\alpha|x)$ is in effect linear should be imposed (Chamberlain, 1984).

representing a vector of means for the time-varying individual characteristics (\bar{x}_i). Therefore, with a slight notational change:

$$y_{it}^* = x'_{it}\beta + \xi y_{it-1} + \gamma_1 \bar{x}_i + \gamma_i + u_{it}. \quad (5)$$

where the correlation between two error terms in the above model is given by

$$\text{corr}(\varepsilon_{it}, \varepsilon_{it-1}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_u^2}. \quad (6)$$

A further problem that often arises in dynamic panel data models is the treatment of the initial observation, y_{i0} , when experimental data is not used. When the start of the stochastic process that generates the data does not coincide with the beginning of the sample, the data suffers from a correlation between the initial observation and the unobserved effect. This initial conditions problem arises because the great majority of individuals in the sample have been active in the labour market prior to the initial period of observation and hence their propensity to engage in work-related training may be influenced by either true state-dependence and/or unobserved factors.

Heckman (1981b) proposes a solution to this initial conditions problem which involves specifying a reduced form equation for the initial observation as a first step followed by a linear regression of the latent variable in the second. The Heckman two-step procedure is widely applied in the literature yet computationally burdensome since it requires special software to be written.

Wooldridge (2002) proposes an alternative approach to the initial conditions problem, which does not require the analyst to specify or even approximate the joint distribution of (y_{i0}, \dots, y_{iT}) given \mathbf{x}_i . He suggests obtaining the joint distribution of (y_{i1}, \dots, y_{iT}) conditional on (y_{i0}, \mathbf{x}_i) and applying standard conditional maximum likelihood estimation methods. This approach alleviates the source of difficulty in applying the approach suggested by Heckman (1981b) by not requiring the specification of the distribution of the initial conditions conditional on the observed covariates and the individual heterogeneity. In order to obtain the density of (y_{i1}, \dots, y_{iT}) conditional on (y_{i0}, \mathbf{x}_i) the density of the unobserved effect conditional on the initial observation and the observed covariates i.e. $f(y_{i1}, \dots, y_{iT} | y_{i0}, \mathbf{x}_i)$ need be specified. Chamberlain (1980)

proposed a similar estimation method for the static probit model with unobserved effects, the main difference between the Chamberlain (1980) approach and the Wooldridge (2002) approach being the additional conditioning on y_{i0} .

Wooldridge (2002) begins from a general dynamic unobserved effects model, where D could be the logit or probit function

$$P(y_{it} = 1 | y_{it-1}, \dots, y_{i0}, \mathbf{x}_i, \alpha_i) = D(\mathbf{x}_{it}\boldsymbol{\beta}' + \rho y_{it-1} + \alpha_i). \quad (7)$$

Consistent estimation of this model with fixed T is impossible given the presence of the unobserved effect. However, the problem can be resolved and consistent estimation can be achieved if the unobserved effect is integrated out of the likelihood function.

Let the density of the unobserved effect (α_i) conditional on the initial observations and \mathbf{x} be $g(\alpha | y_0, \mathbf{x}; \boldsymbol{\zeta})$, then³

$$f(y_1, \dots, y_T | y_0, \mathbf{x}; \boldsymbol{\theta}) = \int_{-\infty}^{\infty} f(y_1, \dots, y_T | y_0, \mathbf{x}, \alpha; \boldsymbol{\beta}) g(\alpha | y_0, \mathbf{x}; \boldsymbol{\zeta}) d\alpha. \quad (8)$$

When $D = \Phi$ the convenient choice for $g(\alpha | y_0, \mathbf{x}; \boldsymbol{\zeta})$ is the normal distribution with mean $\beta_1 + \beta_0 y_{i0} + \boldsymbol{\xi} \mathbf{x}_i$ and variance σ_ε^2 , which follows from $\alpha_i = \beta_1 + \beta_0 y_{i0} + \boldsymbol{\xi} \mathbf{x}_i + \varepsilon_i$ with $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ and independent of (y_{i0}, \mathbf{x}_i) . The model, in indicator function form, is

$$y_{it} = 1[\beta_1 + \mathbf{x}_{it}\boldsymbol{\beta}' + \xi y_{it-1} + \beta_0 y_{i0} + \mathbf{x}_i \boldsymbol{\xi} + \varepsilon_i + v_{it} > 0]. \quad (9)$$

So, since y_{it} given $(y_{it-1}, \dots, y_{i0}, \mathbf{x}_i, \varepsilon_i)$ follows a probit model, and $\varepsilon_i | y_{i0}, \mathbf{x}_i \sim N(0, \sigma_\varepsilon^2)$, the density of (y_{i1}, \dots, y_{iT}) given (y_{i0}, \mathbf{x}_i) is

³ For random vectors \mathbf{y}, \mathbf{x} and \mathbf{z} , let $p(\mathbf{y} | \mathbf{x}, \mathbf{z})$ be the conditional density of $\mathbf{y} | \mathbf{x}, \mathbf{z}$ and $p(\mathbf{x} | \mathbf{z})$ the conditional density of $\mathbf{x} | \mathbf{z}$. The density of (\mathbf{y}, \mathbf{x}) given \mathbf{z} will then be $p(\mathbf{y}, \mathbf{x} | \mathbf{z}) = p(\mathbf{y} | \mathbf{x}, \mathbf{z}) p(\mathbf{x} | \mathbf{z})$, Wooldridge (2002) pp 419.

$$f(y_1, \dots, y_T | \mathbf{z}_i; \boldsymbol{\theta}) = \int_{-\infty}^{\infty} \left[\prod_{t=1}^T f(y_t | \mathbf{z}_{it}, \varepsilon; \boldsymbol{\beta}) \right] \left(\frac{1}{\sigma_\varepsilon} \right) \phi \left(\frac{\varepsilon}{\sigma_\varepsilon} \right) d\varepsilon, \quad (10)$$

where $f(y_t | \mathbf{z}_t, \varepsilon; \boldsymbol{\beta}) = \Phi(\mathbf{x}_t \boldsymbol{\beta} + \varepsilon)^{y_t} [1 - \Phi(\mathbf{x}_t \boldsymbol{\beta} + \varepsilon)]^{1-y_t}$ and $\mathbf{z}_{it} = (1, \mathbf{x}_{it}, y_{it-1}, y_{i0}, \mathbf{x}_i)$ and $\boldsymbol{\theta}$ contains $\boldsymbol{\beta}$ and σ_ε^2 . This model can be estimated by standard random effects probit software.

IV The data

Data from the first seven waves of the British Household Panel Survey (BHPS hereafter), a longitudinal survey of randomly selected households in Great Britain, is used in the current study. The interviews for the first wave of the BHPS were conducted between September and December 1991 and annually thereafter⁴. The sample comprises an unbalanced panel and includes men and women of working age who are present and in employment as employees in the first wave (1991) but who may subsequently drop out of the sample as a result of missing information, attrition or having moved out of scope⁵. Thus the sample excludes self-employed individuals, the unemployed, those in full time education and members of the armed forces⁶.

Over our sample period, the BHPS contains two variables that relate to an individual's participation in training during the twelve months prior to the interview date. The first of these variables records the incidence of formal on-the-job training undertaken as part of the individual's present employment⁷ whilst the second question records any other education or training that was undertaken that enhances skills for current or future employment. The training referred to in this latter respect is, at least potentially, work-related, excluding any education or

⁴ For more details see Taylor, M.F et al (2006).

⁵ Individuals were not allowed to enter or re-enter the sample after the first wave.

⁶ The estimation was carried out for a balanced sample as well (individuals for whom complete BHPS histories are available) with the main findings reported here remaining unaltered.

⁷ In Wave One, only the employed were asked this. At Wave Two, this was extended to all currently working. The scope of the question was widened to include education or training courses.

training undertaken as a pastime, hobby or solely for general interest. In the current analysis we combine responses from both questions in our definition of work-related training⁸.

An inspection of raw-data and conditional probabilities of training participation reveals that individuals in the sample who experience a training spell at any point in the reference period are more likely to engage in training at a later future point. Table 1 presents these probabilities for the overall reference period and a time-wise disaggregation by year.

[table 1 here]

These figures suggest that an individual is less likely to experience transition into work-related training in comparison to someone who has already experienced an event the previous year. The probability of non-participants to continue abstaining from skills investments appears greater than the probability of past participants to enter into further training. This highlights the problem of sustained skills degradation with the corresponding lowering of pay and living standards. This observation is consistent across the sample years and necessitates further investigation.

The figures presented in Table 1 are comparable with reported training figures from such surveys as the Eurostat's continuing Vocational Training Survey, which in 1999 reports training participation of 49% and/or the OECD's International Adult Literacy Survey, which reports training participation figures of 45% for the year prior to the interview date. The Labour Force Survey, on the other hand, reports much lower figures of training participation, around 14%, but that refers to education or training undertaken in the 4-week period prior to the interview date and, in addition, lacks the panel characteristic of the aforementioned surveys and the BHPS, of course. Disparities may in this case arise from the different time horizons and structure of the surveys.

⁸ The relevant questions in the 1991 BHPS are D23 and E17.

Of course, the pattern of behaviour reported in Table 1 may occur for a variety of reasons, not least the occurrence of overlapping observations in the sample, where training lasts more than one observation period. However, whilst this may occur for some individuals in practice it is difficult to believe that this provides the explanation for the observation in general. Training spells where they are documented typically extend for no more than a few days or weeks. Information on the duration of training events in the first seven waves of the BHPS is fragmented and does not allow the construction of a comprehensive measure. Different questions on time spent on training are asked on different occasions. For example, in wave 1 respondents are asked the number of days they spent on training, the mean reported value is 40.77. In wave 2, respondents are asked the number of hours they spent on training per week and the mean reported value is 13.5. From wave 3 onwards, respondents are asked to state the unit of measurement for the training duration i.e. hours, days, weeks or months. From those individuals reporting undertaking training in the reference period only 0.5% report that the training spell lasted ‘months’ on average over the remaining waves. These figures increase our confidence that there is no overlap between training spells in the reference periods, however, a remote possibility remains and should be noted. In addition, since the survey questions are conducted retrospectively, some errors of recollection are also to be expected. In our view this problem is likely to be smaller in relation to the recollection of an event occurring than with respect to the detail of the event, such as length and nature of the training episode.

In modelling work-related training we include a set of variables that reflect individual characteristics such as age, indicators of prior educational attainment, race, and occupation, employment and employer characteristics such as job permanency, part-time, full-time status, hierarchical position within the firm, trade union presence and firm size together with an indicator of training history. Finally, the relationship is estimated separately for men and for women in order to explore gender differences in the response to state dependence.

Pooling the sample across males and females may conceal valuable information about the gender decomposition of the persistence effect that might facilitate more efficient policy interventions.

By decomposing the dataset we can elicit information about the homogeneity or otherwise of the coefficients of the disaggregate units per regressor. Proceeding from the theoretical justification of the sample disaggregation, , we test the pooling assumption using Chow-type likelihood-ratio tests for the composite models. The tests are carried out under the null hypothesis that the coefficients of the model do not differ between different sections of the covariate space. Likelihood-ratio tests are more precise than a Wald version of the Chow test since the latter is based on the inverse of the information matrix and is thus based on a quadratic approximation to the likelihood function (Gutierrez et al, 2001). The tests invalidate the pooling assumption for males and females [$\chi^2(52) = 122.61, p - \text{value} = 0.0000$] suggesting that the pooled effect may be misleading. Furthermore, significance criteria for the dummy terms in the pooled sample regressions support the conclusions of the LR tests. Summary statistics for the main variables are given in Table A2 in the appendix.

The sample is almost equally balanced between men and women and the average age is 38 years old irrespective of gender. Men on average tend to be better qualified and dominate the skilled manual and professional groups. As might be expected more women are employed part-time and are also more strongly represented in the public administration, education and health sectors. A cursory examination of the incidence of training across the occupational groups (Table 2) illustrates the conventional finding that training is heavily skewed away from the unskilled.

[table 2 here]

V Econometric Results

Table 3 presents marginal effects for the random effects probit models for men and women (columns 3-8), where state-dependence and initial conditions effects are allowed for, together with estimates of the comparison models (columns 1 and 2) where the initial observation is

treated as exogenous⁹. Table A1 in the appendix presents coefficient estimates¹⁰. Our dataset allows for a detailed disaggregation of educational attainments between respondents, however, we opt for dividing the sample between those with university/college education and those without. Completion of the A-levels requirement is equivalent to holding a US High School diploma and the sample split is therefore effectively between High School graduates and College graduates (those with university degrees). The models were also estimated with a full set of educational attainment variables included but the results remained unaltered to the third decimal place. The results from these estimations are available upon request. .

All models were estimated by including time means of the time varying covariates in line with Chamberlain's (1984) suggestion, to control for possible correlation between the explanatory variables and the random effect. Year dummies were also included to pick up macroeconomic effects, as were regional dummies to account for any regional disparities. The remainder of this section considers the estimates from the random effects probit models for both men and women.¹¹

The state-dependence effect in the random effects models is found to be highly significant for both High School and College graduates. For the latter the estimated effect is higher as expected, by the tendency, well-documented in the empirical literature (see references in the introduction), of more educated people to train more. Across the sample as a whole, treating the initial observation as exogenous appears to overestimate the contribution of state dependence in the participation probability by approximately 0.07. The assumption of exogenous initial conditions is conclusively rejected by the significance of the estimated coefficient on the auxiliary period-zero term included in the random effects model and, as indicated previously, the pooling assumption for men and women is also rejected by the data. Interestingly, at the sub-group level the estimated effect of state dependence is almost the same between male high school and

⁹ Estimation was performed using STATA 10.2, StataCorp (2006). The random effects probit estimator uses a twelve-point quadrature approximation for the likelihood integral.

¹⁰ Only a subset of the variables included in the estimation are presented in the tables for brevity. A full set of results is available from the authors upon request.

¹¹ Note that there is no restriction on the size of the coefficient imposed in a probit model, as would be required for stationarity in a time-series model.

college graduates and it is only slightly, but significantly, higher for female college graduates in comparison to their high school counterparts¹².

Turning to the conventional determinants, age appears to have a statistically significant, inverse U-shaped profile for female college graduates with training increasing to age 41 before declining. Although this may relate to interrupted employment spells for women we could find no effect associated with dependent children. In contrast, age has a statistically insignificant effect in the model for male employees, suggesting that the training profile for men is flat over their working life. Racial background (white) is also found to have a positive and statistically significant impact for men but not for women, which could be an indication of discrimination operating against non-white males or may simply reflect the supply-side effect that male employees from among the ethnic minorities invest less in human capital. Private sector affiliation reduces training intensity for men and women alike, as does part-time work, and temporary work. Trade union coverage is also associated with a significantly positive coefficient.

Men and women in less skills-intensive occupations also appear less likely to participate in work-related training. Hierarchical effects also appear to be present with the model suggesting a significantly positive effect for both men and women in professional, managerial and skilled non-manual occupations. Supervisory and managerial roles increase the likelihood of training for women similarly.

[table 3 here]

VI Predicted probabilities

¹² Models including a two-year lag were also estimated with substantially the same effect.

The random effects probit models may be used to calculate work-related training probabilities for men and women that may be compared to the raw aggregate probabilities. These calculations are detailed in Table 4a in the case of men and Table 4b for women. In each Table the first panel presents raw data probabilities of experiencing training in the current period, conditional upon the individual having experienced a training episode in the 12-month period prior. The second panel records the predicted probabilities from the random effects probit model. In this calculation the effect of state dependence is calculated as the difference between the predicted probability obtained by setting the individual unobserved effect to zero and conditioning, first on the individual having participated in work-related training in the previous period and second, on the individual not having participated in work-related training in the previous period. The calculations therefore show the probability of a randomly chosen individual being observed participating in any form of work-related training in the current period; holding characteristics constant, and conditional upon having participated or not in work-related training in the previous period. The difference is reported in row (6) of each table. The results suggest that for men, state dependence on average accounts for approximately 53% of the probability of training the current period, conditional on having experienced some form of work-related training in the previous period. For women the corresponding figure is considerably lower, at approximately 38%, suggesting substantially greater state dependence among male workers. The estimate appears stable and consistent across the BHPS waves suggesting that state dependence plays a significant role in determining the incidence of training for both men and women comparable with that associated with formal educational qualifications.

[tables 4a and 4b here]

VI Summary and Conclusions

This study has examined whether an individual's participation in work-related training is characterised by state dependence, with current participation depending in part upon the worker having previously participated in a training episode. Training participation is modelled as a

dynamic random effects probit model where the effects of unobserved heterogeneity and initial conditions are accounted for in a fashion consistent with methods proposed by Chamberlain (1984) and Wooldridge (2002) respectively. Within this framework, the effects of state dependence on the probability of training are captured by the inclusion of the lagged outcome variable. The evidence suggests that not only does state dependence exist for training in the UK but that the effect is important and of a magnitude comparable with that associated with formal educational qualifications. Such a relationship clearly has important implications for labour market behaviour and outcomes. Our evidence suggests that skills investment early in the life cycle will have greater effect than investment undertaken later simply as a result of state dependence. Training today makes it much more likely that the individual will train tomorrow and this provides an important deviation-amplifying mechanism that may underpin wage dispersion and differentiated growth among individual workers and firms. The result also has implications for public policy lifelong learning agendas which again, on the evidence, would appear to be most effectively served by targeting training on the young.

VII References

- ALMEIDA-SANTOS, F., and K. MUMFORD (2004): "Employee Training and Wage Compression in Britain," Institute for the Study of Labor (IZA).
- ARULAMPALAM, W. (1998): "A Note on Estimated Coefficients in Random Effects Probit Models," *The Warwick Economics Research Paper Series (TWERPS)*.
- ARULAMPALAM, W., A. L. BOOTH, and M. P. TAYLOR (2000): "Unemployment Persistence," *Oxford Economic Papers*, 52, 24-50.
- BLACK, S. E., and L. M. LYNCH (1996): "Human-Capital Investments and Productivity," *The American Economic Review*, 86, 263-267.
- BLUNDELL, R., L. DEARDEN, and C. MEGHIR (1996): "The Determinants and Effects of Work-Related Training in Britain," *The Institute for Fiscal Studies*.
- (1999): "Work-Related Training and Earnings," *mimeo, Institute of Fiscal Studies*.
- BLUNDELL, R., L. DEARDEN, and B. SIANESI (1999): "Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy," *Fiscal Studies*, 20 (1), 1-24.
- BOHEIM, R., and A. L. BOOTH (2004): "Trade Union Presence and Employer-Provided Training in Great Britain," *Industrial Relations*, 43 (3), 520-545.
- BOOTH, A. L. (1991): "Job-Related Formal Training: Who Receives It and What Is It Worth?," *Oxford Bulletin of Economics and Statistics*, 53 (3), 281-94.

- BOOTH, A. L., and M. L. BRYAN (2002): "Who Pays for General Training? New Evidence for British Men and Women," Institute for the Study of Labor (IZA).
- CHAMBERLAIN, G. (1984): "Panel Data," *Handbook of Econometrics*, S. Griliches and M. Intriligator (eds), Amsterdam, North-Holland, 1247-318.
- CUHNA, F., and J. HECKMAN (2007): "The Technology of Skill Formation," *IZA DP No.2550*.
- DE LA FUENTE, A., and A. CICCONE (2003): "Human Capital in the Global Knowledge Based Economy: Final Report for the European Commission, Dg for Employment and Social Affairs," *Luxumberg*.
- DEARDEN, L., S. MACHIN, H. REED, and D. WILKINSON (1997): "Labour Turnover and Work-Related Training," *The Institute for Fiscal Studies*, London.
- DEARDEN, L., H. REED, and J. VAN REENEN (2000): "Who Gains When Workers Train? Training and Corporate Productivity in a Panel of British Industries," *Working Paper No. 00/04*, Institute for Fiscal Studies.
- (2005): "The Impact of Training on Productivity and Wages: Evidence from British Panel Data," *IFS Working Paper*, WP05/16 London.
- GREEN, F., S. MACHIN, and D. WILKINSON (1999): "Trade Unions and Training Practices in British Workplaces," *Industrial and Labor Relations Review*, 52 (2), 179-95.
- GREEN, F., and L. ZANCHI (1997): "Trends in the Training of Male and Female Workers in the United Kingdom," *British Journal of Industrial Relations*, 35, 635-644.
- GUTIERREZ, R. G., S. L. CARTER, and D. M. DRUKKER (2001): "On Boundary-Value Likelihood-Ratio Tests," *Stata Technical Bulletin*, 60: 15-18. Reprinted in *Stata Technical bulletin Reprints*, vol. 10, pp. 269-273.
- HECKMAN, J. J. (1981a): "Statistical Models for Discrete Panel Data," in C.F. Manski and D. McFadden (eds), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, 114-78. 24
- (1981b): "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process," in C.F. manski and D. McFadden (eds), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, 115-78.
- KEATING, D., and C. HERTZMAN (1999): "Developmental Health and the Wealth of Nations: Social, Biological and Educational Dynamics," *New York: The Guilford Press*.
- O.CONNELL, P. (1999): "Adults in Training: An International Comparison of Continuing Education and Training," *The Economics and Social Research Institute*, Dublin.
- OECD (1999): "Training of Adult Workers in OECD Countries: Measurement and Analysis," *Employment Outlook*.
- (2003): "Upgrading Workers' Skills and Competencies," *OECD Employment Outlook: Towards More and Better Jobs*, 237-296.
- SOUSOUNIS, P. (2008): "State dependence in work-related training participation in Britain: A comparison of different methods of estimation," Keele University, *mimeo*.
- TAYLOR, M. F. E., J. BRICE, N. BUCK, and E. PRENTICE-LANE (2006): "British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices," *Colchester: University of Essex*
- WOOLDRIDGE, J. M (2002): "Econometric Analysis of Cross Section and Panel Data," *The MIT Press*.
- (2005): "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel

Data Models with Unobserved Heterogeneity," *Journal of Applied Econometrics*, 20, 39-54.

Table 1 Raw-data and conditional probabilities of training participation

	Overall	1992	1993	1994	1995	1996	1997
$P(T = 1)$	0.4244	0.3958	0.4026	0.4129	0.4143	0.4337	0.4097
$P(T = 0)$	0.5756	0.6042	0.5974	0.5871	0.5857	0.5663	0.5903
$P(T = 1 T_{t-1} = 1)$	0.6254	0.6039	0.6522	0.6768	0.6699	0.6780	0.6329
$P(T = 0 T_{t-1} = 1)$	0.3746	0.3961	0.3478	0.3232	0.3301	0.3220	0.3671
$P(T = 1 T_{t-1} = 0)$	0.2815	0.1919	0.2347	0.2401	0.2388	0.2576	0.2434
$P(T = 0 T_{t-1} = 0)$	0.7185	0.8081	0.7653	0.7599	0.7612	0.7424	0.7566

Table 2 Training Participation across occupational groups (%)

	Males		Females	
	Freq.	%	Freq.	%
professional occupations	473	11.06	148	4.33
managerial & technical occupations	1,741	40.70	1,722	50
skilled non-manual occupations	708	16.55	1,055	31
skilled manual occupations	1,003	23.45	239	6.99
partly skilled occupations	304	7.11	218	6
unskilled occupations	49	1.15	37	1.08
Total	4,278	100	3,419	100

Table 3 Marginal Effects from the Random Effects Probit Models

	Comparison Models		Random Effects Models					
	High Sch. Grad.	College Grad.	High School Graduates			College Graduates		
	(1)	(2)	All	Males	Females	All	Males	Females
Trained $t - 1$	0.244*** (0.00939)	0.276*** (0.0118)	0.172*** (0.0149)	0.213*** (0.0222)	0.135*** (0.0198)	0.205*** (0.0186)	0.211*** (0.0252)	0.198*** (0.0276)
Sex (Female)	-0.00523 (0.0108)	0.0472*** (0.0141)	0.00962 (0.0151)			0.0572*** (0.0208)		
Age	-0.000581 (0.00274)	-0.000708 (0.00469)	0.00195 (0.00390)	-0.00193 (0.00571)	0.00348 (0.00541)	0.00300 (0.00681)	-0.0184* (0.00961)	0.0169* (0.00954)
Age ²	-5.28e-05 (3.40e-05)	-2.93e-05 (5.74e-05)	-7.53e-05 (4.78e-05)	-4.18e-05 (7.01e-05)	-7.96e-05 (6.62e-05)	-7.08e-05 (8.29e-05)	0.000150 (0.000116)	- (0.000117)
Race (White)	0.0830*** (0.0291)	0.0550* (0.0301)	0.0737* (0.0405)	0.153*** (0.0455)	-0.00202 (0.0661)	0.0775* (0.0458)	0.0780 (0.0616)	0.1000 (0.0671)
Marital Status (Single)	-0.00390 (0.0113)	-0.0145 (0.0152)	-0.00538 (0.0151)	-0.0318 (0.0225)	0.00467 (0.0202)	-0.0140 (0.0215)	- (0.0307)	0.0325 (0.0293)
Professional	0.205*** (0.0514)	0.118* (0.0716)	0.127* (0.0664)	0.122 (0.0874)	0.278* (0.143)	-0.0282 (0.109)	-0.164 (0.181)	0.0702 (0.125)
Managerial & technical	0.201*** (0.0293)	0.137* (0.0773)	0.159*** (0.0389)	0.204*** (0.0681)	0.146*** (0.0483)	0.00388 (0.105)	-0.122 (0.178)	0.0907 (0.131)
Skilled non-manual	0.151*** (0.0248)	0.114 (0.0717)	0.122*** (0.0328)	0.166** (0.0647)	0.118*** (0.0374)	-0.0143 (0.107)	-0.0853 (0.184)	0.0405 (0.126)
Skilled manual	0.0845*** (0.0264)	0.0757 (0.0747)	0.0935*** (0.0349)	0.103* (0.0551)	0.0790 (0.0492)	-0.0413 (0.109)	-0.168 (0.179)	0.0258 (0.132)
Partly skilled	0.0455* (0.0257)	0.0347 (0.0789)	0.0339 (0.0328)	0.0151 (0.0561)	0.0575 (0.0416)	-0.0766 (0.112)	-0.222 (0.177)	0.0202 (0.130)
Private sector	-0.0790*** (0.0175)	-0.0842*** (0.0226)	- (0.0229)	- (0.0354)	-0.0704** (0.0303)	-0.0742** (0.0306)	-0.0613 (0.0434)	-0.0814* (0.0438)
Permanent job	0.0386* (0.0208)	0.0594** (0.0270)	0.0394 (0.0266)	0.0457 (0.0430)	0.0318 (0.0343)	0.102*** (0.0362)	0.109* (0.0632)	0.104** (0.0441)
Working PT	-0.0636*** (0.0117)	-0.112*** (0.0199)	- (0.0150)	-0.0787* (0.0451)	- (0.0172)	-0.111*** (0.0273)	-0.206*** (0.0610)	- (0.0304)
TU coverage	0.0631*** (0.0104)	0.0861*** (0.0156)	0.0446*** (0.0137)	0.0279 (0.0199)	0.0608*** (0.0188)	0.0777*** (0.0214)	0.0733*** (0.0277)	0.0735** (0.0334)
γ_1			0.141*** (0.0147)	0.149*** (0.0212)	0.136*** (0.0203)	0.170*** (0.0211)	0.167*** (0.0279)	0.152*** (0.0314)
NT	12544	7457	10194	4564	5630	6389	3417	2972
N			2282	1036	1246	1408	743	665

Notes: Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4a Raw Data Probabilities and Predicted Probabilities for Men

	Wave 2 [1992]	Wave 3 [1993]	Wave 4 [1994]	Wave 5 [1995]	Wave 6 [1996]	Wave 7 [1997]
Raw Data Probabilities:						
(1) Training the previous period	0.6239	0.6578	0.6553	0.6775	0.6726	0.6071
(2) No training the previous period	0.1818	0.2490	0.2541	0.2564	0.2395	0.2103
(3) (1)-(2)	0.4421	0.4088	0.4012	0.4211	0.4331	0.3968
Predicted Probabilities holding characteristics constant						
(4) Trained at $t-1$	0.4808	0.5220	0.5133	0.5255	0.5285	0.4604
(5) Not trained at $t-1$	0.2606	0.2952	0.2878	0.2983	0.3009	0.2442
(6) State Dependence	0.2202	0.2268	0.2255	0.2272	0.2276	0.2162
As % of (3)	49.8	55.4	56.2	53.9	52.5	54.4

Table 4b Raw Data Probabilities and Predicted Probabilities for Women

	Wave 2 [1992]	Wave 3 [1993]	Wave 4 [1994]	Wave 5 [1995]	Wave 6 [1996]	Wave 7 [1997]
Raw Data Probabilities						
(1) Training the previous period	0.5843	0.6470	0.6980	0.6627	0.6833	0.6572
(2) No training the previous period (%)	0.2007	0.2217	0.2280	0.2224	0.2738	0.2744
(3) (1)-(2)	0.3836	0.4253	0.4700	0.4403	0.4095	0.3828
Predicted Probabilities holding characteristics constant						
(4) Trained at $t-1$	0.4210	0.4298	0.4486	0.4584	0.5002	0.5033
(5) Not trained at $t-1$	0.2682	0.2757	0.2918	0.3004	0.3379	0.3407
(6) State Dependence	0.1528	0.1541	0.1568	0.1580	0.1623	0.1626
As % of (3)	39.8	36.2	33.3	35.8	39.6	42.4

Appendix

Table A1 Coefficient Estimates from the Random Effects Probit Models

	Comparison Models		Random Effects Models					
	High Sch. Grad.	College Grad.	High School Graduates			College Graduates		
	(1)	(2)	All (3)	Males (4)	Females (5)	All (6)	Males (7)	Females (8)
Trained $t - 1$	0.673*** (0.0257)	0.723*** (0.0319)	0.508*** (0.0414)	0.626*** (0.0616)	0.405*** (0.0562)	0.532*** (0.0486)	0.540*** (0.0655)	0.528*** (0.0726)
Sex (Female)	-0.0149 (0.0309)	0.123*** (0.0369)	0.0298 (0.0469)			0.150*** (0.0547)		
Age	-0.00166 (0.00781)	-0.00184 (0.0122)	0.00603 (0.0121)	-0.00594 (0.0176)	0.0109 (0.0169)	0.00783 (0.0177)	-0.0468* (0.0245)	0.0458* (0.0259)
Age ²	-0.000151 (9.7e-05)	-7.61e-05 (0.00015)	-0.000233 (0.00015)	-0.000129 (0.00022)	-0.000249 (0.00021)	-0.000185 (0.00022)	0.000383 (0.00029)	- (0.00032)
Race (White)	0.254*** (0.0967)	0.141* (0.0762)	0.249* (0.151)	0.591** (0.237)	-0.00630 (0.206)	0.198* (0.116)	0.197 (0.154)	0.261 (0.171)
Marital Status (Single)	-0.0111 (0.0323)	-0.0375 (0.0392)	-0.0167 (0.0469)	-0.0997 (0.0719)	0.0146 (0.0627)	-0.0364 (0.0558)	-0.213*** (0.0774)	0.0892 (0.0812)
Professional	0.535*** (0.129)	0.321 (0.205)	0.358** (0.174)	0.344 (0.230)	0.742** (0.360)	-0.0729 (0.280)	-0.415 (0.462)	0.198 (0.371)
Managerial & technical	0.538*** (0.0760)	0.356* (0.202)	0.456*** (0.105)	0.576*** (0.183)	0.422*** (0.132)	0.0101 (0.274)	-0.313 (0.459)	0.244 (0.349)
Skilled non-manual	0.420*** (0.0679)	0.307 (0.203)	0.366*** (0.0959)	0.469*** (0.173)	0.371*** (0.118)	-0.0371 (0.276)	-0.215 (0.462)	0.112 (0.353)
Skilled manual	0.234*** (0.0717)	0.202 (0.205)	0.277*** (0.0997)	0.312* (0.165)	0.233* (0.138)	-0.107 (0.278)	-0.425 (0.458)	0.0711 (0.367)
Partly skilled	0.127* (0.0706)	0.0912 (0.210)	0.103 (0.0977)	0.0461 (0.169)	0.174 (0.122)	-0.196 (0.283)	-0.565 (0.472)	0.0553 (0.360)
Private sector	-0.220*** (0.0477)	-0.220*** (0.0594)	-0.222*** (0.0668)	-0.237** (0.100)	-0.215** (0.0905)	-0.194** (0.0806)	-0.157 (0.112)	-0.219* (0.117)
Permanent job	0.113* (0.0629)	0.152** (0.0684)	0.127 (0.0895)	0.147 (0.146)	0.103 (0.115)	0.259*** (0.0911)	0.275* (0.159)	0.272** (0.113)
Working PT	-0.186*** (0.0352)	-0.286*** (0.0503)	-0.211*** (0.0504)	-0.265 (0.168)	-0.248*** (0.0559)	-0.284*** (0.0689)	-0.524*** (0.161)	-0.237*** (0.0797)
TU coverage	0.180*** (0.0297)	0.223*** (0.0403)	0.139*** (0.0426)	0.0862 (0.0616)	0.190*** (0.0591)	0.202*** (0.0554)	0.187*** (0.0705)	0.197** (0.0888)
Auxiliary Parameters								
γ_1			0.429*** (0.0448)	0.455*** (0.0654)	0.414*** (0.0616)	0.438*** (0.0545)	0.424*** (0.0716)	0.405*** (0.0828)
Intercept	-190.0*** (20.55)	-111.3*** (27.93)	-2.571*** (0.470)	-1.707** (0.702)	-3.191*** (0.652)	-0.413 (0.638)	1.201 (0.900)	-1.670* (0.928)
r			0.2164	0.1750	0.2368	0.2178	0.1813	0.2245
Log- \mathcal{L}			-5271.63	-2318.67	-2912.92	-3679.18	-1999.33	-1649.89
Log- \mathcal{L} ($r = 0$)			-5334.50	-2335.59	-2953.15	-3720.40	-2015.29	-1669.03

<i>NT</i>	12544	7457	10194	4564	5630	6389	3417	2972
<i>N</i>			2282	1036	1246	1408	743	665

Notes: All models include employer size variables, job role variables, regional dummy variables, time dummies and dummies for the Standard Industrial Classification of the current job. In addition, all RE probit models include time means of the time-varying covariates to allow for correlation between the latter and the unobserved heterogeneity and allow for endogenous initial conditions. Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2 Summary Statistics

	Men		Women	
	mean	sd	mean	sd
Training	0.444	0.497	0.436	0.496
Age	38.25	10.42	38.69	10.45
Race	0.032	0.175	0.026	0.161
Higher degree	0.028	0.165	0.016	0.126
First degree	0.113	0.316	0.101	0.301
Teaching qualification	0.014	0.119	0.043	0.202
Other higher qualification	0.259	0.438	0.141	0.348
Nursing qualification	0.002	0.048	0.039	0.193
GCE A-level	0.154	0.360	0.109	0.312
GCE O-level or equivalent	0.195	0.396	0.282	0.450
Commercial qf no O-levels	0.002	0.044	0.057	0.231
CSE 2-5, Scottish grade 4-5	0.054	0.226	0.034	0.181
Apprenticeship	0.026	0.158	0.004	0.064
Other qualification	0.006	0.076	0.008	0.089
No qualification	0.147	0.354	0.167	0.373
Part-time	0.035	0.184	0.356	0.479
Private sector	0.769	0.426	0.584	0.493
Trade union coverage	0.534	0.498	0.556	0.497
Mining	0.377	0.485	0.143	0.350
Construction	0.059	0.236	0.007	0.083
Wholesale, retail, hotels, transport	0.226	0.418	0.242	0.428
Financial services, real estate	0.120	0.325	0.133	0.340
Public administration, education, health & social	0.181	0.385	0.430	0.495
Other Community, Social	0.021	0.142	0.041	0.198
Agriculture	0.015	0.122	0.004	0.065
Professional	0.088	0.283	0.026	0.159
Managerial and technical	0.324	0.468	0.323	0.468
Skilled non-manual	0.137	0.343	0.389	0.488
Skilled manual	0.309	0.463	0.079	0.271
Partly skilled	0.121	0.326	0.142	0.349
Unskilled	0.021	0.144	0.041	0.199

Establishment < 50	0.387	0.487	0.521	0.500
Establishment 50-99	0.141	0.348	0.125	0.331
Establishment 100-499	0.271	0.445	0.213	0.409
Establishment 500+	0.200	0.406	0.141	0.348
