Deutsches Institut für
Wirtschaftsforschung

DIW BERLIN

SOEP
The German
Socio-Economic
Panel Study

SOEPpapers
on Multidisciplinary Panel Data Research

207

Ken Clark • Nikolaos C. Kanellopoulos

Low Pay Persistence in European Countries

Berlin, June 2009

SOEPpapers on Multidisciplinary Panel Data Research
at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel Study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPpapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPpapers are available at
http://www.diw.de/soeppapers

Low Pay Persistence in European Countries

Ken Clark*
(University of Manchester and IZA, Bonn)

and

Nikolaos C. Kanellopoulos**
(Centre for Planning and Economic Research, Athens, Greece)

* Ken Clark, Economics, School of Social Sciences, Arthur Lewis Building, University of Manchester, Manchester M13 9PL, UK, ken.clark@manchester.ac.uk
** Nikolaos C. Kanellopoulos, Centre for Planning and Economic Research (KEPE), 11 Amerikis Str, 106 72 Athens, Greece, nkanel@kepe.gr

Abstract

Using panel data for twelve European countries over the period 1994-2001 we estimate the extent of state dependence in low pay. Controlling for observable and unobservable heterogeneity as well as the endogeneity of initial conditions we find positive, statistically significant state dependence in every single country. The magnitude of this effect varies by country, however this variation is not systematically related to labour market institutions.

JEL codes: C23, C25, J31, J69

Keywords: Low pay, low pay persistence, state dependence, initial conditions, dynamic random effects probit models.

1. Introduction

Low-wage employment has increasingly become a focus for policymakers in Europe. Against a background of widening earnings and income inequality since the late 1970s, it has been recognised that a substantial proportion of those in poverty are in employment but earning low wages. Bardone and Guio (2005), for example, suggest that in 2001, 7% of all workers in EU-15 countries were living in households classed as poor, representing a total of 11 million workers. The recognition of the working poor as a distinct group led the European Commission (2003) to include a measure of in-work poverty in its list of indicators of social exclusion and a number of policy responses including increases in minimum wages and social benefits targeted at workers have been proposed.

A key issue in the discussion of low-wage employment is whether having a low-paid job is a persistent or transitory state. Some workers will experience low pay as a random event at some point in their working lives, however for others low-paid employment can be a "trap" from which it is difficult to escape. This raises the question of how we interpret the observation that it is frequently the same workers who are observed in low-wage employment in successive time periods. That is, what explains *low pay persistence*? Does this imply something about those workers which makes them prone to low pay or does the very fact of having experienced low pay increase the likelihood of low pay in the future? In the language of the applied econometrics literature this reflects the distinction between heterogeneity and state dependence. Heterogeneity refers to observable and unobservable differences between workers which predispose them to low pay, while state dependence occurs if low pay in a particular time period has a directly causal effect on the probability of being low paid in the future. This occurs when there is a depreciation in human

capital or other earnings-enhancing characteristics while in a low-paid job. Known as "scarring", this is more commonly associated with unemployment spells however it may also apply to the experience of low-wage employment. Furthermore, employers may select against those who have a history of low-wage employment providing another route through which low pay in the past can cause low pay in the present. Depending on whether heterogeneity or state dependence is the correct explanation for low pay persistence, the policy implications could be quite different.

Previous econometric studies of low pay persistence have typically examined one or two countries and have sought to measure the extent to which workers are trapped in low pay as well as to establish the characteristics of those who are likely to be found persistently at the lower end of the pay distribution. Studies generally do find systematic differences in the probability of low pay depending on the personal and job characteristics of workers. For example, Sloane and Theodossiou (1998), using two waves of panel data from the UK, find that younger, better educated and married individuals, as well as those who work for a large firm have a greater chance of escaping from low-paid jobs. However, short tenure is found to increase the probability of remaining low paid. Cappellari (2000) argues that in Italy job characteristics such as occupation and firm size rather than individual characteristics are the most important factors in explaining low pay persistence. Likewise Asplund *et al*. (1998) examine low wage earners in Denmark and Finland and find that the effect of individual characteristics is not as important as the type of occupation. On the contrary Stewart and Swaffield (1998) find that in the UK women are more likely to be low paid than men but those working in large firms and those who are union members are less likely to remain or enter low pay. In common with a number of studies, they also find significant persistence in low pay.

4

In this paper, we extend the evidence base by providing estimates of the nature and extent of low pay persistence in a sample of twelve European countries using panel data from the period 1994-2001. As well as documenting the factors associated with low paid status we estimate how far low pay persistence reflects true state dependence. Measures of state dependence are then related to a range of institutional features of the European countries to attempt to explain the patterns observed.

The paper proceeds as follows. The next section describes the data we use, while section 3 discusses the econometric techniques employed. The estimates of low pay persistence are presented in section 4. We then discuss low pay persistence and labour market institutions in section 5 and conclude in section 6.

2.   Data and descriptives

We analyse low pay persistence using data from the European Community Household Panel (ECHP), the compilation of which was directed by Eurostat. The ECHP is a standardized, annual, longitudinal survey providing information about income, employment, housing, health, human capital and many other social indicators for a large and representative sample of households and individuals in the European Union. It is the only data set which allows us to investigate, on a consistent basis, the nature and extent of low pay persistence in European countries across this time period. When the ECHP started, Eurostat planned a minimum of 1000 households per country except for the four largest countries (Germany, UK, France and Italy) and the four poorest (Greece, Spain, Portugal and Ireland) where the sample would include 2000 households. The first wave in 1994 contained Belgium, Britain, Denmark, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal and Spain. In the second wave Austria was added, in the third Finland and in 1997

Sweden. In the current paper we use data for all these countries with the exception of Luxembourg where sample sizes are low, and Sweden and Finland, which joined the panel relatively late. Further details about the purpose and quality of ECHP data are available in Eurostat (2003a, 2003b, 2004).

We use data from all available waves of ECHP between 1994 and 2001, thus there are eight waves for each country except Austria for which there are seven. The analysis is restricted to males aged 18 to 65, who are full-time employees having left full-time education and who normally work more than 30 hours per week. Women are excluded to avoid issues relating to endogenous female labour supply and the self-employed are not included as their earnings are more difficult to model due to theoretical issues regarding their constitution. The unit of analysis is the individual worker and the key variable, which is used to construct the low pay indicator, is the net hourly wage after deducting any income taxes and social security contributions[1]. The use of net wages is intended to facilitate international comparisons by abstracting from differences between taxation regimes in different countries[2]. Hourly wages are constructed by dividing the previous month's net wages by hours of work.

Sloane and Theodossiou (1998) note that the lack of a consistent definition of low pay in previous studies of earnings mobility has prevented meaningful international comparisons being made. Some authors (e.g. Stewart and Swaffield, 1999) have responded to this by comparing different indicators of low pay within the same study. Given the large number of countries in our sample, such an approach would be unwieldy hence we have opted to use a single definition of low pay which

---

[1] The exception is France where the available information from Eurostat is on gross wages.

[2] There is no consensus, even between studies focusing on single countries, on what is the appropriate measure of earnings. For example, Cappellari (2000, 2002) uses net annual earnings for Italian workers while Stewart and Swaffield (1998) use gross hourly earnings for the UK.

we can measure consistently across the twelve countries. The choice is essentially between a fixed threshold or a measure which is relative to the distribution of earnings in each country. Given disparities in income across the countries studied, the latter approach is preferred and we classify a worker as being low paid if he earns less than two thirds of the mean hourly earnings in each country and year[3].

Table 1 presents some preliminary evidence on the distribution of low pay across European countries and its persistence over time. The probability of being low paid (column 1) varies between 0.065 (Belgium) and 0.194 (Portugal). As well as Portugal, more than 15% of male workers are estimated to be low paid in France, Greece, Ireland, Spain and the UK. However, these are raw probabilities, which do not take into account the status of the worker in the previous time period. Conditional probabilities of being low paid are presented in the next three columns. Column 2 shows the probabilities of remaining in low pay and these are much higher than the unconditional probabilities in column 1 as well as the probability of entering low pay, which is given in column 3. In particular the probability of a worker staying low paid between two successive years varies between 0.387 and 0.650, while for those who are not low paid the inflow probability is never greater than 0.071. Furthermore the ranking of the probabilities of staying in low pay is broadly similar to the ranking of the raw low pay probabilities – in other words, countries with relatively large numbers of workers in low pay are also those where it is hardest to get out of low pay. The last two columns formalise this idea by presenting two descriptive measures of low pay persistence. The first shows how much more likely (in terms of "probability points") those in low pay are to stay in low pay than those out of low pay are to enter low pay,

---

[3] We experimented with other measures of low pay based on the median of the earnings distribution. Full details are available on request.

while the last column shows how many times more likely it is for those in low pay to remain in low pay than for other workers to enter. Both measures demonstrate that workers who were low paid in period t are much more likely to be low paid in $t + 1$ than those who were not low paid in t. Thus, on the basis of the raw data there appears to be strong persistence in low pay across these twelve European countries. Finally, column 4 presents information about the outflow from low pay. This probability is also high but it is smaller for countries with high persistence in low pay[4]. All these measures suggest that there is strong persistence in low pay and the question is how much of it is due to workers' (either observed or unobserved) characteristics or stems from true state dependence.

< Table 1 near here >

To provide some preliminary evidence on how low pay status is distributed across workers, Table 2 shows the probability of being low paid in each country for workers with given characteristics. Unsurprisingly, better educated and older workers are less likely to be low paid. There is however less of a clear pattern regarding whether post-education training reduces the likelihood of being low paid. Agricultural and public sector workers are more likely to be low paid than workers in other sectors and occupations. One of the starkest contrasts in the Table 2 is between those on permanent and non-permanent contracts, with the latter having considerably higher rates of low pay.

< Table 2 near here >

3. Econometric analysis

---

[4] This is expected because $\Pr\left(LP_t = 1 \mid LP_{t-1} = 1\right) = 1 - \Pr\left(LP_t = 0 \mid LP_{t-1} = 1\right)$.

Identifying true state dependence, as opposed to heterogeneity, suggests a modelling approach which incorporates both unobservable and observable influences on low paid status. Since being low paid or not is a discrete dependent variable, the dynamic random effects probit framework represented by equation (1) below is the most appropriate[5].

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

The subscript $i = 1, 2, \ldots, N$ denotes individuals that are included in our sample and the subscript $t = 2, 3, \ldots, T$ represents the time periods for which the model is estimated. L_{it} is a dummy variable for being low paid which is equal to one when L_{it}^*, a latent measure of low-paid status, exceeds some threshold. The vector x_{it} contains (assumed) strictly exogenous explanatory variables. Obviously L_{it-1} is the low pay status of individual i in the previous year. The random error term in this model is composed of two terms. The individual-specific error terms ε_i, captures unobserved heterogeneity which differs between individuals but remains constant for each examined individual, while u_{it} is the "usual" error term with the properties that it is zero mean, uncorrelated with itself, uncorrelated with x_{it}, uncorrelated with ε_i and homoscedastic.

---

[5] The choice of random effects comes from the fact that in non-linear models fixed effects are problematic. MLE estimator is inconsistent in probit models with fixed effects because it suffers from incidental parameter problem (Neyman and Scott, 1948). Chamberlain (1980) proposed a conditional static logit model and later Chamberlain (1984) a dynamic one under the restriction that observable heterogeneity stays time-invariant. Later Honore and Kyriazidou (2000) proposed a conditional dynamic logit model which requires very strong distributional assumptions about the observable heterogeneity over time and has a rate of convergence slower than $\sqrt{N}$.

The assumption that ε_i is uncorrelated with x_{it} for all i and in every t is frequently thought of as unrealistic thus following Mundlak (1978) and Chamberlain (1984) we specify a relationship between the unobserved heterogeneity ε_i and the means (within i, over T) of all time-varying explanatory variables. This implies that $\varepsilon_i = \overline{x}_i'\delta + \alpha_i$ and with the assumption that $\alpha_i \sim iid \ N(0,1)$ and is independent of x_{it} and u_{it} for all i and in all t periods, this leads to a correlated random effects probit model, with extra regressors which are the means of all time-varying variables. Substituting into (1) we get:

$$L_{it}^* = \gamma L_{it-1} + x_{it}'\beta + \overline{x}_i'\delta + \alpha_i + u_{it} \qquad (2)$$

Estimation of this model using Gauss-Hermite quadrature is straightforward however the resulting parameter estimates are only consistent if we are prepared to make a further assumption. This involves whether the initial observation of the dependent variable L_{i1} and the unobserved heterogeneity ε_i are correlated or not. This "initial conditions problem" arises when the beginning of the estimation period does not coincide with the beginning of the stochastic process which generates low pay status. Consider the following data generation process. For periods after the first period from (2),

$$L_{it}^* = \gamma L_{it-1} + x_{it}'\beta + \overline{x}_i'\delta + \alpha_i + u_{it} \text{ for } t = 2,...,T$$

while for the initial period we have:

$$L_{i1}^* = x_{i1}'\lambda + \eta_i$$

10

Notice that no information regarding L_{i0} is available. If η_i and α_i are correlated then L_{i1}^* is correlated with α_i in $L_{i2}^* = \gamma L_{i1} + x_{i2}^{'} \beta + \bar{x}_i^{'} \delta + \alpha_i + u_{i2}$ and γ and β can not be estimated consistently.

A general solution to this is to jointly estimate the random effects probit for $t > 1$ and the probit for $t = 1$. This is the approach taken by Heckman (1981a, 1981b), who specifies a reduced form equation for the initial observation:

$$L_{i1}^* = z_i^{'} \lambda + \eta_i \qquad (3)$$

where z_i is a vector of strictly exogenous variables, which affect L_{i1}^*, $\mathrm{var}(\eta_i) = \sigma_\eta^2$ and $corr(\alpha_i, \eta_i) = \rho$. To ensure identification of the system of equations described by equations (2) and (3) z_i is a superset of x_i and includes instruments which affect only the probability of being low paid in the first period. Presample information, which is assumed to influence L_{i1}^* can be included in z_i as well as the vector of means $\bar{x}_i$. To allow for a non-zero ρ, a linear specification is introduced, in terms of orthogonal error components:

$$\eta_i = \theta \alpha_i + u_{i1} \qquad (4)$$

By construction α_i and u_i are orthogonal to one another with $\theta = \rho \sigma_\eta / \sigma_\alpha$ and $\mathrm{var}(u_{i1}) = \sigma_\eta^2 (1 - \rho^2)$. Furthermore, it is assumed that the initial observation of L is not correlated with u_{it}, i.e. $\mathrm{E}(u_{it}, L_{i1}) = 0$ and also it is not correlated with x_{it} for all i and in all $t = 2, \ldots, T$.

If now equation (4) is replaced into equation (3) equation (5) emerges

$$L_{i1}^* = z_i^{'} \lambda + \theta \alpha_i + u_{i1} \qquad (5)$$

11

which in combination with equation (2) constitute the following full specification of the Heckman's model:

$$\left\{ \begin{array}{l} L_{i1}^* = z_i'\lambda + \theta\alpha_i + u_{i1} \\ L_{it}^* = \gamma L_{it-1} + x_{it}'\beta + \overline{x}_i'\delta + \alpha_i + u_{it}, \quad i = 1, 2, \ldots, N \quad \text{and} \quad t = 2, \ldots, T \end{array} \right\} \quad (6)$$

According to Heckman (1981a) and Heckman (1981b) under the assumption that $\alpha_i \sim IN\left(0, \sigma_\alpha^2\right)$, is independent of u_{it} and that the distribution of L_{it}^* conditional on α_i, x_{it} and L_{it-1} is independent normal this model can be estimated by maximum likelihood, however it is computationally time-consuming and the procedure is not guaranteed to converge.

Orme (1996) suggests a two-step estimation method, which is easy to estimate using standard random effects probit software.[6] Orme follows Heckman by assuming that the model is fully specified by equations (2) and (3). He then suggests a linear specification, in terms of orthogonal error components in such a way that again $\rho \neq 0$:

$$\alpha_i = \kappa\eta_i + w_i \quad (7)$$

By construction η_i and w_i are orthogonal to each other, $\kappa = \rho\sigma_\alpha / \sigma_\eta$ and $\text{var}\left(\kappa\right) = \sigma_\alpha^2\left(1 - \rho^2\right)$.

The next step is to substitute (7) into (2) and get:

$$L_{it}^* = x_{it}'\beta + \gamma L_{it-1} + \delta\overline{x}_i + \kappa\eta_i + w_i + u_{it} \quad (8)$$

---

[6] An alternative, simple estimator is also provided by Wooldridge (2005).

where $i = 1, 2, \ldots, n$ and $t = 2, \ldots, T$.

In this "new" random effects probit, there are two individual specific random effects, η_i and w_i and the assumption of bivariate normality of (η_i, α_i) implies that

$$E(w_i \mid L_{i1}) = 0 \text{ and } E(\eta_i \mid L_{i1}) = e_i, \text{ where } e_i = \frac{(2L_{i1} - 1)\phi(\lambda' z_i)}{\Phi\left[(2L_{i1} - 1)\lambda' z_i\right]} \text{ by construction.}$$

Since u_{it} is assumed to be orthogonal to the regressors, w_i can be treated as the common error component in a random effects probit, as long as we take care of the unobservable η_i. Taking into consideration that e_i is derived from a probit model from equation (3)[7], it is reasonable to substitute η_i by its conditional expectation $E(\eta_i \mid L_{i1}) = e_i$. Thus, equation (8) becomes a random-effects probit with an extra regressor e_i which can easily be generated from the parameter estimates from the first stage probit model: $e_i = (2L_{i1} - 1)\phi(\hat{\lambda}' z_i) \big/ \Phi\left[(2L_{i1} - 1)\hat{\lambda}' z_i\right]$. A test of the null hypothesis that $\rho = 0$ can be obtained by a simple t-test on the coefficient on e_i.

Orme (1996) and Arulampalam and Stewart (forthcoming) provide Monte Carlo evidence which suggests that Orme's approach can provide estimates which are no worse than the Heckman procedure, but at a much lower cost in terms of computation time, so long as T ≥ 6 and N ≥ 800.[8] These criteria are met in our empirical implementation and in what follows we report results computed using the procedure proposed by Orme (1996).

---

[7] Remember that $E(\eta_i \mid L_{i1}) = e_i$

[8] All the estimations were carried out using Stata version 8.2/IC (StataCorp (2003)). Estimation of the Orme model took between one and four minutes. Wooldridge's (2005) estimator was slower taking from three to eleven minutes while the Heckman estimator itself was substantially more time-consuming with convergence times between two to ten hours.

4. Empirical results

4.1. *Impact of explanatory variables*

The estimates of the model using the Orme procedure are contained in Tables 3 and 4. Before discussing the main parameter of interest, γ, which measures the extent of low pay persistence in each country we briefly consider the estimates of the other parameters in the model, those relating to the explanatory variables. These are shown in Table 3 and are derived from the second stage of the Orme procedure, i.e. equation (8). In the first stage equation, we included additional regressors denoting whether the individual was unemployed during the last five years before joining the survey and two variables indicating whether his most frequent activity a year before joining the survey was employed or unemployed. Such pre-sample information is intended to identify the extent of true state dependence by proxying the stochastic process determining low pay prior to the initial observations of low paid status. Results from the first stage are not reported here, but are available on request.

< Table 3 near here >

Table 3 contains marginal effects of the explanatory variables on the probability of being low paid[9]. The signs of these are generally in line with what one would expect according to human capital theory, and the huge empirical literature on earnings functions. Having a larger endowment of formal education reduces the probability of low pay. For example, having completed university level education is both economically and statistically significant in every country with marginal effects

---

[9] There is a slight complication in interpretation of marginal effects in a two stage model like this which arises from the fact that the initial conditions correction terms e_i above is also affected by marginal changes in the explanatory variables. The marginal effects presented here are thus best interpreted as conditional on the values of the explanatory variables in the initial period, T=1.

14

ranging from -0.016 in Belgium to -0.095 in Portugal. Further training, defined as any form of vocational or training course other than general or higher education, also reduces the probability of being low paid in the majority of countries and is statistically significant at 10% or lower in five of them. With the sole exception of Germany, older workers are generally less likely to be found in low pay, while marriage has a negative and significant marginal effect in six of the twelve countries.

As well as these individual variables, job-related characteristics are also included in model. Of these, the strongest finding is for type of work contract where those on permanent contracts have a lower probability of being low paid. The magnitude of this effect ranges from a reduction of 0.073 in France to a reduction of 0.012 in Belgium and it is statistically significant in every country except Belgium. Working in the public sector is significant only for Denmark, France, Italy and the UK and in all cases its marginal effect is positive. Wherever working in supervisory position is significant, it reduces the likelihood of low pay.

4.2.   Initial conditions and state dependence

Turning now to the issue of state dependence, the key estimation problem is the potential endogeneity of the intitial conditions. The Orme procedure allows a simple test of the null hypothesis of endogeneity, which is a test of the significance of the "correction" variable e_i. The estimates of the parameter on this variable together with its estimated standard error are presented in Table 3 in the row labelled "Initial Conditions". Clearly this term is highly significant (at 1% or lower) for each of the

countries. The effect that controlling for initial conditions has on the estimates of the magnitude of state dependence will be discussed below.[10]

The estimates of the marginal effect associated with the parameter γ are presented in the first row of Table 3.[11] We call this the dynamic marginal effect and we argue that this measures true state dependence in low pay. It is clear that, after controlling for observed and unobserved heterogeneity, and for each of the countries in our sample, being low paid in period t-1 has a positive and statistically significant effect on the probability of being low paid in period t. There is therefore state dependence in low pay in these European countries. The magnitude of this lies between 0.066 and 0.237 with marginal effects of greater than 0.1 in 7 of the twelve countries (France, Germany, Greece, Ireland, Netherlands, Portugal and the UK)[12].

Table 4 provides further information on the extent of state dependence. The marginal effects from the first row of Table 3 are reproduced in the fourth column while the first three columns contain other measures of the same parameter. These are, respectively the probability of remaining in low pay from the raw data (column 5 of table 1), an estimate of the marginal effect calculated from the coefficient on a lagged dependent variable in a pooled probit model (equation (1) without the ε_i term), and the estimate of the same parameter from a dynamic random effects probit (equation (1)). Thus, moving across the columns from left to right in Table 4, we

[10] To our knowledge this estimator has not been used before to model persistence in low pay. Arulampalam *et. al* (2000) used the Orme approach to estimate unemployment persistence in the UK.

[11] This is the effect of low pay status last period on the probability of being low paid in period t and is given by: $\Pr(L_{it} = 1 \mid L_{it-1} = 1) - \Pr(L_{it} = 1 \mid L_{it} = 0)$. We have computed this taking account of the issues raised in Wooldridge (2005) and Arulampalam (1999).

[12] To confirm the robustness of our results we also estimated a dynamic random effects probit controlling for the initial conditions, using the methodology suggested by Wooldridge (2005). The results were very similar with the dynamic marginal effects lying between 0.046 and 0.226 and with 4 out of 12 countries reporting a value greater than 0.1 (France, Ireland, Portugal and the UK).

control for observed heterogeneity (column (1) to column (2)), observed and unobserved heterogeneity (column (3)) and finally heterogeneity plus initial conditions (column 4).

< Table 4 near here>

It is clear that as we move from left to right the estimated extent of low pay persistence declines as we control for progressively more factors. Between columns (1) and (2), which could be thought of as controlling for observables, the reduction in the average measure of persistence is 29%, controlling for unobservables reduces the raw level a further 18% while taking account of initial conditions leads to another reduction of 29% of the initial estimate. Overall, the extent of low pay persistence estimated using the Orme model is 76% lower on average than that from the raw data. Controlling for heterogeneity and initial conditions is thus extremely important when trying to establish the level of true state dependence in low pay, and the effect of each of these factors varies by country. However there remains a fairly strong positive correlation between the various measures, as the scatter matrix in Figure 1 shows. The lowest correlation in these graphs is $r = 0.72$ ($p = 0.01$) for the relationship between the final (Orme) estimates and the pooled probit. The correlation between the estimates from the raw data and the final estimates is $r = 0.81$ ($p = 0.00$). In other words, countries which score highly on the raw measure of persistence also tend to be those where true state dependence is estimated to be high.

< Figure 1 near here >

To put these results in context it would useful to compare our results with those from other researchers. However, as previously noted, studies of low pay persistence have used different definitions of low pay and different methodologies and this should be borne in mind. It is however worth noting that both Stewart and

Swaffield (1999) for the UK and Cappellari (2000 and 2002) for Italy find a much higher proportion of the raw low pay persistence to be due to genuine state dependence.

5.   Further discussion

In the previous section we have demonstrated the existence of positive, statistically significant state dependence in low pay in a sample of European countries. It is interesting to speculate about the variation in the magnitude of this effect. For example, Portugal is the country with the strongest low pay persistence which could be related to the fact that Portugal is a highly regulated labour market characterized by high strictness of employment protection legislation which works as a safeguard for those in employment, (OECD, 1997). Even though the Portuguese and other Mediterranean labour markets are not characterized by high levels of union density their segmented labour markets might work as a safeguard for the workers in the internal labour markets. Comparing Portugal with the other Southern European Union countries (Greece, Italy, Spain) it seems that Portuguese economy is much less "mobile" than the other Mediterranean economies. Among these countries Spain displays the lowest persistence while Greece has the highest.

On the other hand Denmark is, almost in all cases, the country with the smallest dynamic marginal effect, suggesting that low pay persistence is minimal and Danish workers are not trapped into low wages. In this context it is worth noting that Denmark is a country with higher participation and employment rates in Europe and also during the period studied here Denmark recorded very low unemployment rates. This is mainly a result of a combination of modest employment protection legislation

with a social security net that efficiently helps the Danish unemployed. Denmark is an excellent example of a labour market characterized by "flexicurity", which successfully combines high job mobility, high flexibility in the labour market and high unemployment benefits.

Two countries with relatively high dynamic marginal effect are France and Germany which are also economies with regulated labour markets. Moreover, for Germany it is worth mentioning that there is high union density and very strict compliance with the collective bargaining. On the other hand in France the high levels of minimum wage might increase low pay persistence.

An interesting finding is that UK and Ireland turn out in most cases with higher than the overall average low pay persistence. These two economies were expected more flexible because of relatively low levels of employment regulations. As there is relatively low public interference and weak employment protection legislation it might be expected that for these γ would be lower than in other countries. However, they show a rather high "degree of immobility" as it is measured by γ. More specifically UK and Ireland have a higher γ than the Mediterranean countries and the "continental" countries.

To consider these ideas a little more formally, Figure 2 presents the results of correlating the dynamic marginal effect from each country with various measures of labour market institutions. Institutional aspects of developed country labour markets have been seen by researchers as important determinants of labour market outcomes such as employment and unemployment rates (see, for example, Blanchard and Wolfers, 2000 or Belot and Van Ours, 2004).

< Figure 2 near here >

19

The data on institutions are taken from Bassanini and Duval (2006) and the institutional variables are briefly defined in the appendix.[13] Figure 2 suggests that low pay persistence is negatively correlated with the unemployment rate, replacement rate, tax wedge, union density and coverage and active labour market programmes, while positive correlations exist for employment protection and product market regulation. It should be noted however that some of these correlations, both positive and negative, are very low, and only the tax wedge is statistically significant at conventional levels. In standard models of unemployment (for example, the wage-setting, price-setting model) the tax wedge is thought to increase equilibrium unemployment, however it is less clear how, from a theoretical perspective, persistence in low pay would be affected by the wedge. Indeed, the overall impression from Figure 2 is of no clear or simple explanation of how the magnitude of state dependence in low pay varies between countries.

6.   Conclusions

In conclusion, using the European Community Household Panel data we have demonstrated the existence of positive, statistically significant state dependence in low pay in a sample of twelve European countries. Our results suggest that heterogeneity, both observed and unobserved, explains a substantial majority of the overall tendency for outflows from low pay to be relatively small. However, a role for true state dependence – the idea that being low paid permanently affects the future likelihood of exiting low pay – remains. To the extent that persistence in low pay is not the result of genuine state dependence but reflects differences between workers in productive abilities, there is scope for policy to enhance human capital to free

---

[13] Greece is excluded from this analysis as no institutional data were available.

European workers from the low pay trap. Regarding true state dependence, there is substantial variation between European countries. The explanation of this variation requires further theoretical and empirical work.

References

Arulampalam, W., Booth, A. L., Taylor, M. P., 2000. Unemployment persistence. Oxford Economic Papers 52, 24-50.

Arulampalam, W., Stewart, M., 2009. Simplified implementation of the Heckman estimator of the dynamic probit model and a comparison with alternative estimators. Oxford Bulletin of Economics and Statistics. , forthcoming.

Asplund, R., Bingley, P., Westergard Nielsen, N., 1998. Wage mobility for low-wage earners in Denmark and Finland. In: Asplund, R., Sloane, P. J., Theodossiou, I. (Eds.), Low Pay and Earnings Mobility in Europe. Edward Elgar Publishing Limited, Cheltenham, U.K.

Bardone, L., Guio, A.-C., 2005. In-work poverty. New commonly agreed indicators at the EU level. Statistics in Focus. Population and Social Conditions. Luxembourg, Eurostat. 5/2005: 1-11.

Bassanini, A., Duval, R., 2006. Employment patterns in OECD countries: Reassessing the role of policies and institutions. Working Paper No. 486. Paris, OECD.

Belot, M., Van Ours, J., 2004. Does the recent success of some OECD countries in lowering their unemployment rates lie in the clever design of their labour market reform? Oxford Economic Papers 56 (4), 621-642.

Blanchard, O., Wolfers, J., 2000. The role of shocks and institutions in the rise of European unemployment: The aggregate evidence. The Economic Journal 110, 1-33.

Cappellari, L., 2000. Low-wage mobility in the Italian labour market. International Journal of Manpower 21 (3-4), 264-90.

Cappellari, L., 2002. Do the 'working poor' stay poor? An analysis of low pay transitions in Italy. Oxford Bulletin of Economics and Statistics 64 (2), 87-110.

Chamberlain, G., 1980. Analysis of covariance with qualitative data. Review of Economics Studies (47), 225-238.

Chamberlain, G., 1984. Panel data. In: Griliches, S., Intriligator, M. (Eds.), Handbook of Econometrics. North Holland, Amsterdam.

European Commission, 2003. Joint Report on Social Inclusion. Luxembourg.

Eurostat, 2003a. ECHP UDB Description of Variables, PAN 166.

Eurostat, 2003b. ECHP UDB Manual - European Community Household Panel Longitudinal User's Database - Waves 1 to 8 - Survey years 1994 to 2001.

Eurostat. (2004). ECHP Documents. Retrieved September, 2005, from http://forum.europa.eu.int/Public/irc/dsis/echpanel/home.

Heckman, J. J., 1981a. Heterogeneity and state dependence. In: Rosen, S. (Ed.), Studies in Labor Markets. Chicago Press, Chicago, IL.

Heckman, J. J., 1981b. The incidental parameters problem and the problem of initial conditions in estimating a discrete time - discrete data stochastic process. In: Manski, C. F., McFadden, D. (Eds.), Structural Analysis of Discrete Data with Econometric Application. MIT Press, Cambridge.

Honore, B. E., Kyriazidou, E., 2000. Panel data discrete choice models with lagged dependent variables. Econometrica 68, 839-874.

Mundlak, Y., 1978. On the pooling of time series and cross section data. Econometrica 46, 69-85.

Neyman, J., Scott, E. L., 1948. Consistent estimates based on partially consistent observations. Econometrica 16, 1-32.

OECD, 1997. Employment Outlook. OECD, Paris.

Orme, C. D., 1996. The initial conditions problem and two-step estimation in discrete panel data models. Discussion Paper Series No 9633. The University of Manchester.

Sloane, P. J., Theodossiou, I., 1998. An econometric analysis of low pay and earnings mobility in Britain. In: Asplund, R., Sloane, P. J., Theodossiou, I. (Eds.), Low pay and earnings mobility in Europe. Cheltenham, U.K. Edward Elgar Publishing Limited,

StataCorp., 2003. Stata Statistical Software: Release 8.0. StataCorp. College Station, TX, USA.

Stewart, M. B., Swaffield, J. K., 1998. The earnings mobility of low-paid workers in Britain. In: Asplund, R., Sloane, P. J., Theodossiou, I. (Eds.), Low pay and

earnings mobility in Europe. Cheltenham, U.K. Edward Elgar Publishing Limited,

Stewart, M. B., Swaffield, J. K., 1999. Low pay dynamics and transition probabilities. Economica 66 (261), 23-42.

Wooldridge, J. M., 2005. Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. Journal of Applied Econometrics 20, 39-54.
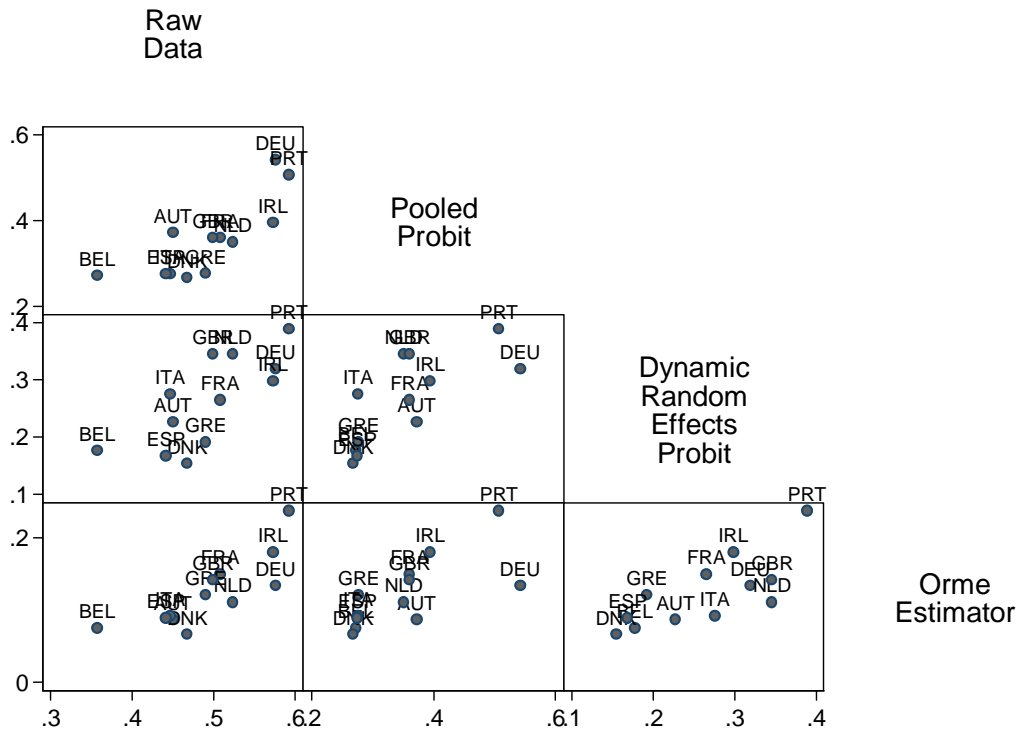
Fig. 1. Scatter matrix of persistence measures. Note: See Table 4 and the text for definitions of the alternative measures of low pay persistence. The countries featured are: Austria (AUT), Belgium (BEL), Denmark (DNK), France (FRA), Germany (DEU), Greece (GRE), Ireland (IRL), Italy (ITA), Netherlands (NLD), Portugal (PRT), Spain (ESP), Britain (GBR).
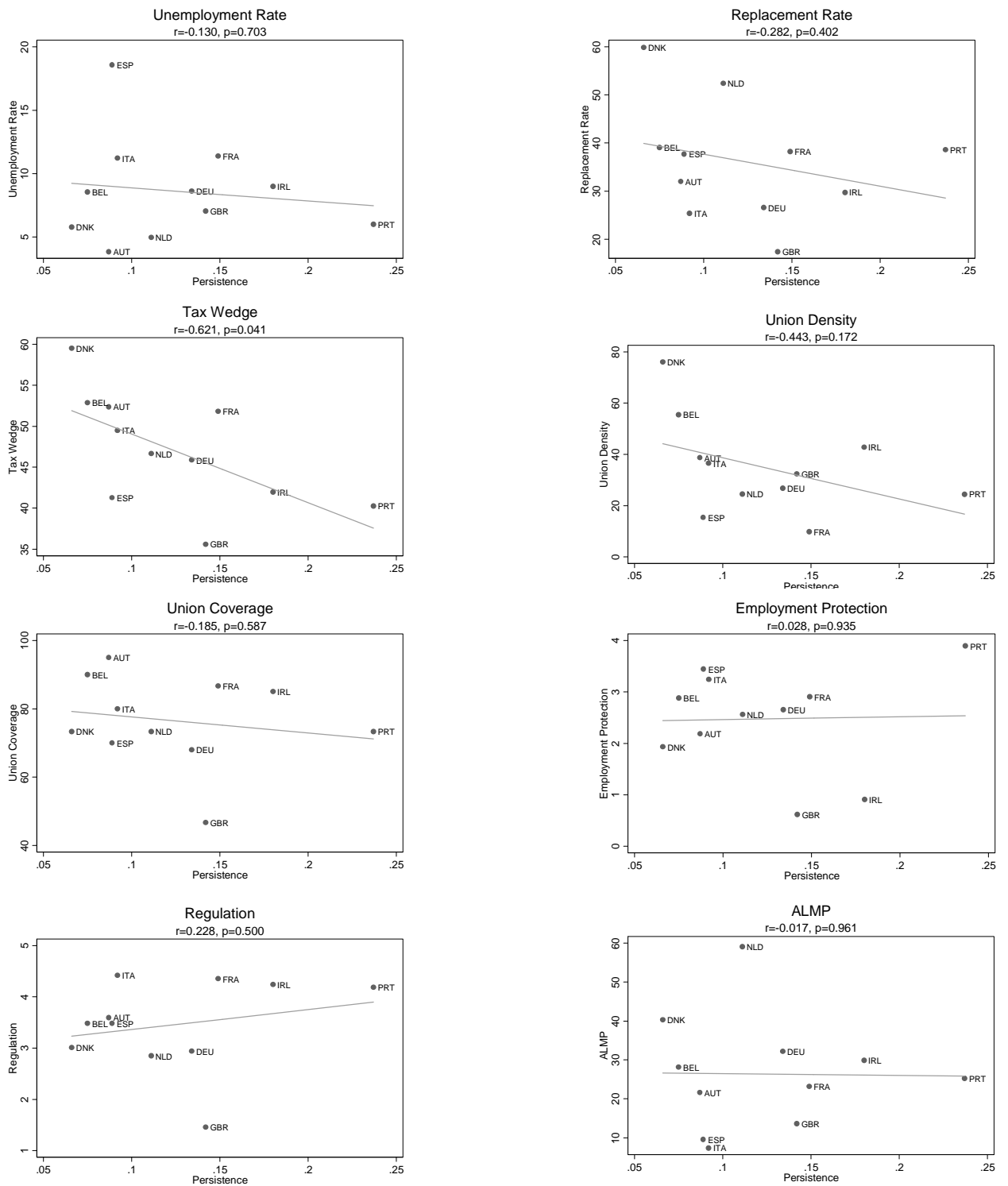
Fig. 2. Low pay persistence and institutions. Note: Persistence is measured here by the estimated dynamic marginal effect from the Orme model. For definition of other variables used see Appendix. Country acronyms are defines in the notes to Fig. 1.

26

Table 1
Transition Probabilities of Low Pay

	Pr(LP=1)	Pr(LP$_t$=1|LP$_{t-1}$=1)	Pr(LP$_t$=1|LP$_{t-1}$=0)	Pr(LP$_t$=0|LP$_{t-1}$=1)	State Dependence	Ratio
	[1]	[2]	[3]	[4]	[5]=[2]–[3]	[6]=[2]/[3]
Austria	0.084	0.471	0.021	0.529	0.450	22.327
Belgium	0.065	0.387	0.031	0.613	0.357	12.654
Denmark	0.080	0.492	0.025	0.508	0.467	19.618
France	0.166	0.561	0.053	0.439	0.508	10.599
Germany	0.131	0.620	0.044	0.380	0.576	13.971
Greece	0.160	0.550	0.060	0.450	0.490	9.106
Ireland	0.189	0.615	0.042	0.385	0.573	14.599
Italy	0.074	0.473	0.026	0.527	0.446	17.970
Netherlands	0.095	0.553	0.029	0.448	0.524	19.118
Portugal	0.194	0.650	0.058	0.350	0.592	11.242
Spain	0.188	0.515	0.074	0.485	0.441	6.962
UK	0.155	0.555	0.056	0.445	0.499	9.905

Note: The table reports: 1. Probability of being low paid, 2. Probability of being low paid in year *t* conditional on being low paid in year *t-1*, 3. Inflow to low pay, i.e. probability of being low paid in year *t* conditional on not being low paid in year *t-1*, 4. Outflow of low pay, i.e. probability of not being low paid in year *t* conditional on being low paid in year *t-1*

Table 2
Low Pay Probabilities by characteristics

	Austria	Belgium	Denmark	France	Germany	Greece
University	0.026	0.041	0.024	0.080	0.081	0.056
High school	0.062	0.069	0.075	0.157	0.147	0.168
Primary school	0.207	0.083	0.179	0.195	0.144	0.214
Age 18-24	0.237	0.234	0.449	0.571	0.391	0.503
Age 25-34	0.062	0.090	0.056	0.204	0.129	0.204
Age35-44	0.060	0.048	0.042	0.104	0.110	0.083
Age 45 +	0.047	0.034	0.026	0.084	0.098	0.078
Married	0.048	0.042	0.031	0.093	0.095	0.087
Not Married	0.137	0.119	0.143	0.286	0.212	0.318
Training	0.054	0.047	0.090	0.157	0.393	0.071
Not Training	0.090	0.067	0.071	0.166	0.121	0.164
Agriculture	0.370	0.139	0.190	0.430	0.423	0.396
Industry	0.076	0.062	0.069	0.155	0.106	0.188
Services	0.080	0.070	0.073	0.164	0.139	0.135
Public	0.098	0.080	0.086	0.196	0.141	0.228
Private	0.044	0.048	0.063	0.085	0.097	0.047
Permanent Contract	0.060	0.053	0.038	0.124	0.114	0.108
Temporary Contract	0.161	0.120	0.146	0.422	0.254	0.344

	Ireland	Italy	Netherlands	Portugal	Spain	UK
University	0.089	0.023	0.030	0.015	0.075	0.104
High school	0.200	0.051	0.080	0.103	0.169	0.168
Primary school	0.202	0.100	0.117	0.223	0.249	0.222
Age 18-24	0.486	0.258	0.661	0.396	0.478	0.450
Age 25-34	0.169	0.094	0.118	0.164	0.218	0.137
Age35-44	0.079	0.035	0.040	0.127	0.123	0.076
Age 45 +	0.059	0.039	0.037	0.157	0.106	0.102
Married	0.057	0.042	0.042	0.138	0.131	0.087
Not Married	0.365	0.145	0.219	0.302	0.307	0.246
Training	0.246	0.035	0.206	0.079	0.128	0.097
Not Training	0.181	0.076	0.087	0.197	0.196	0.180
Agriculture	0.624	0.246	0.301	0.570	0.535	0.383
Industry	0.184	0.073	0.104	0.193	0.166	0.128
Services	0.156	0.056	0.085	0.142	0.173	0.172
Public	0.247	0.094	0.110	0.228	0.222	0.175
Private	0.046	0.035	0.046	0.075	0.075	0.062
Permanent Contract	0.132	0.046	0.071	0.159	0.104	0.142
Temporary Contract	0.352	0.234	0.336	0.331	0.343	0.366

Note: Low Pay probabilities conditional on level of education, age, marital status, post education training, industry and sector of employment and type of employment contract.

Table 3
Low Pay Probability model – Orme Estimator

	Austria	Belgium	Denmark	France	Germany	Greece
Low Paid t-1	0.087***	0.075***	0.066***	0.149***	0.134***	0.121***
	[0.022]	[0.024]	[0.022]	[0.017]	[0.017]	[0.016]
Initial Conditions	0.031***	0.028***	0.024***	0.057***	0.094***	0.038***
	[0.004]	[0.005]	[0.004]	[0.004]	[0.005]	[0.005]
University	-0.033***	-0.016**	-0.025***	-0.071***	-0.046***	-0.058***
	[0.007]	[0.008]	[0.006]	[0.007]	[0.013]	[0.009]
High School	-0.026***	0.005	-0.007	-0.003	0.023**	-0.023***
	[0.008]	[0.007]	[0.006]	[0.006]	[0.011]	[0.008]
Training	-0.011*	-0.014**	0.005	0.018	0.030	-0.032*
	[0.006]	[0.007]	[0.004]	[0.011]	[0.025]	[0.019]
Age 25-34	-0.006	-0.007	-0.014*	-0.046***	-0.040***	-0.069***
	[0.007]	[0.013]	[0.007]	[0.013]	[0.013]	[0.010]
Age 35-44	0.003	-0.016	-0.006	-0.075***	-0.006	-0.087***
	[0.009]	[0.014]	[0.009]	[0.014]	[0.015]	[0.012]
Age 45+	-0.009	-0.033***	-0.018**	-0.086***	0.035**	-0.091***
	[0.009]	[0.012]	[0.009]	[0.015]	[0.018]	[0.012]
Public Sector	0.004	0.003	0.021*	0.040*	0.019	0.021
	[0.016]	[0.013]	[0.013]	[0.024]	[0.017]	[0.019]
Permanent Contract	-0.045***	-0.012	-0.018**	-0.073***	-0.021*	-0.046***
	[0.012]	[0.011]	[0.009]	[0.014]	[0.012]	[0.009]
Supervisory	0.008	0.001	-0.008	-0.024*		-0.051**
	[0.013]	[0.014]	[0.008]	[0.013]		[0.023]
Intermediate	-0.004	-0.004	-0.002	-0.007		0.041*
	[0.007]	[0.009]	[0.008]	[0.009]		[0.024]
Industry	-0.064*	0.007	-0.010	0.116*	-0.109***	-0.004
	[0.039]	[0.036]	[0.026]	[0.064]	[0.038]	[0.051]
Services	-0.043	0.030	0.011	0.103*	-0.089***	0.018
	[0.034]	[0.034]	[0.027]	[0.054]	[0.034]	[0.052]
Married	0.003	-0.006	-0.000	-0.042**	-0.025	-0.039
	[0.016]	[0.019]	[0.010]	[0.018]	[0.017]	[0.026]
LogL.	-999.79	-752.50	-697.40	-2360.53	-2768.81	-1906.41
N	7961	5274	6441	10787	10269	8056

Table 3 *(continued)*

	Ireland	Italy	Netherlands	Portugal	Spain	UK
Low Paid t-1	0.180***	0.092***	0.111***	0.237***	0.089***	0.142***
	[0.023]	[0.016]	[0.021]	[0.016]	[0.010]	[0.016]
Initial Conditions	0.041***	0.022***	0.034***	0.062***	0.058***	0.057***
	[0.005]	[0.002]	[0.003]	[0.005]	[0.004]	[0.004]
University	-0.051***	-0.018***	-0.021***	-0.095***	-0.069***	-0.054***
	[0.009]	[0.006]	[0.006]	[0.017]	[0.007]	[0.007]
High School	-0.027***	-0.019***	-0.008**	-0.056***	-0.044***	-0.033***
	[0.007]	[0.003]	[0.004]	[0.009]	[0.006]	[0.007]
Training	-0.004	-0.006	-0.003	0.012	-0.034***	-0.023***
	[0.011]	[0.008]	[0.006]	[0.028]	[0.009]	[0.006]
Age 25-34	-0.020**	-0.016***	-0.045***	-0.041***	-0.040***	-0.065***
	[0.009]	[0.005]	[0.008]	[0.008]	[0.008]	[0.008]
Age 35-44	-0.022*	-0.029***	-0.071***	-0.043***	-0.075***	-0.076***
	[0.012]	[0.005]	[0.011]	[0.010]	[0.009]	[0.008]
Age 45+	-0.024*	-0.028***	-0.072***	-0.044***	-0.089***	-0.059***
	[0.013]	[0.006]	[0.011]	[0.010]	[0.009]	[0.009]
Public Sector	0.013	0.014**	0.015	0.021	0.003	0.070***
	[0.024]	[0.007]	[0.009]	[0.015]	[0.017]	[0.014]
Perm. Contract	-0.021*	-0.032***	-0.032***	-0.038***	-0.053***	-0.068***
	[0.011]	[0.006]	[0.009]	[0.008]	[0.006]	[0.016]
Supervisory	-0.003	-0.008	-0.017**	0.011	-0.022	-0.038***
	[0.018]	[0.008]	[0.008]	[0.034]	[0.016]	[0.011]
Intermediate	0.002	-0.008	-0.008	0.006	-0.005	-0.025***
	[0.013]	[0.005]	[0.006]	[0.021]	[0.009]	[0.009]
Industry	-0.073**	-0.017	0.023	-0.042*	-0.032	-0.051*
	[0.031]	[0.013]	[0.028]	[0.024]	[0.020]	[0.029]
Services	-0.069**	-0.008	0.002	-0.028	-0.010	-0.020
	[0.032]	[0.013]	[0.024]	[0.024]	[0.020]	[0.032]
Married	-0.022	-0.021*	-0.031**	-0.081***	-0.034**	-0.038**
	[0.025]	[0.011]	[0.013]	[0.016]	[0.017]	[0.015]
LogL.	-1440.90	-1814.19	-1643.26	-3578.31	-4042.66	-2939.53
N	7028	15858	12009	13626	14801	12054

Note: The table reports marginal effects from the Orme model; for estimation methods see text. Standard errors are in brackets. All models also contain averages of time varying variables and year dummies (not reported for brevity). LogL and sample sizes refer to periods 2 to T. * *p*-value less than 0.10, ** *p*-value less than 0.05, *** *p*-value less than 0.01

Table 4
Dynamic Marginal Effects of Previous Low Pay – Alternative Estimators

	Raw Data [1]	Pooled Probit [2]	D.R.E. Probit [3]	Orme Model [4]
Austria	0.450	0.372	0.227	0.087
Belgium	0.357	0.273	0.177	0.075
Denmark	0.467	0.268	0.154	0.066
France	0.508	0.361	0.265	0.149
Germany	0.576	0.542	0.319	0.134
Greece	0.490	0.277	0.192	0.121
Ireland	0.573	0.395	0.298	0.180
Italy	0.446	0.276	0.275	0.092
Netherlands	0.524	0.351	0.346	0.111
Portugal	0.592	0.507	0.389	0.237
Spain	0.441	0.275	0.168	0.089
UK	0.499	0.361	0.346	0.142

Note: All Dynamic Marginal Effects are significant at 1%. For estimation methods see text.

Appendix A

Definition of variables

Variable	Definition
Low Pay	= 1 if low paid in current year, = 0 otherwise
University	= 1 if highest level of education completed university, = 0 otherwise
High School	= 1 if highest level of education completed high-school, = 0 otherwise
	(Base category = no qualifications or highest level of education is primary)
Training	= 1 if training course was undertaken the last year, = 0 otherwise
Age 25-34	= 1 if aged 25-34 years, = 0 otherwise
Age 35-44	= 1 if aged 35-44 years, = 0 otherwise
Age 45-65	= 1 if aged 45-65 years, = 0 otherwise
	(Base category = aged 18-24 years)
Public	= 1 if current job is in the public sector, = 0 otherwise
Permanent Employment	= 1 if holds a permanent employment contract, = 0 otherwise
Supervisory	= 1 if current job's status is supervisory, = 0 otherwise
Intermediate	= 1 if current job's status is intermediate, = 0 otherwise
	(Base category = non-supervisory)
Unemployed 5 years ago	= 1 if unemployed within the last five years, = 0 otherwise
Last employed	= 1 if employed one year ago, = 0 otherwise
Last unemployed	= 1 if unemployed one year ago, = 0 otherwise
	(Base category = inactive one year ago)
Industry	= 1 if current job is in the industry sector, = 0 otherwise
Services	= 1 if current job is in the services sector, = 0 otherwise
	(Base category = agricultural sector)
Married	= 1 if married, = 0 otherwise
Unemployment Rate	Unemployed workers as share of the labour force
Replacement Rate	Average unemployment benefit replacement rate across different income situations, family situations and unemployment durations
Tax Wedge	The sum of personal income tax and all social security contributions as a percentage of total labour cost.
Union Density	The share of workers affiliated to a trade union, in %.
Union Coverage	The share of workers covered by a collective agreement, in %.
Employment Protection	OECD summary indicator of the stringency of Employment Protection Legislation.
Regulation	OECD summary indicator of regulatory impediments to product market competition in seven non-manufacturing industries.
ALMP	Public expenditures on active labour market programmes per unemployed worker as a share of GDP per capita (or public expenditures on active labour market programmes as a share of GDP, depending on econometric specifications), in %.