The Inter-related Dynamics of Unemployment and Low Pay*

Mark B. Stewart
University of Warwick

January 2002

*I am grateful to Wiji Arulampalam, Keith Cowling, Paul Gregg, Mary Gregory, Richard Jackman, Steve Jones, Alan Manning, Robin Naylor, Steve Nickell, Odbjorn Rauum, Sherwin Rosen, Jeremy Smith, Ian Walker and seminar participants at LSE, Warwick and the EALE/SOLE joint World Congress in Milan for helpful comments and discussions and to the Leverhulme Trust for financial support. The British Household Panel Survey data used in the paper were collected by the ESRC Research Centre on Micro-social Change and made available through The Data Archive.

Abstract

This paper examines the extent of state dependence in individual unemployment and low paid employment and the inter-related dynamics between the two. Evidence is presented that (after controlling for unobserved heterogeneity and initial conditions) the low paid are more likely to become unemployed and the unemployed more likely to be low paid on re-entry to employment. The impacts of unemployment and low pay are found to be insignificantly different from one another, both on the probability of unemployment in the next period and on the probability of low pay. Evidence is presented that low paid jobs act as the main conduit for repeat unemployment and considerably increase its probability. Those who get a better paid job eliminate the increased risk of repeat unemployment.

JEL classification: J64, J31.

Key words: unemployment, low pay, dynamics, state dependence, unobserved heterogeneity, wage mobility, dynamic random effects models, repeat unemployment, recurrent unemployment.

Address for correspondence: Mark B. Stewart, Economics Department, University of Warwick, Coventry CV4 7AL, UK.
E-mail: Mark.Stewart@warwick.ac.uk

1 Introduction

A significant proportion of the unemployed who get re–employed lose their jobs again fairly quickly.[1] For Britain the existing evidence suggests that this is in large part due to strong state dependence in individual unemployment experience.[2] That is to say, experiencing unemployment of itself makes future unemployment more likely. However we know little about the mechanism that lies behind this observed persistence. This paper argues that a large part of the reason for the persistence is the quality of the jobs that re-entrants get, which tend to be low paid and unstable, what might be labelled "dead–end" jobs. This means that a fuller picture of the dynamics of unemployment is provided by examining it in conjunction with the type of job taken when an exit from unemployment occurs.

There is also evidence of similarly strong persistence in low pay.[3] Someone who is low paid in one period is far more likely to be low paid in the following period than someone higher up the pay distribution, even after controlling for other observable factors and for the endogenous selection or initial conditions problem that this comparison involves. As with the persistence in unemployment, we know less about the mechanism that lies behind this observed persistence. As well as this persistence in both unemployment and low pay, there is also evidence of a link between them giving a "low pay – no pay cycle".[4] The low paid are more likely to become unemployed in the future, the unemployed are more likely to be low paid on re–entry to employment, and this probability of being low paid rises even further if the individual was low paid before becoming unemployed.[5] How much of this is due to heterogeneity remains an open question. This paper studies these inter-related dynamics of low pay and unemployment.

Any "scarring" effects of unemployment are generally viewed as resulting from the non–accumulation of new (and the deterioration of existing) human capital during an unemployment spell and from adverse selection. Stewart (1999c) puts forward the hypothesis, on the basis of the evidence described in the previous paragraph, that certain low paid jobs may have similar "scarring" effects to unemployment on an individual's future labour market prospects.

[1] See for example Clark and Summers (1979) and Layard et al. (1991). The unemployed claimant count figures for Britain indicate that about half of those leaving the count return within a year (Sweeney, 1996). This figure has remained fairly constant throughout the 1980s and 1990s (Teasdale, 1998).

[2] See Arulampalam et al. (2000) and Narendranathan and Elias (1993). In contrast the evidence for the US suggests a lack of state dependence. See Heckman and Borjas (1980) and Corcoran and Hill (1980).

[3] See Stewart and Swaffield (1997, 1999). There is also an extensive literature on wage persistence and wage dynamics in general, not focusing specifically on the bottom end of the distribution. See Atkinson et al. (1992) for a survey and Moffitt and Gottschalk (1993), Baker (1997) and Dickens (2000) for recent contributions.

[4] Stewart and Swaffield (1997), Stewart(1999a), Gregg and Leonardi (1999).

[5] For the US there is considerable evidence of significant and long-lasting earnings losses associated with job displacement. (See Jacobson et al., 1993, Stevens, 1997 and Kletzer, 1998, *inter alia.*) In addition Stevens in particular stresses the likelihood and important impact of additional job losses subsequent to the initial displacement. For Britain Gregg and Wadsworth (1996), Gregory and Jukes (1997) and Gregg et al. (1999) present evidence of a negative impact of unemployment on subsequent earnings.

While the hypothesis as stated above views jobs only in terms of their pay levels, distinguishing between low paid and higher paid jobs, this may be viewed as just one dimension of the "good jobs / bad jobs" debate (Burtless et al., 1990, Acemoglu, 1997). Similarly Layard et al. (1990), who make extensive use in their analysis of a two–sector model, argue that "employers offering good jobs may well use a person's current position as a screening device. While unemployment is a bad signal, being in a low–quality job may well be a worse one" (Layard et al., 1990, page 249). The two sectors, or the good job / bad job divide, may be viewed as resulting from efficiency wages or an insider / outsider context. In the segmented labour market literature it is hypothesised that "bad" jobs may have a "scarring" effect on future employment prospects, allowing skills to deteriorate and thereby reducing future earnings (see Taubman and Wachter, 1986). McCormick (1990) terms these "stigmatized" jobs and argues that faced by uncertainty about worker quality, firms use job type held (be it "good" vs. "bad" or skilled vs. unskilled), alongside unemployment duration (Blanchard and Diamond, 1994), as a cheap indicator of future productivity. This works in equilibrium because high productivity workers find "bad" jobs more costly and so are prepared to invest in moving quickly to a new "good" job.

This paper examines the extent to which "bad" jobs have "scarring" effects on future employment and wage prospects, alongside those of unemployment. The econometric models used allow for unobserved (as well as observed) heterogeneity and inter-related dynamics and explicitly model initial conditions.

2 Probabilities of unemployment and low pay

This section describes the data used in the paper and conducts a preliminary examination of the conditional probabilities of being unemployed and of being in a low paid job that the data display. The data used in the paper are from the first six waves of the British Household Panel Survey (BHPS), covering 1991–1996. The BHPS contains a nationally representative sample of households whose members are re-interviewed each year.[6] The sample used here contains only Original Sample Members (OSMs), is restricted to those aged between 18 and the state retirement age[7] and excludes full-time students.

The raw (or aggregate) probabilities of unemployment and low pay for various groups distinguished by status at $t-1$ are presented in Table 1. The unemployment indicator used is based on the ILO/OECD definition of unemployment. Under this definition a person is unemployed if he or she does not have a job, but had looked for work in the past four weeks and is available for work.[8] The table shows that there is considerable persistence in unemployment in terms of the raw data. The conditional probabilities in the first column of Table 1, based on data pooled over the years

[6] See Taylor (1996) for details.

[7] 65 for men and 60 for women.

[8] An individual is classified as having a job if they did any work for pay or profit during the week prior to interview (for any number of hours and including casual work) - either as an employee or self employed. They are also classified as employed if they had a job that they were away from that week or were waiting to take up a job already obtained. Looking for work includes "registration at employment agency, approaching employers, checking ads, making enquiries" and also includes looking for a government scheme.

1992 to 1996, exhibit a strong association between present and past unemployment. The first row of the table gives the *unconditional* probability of being unemployed at a point in time.[9] The aggregate probability is about 0.07. This is made up of a probability of about 0.08 for men and about 0.06 for women. The remainder of the table gives conditional probabilities on the basis of status at $t-1$ (i.e. at the previous interview – roughly 12 months previously[10]).

The second and third rows give the conditional probabilities for those out of and in the labour force respectively at $t-1$. It is clear that the probability of being unemployed currently is much higher for those who were out of the labour force at the previous interview (roughly 12 months previously) than for those who were in the labour force (employed or unemployed).[11] For those out of the labour force at $t-1$ (but in at t) the probability of being unemployed at t is 0.37, compared with 0.05 for those who were in the labour force (employed or unemployed) at $t-1$.

Among those in the labour force at $t-1$, the next two rows of the table then partition the sample into those who were unemployed at the previous survey and those who were employed. The conditional probability of current unemployment is considerably higher for the former group than the latter (0.46 against 0.03). The final row of this block of the table gives the ratio of these two probabilities. Someone who was unemployed at the previous survey date is 16.8 times as likely to be unemployed at the date of the current interview as someone who was employed at the previous interview. This ratio of conditional probabilities can be viewed as a measure of the persistence in the probability of unemployment. The figures for men and women are fairly similar (16.4 and 15.9 respectively), despite both conditional probabilities being quite a bit lower for women.

Part, or even all, of this persistence may be due to heterogeneity. The probability of unemployment is higher for the young, for those without qualifications, for those with poor health, etc. Even if there were no structural persistence for individuals, this heterogeneity would cause the group of individuals unemployed at the previous date to have a higher aggregate probability of unemployment at the current date than those who were employed.

By way of preliminary analysis, the first column of Table 2 presents the results of fitting a simple pooled Probit model (1992–1996) including the previous state as one of the covariates. Unobserved heterogeneity, considered in the next section, is ignored here. Defining a binary variable $y_{it} = 1$ if individual i is unemployed at t and $y_{it} = 0$ if employed at t the probabilities of unemployment and employment are specified as:

$$P\left[y_{it}|x_{it}, y_{it-1}\right] = \Phi\left[(x_{it}'\beta + \gamma y_{it-1})(2y_{it} - 1)\right] \tag{1}$$

where Φ is the cumulative distribution function of a standard normal random variable. As additional covariates I include years of education, whether the individual has any qualifications, a quadratic in years of experience, gender and marital status (and the interaction between them), health status, whether resident in London or the South East and the local unemployment-vacancy ratio.

[9] Conditional on being in the labour force (i.e. either employed or unemployed) at t, but unconditional in the sense of not being conditioned on status at $t-1$.

[10] 72% of the pooled sample are interviewed less than 30 days either side of the anniversary of the previous interview, 91% within 2 months.

[11] Again the sample for these probabilities is restricted to those who are in the labour force at t.

This model can be used to calculate predicted probabilities of current unemployment given unemployment at the previous survey date and given employment at the that date, but with otherwise identical observed characteristics (i.e. x's). There are a number of ways of calculating simple summary measures of these probabilities (the usual index number problem). The one used here, though non-standard, both has intuitive appeal and is easily extendable to the models described in Section 3 below. The two probabilities of interest are evaluated at a value of $x'\beta$ that gives the probability of unemployment at t given employment at $t-1$ equal to the corresponding sample proportion. More formally, define $\overline{p}_0$ to be the sample proportion who have $y_{it} = 1$ among those with $y_{it-1} = 0$. Then the evaluation can be conducted at $x'\beta = \Phi^{-1}(\overline{p}_0)$. For someone with characteristics corresponding to this value, the predicted probability given employment at $t-1$ is simply $\overline{p}_0$ and the predicted probability given unemployment at $t-1$ is given by

$$\hat{p}_1 = \Phi[\hat{\gamma} + \Phi^{-1}(\overline{p}_0)]. \tag{2}$$

These two predicted probabilities are also given in the first column of Table 2 along with the ratio

$$R = \frac{\hat{p}_1}{\overline{p}_0}. \tag{3}$$

For the probability of unemployment, this gives an estimated probability ratio of 15.5. (The table also gives the corresponding unadjusted ratio of conditional probabilities on the same sample.[12]) Thus, despite their importance in terms of statistical and numerical significance, controlling for the covariates listed above reduces the extent of the measured persistence in unemployment very little. Someone unemployed at the previous survey date is still 15.5 times as likely to be unemployed currently as someone who was employed at the previous date, but otherwise has the same values of the characteristics listed.

Turning to the probability of being low paid, the second column of Table 1 presents raw (or aggregate) conditional probabilities based on a low pay threshold of £3.50 per hour in April 1997 terms.[13] This threshold classifies around 10% of employees as low paid.[14] The lower block of the column indicates that the probability of being low paid currently is much higher for those who were low paid at the time of the previous interview (roughly 12 months previously) than for those who were paid above the threshold (0.57 against 0.03). Those low paid at $t-1$ are 16.8 times as likely to be low paid at t as someone higher paid at $t-1$. As for unemployment, the impact of observed heterogeneity on this measure of persistence is examined using a pooled Probit model. The results are given in the third column of Table 2. Once again controlling for the same covariates reduces the extent of the measured persistence only very slightly. Someone low paid at the previous survey date is still 14.1 times as likely to be low paid currently as someone who was higher paid at the previous date, but otherwise has the same values of the characteristics listed at the foot of Table 2.

[12] The value differs slightly from that in Table 1, since the samples used in Table 2 exclude observations with missing values on the covariates listed above (and at the foot of the table).

[13] Pay is defined in terms of average hourly earnings and is adjusted to April 1997 terms using the Average Earnings Index. This threshold is the main definition of low pay used by the Low Pay Commission in their analysis of low pay in the UK (LPC, 1998).

[14] A national minimum wage was introduced in April 1999 at £3.60 per hour for those aged 22 and over. There are lower rates of £3.00 for those aged 18–21 (inclusive) and £3.20 for those aged 22 and over on accredited training within the first 6 months of a new job with a new employer.

Summary descriptive statistics to illustrate the “low pay – no pay cycle” are also presented in Tables 1 and 2. The lower block of the first column of Table 1 indicates that employees who were low paid at $t-1$ are 2.7 times as likely to be unemployed at t as those who were higher paid at $t-1$. The second column of Table 2 shows that controlling for the same characteristics as before changes this ratio very little.

The upper block of the second column of Table 1 indicates that those unemployed at $t-1$ are 3 times as likely to be low paid at t (given employment) as those employed at $t-1$. This means that they are about half as likely as those who were low paid at $t-1$ and about 8 times as likely as those who were higher paid at $t-1$. As above, these figures are altered little by controlling for the same characteristics as before – results of which are given in the fourth column of Table 2.

Thus in terms of aggregate conditional probabilities, there is considerable persistence in both unemployment and low paid employment. In addition the aggregate dynamic processes are inter-related, in the sense that the current probability of each depends on the past occurrence of the other, meaning that in terms of aggregate probabilities there is evidence of a “low pay – no pay cycle”. None of the probability ratios are reduced much by allowing for observed heterogeneity in some of the more obvious variables that influence these probabilities.

3 Unobserved heterogeneity and initial conditions

The results of the preliminary data examination in the previous section do not take account of unobserved heterogeneity or initial conditions. If the unobserved heterogeneity exhibits persistence over time, then ignoring it will lead to an overstatement of the true persistence or state dependence in unemployment and low pay. This issue is the focus of attention in this section. Two contrasting models and estimation methods are used to address the problem. Section 3.1 outlines a dynamic random effects probit model for a binary outcome and an estimation approach proposed by Heckman (1981) and used recently by Arulampalam et al. (2000). In Section 4 this model is applied sequentially to the probability of being unemployed and the probability of being low paid (given employment). Applying the model in this way ignores the possible selection bias in conditioning on employment when analysing the probability of being low paid. This is analogous to the common practice of treating the selection on employment as exogenous when estimating a wage equation. This issue is discussed further when the results are presented in Section 4 below.

The dynamic random effects probit model outlined in Section 3.1 requires an auxiliary distributional assumption on the individual-specific effects. To investigate the potential sensitivity of the results to this assumption an alternative approach is also used involving GMM estimation of a dynamic linear probability model. This is outlined in Section 3.2.

3.1 A dynamic random effects probit model

The latent linear equation for this model is specified as

$$y_{it}^* = x_{it}'\beta + \gamma y_{it-1} + \varepsilon_i + u_{it} \tag{4}$$

$(i = 1, \ldots, N; t = 2, \ldots, T)$ with $u_{it} \sim N(0, \sigma_u^2)$, where y_{it}^* is a measure of the propensity toward a particular labour market state (unemployed vs. employed or low pay vs. higher pay) and y_{it} is an indicator variable for being in that state:

$$y_{it} = \begin{cases} 1 & \text{if } y_{it}^* \geq 0 \\ 0 & \text{else.} \end{cases} \tag{5}$$

The subscript i indexes individuals and t time periods. N is large, but T is small and fixed. Asymptotics are on N alone. The composite error term contains two components, ε and u. Although the u_{it} are assumed *iid*, the composite error term will be correlated over time due to the individual–specific time–invariant ε terms.

In the basic form of the model it is assumed that $\varepsilon_i \sim iidN$ and independent of u_{it} for all i and t. In the standard (uncorrelated) random effects model ε_i is also assumed uncorrelated with x_{it}. Alternatively, following Mundlak (1978) and Chamberlain (1984), the possibility that the unobserved ε_i are correlated with the observed characteristics in the model can be allowed for by assuming a linear relationship between ε and either the time means of the x-variables or a combination of their lags and leads, e.g.:

$$\varepsilon_i = \overline{x}_i' a + \alpha_i \tag{6}$$

where $\alpha_i \sim iidN(0, \sigma_\alpha^2)$ and independent of x_{it} and u_{it} for all i and t. Thus the latent linear equation may be written as

$$y_{it}^* = x_{it}'\beta + \gamma y_{it-1} + \alpha_i + u_{it} \tag{7}$$

$(i = 1, \ldots, N; t = 2, \ldots, T)$ where, to reduce notation changes, the time means or lags and leads have been subsumed into the x-vector.

If the composite error in this equation is written as $v_{it} = \alpha_i + u_{it}$, the individual-specific random effects specification adopted implies that the correlation between the v_{it} in any two (different) periods will be the same:

$$\rho = Corr(v_{it}, v_{is}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_u^2} \qquad t, s = 2, \ldots, T; t \neq s \tag{8}$$

To estimate the model an assumption is required about the initial observations, y_{i1}, and in particular about their relationship with the unobserved heterogeneity component, α_i. The assumption giving rise to the simplest form of model for estimation is to take the initial conditions, y_{i1}, to be exogenous. This would be appropriate if the start of the process coincided with the start of the observation period for each individual, but will not be valid for the type of process being considered in this paper. Under this assumption the likelihood can be decomposed into two independent factors, the first giving the joint probability for $t = 2, \ldots, T$ and the second that for $t = 1$, and the first can be maximised without reference to the second. In this case estimation can be conducted using a standard Random Effects Probit program. If the initial conditions are correlated with the α_i, as we would expect in the current context, this method of estimation will overstate the degree of state dependence, γ (see for example Chay and Hyslop, 1998).

The approach to the initial conditions problem proposed by Heckman (1981) and adopted here involves specifying a linearised reduced form equation for the initial value of the latent variable:

$$y_{i1}^* = z_{i1}'\pi + \eta_i \tag{9}$$

where z_{i1} is a vector of exogenous instruments (and includes x_{i1}) and η_i is correlated with α_i, but uncorrelated with $u_{it}, t \geq 2$. Using an orthogonal projection, it can be written as

$$\eta_i = \theta\alpha_i + u_{i1} \tag{10}$$

with α_i and u_{i1} independent of one another. It is also assumed (following Arulampalam et al., 2000) that u_{i1} satisfies the same distributional assumptions as u_{it} for $t = 2, \ldots, T$. The linearised reduced form for the latent variable for the initial time period is therefore specified as

$$y_{i1}^* = z_{i1}'\pi + \theta\alpha_i + u_{i1} \tag{11}$$

$(i = 1, \ldots, N)$. The z-vector will include period 1 values of the x-variables together with pre-sample variables. (Heckman (1991) also suggests using polynomial functions of variable values in the pre-sample periods.) Parental variables and pre-first-wave variables related to labour market entry are used below.

Since y is a binary variable, a normalisation is required. A convenient one is that $\sigma_u^2 = 1$. If u_{it} is normally distributed, the transition probability for individual i at time t given α_i is given by

$$P\left[y_{it}|x_{it}, y_{it-1}, \alpha_i\right] = \Phi\left[(x_{it}'\beta + \gamma y_{it-1} + \alpha_i)(2y_{it} - 1)\right]. \tag{12}$$

The joint probability of the observed binary sequence for individual i in the Heckman approach is thus

$$\prod_{t=2}^{T} \Phi\left[(x_{it}'\beta + \gamma y_{it-1} + \alpha_i)(2y_{it} - 1)\right] \Phi\left[(z_{i1}'\pi + \theta\alpha_i)(2y_{i1} - 1)\right]. \tag{13}$$

For a random sample of individuals the likelihood to be maximised is then given by

$$\prod_i \int_{\alpha^*} \left\{ \prod_{t=2}^{T} \Phi\left[(x_{it}'\beta + \gamma y_{it-1} + \sigma_\alpha\alpha^*)(2y_{it} - 1)\right] \Phi\left[(z_{i1}'\pi + \theta\sigma_\alpha\alpha^*)(2y_{i1} - 1)\right] \right\} dF(\alpha^*) \tag{14}$$

where F is the distribution function of $\alpha^* = \alpha/\sigma_\alpha$. Under the normalisation used, σ_α can be written as $\sigma_\alpha = \sqrt{\rho/(1-\rho)}$. If α is taken to be normally distributed, the integral over α^* can be evaluated using Gaussian–Hermite quadrature (Butler and Moffitt, 1982). Maximum Likelihood estimates of the parameters of this model applied to both the probability of being unemployed and the probability of being low paid (given employment) are given in the Section 4.

3.2 GMM estimation of a dynamic LPM

A potential drawback to the dynamic random effects probit model considered in the previous sub-section is the possible sensitivity of its estimates to the auxiliary distributional assumption on the individual-specific effects. The model in this subsection takes an alternative approach in the context of a linear probability model (LPM) and uses a GMM estimator. The LPM is specified as follows

$$y_{it} = x'_{it}\mu + \lambda y_{it-1} + f_i + g_t + \omega_{it} \tag{15}$$

for $t = 2, \ldots, T$ and $i = 1, \ldots, N$. Differencing removes the individual-specific effects:

$$\Delta y_{it} = \Delta x'_{it}\mu + \lambda \Delta y_{it-1} + \Delta g_t + \Delta\omega_{it} \tag{16}$$

for $t = 3, \ldots, T$ and $i = 1, \ldots, N$. Even if the ω_{it} are serially independent, Δy_{it-1} and $\Delta\omega_{it}$ will be correlated and OLS applied to this differenced equation will be biased and inconsistent. A range of Instrumental Variable estimators have been proposed to provide consistent estimation. The now commonly used Arellano and Bond (1991) GMM estimator, involving a different number of instruments in each time period, is based on the moment conditions

$$E(y_{it-s}\Delta\omega_{it}) = 0, \qquad \text{for } t = 3, \ldots, T \text{ and } s \geq 2. \tag{17}$$

This gives $(T-1)(T-2)/2$ orthogonality conditions in general and 10 in the current context. This estimator provides efficiency gains over the simpler IV estimators proposed by Anderson and Hsiao (1981), which use for example either y_{it-2} or Δy_{it-2} to instrument Δy_{it-1}. However the results from using these estimators are also provided for comparison in Section 4 below.

The consistency of all these estimators requires the ω_{it} to be serially uncorrelated. Arellano and Bond (1991) propose a test on the second order residual correlation coefficient which will be used to test this below, along with a Sargan test of the over-identifying restrictions.

In situations where T is small and λ large the standard linear GMM estimator described above has been found to have poor finite sample properties. Lagged levels provide weak instruments for first differences in this case. Blundell and Bond (1998) propose a "system" GMM estimator based on a stacked system comprising the $T-2$ equations in first differences together with the $T-2$ equations in levels corresponding to periods $3, \ldots, T$ for which instruments are observed. The Arellano-Bond instruments are used for the first-differenced equations. The differences $\Delta y_{i2}, \ldots, \Delta y_{iT-1}$ are used as instruments for the levels equations as suggested by Arellano and Bover (1995). This estimator encompasses the non-linear GMM estimator of Ahn and Schmidt (1995) and Blundell and Bond demonstrate that there are substantial asymptotic efficiency gains relative to that estimator. Both standard Arellano-Bond GMM estimates and Blundell-Bond "system" GMM estimates are presented in the next section.

4 Results

The dynamic random effects probit model of Section 3.1 is applied separately to the probability of being unemployed at t and the probability of being low paid at t (conditional on employment). The four probit models reported in Table 2 and described in section 2 above are taken as the starting point of the analysis. The models reported there ignore unobserved heterogeneity and initial conditions and hence are likely to overstate the true state dependence in unemployment and low pay. GMM estimates of the dynamic LPM are provided for comparative purposes.

4.1 The probability of unemployment

Estimates from the dynamic random effects probit model for the probability of being unemployed at t are given in Column 2 of Table 3. The x-vector is as used in Table 2. The corresponding pooled probit model (without random effects) estimated on the same sample is given in the first column for comparison.

As indicated in the previous section, parental variables and pre-first-wave variables related to labour market entry are used as instruments. The instruments used include most of those used by Arulampalam et al. (2000), together with a number of additions. Dummy variables for father's broad SEG at the time the respondent was 14 are included in z together with dummies for father not working and father deceased. Similar variables in relation to the respondent's mother at the same date are also included. An indicator for whether or not the first labour market spell after leaving full-time education was an employment spell is included. Dummy variables for the broad SEG of the first job held (after leaving full-time education) are included. Also included are an indicator of whether this first job was a temporary job (including casual, seasonal and fixed contract jobs) and an indicator of whether the individual left this first job due to redundancy. In the estimated linearised reduced form for the initial condition this set of instruments (i.e. the variables in z excluding the period 1 values of the x variables) are jointly highly significant.[15]

Care must be taken when comparing coefficient estimates across the columns in Table 3. The dynamic random effects probit model and the pooled probit model involve different normalisations (see Arulampalam, 1999). The random effects probit coefficient estimates are normalised on $\sigma_u^2 = 1$ (see section 3.1 above), while the pooled probit coefficient estimates are normalised on $\sigma_v^2 = 1$. To put it another way, the random effects probit estimation provides an estimate of γ/σ_u, whereas the pooled probit estimation provides an estimate of γ/σ_v. Thus to make comparisons between them the former needs to be multiplied by an estimate of $\sigma_u/\sigma_v = \sqrt{1-\rho}$. The renormalised estimate of γ (the coefficient on unemployment at $t-1$) in Column 2 of Table 3 is 0.920. Thus, compared with the pooled probit estimate (without random effects and without allowance for initial conditions) in Column 1, the estimate of γ is reduced by a half in the random effects model, but remains strongly significant.

Predicted conditional probabilities can be calculated for this model by a comparable method to that used above for the pooled probit. Equation (12) in Section 3.1 gives the conditional probabilities given x_{it}, y_{it-1} and α_i. The probability that $y_{it} = 1$ given $y_{it-1} = 0$ (and x_{it}, α_i) is evaluated at the value of $(x'_{it}\beta + \alpha_i)$ that gives the sample proportion given $y_{it-1} = 0$, i.e. it is evaluated at $\Phi^{-1}(\overline{p}_0)$. The predicted probability given $y_{it-1} = 1$ is then given again by equation (2) and the ratio of conditional probabilities, R, by equation (3). For the sample on which the model is estimated

[15] A $\chi^2(13)$ Wald test statistic of 117.47, giving a p-value <0.0001.

the ratio of raw conditional probabilities indicates that those unemployed at $t-1$ are 21 times as likely to be unemployed at t as those employed at $t-1$. Controlling for the specified observed heterogeneity in the pooled probit model reduces this to 19 times as likely. Controlling for initial conditions and unobserved, as well as observed, heterogeneity cuts this measure of persistence by about a half. An individual with a given set of characteristics (observed and unobserved) is about 9 times as likely to be unemployed at t if they had been unemployed at $t-1$ as if they had been employed at $t-1$.

The estimate of ρ is significantly greater than zero and indicates that the correlation between the composite error term in the latent equation in any two (different) periods for the same individual is about 0.4. Equivalently this means that roughly 40% of the error variance is due to the individual-specific effects. The hypothesis that $\theta = 0$, which corresponds to exogeneity of the initial condition is strongly rejected (an asymptotic t-ratio of 6.65). Rather the estimate of θ is numerically close to, and insignificantly different from, 1 (an asymptotic t-ratio of 0.32). The impact of these individual effects in the linearised reduced form for the latent variable for the initial time period is not significantly different from the impact in the structural form for periods 2–6.[16]

The ratio of raw probabilities in the sample used here is about a quarter higher than that in the sample in the first column of Table 2. This is primarily due to the exclusion of those who were out of the labour force at an earlier date in the sample period. 9% of those in the sample in Table 2 (waves 2–6) were not in the labour force in wave 1. Excluding them raises the proportion of those unemployed at $t-1$ who were unemployed at t and lowers the proportion of those employed at $t-1$ who were unemployed at t, increasing the ratio from 17.0 to 18.5. A further 7% of the observations in Table 2 are for people who were out of the labour force at one of the waves between 2 and t. Excluding them as well further raises the proportion of those unemployed at $t-1$ who were unemployed at t and further lowers the proportion of those employed at $t-1$ who were unemployed at t, increasing the ratio to 20.8.[17]

Strictly the model must be estimated without prior gaps in the panel to give the correct conditioning sequence. However to give an impression of the magnitude of any selection bias from excluding those out of the labour force at an earlier wave, column 3 of Table 3 adds the second group identified above to the sample. (Being in the labour force at wave 1 (so that the initial condition is defined) is retained as a requirement for inclusion in the sample.) This results in only very slight changes in parameter estimates and the ratio of predicted probabilities.

4.2 Continuing spells vs. repeat unemployment

Those who are unemployed at $t-1$ and again at t consist of two rather different groups. First there are those for whom the two points in time are part of a continuing spell

[16] Under the null hypothesis that $\theta = 1$ the correlation between the period 1 reduced form error and the error terms in periods 2–6 is the same as that between any two periods from 2 onwards (as given in section 3 above). The implied correlation between v_1 and v_t for $t \geq 2$ is given by $\theta\rho/\sqrt{\theta^2\rho + 1 - \rho}$. This model gives a value for this correlation of 0.40, as compared with 0.39 within periods 2–6.

[17] A further 0.5% are excluded due to having missing values on one of the instruments.

without employment (possibly a single unemployment spell or possibly containing a period out of the labour force). Second there are those who have an intervening spell of employment (or possibly more than one), but then are unemployed again at t. This second category is what might be labelled *repeat unemployment*.[18]

The implications of these two types of unemployment "persistence" are somewhat different. In the case of continuing spells, $R > 1$ may simply reflect the fact that employment spells tend to last much longer than unemployment spells. To illustrate this consider the following simplified situation. Denote the instantaneous conditional probability of leaving unemployment (the unemployment hazard rate) by λ_u and that for leaving employment by λ_e and suppose that there is no duration dependence in either. Suppose further that there is at most one change of state between the two dates being considered (which, for example, rules out repeat unemployment). Thus someone unemployed at the start of the period either remains unemployed throughout or gets a job at some point and then remains employed. Similarly someone employed at the start of the period either remains employed throughout or loses their job at some point and then remains unemployed.

In this case the conditional probabilities of being unemployed at t given unemployed or employed at $t-1$ are

$$\begin{aligned} P(y_{it} = 1 | y_{it-1} = 1) &= \exp(-\lambda_u) \\ P(y_{it} = 1 | y_{it-1} = 0) &= \frac{\lambda_e}{\lambda_u - \lambda_e} \left[\exp(-\lambda_e) - \exp(-\lambda_u)\right] \end{aligned} \quad (18)$$

The ratio of these two conditional probabilities, R, depends just on the two hazard rates and hence on the mean completed durations of employment and unemployment. To give a very simple numerical illustration, a mean completed employment duration of 20 years and a mean completed unemployment duration of 1 year would give a value of R of 12.0. Note that this value of R results despite there being no linkage at all between spells, i.e. no impact of past history, and no duration dependence within spells.

Thus the implications of continuing spells and repeat unemployment are very different. However incorporating this distinction into the probability model used above is not straightforward. One cannot for example simply partition the y_{it-1} regressor. This would result in a "perfect classifier": $y_{it-1} = 1$ and a continuing spell implies $y_{it} = 1$. One way to analyse this is to proceed as follows. Define a dummy variable $d_{it} = 1$ if individual i spends any time in employment between $t-1$ and t, $d_{it} = 0$ if not. Then

$$\begin{aligned} P[y_{it} = 1 | y_{it-1}] &= P[y_{it} = 1 | y_{it-1} = 0 \text{ or } d_{it} = 1] \left\{1 - P[d_{it} = 0 | y_{it-1} = 1]\right\} \\ &\quad + P[d_{it} = 0 | y_{it-1} = 1] \end{aligned} \quad (19)$$

Thus the unemployment probabilities looked at in the previous subsection can be decomposed by looking at models for $P[y_{it} = 1 | y_{it-1} = 0 \text{ or } d_{it} = 1]$ and

[18] *Repeat unemployment* as defined here requires a period of employment (however short) between the two interview dates. Many of those unemployed at both $t-1$ and t have an intervening period out of the labour force. Sweeney (1996) reports that a third of those leaving claimant unemployment do not enter employment. Those with a period of unemployment that includes $t-1$, followed by a period out of the labour force, followed by a period of unemployment that includes t are classified here as having a continuing spell. This contrasts with Arulampalam et al. (2000), who treat this group as having separate spells.

$P[d_{it} = 0|y_{it-1} = 1]$. Columns 4–6 of Table 3 give the results from estimating models corresponding to those in Columns 1–3 excluding continuous unemployment spells.[19]

Considering first the model without random effects, excluding those continuously unemployed, i.e. those for whom the unemployment at $t-1$ and t are part of the same spell, reduces R from 19.3 to 6.2. The probit coefficient on y_{it-1} is cut by roughly half but remains highly significant. Although the ratio is reduced, those unemployed at $t-1$ who find a job are still more than 6 times as likely to be unemployed again at t as those who were employed at $t-1$.

Turning to the model with unobserved heterogeneity, excluding those continuously unemployed (i.e. without employment) between $t-1$ and t cuts the adjusted estimate of γ by almost two-thirds, although again it remains significantly greater than zero.[20] The ratio of predicted conditional probabilities, R, is reduced from 8.9 to 2.3. Someone unemployed at $t-1$ who finds a job is still 2.3 times as likely to be repeat unemployed at t as someone who was employed at $t-1$, but otherwise has the same observed and unobserved (but fixed) characteristics. Excluding continuing spells and allowing for the initial conditions considerably reduces the degree of persistence exhibited, but it remains significant.

4.3 GMM estimates of a dynamic LPM

The estimated effects in the dynamic random effects probit model used here are potentially sensitive to the auxiliary distributional assumption for the individual-specific effects. To investigate this issue GMM estimates of a dynamic LPM as described in Section 3.2 are also presented. The starting sample used is as in Column 2 of Table 3. Column 1 of Table 4 gives the OLS estimates (with robust standard errors) of the LPM for the probability of being unemployed at t. The control variables are as in Table 3. The results for the LPM are, as expected, very similar to those for the probit model on the same sample (once put on a comparable basis). The estimated value of R is 20.6.

Column 2 gives the Arellano-Bond GMM estimator using only lagged unemployment variables as instruments.[21] Column 3 gives the corresponding estimates using as additional instruments the variables used in the Dynamic Random Effects probit

[19] Arulampalam et al. (2000) also provide estimates for a variant of their model with continuing spells excluded, but define continuation as an elapsed duration exceeding the length of time between interviews. Thus those who leave the labour force and then return to unemployment are retained in their sample. To identify "repeat unemployment", unemployment at $t-1$ and t are defined here to be part of a continuing spell if there was no period of employment between $t-1$ and t.

[20] Scaling by $\sqrt{1-\hat{\rho}}$ gives an adjusted estimate of γ of 0.317, compared with an adjusted estimate for Column 2 of 0.920. This sharp reduction contrasts with Arulampalam et al. (2000), who retain those in a different unemployment spell but without any intervening employment and find a much smaller fall. Corcoran and Hill (1980) find this data "overlap", as they term it, to be an important contributory factor to the aggregate persistence in the US. Once it and unobserved heterogeneity are taken account of, their estimate of the equivalent of γ is insignificantly different from zero.

[21] The 1-step estimates are presented. Simulation evidence suggests the 2-step estimates can perform poorly in hypothesis testing, particularly when the ω_{it} are heteroskedastic. Doornik et al. (1999) advise that "inference based on asymptotic standard errors for the one-step estimators seems to be more reliable". For all the GMM columns in Table 4 the 2-step estimates and their standard errors are very similar to the 1-step estimates presented.

model in Table 3 and described above. Both models pass the Arellano-Bond test on the second order residual correlation coefficient and the Sargan test of over-identifying restrictions.[22] The estimates in Column 3 imply an estimated value of R of 11.5. This is slightly below half the value from the OLS estimation of the LPM on the same sample ($t \geq 3$), a very similar proportional reduction to that with the dynamic random effects probit model.

Although the estimate of λ is not very large, alleviating worries about the weak-instruments problem, results for the Blundell and Bond system GMM estimator are given in Column 4 for comparative purposes. The estimates are fairly similar to those in Column 2. Again the model passes the AR(2) and Sargan tests.[23] In this case however the Sargan test rejects the null hypothesis when the additional z-vector instruments are added (results not shown), although the coefficient estimates are not much altered.

The final two columns of the table give the results from the two Anderson-Hsiao IV estimators described in Section 3.2 for comparison. The estimates of λ are fairly similar to those in Column 2. However the AR(2) test rejects the null when the lagged difference is used as instrument. Since a single instrument is used, the model is just identified in both cases. Interestingly when the additional z-vector instruments are added (results not shown), the Sargan test of the over-identifying restrictions is failed in both cases. There are thus some doubts about the appropriateness of these two IV estimators in the current context. Overall the evidence would seem to support the use of the Arellano-Bond GMM estimator.

GMM estimation of a dynamic LPM produces estimates and evaluations of predicted probabilities fairly similar to those from the dynamic random effects probit model in Column 2 of Table 3. The results in this case do not seem to be driven by the auxiliary distributional assumption on the random effects.

The same conclusion appears to hold for the other columns of Table 3. Adding those with labour force gaps to the sample as in Column 3 of Table 3 gives GMM estimates which increase R to 12.1, a similar change to that based on the random effects probit estimates. Excluding those continuously unemployed between $t-1$ and t as in columns 4 and 5 of Table 3 produces OLS estimates that give $\hat{p}_1 = 0.140$ and $R = 6.8$ and a GMM estimates that give $\hat{p}_1 = 0.040$ and $R = 2.2$, again very similar to those based on the random effects probit estimates.

[22]The assumption of no serial correlation in the ω_{it} is essential for consistency of the estimators used. Tests for first and second order serial correlation in the first-differenced residuals are reported. If the ω_{it} are not serially correlated, there should be evidence of significant negative first order serial correlation in the differenced residuals ($\hat{\omega}_{it} - \hat{\omega}_{it-1}$) and no evidence of second order serial correlation in the differenced residuals. The test statistics presented are asymptotically N(0,1) under the null hypothesis of no autocorrelation. The Sargan instrument-validity test presented is based on the 2-step GMM estimator since only this version of the test is heteroskedasticity-consistent. (See Arellano and Bond (1991) on both procedures.)

[23]The validity of the extra instruments in the Blundell-Bond system estimator can be tested using the difference between the Sargan test statistic for the system estimator and that for the corresponding first-differenced GMM estimator. This gives a $\chi^2(4)$-statistic of 5.37 and a p-value of 0.25. The validity of the extra instruments is supported here.

4.4 The impact of low pay

To examine the first component of the "low pay – no pay cycle", Table 5 adds an indicator variable for being low paid at $t-1$ to the models presented in Table 3. In Table 3 the base group for status at $t-1$ is the employed. In this table those self-employed at $t-1$ are excluded and employees at $t-1$ are partitioned into those low paid and those paid above the threshold, with the latter group becoming the base group. The x- and z-variables are exactly as used in the models presented in Table 3. Low pay at $t-1$ is treated as endogenous, i.e. correlated with α, and instrumented in the same way as unemployment at $t-1$.

Looking first at the results in Columns 1 and 2, the introduction of unobserved heterogeneity and allowance for initial conditions reduces the ratio of predicted probabilities, R, for those unemployed at $t-1$ relative to those employed but low paid at $t-1$, evaluated as in Section 4.1, by a half (from 11.7 to 5.7).

The corresponding ratio for those low paid at $t-1$ relative to those higher paid at $t-1$ remains unchanged. An individual with a given set of characteristics (observed and unobserved) is about 6 times as likely to be unemployed at t if they had been unemployed at $t-1$ as if they had been a low paid employee at $t-1$ and about twice as likely in this latter case as if they had been a higher paid employee at $t-1$.

The estimates of ρ and θ are fairly similar to those in Table 3, exogeneity of the initial condition ($\theta = 0$) is strongly rejected (asymptotic t-ratio of 5.65) and the hypothesis that $\theta = 1$ is not rejected (asymptotic t-ratio of 0.57). Also as in Table 3, the informal examination of the selection bias from excluding those out of the labour force at an earlier wave (Table 5, Column 3) suggests this is relatively minor.

Columns 4–6 of Table 5 give the corresponding results after excluding continuous unemployment spells. The ratio of predicted conditional probabilities, R, unemployed relative to low paid, is reduced from 5.7 to 1.7. Someone unemployed at $t-1$ who finds a job is 1.7 times as likely to be repeat unemployed at t as someone who was an employee on low pay at $t-1$, but otherwise has the same observed and unobserved (but fixed) characteristics. This latter person is in turn 1.7 times as likely as someone who was higher paid. In terms of predicting subsequent unemployment, low paid employment holds a position roughly half way between previous (but not continuous) unemployment and higher paid employment.[24]

However the coefficients on the indicator variables for being unemployed at $t-1$ and being low paid at $t-1$ are not significantly different from one another at conventional significance levels. The Wald test of their equality gives a $\chi^2(1)$ statistic of 1.97, implying a p-value of 0.16.[25] Although someone unemployed at $t-1$ who finds a job is almost twice as likely to be unemployed again at t as someone low paid at $t-1$, the difference is not statistically significant. One cannot reject the hypothesis that the adverse effects of being unemployed at $t-1$ and of being low paid at $t-1$ on the probability of being unemployed at t (excluding those continuously unemployed)

[24] There is some evidence in Column 6 of a downward bias in the coefficient on low paid at $t-1$ from excluding those out of the labour force at an earlier wave. In the model in Column 6 low paid at $t-1$ is much closer to unemployed at $t-1$ than it is to higher paid at $t-1$.

[25] This test statistic halves (and the p-value doubles) in the final column of the table where those out of the labour force at an earlier wave are added to the sample.

are equal to one another.

Corresponding GMM estimates of the dynamic LPM are given in Table 6. In general the GMM estimates give rise to increased predicted probabilities given either low pay at $t-1$ or unemployment at $t-1$. For example in Column 2 the predicted probability given low pay at $t-1$ is double that from the random effects probit model (Table 5, Column 2) and that given unemployment at $t-1$ is up 50% on Table 5. This results in the ratio of predicted probabilities for low pay relative to higher pay more than doubling and the ratio for unemployment relative to low pay falling by about a third.

In Column 5 after those continuously unemployed between $t-1$ and t have been excluded, again the predicted probabilities given low pay at $t-1$ and unemployment at $t-1$ are both higher than in the corresponding column of Table 5, particularly in the case of low pay. The predicted probability given low pay at $t-1$ has moved closer to that given unemployment at $t-1$. On the basis of the GMM estimates, someone unemployed at $t-1$ who finds a job is 1.3 times as likely to be repeat unemployed at t as someone who was an employee on low pay at $t-1$, but otherwise has the same observed and unobserved (but fixed) characteristics. This latter person is in turn 3.3 times as likely as someone who was higher paid. In terms of predicting subsequent unemployment, the results of Table 5 indicate that low paid employment holds a position roughly half way between previous (but not continuous) unemployment and higher paid employment. The GMM estimates in Table 6 shift this position to about two-thirds of the way towards unemployment.

Another potential explanation for the results in this subsection on the differences in the conditional probability of unemployment at t between those low paid and those higher paid at $t-1$ needs to be considered. This difference could reflect a difference in elapsed job duration at time $t-1$ if the low paid typically have shorter elapsed durations than the higher paid and if the probability of job loss is greater for those with shorter durations.

This competing hypothesis is tested by adding to each of the models in Table 5 a variable measuring, for those employed at $t-1$, the length of time they have been in their jobs as at time $t-1$. This variable is treated as endogenous, i.e. correlated with α, and instrumented in the same way as unemployed at $t-1$ and low paid at $t-1$. It has a significant negative effect on the probability of being unemployed at t, but its inclusion alters the coefficients on unemployment at $t-1$ and low paid at $t-1$ very little. In Column 5 the coefficient on unemployed at $t-1$ falls to .403 (.146) and that on low paid at $t-1$ falls to .195 (.106). The corresponding predicted probabilities also fall only very slightly, reducing both ratios, unemployed to low paid and low paid to higher paid, to 1.6. Similarly in Column 6 these coefficients fall to .414 (.109) and .314 (.088) and the ratios to 1.2 and 2.1. Thus differences in elapsed job duration are not responsible for the low pay effects reported in Table 5.

4.5 Low pay as a conduit to repeat unemployment

For those who experience repeat unemployment in the sense examined above, the BHPS unfortunately does not provide information on the hourly rate of pay of the job(s) held between the unemployment at $t-1$ and again at t. An alternative way to investigate this involves looking at the persistence over a 2-year rather than a 1-year period. In this way, for those unemployed at $t-2$ and employed at $t-1$ the impact of their pay level at $t-1$ on their probability of repeat unemployment at t can be estimated.[26]

Estimates of models for the probability of being low paid at t based on alternate waves (specifically waves 1, 3 and 5) are given in Table 7. Columns 1 and 2 give estimates for the equivalent specification to Table 3: y_{it} determined by x_{it} and y_{it-2}. Column 2 gives the dynamic random effects probit estimates and Column 1 gives the pooled probit estimates on the same sample. As might be expected, the degree of measured persistence falls considerably when considered over two years rather than one. The estimates in Column 2 indicate that someone with a given set of characteristics (observed and unobserved) is 4.5 times as likely to be unemployed at t if they had been unemployed at $t-2$ as if they had been employed at $t-2$.

Columns 3 and 4 of Table 7 then partition this effect according to status at $t-1$. The two categories reported are for those low paid at $t-1$ and for those higher paid at $t-1$. Indicator variables are also included for those self-employed or with missing earnings information and for those unemployed or out of the labour force at $t-1$. This latter group will include any continuing spells (as well as others) and so remove them from the base group. A further indicator variable is also included for those employed at $t-2$ and unemployed at $t-1$.[27] This means that the base group is those employed at both $t-2$ and $t-1$ to give the required comparison. The comparison is thus of those unemployed at $t-2$ and employed at $t-1$ with those employed at $t-2$ and $t-1$.

There is a significant positive effect on the probability of unemployment at t of being low paid at $t-1$ after unemployment at $t-2$, but an insignificant (negative but numerically small) effect of being higher paid at $t-1$. An individual unemployed at $t-2$ and employed on low pay at $t-1$ is 3.8 times as likely to be unemployed at t as someone with the same set of characteristics (observed and unobserved), but employed at $t-2$ and $t-1$. In contrast someone employed on higher pay at $t-1$ after being unemployed at $t-2$ has the same probability of being unemployed at t as someone employed at $t-2$ and $t-1$.

As in the previous subsection, the possibility that this difference between low and higher pay reflects a difference in elapsed job duration should be addressed. Those higher paid at $t-1$ may have got their jobs earlier subsequent to the unemployment at $t-2$ than those whose job at $t-1$ is low paid. This hypothesis is tested by adding to the models reported in Columns 3 and 4 of Table 7 a variable measuring, for those unemployed at $t-2$ and employed at $t-1$, the length of time they have been in their jobs as at time $t-1$ (and treating it as endogenous). Its effect is insignificant in both cases (t-statistics of 0.78 and -0.29 respectively) and its inclusion has little effect on the coefficients on low and higher paid at $t-1$.[28] The corresponding ratios

[26]This also provides an alternative approach to the problem of continuing spells, the proportion of which is much reduced when the length between observation points is increased from one year to two.

[27]All the indicator variables listed in this paragraph are treated as endogenous and instrumented in the same way as y_{t-2}.

[28]In the model of Column 3 the coefficient on low paid at $t-1$ is 1.080 (.290) and that on higher paid is .455 (.238). In the model of Column 4 the coefficient on low paid at $t-1$ is .594 (.381) and

of predicted probabilities also change relatively little. In Column 3 to 8.3 and 2.8, in Column 4 to 3.5 and 1.1. Thus differences in elapsed job durations at $t-1$ are not responsible for these effects.

There is an increased probability of being unemployed again at t having been so at $t-2$ if the intervening point at $t-1$ was one of low paid employment, but not if it was one of higher paid employment.[29] Low paid jobs act as a conduit to repeat unemployment. Higher paid jobs eliminate the increased risk.

4.6 The probability of being low paid

To complete the picture estimates of models for the probability of being low paid are given in Table 8. Estimation of the dynamic random effects probit model (Column 2) cuts the ratio of predicted probabilities, R, by three-quarters. Without controlling for unobserved heterogeneity or initial conditions those low paid at $t-1$ are 17 times as likely to be low paid at t as those higher paid at $t-1$. However an individual with a given set of characteristics (observed and unobserved) is about 4 times as likely to be low paid at t if they had been low paid at $t-1$ as if they had been higher paid at $t-1$.

The estimate of ρ is significantly greater than zero and indicates that the correlation between the composite error term in the latent equation in any two (different) periods for the same individual is about 0.6, i.e. that 60% of the error variance is due to the individual-specific effects. The estimate of θ is (as for unemployment) significantly different from zero (t-statistic of 7.22) and insignificantly different from 1 (t-statistic of 0.97).

The ratio of raw probabilities in the sample used here is somewhat higher than that in the sample in Column 3 of Table 2. This is primarily due to the exclusion of those who were not employees at one (or more) of the earlier interview dates. Strictly the model must be estimated without prior gaps in the panel, as pointed out above. However, to give an impression of the magnitude of any selection bias from excluding those not employees at an earlier wave, column 3 adds those not employees at one of the waves between 2 and t. (The requirement of being an employee at wave 1 has to be retained.) Like in the case of unemployment, this raises the ratio of predicted probabilities only slightly.

that on higher paid is .046 (.329).

[29] The equivalent GMM estimation of the dynamic LPM, since T=3, involves estimation of the differenced equation (16) on a single cross-section and with a single instrument (it is equivalent to the Anderson-Hsiao IV estimator with y_{it-2} as instrument). There is thus no possibility to test the essential serial correlation restriction. A better approach involves GMM estimation on waves 4–6 of the differenced equation with suitable interactions between y_{t-2}, y_{t-1} and low pay status at $t-1$ to match the construction above. This gives a coefficient on "unemployed at $t-2$, higher paid at $t-1$" of 0.019 with a standard error of 0.024. Thus again insignificantly different from zero (an asymptotic t–statistic of 0.77). The model passes both the serial correlation and Sargan tests. The predicted probabilities are 0.017 for the base group, 0.035 for unemployed at $t-2$ and then higher paid at $t-1$ (implying $R=2.1$) and 0.094 for unemployed at $t-2$ and then low paid at $t-1$ (implying $R = 5.6$). The predicted probability for those unemployed at $t-2$ and then higher paid at $t-1$ from applying OLS to the same sample (waves 4–6) is 0.093. Thus GMM estimation on waves 4–6 gives a very similar proportional reduction in $\hat{p}_1$ to Dynamic Random Effects Probit estimation on waves 1,3,5 (Columns 3 and 4 of Table 7).

Columns 4 and 5 of Table 8 examine the second component of the "low pay – no pay cycle" by adding those unemployed at $t-1$ to the sample and incorporating an indicator variable for this group in the model. The probability being modelled is that of being low paid (given an employee). Hence those unemployed at t are excluded from the sample used for estimation. Thus in the sample of individuals without prior gaps those unemployed at $t-1$ would be excluded and the effect sought here would not be estimable. Hence this model must be estimated on the sample that includes those with prior gaps, i.e. those who were unemployed rather than an employee at one (or more) of the waves between 2 and t. This means that the estimates in Column 5 should be compared with those in Column 3. The (adjusted) coefficient on low pay at $t-1$ and the ratio of predicted probabilities (low paid at $t-1$ to higher paid at $t-1$) are very similar to those in column 3, which excludes these extra observations (as are those for the pooled probit without the random effects).

The coefficient on unemployed at $t-1$ and the ratio of predicted probabilities (unemployed at $t-1$ to higher paid at $t-1$) are slightly higher than those for low pay at $t-1$. This represents a turn round in the relative importance of unemployment and low pay at $t-1$ compared with the pooled probit model without random effects in Column 4. However the coefficients on low pay at $t-1$ and unemployment at $t-1$ are not significantly different from one another: a $\chi^2(1)$ statistic of 0.45, giving a p-value of 0.50. As far as the probability of being low paid at t is concerned someone with a given set of characteristics (observed and unobserved) would be in the same position if they were low paid at $t-1$ and if they were unemployed at $t-1$. One cannot reject the hypothesis that the adverse effects of being unemployed at $t-1$ and of being low paid at $t-1$ on the probability of being low paid at t (given employment) are equal to one another.[30]

5 Conclusions

This paper examines the extent of state dependence in individual unemployment and low paid employment and the inter-related dynamics between the two. The main findings of the paper are as follows:

- The aggregate persistence in unemployment considerably overstates an individual's risk of repeat unemployment. Over half the measured persistence results from continuing unemployment spells (in the sense of there being no intervening employment) and about a third is removed when unobserved heterogeneity and initial conditions are taken account of. Despite this, an individual unemployed at $t-1$ who finds a job is still 2.3 times as likely to be unemployed again at t as someone who was employed at $t-1$, but otherwise has the same observed and unobserved characteristics; and this difference is statistically significant.
- Low pay at $t-1$ has almost as large an adverse effect as unemployment at $t-1$ on both the probability of employment at t and the probability of being paid

[30] A caveat to the results of this subsection is that they ignore any selection bias from the necessary conditioning on employment when modelling the probability of being low paid, i.e. they ignore the possible correlation between the error terms in the latent equations for unemployment and low pay.

above the threshold if employed. The difference between the effects of past low pay and past unemployment is statistically insignificant in both cases.

- Low paid jobs act as the main conduit for repeat unemployment. Those who manage to get a better job eliminate their increased risk of repeat unemployment. The probability of re-entering unemployment for someone who gets a low paid job after a spell of unemployment is about four times greater than for someone with the same observed and unobserved characteristics originally in employment. The equivalent person who gets a higher paid job after the spell of unemployment has the same probability of becoming unemployed again as the person originally in employment.

In terms of future prospects (the probability of employment and the probability of being paid above the threshold if employed), low paid jobs are closer to unemployment than to higher paid jobs. The results in this paper cast doubt on the belief that all jobs are “good” jobs in the sense of improving future prospects. The results suggest that low paid jobs typically do not lead on to better things.

The policy objective, rather than simply being to get an unemployed individual into any job, should be to get him or her into a “good” job, one with improved future prospects. Typically low paid jobs are not “good” jobs in this sense. The results of the paper are consistent with the hypothesis that a low paid job does not augment a person’s human capital significantly more than unemployment. If unemployed individuals’ employment prospects are to be permanently improved, they need to gain access to jobs where they can augment their skills, raise their productivity and move up the pay ladder.

References

Acemoglu, Daron (1997), "Good Jobs versus Bad Jobs: Theory and Some Evidence", CEPR Discussion Paper 1588.

Ahn, Seung C. and Peter Schmidt (1995), "Efficient estimation of models for dynamic panel data", *Journal of Econometrics*, **68**, 5–27.

Anderson, T.W. and Cheng Hsiao (1981), "Estimation of dynamic models with error components", *Journal of the American Statistical Association*, **76**, 598–606.

Arellano, Manuel and Stephen Bond (1991), "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *Review of Economic Studies*, **58**, 277–97.

Arellano, Manuel and Olympia Bover (1995), "Another look at the instrumental-variable estimation of error-components models", *Journal of Econometrics*, **68**, 29–52.

Arulampalam, Wiji (1999), "A note on estimated coefficients in random effects probit models", *Oxford Bulletin of Economics and Statistics*, **61**, 597–602.

Arulampalam, Wiji, Alison L. Booth and Mark P. Taylor (2000), "Unemployment persistence", *Oxford Economic Papers*, **52**, 24–50.

Atkinson, A.B. F. Bourguignon and C. Morrisson (1992), *Empirical Studies of Earnings Mobility*, Chur: Harwood Academic Publishers.

Baker, Michael (1997), "Growth-rate heterogeneity and the covariance structure of life-cycle earnings", *Journal of Labor Economics*, **15**, 338–75.

Blanchard, Olivier J. and Peter A. Diamond, (1994), "Ranking, unemployment duration, and wages" *Review of Economic Studies*, **61**, 417–34.

Blundell, Richard and Stephen Bond (1998), "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics*, **87**, 115–43.

Burtless, Gary (ed.) (1990), *A Future of Lousy Jobs? The Changing Structure of U.S. Wages*, Washington: Brookings Institute.

Butler, J.S. (1996), "Estimating the correlation in censored probit models", *Review of Economics and Statistics*, **78**, 356–358.

Butler, J.S. and Robert Moffitt, (1982), "A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model", *Econometrica*, **50**, 761–64.

Chamberlain, Gary (1984), "Panel Data", in *Handbook of Econometrics*, ed. by Z. Griliches and M. Intrilligator, Oxford and Tokyo: North-Holland.

Chay, Kenneth Y. and Dean Hyslop (1998), "Identification and Estimation of Dynamic Binary Response Panel Data Models: Empirical Evidence Using Alternative Approaches", CLE, U. of California Berkeley, Working Paper No. 5.

Clark, Kim B. and Lawrence H. Summers (1979), "Labour market dynamics and unemployment: a reconsideration", *Brookings Papers on Economic Activity*, **1**, 13–60.

Corcoran, Mary and Martha S. Hill (1980), "Reoccurrence of Unemployment Among Adult Men", *Journal of Human Resources*, **20**, 165–183.

Dickens, Richard (2000), "The evolution of individual male earnings in Great Britain: 1975–94", *Economic Journal*, **110**, 27–49.

Doornik, Jurgen A., Manuel Arellano and Stephen Bond (1999), "Panel data estimation using DPD for Ox", http://www.nuff.ox.ac.uk/users/doornik/.

Gregg, Paul, Genevieve Knight and Jonathan Wadsworth (1999), "The cost of job loss", in *The State of Working Britain*, ed. by Paul Gregg and Jonathan Wadsworth, Manchester University Press.

Gregg, Paul and Marco Leonardi (1999), "Wage and employment dynamics in the UK 1978–1995", CEP LSE Working Paper No. 1019.

Gregg, Paul and Jonathan Wadsworth (1996), "Mind the gap, please? The changing nature of entry jobs in Britain", CEP LSE Discussion Paper No. 303.

Gregory, Mary and Robert Jukes (1997), "The effects of unemployment on subsequent earnings: A study of British men 1984-94", Oxford Discussion Paper.

Heckman James J. (1981), "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*, Cambridge: MIT Press.

Heckman, James J. (1991), "Identifying the hand of the past: Distinguishing state dependence from heterogeneity", *American Economic Review*, **81**, 75–79.

Heckman, James J. and George J. Borjas (1980), "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence", *Economica*, **47**, 247–83.

Hyslop, Dean (1995), "State Dependence, Serial Correlation and Heterogeneity in Intertemporal Participation Behavior: Monte Carlo Evidence and Empirical results for Married Women", IRS, Princeton University, Working Paper 347.

Jacobson, Louis S., Robert J. LaLonde and Daniel G. Sullivan (1993), "Earnings losses of displaced workers", *American Economic Review*, **83**, 685–709.

Kletzer, Lori G. (1998), "Job displacement", *Journal of Economic Perspectives*, **12**, 115–36.

Layard, Richard, Stephen Nickell and Richard Jackman (1991), *Unemployment: Macroeconomic Performance and the Labour Market*, Oxford: Oxford University Press.

Low Pay Commission (1998), *The National Minimum Wage*, First Report, Cm. 3796, London: The Stationary Office.

McCormick, Barry (1990), "A theory of signalling during job search, employment efficiency and "stigmatized" jobs", *Review of Economic Studies*, **57**, 299–313.

Moffitt, Robert and Peter Gottschalk (1993), "Trends in the covariance structure of earnings in the US: 1969–1987", Brown University Working Paper No. 9/93.

Mundlak, Yair (1978), "On the Pooling of Time Series and Cross Section Data", *Econometrica*, **46**, 69–85.

Narendranathan, Wiji and Peter Elias (1993), "Influences of Past History on the Incidence of Youth Unemployment: Empirical Findings for the UK", *Oxford Bulletin of Economics and Statistics*, **55**, 161–185.

Poirier, Dale J. (1980), "Partial observability in bivariate probit models", *Journal of Econometrics*, **12**, 210–217.

Stevens, Ann Huff (1997), "Persistent effects of job displacement: The importance of multiple job losses", *Journal of Labor Economics*, **15**, 165–88.

Stewart, Mark B. (1999a), "Low Pay in Britain", in *The State of Working Britain*, ed. by Paul Gregg and Jonathan Wadsworth, Manchester University Press.

Stewart, Mark B. (1999b), "Low Pay, No Pay Dynamics", in *Persistent Poverty and Lifetime Inequality*, HM Treasury Occasional Paper No. 10.

Stewart, Mark B. (1999c), "The Dynamics of Low Pay and Low Income", in *The National Minimum Wage, Incomes and the Low Paid*, Low Pay Commission Occasional Paper No. 2.

Stewart, Mark B. and Joanna K. Swaffield (1997), "The Dynamics of Low Pay in Britain", in P. Gregg (ed.), *Jobs, Wages and Poverty: Patterns of Persistence and Mobility in the Flexible Labour Market*, London: CEP.

Stewart, Mark B. and Joanna K. Swaffield (1999), "Low Pay Dynamics and Transition Probabilities", *Economica*, **66**, 23–42.

Sweeney, Kate (1996), "Destinations of Leavers from Claimant Unemployment", *Labour Market Trends*, October, 443–452.

Taubman, Paul and Michael L. Wachter (1986), "Segmented Labor Markets", in *Handbook of Labor Economics*, ed. by Orley Ashenfelter and Richard Layard, Oxford and Tokyo: North-Holland.

Teasdale, Paul (1998), "Incidence and Repeat Spells of Unemployment: An Analysis Using Claimant Data", *Labour Market Trends*, November, 555-562.

Taylor, Marcia Freed (ed.) with John Brice, Nick Buck and Elaine Prentice (1996) *British Household Panel Survey User Manual*, Colchester: University of Essex.

TABLE 1

Conditional Probabilities of Unemployment and Low Pay
(Pooled data 1992–1996)

	$P(U)_t$	$P(L)_t$
Unconditional	.069	.106
Out of labour force at $t-1$	.369	.406
In labour force at $t-1$	.053	.091
Unemployed at $t-1$	.457	.258
Employed at $t-1$	.027	.085
Ratio	16.8	3.0
Employee at $t-1$	.027	.084
Employee at $t-1$, non-missing earnings	.026	.083
Low paid at $t-1$	.059	.569
Higher paid at $t-1$	.022	.034
Ratio	2.7	16.8
Sample size	25,565	17,912

Notes:
1. Pooled data for BHPS waves 2 to 6.
2. Column 1 gives $P(U)_t$, the probability of being unemployed at t.
3. ILO/OECD definition of unemployment used.
4. Sample in column 1 restricted to those in the labour force (i.e. employed or unemployed) at t.
5. Column 2 gives $P(L)_t$, the probability of being low paid at t given an employee.
6. Low pay threshold is £3.50 in April 1997 terms (adjusted using the Average Earnings Index).
7. Sample in column 2 restricted to those who are employees with non-missing pay data at t.
8. Sample sizes given are for the unconditional probabilities.

TABLE 2

Probit Models for Unemployment and Low Pay Probabilities

(Pooled data 1992–1996)

	$P(U)_t$		$P(L)_t$	
	1	2	3	4
ML coefficient estimates				
Unemployed at $t-1$	1.723	1.806		1.144
	(.040)	(.042)		(.065)
Low paid at $t-1$		.395	1.765	1.759
		(.060)	(.042)	(.042)
Predicted probabilities:				
Unemployed at $t-1$	.420	.420		.247
Employed at $t-1$	.027			
Higher paid at $t-1$		.053	.475	.473
Low paid at $t-1$		.022	.034	.034
Ratio of predicted probabilities, R				
Unemployed to employed	15.5			
Unemployed to low paid		7.9		0.5
Low paid to higher paid		2.4	14.1	14.0
Ratio of raw probabilities on same sample				
Unemployed to employed	17.0			
Unemployed to low paid		8.3		0.5
Low paid to higher paid		2.5	16.9	16.9
Sample size	22,598	17,821	15,043	15,592

Notes:
1. Pooled data for waves 2 to 6.
2. Standard errors in brackets.
3. Columns 1 and 2 give results for the probability of unemployment at t (given in the labour force).
4. Columns 3 and 4 give results for the probability of low pay at t (given employee).
5. Unemployment and low pay definitions as in Table 1.
6. Each Probit model also contains years of education, whether individual has any qualifications, years of experience (quadratic), gender and marital status (and the interaction between them), health status, whether resident in London or the South East and the local unemployment-vacancy ratio.

TABLE 3

Dynamic Random Effects Probit Models for the Probability of Being Unemployed at t

(Panel data, t = 1992–1996)

	1	2	3	4	5	6
Unemployed at $t-1$	1.861	1.176	1.210	.900	.374	.352
	(.049)	(.093)	(.083)	(.082)	(.130)	(.109)
Random effects	No	Yes	Yes	No	Yes	Yes
ρ	0	.388	.355	0	.281	.252
		(.051)	(.047)		(.063)	(.056)
θ		1.051	1.066		.882	1.023
		(.158)	(.163)		(.250)	(.277)
Sample:						
Required to be in labour force in all waves?	Yes	Yes	No	Yes	Yes	No
Exclude those continuously unemployed between $t-1$ and t?	No	No	No	Yes	Yes	Yes
Predicted probabilities:						
$\overline{p}_0$	.023	.023	.024	.021	.021	.021
$\hat{p}_1$	.448	.207	.222	.127	.048	.047
Ratio, R	19.3	8.9	9.2	6.2	2.3	2.2
Predicted probabilities from simple probit on same sample:						
$\hat{p}_1$		.448	.435		.127	.111
Ratio, R		19.3	18.0		6.2	5.2
Sample size: waves 2–6	18,752	18,752	19,796	18,145	18,145	19,365
total	18,752	23,491	24,715	18,145	22,745	24,199

Notes:
1. Pooled data for waves 2 to 6.
2. Standard errors in brackets.
3. Unemployment and low pay definitions as in Table 2.
4. Other variables in models as in Table 2.
5. Specification of z-vector (including instruments) as described in text.

TABLE 4

GMM Estimation of Dynamic LPM for the Probability of Being Unemployed at t

(Panel data, t = 1992–1996)

	1	2	3	4	5	6
Unemployed at $t-1$	.455	.216	.214	.224	.208	.191
	(.020)	(.013)	(.028)	(.028)	(.014)	(.033)
Method of estimation	OLS	GMM	GMM	GMM	IV	IV
Instrument set		A	A,Z	B	y_{t-2}	Δy_{t-2}
Test statistics:						
AR(1) (N(0,1))	-3.95	-42.0	-12.0	-13.9	-44.6	-19.1
AR(2) (N(0,1))	3.78	0.90	0.47	0.58	0.78	2.42
Sargan ($\chi^2(d)$)		4.70	25.38	10.07		
(deg. of freedom.(d))		(9)	(22)	(13)		
[p–value]		[.86]	[.28]	[.69]		
Predicted probabilities:						
$\overline{p}_0$	.023	.020	.020	.021	.020	.019
$\hat{p}_1$	.478	.237	.235	.245	.228	.210
Ratio, R	20.6	11.6	11.5	11.6	11.1	11.1
Predicted probability from OLS on same sample:						
$\hat{p}_1$		.482	.482	.463	.482	.470
Ratio, R		23.5	23.5	21.8	23.5	25.0
Sample size	18,752	14,013	14,013	18,067	14,013	9,959

Notes:
1. Pooled data for waves 2 to 6.
2. Robust standard errors in all columns.
3. Unemployment definition as in Table 2.
4. Other variables in models as in Table 2.
5. Instrument sets: A = Arellano-Bond, B = Blundell-Bond, Z = instruments used in Random Effects Probit model in Table 3, as described in text.

TABLE 5

Dynamic Random Effects Probit Models for the Probability of Being Unemployed at t
(Panel data, t = 1992–1996)

	1	2	3	4	5	6
Unemployed at $t-1$	1.961	1.303	1.456	.995	.470	.470
	(.055)	(.113)	(.078)	(.088)	(.146)	(.109)
Low paid at $t-1$	.294	.301	.404	.232	.227	.340
	(.082)	(.102)	(.083)	(.089)	(.107)	(.088)
Random effects	No	Yes	Yes	No	Yes	Yes
ρ	0	.374	.271	0	.282	.240
		(.065)	(.049)		(.075)	(.061)
θ		1.113	1.263		.941	1.073
		(.197)	(.252)		(.305)	(.329)
Sample:						
Required to be in labour force in all waves?	Yes	Yes	No	Yes	Yes	No
Exclude those continuously unemployed between $t-1$ and t?	No	No	No	Yes	Yes	Yes
Predicted probabilities:						
Higher paid at $t-1$	.020	.020	.021	.019	.019	.019
Low paid at $t-1$	.040	.040	.051	.032	.032	.041
Unemployed at $t-1$	.465	.228	.279	.138	.053	.054
Ratio: Unemp to LP	11.7	5.7	5.5	4.3	1.7	1.3
Ratio: LP to HP	2.0	2.0	2.5	1.7	1.7	2.2
Test of equality of coefficients on Unemp and LP:						
$\chi^2(1)$ Wald statistic		48.38	96.86		1.97	0.95
[p–value]		[.00]	[.00]		[.16]	[.33]
Sample size: waves 2–6	13,506	13,506	15,723	13,016	13,016	15,309
total	13,506	17,229	19,994	13,016	16,607	19,494

Notes:
1. Pooled data for waves 2 to 6.
2. Standard errors in brackets.
3. Unemployment and low pay definitions as in Table 2.
4. Other variables in models as in Table 2.
5. Specification of z-vector (including instruments) as described in text.

TABLE 6

GMM Estimation of Dynamic LPM for the Probability of Being Unemployed at t
(Panel data, t = 1992–1996)

	1	2	3	4	5	6
Unemployed at $t-1$	.478	.315	.349	.134	.050	.064
	(.022)	(.041)	(.043)	(.022)	(.027)	(.030)
Low paid at $t-1$	.018	.074	.086	.013	.036	.045
	(.007)	(.014)	(.014)	(.006)	(.010)	(.010)
Method of estimation	OLS	GMM	GMM	OLS	GMM	GMM
Sample:						
Required to be in labour force in all waves?	Yes	Yes	No	Yes	Yes	No
Exclude those continuously unemployed between $t-1$ and t?	No	No	No	Yes	Yes	Yes
Test statistics:						
AR(1) (N(0,1))	-2.53	-9.80	-9.92	0.42	-7.96	-8.68
AR(2) (N(0,1))	3.37	1.11	1.09	1.96	1.34	1.17
Sargan ($\chi^2(d)$)		4.05	4.84		9.74	11.66
(deg. of freedom.(d))		(9)	(9)		(9)	(9)
[p–value]		[.91]	[.85]		[.37]	[.23]
Predicted probabilities:						
Higher paid at $t-1$	.020	.017	.018	.019	.016	.017
Low paid at $t-1$	.038	.092	.104	.031	.051	.061
Unemployed at $t-1$	.499	.333	.367	.153	.066	.081
Ratio: Unemp to LP	13.0	3.6	3.5	4.8	1.3	1.3
Ratio: LP to HP	1.9	5.3	5.7	1.7	3.3	3.7
Sample size	13,506	9,783	10,797	13,016	9,425	10,488

Notes:
1. Pooled data for waves 2 to 6.
2. Robust standard errors in all columns.
3. Unemployment definition as in Table 2.
4. Other variables in models as in Table 2.
5. GMM estimates with Arellano-Bond instrument set.

TABLE 7

Dynamic Random Effects Probit Models for the Probability of Being Unemployed at t
(Panel data for alternate waves only: 1991, 93, 95)

	1	2	3	4
Unemployed at $t-2$	1.442	.769		
	(.072)	(.214)		
Unemployed at $t-2$ and:				
low paid at $t-1$			1.083	.635
			(.269)	(.336)
higher paid at $t-1$			.479	-.018
			(.182)	(.272)
Random effects	No	Yes	No	Yes
ρ		.434		.212
		(.130)		(.129)
θ		.804		1.360
		(.212)		(.765)
Predicted probabilities:				
Employed at $t-2$ (base group)	.030	.030	.023	.023
Unemployed at $t-2$	.328	.132		
Unemployed at $t-2$ (then LP)			.180	.086
Unemployed at $t-2$ (then HP)			.064	.022
Ratios of predicted probabilities:				
Unemployed to employed	11.1	4.5		
Unemp (then LP) to employed			7.9	3.8
Unemp (then HP) to employed			2.8	1.0
Sample size: waves>1	7,626	7,626	7,626	7,626
total	7,626	11,804	7,626	11,804

Notes:
1. Data for alternate waves only: t = 1993, 95.
2. Standard errors in brackets.
3. Unemployment and low pay definitions as in Table 2.
4. Other variables in models as in Table 2.
5. Specification of z-vector (including instruments) as described in text.
6. Also included in Columns 3 and 4 (and treated as endogenous) are indicator variables for unemployed at $t-2$ interacted with self-employment or missing earnings at $t-1$ and interacted with unemployed at $t-1$, plus one for employed at $t-2$ and unemployed at $t-1$.

TABLE 8

Dynamic Random Effects Probit Models for the Probability of Being Low Paid at t
(Panel data, t = 1992–1996)

	1	2	3	4	5
Low paid at $t-1$	1.817	.645	.812	1.781	.823
	(.053)	(.089)	(.085)	(.050)	(.085)
Unemployed at $t-1$				.959	.948
				(.129)	(.170)
Random effects	No	Yes	Yes	No	Yes
ρ	0	.604	.539	0	.525
		(.033)	(.039)		(.039)
θ		1.155	1.241		1.287
		(.160)	(.186)		(.197)
Sample includes those not continuously employed?	No	No	Yes	Yes	Yes
Predicted probabilities:					
Higher paid at $t-1$	.027	.027	.029	.029	.029
Low paid at $t-1$	.458	.101	.141	.457	.143
Unemployed at $t-1$				.176	.173
Ratio: LP to HP	16.8	3.7	4.8	15.5	4.9
Ratio: Unemp to HP				6.0	5.9
Ratios from simple probits on same sample:					
LP to HP		16.8	15.6		15.5
Unemp to HP					6.0
Test of equality of coefficients on Unemp and LP:					
$\chi^2(1)$ Wald statistic				37.56	0.45
[p–value]				[.00]	[.50]
Sample size: waves 2–6	11,425	11,425	12,380	12,534	12,534
total	11,425	14,505	15,688	12,534	15,865

Notes:
1. See notes to Table 3.