Earnings Mobility among Italian Low Paid Workers§

Lorenzo Cappellari
Institute for Social and Economic Research
University of Essex
Colchester, UK CO4 3SQ
e-mail: lorenzo@essex.ac.uk

June 2001

Abstract
This paper uses Italian panel data to analyse transition probabilities at the bottom of the earnings distribution during the 1990s. The analytical framework is characterised by the ability to account for endogeneity of initial conditions and earnings attrition. Results show that both are endogenous for the estimation of low pay transitions. In particular it is found that the low paid are more likely to exit from the earnings distribution compared to the higher paid, revealing higher employment instability. The data also reveal considerable state dependence, i.e. the probability of experiencing low pay depends upon past low pay experiences rather than on personal attributes. Extensions of the model to longer term transitions suggest that state dependence effects are concentrated at the beginning of low pay spells, while subsequent low pay experiences contribute to a lesser extent.

keywords: low pay, earnings mobility, initial conditions, earnings attrition
JEL code: C23, C35, D31, J31

---

§ This paper extends Chapter Five of my PhD dissertation at the University of Warwick. Thanks go to W. Arulampalam, C. Dell'Aringa, S. Jenkins, C. Lucifora and especially M. Stewart for comments and advice. Seminar participants at the 1999 EALE conference and the Catholic University of Milan provided useful remarks. Data availability from the Bank of Italy and financial support from the Catholic University of Milan and the Nuffield Foundation under the New Career Development Scheme are gratefully acknowledged. Usual disclaimers apply.

NON-TECHNICAL SUMMARY

The issue of low paid employment has received considerable attention from researchers and policy makers in recent years. The rise of earnings inequality occurred in many industrialised economies has placed a growing proportion of the employed labour force below pre-determined “decency thresholds” in the earnings distribution, raising equity concerns and revitalising interest in redistributive tools such as minimum wages. Authors have also stressed that the diffusion of low paid employment might lead to efficiency losses.

Recent research on low paid employment underlines the need of a dynamic analysis of the phenomenon. Shifting the focus of analysis from “who is low paid at a point in time?” to “who remains low paid from one period to another?” or “who becomes low paid?” helps in assessing the causes of low paid employment, rather than its symptoms. In other words, evidence on the degree of mobility across the low pay threshold from one period to another can reveal to what extent low pay is a transitory or prolonged episode of earnings careers, yielding insights into the urgency of income support policies for the low paid. Moreover, a proper dynamic treatment enables researchers to disentangle between competing explanations of low pay persistence, namely heterogeneity versus state dependence. Under the first circumstance, individuals are entrapped into low pay because they lack some fundamental characteristic which favour earnings growth, say education. If this is the case, then a policy indication to fight low pay is to focus policy interventions on those low paid employees with low educational endowment. At the opposite extreme, low pay persistence could arise from state dependence. In this case, low pay persistence has nothing to do with personal characteristics, but with the experience of low pay in itself. Even in a world of identical individuals, experiencing low pay might worsen earnings careers if – for example – current low pay is used by prospective employers as a signal of the quality of the employee. If state dependence is the cause of low pay traps, then policies targeted on “problem groups” within the pool of low paid employees might be misleading, and more generalised measures could be needed.

This paper analyses the earnings mobility of low paid Italians while treating two econometric problems inherent to this kind of exercise. Panel data on earnings from the Bank of Italy’s Survey on Household Income and Wealth (SHIW) for the 1990s have been used. The first is the so-called initial conditions problem. Analysing mobility requires researchers to estimate the impact of past low pay on current low pay. As long as past low pay is not randomly distributed in the sample, such an exercise might lead to biased inference unless the econometric model is not properly adjusted. Similarly, earnings mobility can be analysed only among those employees for which earnings are observed in more than one time period. Individual experiencing earnings attrition, i.e. exiting from the earnings distribution from one period to another, do not contribute to estimation. Again, as long as attrition propensities are not randomly distributed, estimating mobility on the so-called balanced sample can lead to biased results and a proper adjustment of the econometric model is needed. The econometric model of this paper is robust to the presence of non-random initial conditions and earnings attrition.

Results first of all indicate the need of allowing for non-random initial conditions and earnings attrition, the coefficients measuring their impact being statistically different from zero.

The analysis of the relationship between personal attributes and low pay transition probabilities has shown that employees with low educational qualifications, female employees and southern workers are exposed to the risk of being trapped into low pay. The probability of dropping into low pay, on the other hand, appears to be associated with manual jobs and with jobs in the metal-manufacturing industry and in small firms.

A statistical test for state dependence indicate that its absence cannot be ruled out. The estimated model has been used to compute the contribution of state dependence to aggregate persistence, finding that the former accounts for more than 50% of aggregate persistence. This results indicates that low pay persistence is to a large extent independent of personal attributes, suggesting that policies targeted on the whole pool of low paid should be appropriate.

An extensions of the core model aimed at analysing longer term mobility has also been developed. Results show that the incidence of state dependence is lower in the longer run, suggesting that the factor causing state dependence are effective as soon as individuals are exposed to low pay.

Introduction

The issue of low paid employment has received considerable attention from researchers and policy makers in recent years (see OECD, 1993 and 1996). The rise of earnings inequality occurred in many industrialised economies has placed a growing proportion of the employed labour force below pre-determined “decency thresholds” in the earnings distribution, raising equity concerns and revitalising interest in redistributive tools such as minimum wages (see Freeman, 1996). Authors have also stressed that the diffusion of low paid employment might lead to efficiency losses if it is concentrated in industries with monopsonistic labour markets in which rising inequality translates into a widening gap between equilibrium wages and their perfectly competitive level (Lucifora, 1998).

Recent research on low paid employment underlines the need of a longitudinal analysis of the phenomenon (Stewart and Swaffield, 1999; Dickens, 2000). Evidence on the degree of mobility across the low pay threshold from one period to another can reveal to what extent low pay is a transitory or prolonged episode of earnings careers, yielding insights into the urgency of income support policies for the low paid. Stewart and Swaffield (1999) have also shown for Britain that low pay might be state dependent, i.e. low pay in one period raises the *coeteris paribus* probability of low pay in the future, a fact which should be taken into account when defining target groups for policy.

The present paper addresses these issues using Italian panel data on individual earnings in the 1990s. The econometric framework allows for initial conditions endogeneity along the lines set out by Stewart and Swaffield (1999). As shown in their paper, analysing low pay mobility requires current low pay probabilities to be conditioned on their lagged values, which is endogenous as long as there is some serially correlated unobserved earnings component. In addition I also control for endogenous earnings attrition. Mobility within the classes of the earnings distribution over time can be observed only for employees with valid earnings at each point in time, while individuals who exit from the sample of earnings recipients due to either panel attrition or movement out of the labour market will not contribute to estimation. If these individuals have unobserved characteristics which are correlated with their propensity to move across the low pay threshold (say because they have weaker labour market attachment), estimating low pay dynamics on the balanced earnings panel will yield biased estimates. Bingley et al. (1995) found that attrition is not ignorable when estimating mobility on Danish data. Finally, I investigate longer term mobility by extending the set-up above to account for earnings transitions over three consecutive panel waves, thus analysing higher order dynamics.

The analytical framework relies on estimation of multivariate probability models with which low pay transition probabilities equations can be estimated while controlling for the probability of lagged low pay and the probability of having valid earnings. Simulated maximum likelihood (SML) techniques are used to estimate multi-dimensional integrals.

Results indicate that the exogeneity hypothesis can be rejected for both initial conditions and attrition. In particular, earnings attrition probabilities and low pay probabilities at the beginning of the transition are positively associated, suggesting that low paid jobs are more unstable compared to higher pay. Moreover, I find that while observed attributes have some impact in shifting the probability of crossing the low pay threshold, a relevant share of earnings persistence is accounted for by state dependence effects, suggesting that policies targeted on "problem groups" among the low paid might have a limited impact, as low pay experiences affect, by themselves, earnings progressions, irrespective of personal attributes. Endogeneity of initial conditions and attrition appear to be relevant also in the case of higher order dynamics. Results from this latter model suggest that state dependence effects are concentrated at the beginning of a low pay spell, while subsequent low pay experiences contribute to a lesser extent.

I. The SHIW data

The data used in this study originate from the panel component of the Survey on Households Income and Wealth (SHIW), administered by the Bank of Italy since 1977. Interviews have been carried out on an annual basis until 1987 and biannually afterwards until 1995, while the latest wave refers to 1998. The sampling unit is the household; however, detailed information is available also at the individual level. Although originally designed as a repeated cross-sections sample, the survey includes a panel sub-sample since 1989. While initially fairly small, the proportion of panel-households (i.e. households sampled in at least two consecutive waves) has increased in recent waves, being approximately 40% since 1993.[1]

This study utilises the last three waves of the survey, i.e. 1993, 1995 and 1998. Apart from the aforementioned limited size of the panel sub-sample before 1993, data limitations prevented us from extending the analysis to earlier waves. In particular, information on individuals' parental background (education and occupation) for the head of household and the spouse has been introduced in the survey only since 1993. As we will see below, these variables play a crucial role in the econometric analysis, implying that the model cannot be estimated on waves preceding 1993. In addition, the structure of the questionnaire changed over time, in particular

[1] The allocation of households to the panel sub-group is carried out on a random basis among households who report availability for re-interview. Roughly 90% of households were available for re-interview in 1993 and 1995.

for what concerns labour market variables, and the selected waves provide good degree of homogeneity in the available information.

[Table 1 around here]

For the purposes of this analysis I select full-time employees aged 18-58 if female and 18-60 if male who were members of panel households. Table 1 provides a description of the SHIW sample in 1993 and 1995, the two years in which the starting point of an earnings transition can be observed. Column 1 reports the composition of the selected sample with respect to a set of personal and job characteristics.[2] Reported figures show that SHIW data reproduce the characteristics of the whole population of Italian employees. Column 2 reports the same statistics as column 1 but computed on the whole SHIW cross-sectional sample, i.e. it also includes employees who are members on non-panel households. A comparison between the two columns shows that there are no relevant differences in sample structure between the cross-sectional sample and the sub-group of employees from panel households. Column 3 of the Table restricts the attention to employees from panel households with valid earnings in two consecutive waves, i.e. the balanced sample of earnings recipients; it is only for this sub-sample that transitions across the low pay threshold can be observed. Comparisons with column 1 and 2 reveal that employees in the balanced sample tend to be more educated, more likely to hold non-manual jobs, to work for large employers and to be affiliated to the public sector when compared to the cross-sectional sample, all features which may indicate a strong labour market attachment.[3] Column 4 investigates the features of the balanced panel by a different perspective, i.e. by reporting the probability of being in the balanced sample conditional on observed characteristics. First, we can observe that such a probability is lower for the more experienced group of workers, who leave the earnings distribution to enter retirement. Also, column 4 confirms that the probability of being in the balanced panel is higher for more educated individuals, those in non-manual occupations, in the public sector and in larger firms. The Table also reveals that the rate of exit from the earnings distribution is larger between 1995 and 1998 than it is for the 1993-95 transition. This can be explained by the longer width of the former observation window, implying a larger chance of leaving the sample of wage earners, and by the fact that both sample size and the proportion of panel households were slightly lower in 1998 compared to the previous waves (see D'Alessio and Faiella, 2000).

[2] Some of the observed characteristics are amalgamated at a rather aggregate level, for example in the case of education or occupation. The choice of the level of aggregation is aimed at avoiding small cells size problems, which are particularly likely to arise in a model of low pay transitions where some of the parameters of interest are estimated conditionally on being low paid.

[3] Information on employer size is available only for private sector employees.

II. Low pay definitions and aggregate transition probabilities

Several definitions of low pay threshold have been proposed by previous studies, with alternatives ranging from some legally set minimum pay (Smith and Varvricheck, 1992) to fixed proportions of median or mean earnings (Stewart and Swaffield, 1999) or to relative definitions based upon quantiles (Gregory and Elias, 1994; OECD, 1996). In this study I follow this latter approach. In particular, in order to assess the sensitivity of results to the use of different thresholds, two low pay cut-offs have been analysed in parallel, namely the bottom quintile and the third decile of the distribution of hourly net earnings for full-time employees aged 18-58 if female and 18-60 if male.[4] The use of order statistics guarantees that thresholds are robust to outliers and can be easily updated over time. Moreover, it should be stressed that low-pay cut-off points have been computed from the whole SHIW cross-sectional sample (i.e. employees from panel and non-panel households) but have then been applied to analyse transition probabilities of a sub-sample, namely employees from panel households in the balanced panel of earnings recipients. Hence an individual moving out of – say – the poorest fifth does not need to be replaced by another individual moving into low pay, as would be the case if low pay thresholds were computed from the balanced sample of earnings recipients. In this sense, the thresholds utilised in this study combine the absolute and relative approaches.

<u>[Table 2 around here]</u>

Table 2 reports summary statistics of the distribution of nominal hourly earnings in the three SHIW waves. It can be observed that mean earnings exceed the 50th percentile of the distribution, a symptom of distribution (positive) skewness. Also, it can be observed that nominal mean earnings grew rather slowly between 1993 and 1995 (approximately 2 percentage points per year), and faster afterwards (an average increase of 3.5 percentage points per year between 1995 and 1998). If contrasted with the evolution of the CPI (also reported) these figures show that real earnings have been declining between 1993 and 1995 and recovering afterwards, so that by the end of the 1990 decade the growth of nominal earnings is almost in line with that

[4] The earnings information available in the SHIW refers to yearly earnings, inclusive of extra-time compensations and fringe benefits, net of income taxes and social security contributions. On the working time side, the survey reports the number of months worked in the year and the number of hours worked on average in a week, including extra-time. No information is available on the number of weeks worked on average in a month. In order to derive hourly earnings, I have assumed that each individual worked 52/12 weeks per month. Cappellari (2000) analyses low pay transitions using monthly and hourly earnings in parallel, showing that there are not dramatic differences in results between the two cases.

of consumer prices.[5] The Table also reports some measures of earnings dispersion in the three years, namely the standard deviation of earnings levels and the log-ratio between selected percentiles; estimates suggest that the earnings distribution has been stable during this period.

[Table 3 around here]

Low pay transition rates are reported in Table 3. The first two rows in the Table report the transition matrix of hourly earnings for the 1993-95 and 1995-98 transitions respectively. The probability of persisting in low pay is 56.2 percent between 1993 and 1995, while the corresponding figure for the 1995-98 transition is slightly higher at 57.9%. Since we should expect three year persistence rates to be lower that two year ones - the chance of moving out of low pay being larger over wider time windows – the estimates suggest that the earnings rigidity at the bottom end of the distribution increased in the second half of the 1990s. On the other hand the probability of falling into low pay from higher pay was 6.2% and 6.9% for the two and three year transition, respectively.

The model of low pay transition probabilities will be estimated pooling data from the two transitions following the approach of Stewart and Swaffield (1999). The third and fourth rows of Table 3 provide aggregate transition rates for the two low pay thresholds obtained after pooling transitions. Raw low pay persistence is estimated to be 56.9% when the threshold is the bottom quintile and 67.7% when the third decile is used. On the other hand, the probability of falling into low pay is 6.5% and 9%, respectively. These figures show that the conditional low pay probability varies considerably according to the conditioning starting state, i.e. the probability of experiencing low pay is characterised by state dependence. Using the difference prob($L_t|L_{t-s}$)-prob($L_t|H_{t-s}$) (with L_t and H_t indicating low and high pay in year t, respectively) as a measure of raw state dependence, Table 3 indicates that conditional low pay probabilities rise by 50% points when the starting state changes from high pay to low-pay and low pay is set at the bottom quintile of the distribution; the corresponding figure for the third decile is 59%.

State dependence in aggregate transition rates could arise from individual *heterogeneity* or *genuine state dependence* (GSD).[6] In the first case, the larger conditional low pay probability characterising the initially low paid is due to observed and unobserved persistent factors which affect low pay propensities and differ between workers above and below the low pay threshold. In this case, policies targeted according to the factors causing persistence can reduce entrapment

[5] Major changes in the system of wage indexation took place at the beginning of the 1990s, whereby ex-ante wage compensations for inflation where substituted by bargained ex-post compensations. The figures reported in Table 2 suggest that this system was not entirely effective in protecting real wages against inflation.

[6] See Heckman (1981a) for a general discussion and Stewart and Swaffield (1999) for an illustration in the context of low pay transitions.

into low paid jobs. In the case of GSD, on the other hand, it is the experience of low pay which, by itself, modifies individual tastes and constraints, increasing the probability of future low pay experiences.[7] This implies that policies targeted on "problem groups" might be misplaced and the whole pool of low paid employees should form the focus for intervention. The model of the next Section will test for and measure the extent of GSD in low pay transition probabilities.

Row 5 of Table 3 enlarges the sample for the analysis of transition probabilities by including also those employees who exit from the earnings distribution during the transition. As can be seen, the impact of this inclusion is substantive: 30% of those who earn above the low pay threshold in the starting year leave the distribution during the transition, and the figure rises to 45% when the initially low paid are taken into account. Including these exits consequently changes conditional low pay probabilities in the arrival year, which are now remarkably lower. Overall, the average (over transitions and starting states) rate of exits from the distribution of earnings is approximately 33%.[8] Additional insights on patterns of attrition from the earnings distribution are provided in row 6 of Table 3, where destination states of those who exits from the earnings distribution are specified. The estimates indicate that employees who start from low pay and exit the distribution are more likely to end up in part-time or self-employment, unemployment or to exit from the SHIW sample, when compared to workers initially high paid. These figures seem to suggest that low pay jobs are characterised by larger instability compared to high pay jobs. In particular, the evidence about entry rates into unemployment is consistent with the presence of cycles of low pay and unemployment as singled out by Stewart (1999). On the other hand, higher entry rates into retirement from high pay compared to low pay may reflect the life cycle of earnings.

III. A model of low pay transition probabilities

The estimation of an econometric model for low earnings mobility requires researchers to tackle two potential sources of endogenous sample selection inherent to this kind of problem; this Section lays out an analytical framework that enables us to analyse earnings mobility while controlling for both of them.

A first source of endogeneity arises from the so-called *initial conditions problem* (see Heckman, 1981b, and Stewart and Swaffield, 1999, for the case of low pay transitions). The

[7] As discussed in Stewart and Swaffield (1999) genuine state dependence might, for example, result from bad signalling, if employers use salary histories to assess the quality of prospective employees. Also human capital depreciation or alterations of search behaviour could cause past low pay to raise future low pay probabilities.

[8] As pointed out when commenting Table 1, the bulk of exits from the earnings distribution occurs between 1995 and 1998, with an overall exit rate of 46%.

problem is one of endogeneity of the lagged dependent variable in a dynamic panel data model. Estimating conditional low pay probabilities requires conditioning current low pay on past low pay. Unobservability of the initial conditions of the earnings process and serial correlation of earnings unobservables (due to unobserved heterogeneity and/or shocks persistence) imply that a common component – the initial condition – will be present in unobervables at each time period, causing past low pay to be endogenous with respect to current low pay. Stewart and Swaffield (1999) show that the problem can be handled as an endogenous selectivity one, in which transition probabilities and the probability of selection into the initial state are simultaneously estimated.

The second source of endogenous selection is due to non-random attrition from the earnings distribution. As long as individuals exiting from the sample of earnings recipients have unobservable characteristics which are correlated with their propensity to move across the low pay threshold, estimating the model of earnings transition on the balanced sample will produce a sample selection bias. Again, the problem can be solved by modelling the probability of selection into the balanced sample and earnings transition probabilities simultaneously. This kind of approach has been applied by Bingley et al (1995) to model wage mobility in Denmark.

The present paper adopts a three-variate probit set-up to simultaneously model the probability of selection into the balanced panel, the probability of selection into the initial pay state and the probability of transition across the low pay threshold. The model extends the one of Stewart and Swaffield (1999) by adding an attrition equation and will be estimated pooling transitions from the SHIW.[9] Let us assume that at the start of a transition (period *t-s*) earnings can be observed for a random sample of *N* employees and can be written as:

$$(1) \qquad \begin{aligned} & g(y_{it-s}) = \boldsymbol{\delta}'\mathbf{x}_{it-s} + u_{it-s} \\ & i = 1 \ldots N \end{aligned}$$

where in the SHIW sample *s* is either equal to 2 or 3, y_{it-s} is a measure of earnings for individual *i* in period *t-s*, **x** is a column vector including a constant term and observed attributes, $\boldsymbol{\delta}$ is an associated parameter vector and u_{it-s} is an error term. Moreover, following Stewart and Swaffield (1999), I assume that there exists a monotone transformation *g*(.) such that the unobserved earnings component is standard normal distributed. Let λ_{t-s} be the low pay threshold

[9] The three equations structure resembles the one in Bingley et al. (1995). In that paper, however, the main equation is an ordered probit for the direction of movements across wage deciles, rather than a probit for low pay transition probabilities. Moreover, while Bingley et al. include in the attrition equation also employees who enter the earnings distribution during the transition, here I follow the approach of other attrition studies and only consider exits (see, for example, Lillard and Panis, 1998). The inclusion of entries implies that personal attributes can be observed only after the “decision” to remain the sample has taken place, while for exits they are observed before such decision takes place.

in period *t-s* and L_{it-s} be an indicator variable for the low pay event, $L_{it-s} = I(y_{it-s} \leq \lambda_{t-s})$, where *I*(*A*) equals 1 whenever *A* is true and 0 otherwise. The probability that an individual will be low paid in period *t-s* is:

$$\text{(2)} \qquad \begin{aligned} &\text{prob}(L_{it-s}=1) = \text{prob}(y_{it-s} \leq \lambda_{t-s}) = \text{prob}(g(y_{it-s}) \leq g(\lambda_{t-s})) = \\ &\Phi(g(\lambda_{t-s}) - \boldsymbol{\delta}'\mathbf{x}_{it-s}) = \Phi(\boldsymbol{\beta}'\mathbf{x}_{it-s}) \end{aligned}$$

where $\Phi(.)$ is the standard normal cumulative density function (c.d.f.), the new constant term in $\boldsymbol{\beta}$ subsumes the difference between $g(\lambda_{t-s})$ and the old constant in $\boldsymbol{\delta}$ and the coefficients associated with individual characteristics in $\boldsymbol{\beta}$ are the same as in $\boldsymbol{\delta}$, but with opposite sign. It has to be stressed that the use of the specification in (1) does not require any distributional assumption on wages or log-wages. Moreover, the non-linear treatment of the wage variable implicit in (1) corresponds to the idea that the wage process is not continuous, but some break occurs in correspondence of the low-pay threshold. In this way the effect of workers attributes on low-pay probabilities can be estimated directly; to obtain similar effects from usual (log-) wage regressions, distributional assumptions would be needed (see Lillard and Willis, 1978).

Next let r^*_{it} be some latent propensity of retention into the earnings distribution between periods *t-s* and *t*:

$$\text{(3)} \qquad r^*_{it} = \boldsymbol{\psi}'\mathbf{w}_{it-s} + v_{it}$$

where the error term v_{it} is distributed as in (1) and $\boldsymbol{\psi}$ and $\mathbf{w}$ are column vectors. If r^*_{it} is lower than some threshold (which can be set to 0 without loss of generality) individuals exit from the earnings distribution between *t-s* and *t*; otherwise they remain into the distribution so that their transition can be observed. Let R_{it} be a dummy indicator of the retention outcome: $R_{it}=I(r^*_{it}>0)$.

Finally, let us specify the earnings distribution of year *t* conditionally on the outcomes of both initial low pay and retention:

$$\text{(4)} \qquad h(y_{it}) = \begin{cases} \boldsymbol{\gamma}_1'\mathbf{z}_{it-s} + \varepsilon_{it} & \text{if} \quad L_{it-s}=1 \quad \text{and} \quad R_{it}=1 \\ \boldsymbol{\gamma}_2'\mathbf{z}_{it-s} + \varepsilon_{it} & \text{if} \quad L_{it-s}=0 \quad \text{and} \quad R_{it}=1 \end{cases}$$

where the $\boldsymbol{\gamma}$s and $\mathbf{z}_{it-s}$ are column vectors and *h*(.) is a monotonic unspecified transformation such that the error term ε_{it} is standard normally distributed.[10] The parameter vector in (4) switches according to the outcomes of initial low pay, i.e. the $\boldsymbol{\gamma}$s parameterise earnings transitions. Also, period *t* earnings distribution can not be observed if individuals exit from the sample of earnings

[10] Observed attributes are measured at the beginning of the transition in order to avoid simultaneity between changes in attributes and changes in wages. Note that since this equation refers to earnings conditional on lagged pay states and attrition, the error term differs from the one for unconditional earnings in (1).

recipients during the transition, i.e. when R_{it}=0. By applying a transformation similar to the one used in (2) for period *t-s* earnings, period *t* low pay probabilities may be written as follows:

$$\text{(5)} \qquad \text{prob}(L_t) = \begin{cases} \Phi(\boldsymbol{\eta}_1 ' \mathbf{z}_{it-s}) & \text{if} \quad L_{it-s} = 1 \quad \text{and} \quad R_{it} = 1 \\ \Phi(\boldsymbol{\eta}_2 ' \mathbf{z}_{it-s}) & \text{if} \quad L_{it-s} = 0 \quad \text{and} \quad R_{it} = 1 \end{cases}$$

In order to derive the likelihood function of the model some assumption on the joint distribution of the errors of (1), (3) and (4) is needed. Here I allow them to be jointly distributed as a three-variate normal (denoted by N_3) with zero means, unit variances and free correlation:

$$\text{(6)} \qquad (u_{it-s} \quad v_{it} \quad \varepsilon_{it}) \sim N_3[0 \quad 0 \quad 0; \;\; 1 \;\; 1 \;\; 1, \;\; \rho_1 \quad \rho_2 \quad \rho_3]$$

Correlation across equations allows for individual specific unobserved heterogeneity. Testing the statistical significance of the correlation coefficients in (6) provides a test for the exogeneity of the two selection mechanisms. In particular, ρ_1 measures correlation in unobservables between initial low pay and retention, indicating whether the initially poor are more or less likely to be in the balanced sample compared to those initially highly paid. ρ_2 measures correlation of unobservables between initial low pay and conditional low pay, showing whether the initially poor are more or less likely to persist or fall into low pay compared to the initially highly paid. Finally, ρ_3 measures correlation in unobservables between retention propensities and low pay transitions, indicating whether those in the balanced sample are more or less likely to persist in low pay or to fall into it compared to those exiting the distribution.

To summarise, the model consists of a low pay probit equation for period *t* with endogenous switching on the outcomes of period *t-s* low pay probit and endogenous partial observability depending upon the outcomes of the retention probit. Note that multiple selectivity (into initial low pay and into the balanced panel) takes place simultaneously, i.e. no assumption has been made about nesting sequences between the two selection equations. Moreover neither selection equation is conditioned on the other, a feature whose relevance will be clearer later on.

The likelihood contribution of individual *i* can be written as:

$$\text{(7)} \qquad \begin{aligned} \ell_i = {} & \Phi_3(k_i \boldsymbol{\eta}_1 ' \mathbf{z}_{it-s}, m_i \boldsymbol{\psi}' \mathbf{w}_{it-s}, q_i \boldsymbol{\beta}' \mathbf{x}_{it-s}; k_i m_i \rho_3, k_i q_i \rho_2, m_i q_i \rho_1)^{Lit\text{-}sRit} \times \\ & \Phi_3(k_i \boldsymbol{\eta}_2 ' \mathbf{z}_{it-s}, m_i \boldsymbol{\psi}' \mathbf{w}_{it-s}, q_i \boldsymbol{\beta}' \mathbf{x}_{it-s}; k_i m_i \rho_3, k_i q_i \rho_2, m_i q_i \rho_1)^{(1\text{-}Lit\text{-}s)Rit} \times \\ & \Phi_2(m_i \boldsymbol{\psi}' \mathbf{w}_{it-s}, q_i \boldsymbol{\beta}' \mathbf{x}_{it-s}; m_i q_i \rho_1)^{(1-Rit)} \\ & k_i = 2L_{it} - 1; \quad m_i = 2R_{it} - 1; \quad q_i = 2L_{it-s} - 1 \end{aligned}$$

where $\Phi_j(.)$ is the j-variate normal c.d.f.. To solve the computational problem posed by the presence of three-variate normal integrals I utilise simulated maximum likelihood (SML) estimation, so that in estimation $\Phi_3(.)$ is replaced by its simulated counterpart $\tilde{\Phi}_3(.)$. In particular, I adopt the so-called *GHK simulator*.[11]

Note that although the $\boldsymbol{\eta}$ vectors in (7) are estimated conditioning on initial low pay, the whole expression refers to the joint probability of the data. Transition probabilities can be computed as:

$$\hat{\text{prob}}(L_{it} = 1 \mid L_{it-s} = 1, R_{it} = 1) = \frac{\tilde{\Phi}_3(\hat{\boldsymbol{\eta}}_1'\mathbf{z}_{it-s}, \hat{\boldsymbol{\psi}}'\mathbf{w}_{it-s}, \hat{\boldsymbol{\beta}}'\mathbf{x}_{it-s}; \hat{\rho}_3, \hat{\rho}_2, \hat{\rho}_1)}{\Phi_2(\hat{\boldsymbol{\psi}}'\mathbf{w}_{it-s}, \hat{\boldsymbol{\beta}}'\mathbf{x}_{it-s}; \hat{\rho}_1)}$$

(8)

$$\hat{\text{prob}}(L_{it} = 1 \mid L_{it-s} = 0, R_{it} = 1) = \frac{\tilde{\Phi}_3(\hat{\boldsymbol{\eta}}_2'\mathbf{z}_{it-s}, \hat{\boldsymbol{\psi}}'\mathbf{w}_{it-s}, -\hat{\boldsymbol{\beta}}'\mathbf{x}_{it-s}; \hat{\rho}_3, -\hat{\rho}_2, -\hat{\rho}_1)}{\Phi_2(\hat{\boldsymbol{\psi}}'\mathbf{w}_{it-s}, -\hat{\boldsymbol{\beta}}'\mathbf{x}_{it-s}; -\hat{\rho}_1)}$$

where hats denote estimates.

In order identify the model, exclusion restrictions are needed in terms of variables entering the **x** or **w** vectors but not the **z** one, i.e. variables affecting either initial low pay or retention but, conditional on these, with no effect on low pay transition.[12] Heckman (1981b) suggests that initial conditions can be instrumented by using information prior to labour market entry. Stewart and Swaffield (1999) use indicators of parental education and occupation. Since 1993 the SHIW has included questions on occupation and education of the household head's and spouse's parents.[13] A set of 10 dummies for manual occupation, non employment, education equal to or greater than high school and missing information on education or occupation was derived for each parent and has been used as instrument. In addition, as pointed out by Stewart and Swaffield (1999), a quadratic term in experience (which enters the equation for initial low pay) can be excluded from the equation for low pay transition given its interpretation of wage change equation. Based on this assumption, the equation for initial low pay is over-identified and the validity of parental background as instruments can be tested. Identification of the retention equation requires variables affecting employment probabilities plus information on participation into the survey, the implementation of interviews and personal characteristic of the interviewer (see, for example, Zabel, 1998). While there is no clear *a priori* on variables of the first kind that can be excluded from the transition equation, information of the second kind is not available in the SHIW. However, as pointed out above, neither of the two selection equations is conditional

---

[11] Geweke-Hajivassiliou-Keane. See Hajivassiliou and Ruud (1994) and Gourieroux and Monfort (1996, pp. 93-107) for discussions of simulation methods and their application to maximum likelihood estimation of multivariate limited dependent variable models. The simulator is not used for bivariate c.d.f.'s which are normally available within statistical packages.

[12] Alternatively, one could rely on the functional form of the model.

[13] For those employees who were "child" in the interviewed household, the information has been recovered from the household questionnaire, while for "other relatives" or "non relatives" information has been coded as missing.

on the other, implying that the retention equation can be identified using the same set of instruments used for initial conditions.

The endogenous switching structure of the model allows us to investigate the issue of state dependence. First of all, a measure of aggregate state dependence (ASD) can be computed from estimated parameters as the difference in average conditional low pay probabilities, with averages taken over the samples of the initially low paid and high paid in the balanced sample:

$$\begin{aligned}\text{ASD} = & \{\textstyle\sum_{\{i:Lit-s=1,Rit=1\}} \hat{\text{prob}}(L_{it}=1 \mid L_{it-s}=1, R_{it}=1) / \sum_{\{i:Lit-s=1,Rit=1\}} L_{it-s}\} - \\ & \{\textstyle\sum_{\{i:Lit-s=0,Rit=1\}} \hat{\text{prob}}(L_{it}=1 \mid L_{it-s}=0, R_{it}=1) / \sum_{\{i:Lit-s=0,Rit=1\}} (1-L_{it-s})\}\end{aligned} \tag{9}$$

Secondly the hypothesis of absence of genuine state dependence can be formulated as $H_0: \eta_1 = \eta_2$, i.e. the impact of personal attributes on conditional low pay probabilities does not depend upon past low pay experiences.[14] Finally, an indicator of GSD may be defined as the difference in conditional low pay probabilities an average individual would have experienced had she started the transition from below or above the low pay threshold, the average being taken over the balanced sample of earnings recipients:

$$\text{GSD} = (\textstyle\sum_{i} R_{it})^{-1} \sum_{i:Rit=1} \left\{ \dfrac{\tilde{\Phi}_3(\hat{\boldsymbol{\eta}}_1'\mathbf{z}_{it-s}, \hat{\boldsymbol{\psi}}'\mathbf{w}_{it-s}, \hat{\boldsymbol{\beta}}'\mathbf{x}_{it-s}; \hat{\rho}_3, \hat{\rho}_2, \hat{\rho}_1)}{\Phi_2(\hat{\boldsymbol{\psi}}'\mathbf{w}_{it-s}, \hat{\boldsymbol{\beta}}'\mathbf{x}_{it-s}; \hat{\rho}_1)} - \dfrac{\tilde{\Phi}_3(\hat{\boldsymbol{\eta}}_2'\mathbf{z}_{it-s}, \hat{\boldsymbol{\psi}}'\mathbf{w}_{it-s}, -\hat{\boldsymbol{\beta}}'\mathbf{x}_{it-s}; \hat{\rho}_3, -\hat{\rho}_2, -\hat{\rho}_1)}{\Phi_2(\hat{\boldsymbol{\psi}}'\mathbf{w}_{it-s}, -\hat{\boldsymbol{\beta}}'\mathbf{x}_{it-s}; -\hat{\rho}_1)} \right\} \tag{10}$$

[Table 4 around here]

IV. Results

Results obtained by estimating the simulated three-variate probit of the last Section on the pooled transitions sample are presented in Table 4.[15] Explanatory variables for the transition equation included in the **z** vector are a gender dummy, potential labour market experience, an education dummy, occupational dummies, dummies for industrial affiliation, employer size dummies, regional dummies and a dummy for the 1995-98 transition. The **x** vector includes all variables in the **z** vector plus a quadratic in potential labour market experience and the set of parental background dummies. Finally, following the discussion about identification of the

[14] In a dynamic random effect probit in which the effect of lagged states is subsumed into a dummy variable genuine state dependence is tested by testing the significance of the estimated coefficient on that dummy, see e.g. Arulampalam et al. (2000). The test proposed in this paper generalises that framework to the case in which the whole parameter vector associated to personal characteristics switches according to lagged states.

[15] I assume that observations from the two transitions are independent. I also experimented with a robust variance estimator which accounts for repeated observation on the same individual in the two transitions and found differences in results to be irrelevant.

retention equation in the previous Section, the **w** vector is set equal to **x**; I will refer to a unique **x** vector from now onwards.

Results are presented in terms of “marginal effects” of explanatory variables on the conditional low pay probabilities given by (8). A change in an element of **z** also implies a change in the corresponding element of the **x** vector, thus changing not only the conditional probability, but also the conditioning ones. In order to hold conditioning events constant when computing marginal effects on transition probabilities I proceed as follows.[16] I compute predicted probabilities for the two conditioning events (using estimated parameters from the three-variate probit, the **x** vector and univariate normal c.d.f.’s) and average them over the relevant samples, i.e. observations in the balanced sample for the retention probability and observations in the balanced sample and below or above initial low pay for the probabilities of initial low or high pay. I next compute the arguments of these average predicted probabilities and use them into the multivariate normal c.d.f.’s of (8), thus holding the probabilities of the conditioning events fixed. Finally, each marginal effect is calculated as the change in the conditional probabilities of (8) induced by a change in an element of **z** with respect to a base category. The base category is given by an employee with 20 years of potential labour market experience and a value of 0 in all the dummy variables in **z**. For dummy variables the effect is the change in transition probabilities with respect to the base category when the dummy changes from 0 to 1. For potential labour market experience the effect is the one determined by a change of the variable to 30 years.

It is instructive to begin our discussion of results by considering the estimated covariance matrix of error terms. The three correlation coefficients are statistically significant at usual confidence levels: thus both initial conditions and retention are endogenous for the estimation of low pay transitions and should not be ignored. The correlation between unobservables of initial low pay and retention is negative reflecting the higher propensity to exit from the balanced panel of the low paid compared to the higher paid. Correlation of unobservables between initial conditions and conditional low pay probabilities is also negative, meaning that those who begin the transition below the low pay threshold are less likely to experience a small earning change -and thence to be low paid at the end of the transition- compared to the higher paid, a finding reflecting Galtonian regression towards the mean. Finally, correlation in unobservables between sample retention and conditional low pay probability is positive. Individuals in the balanced earnings sample have a higher probability to either persist in low pay (if they are low paid at the start of the transition) or to fall into it (if they are initially high paid). This last finding combines evidence from the two other correlation coefficients. Given that the low paid have a lower retention probability (ρ_1<0) and a lower conditional low pay probability (ρ_2<0) compared to the higher paid, then the conditional low pay probability will be higher in the balanced sample compared to what it would had been in the absence of earnings attrition.

The bottom panel of Table 4 also reports results from tests for the validity of parental background variables as instruments. While these variables do not appear to be significant in the equations for conditional low pay probabilities, their simultaneous exclusion from the two selection equations is rejected. These results support the use of parental background variables as instruments for the multiple selectivity equations.

The average (over the balanced earnings sample) predicted conditional low pay probability is reported at the top of the Table. The model replicates the aggregate transition rates of Table 3. The stylised individual used as a reference for the computation of marginal effects has low pay persistence and entry rates higher then the sample average ones. Comparing the reference individual with a female employee with otherwise similar characteristics shows that the latter experiences larger conditional low pay probabilities, between 5 and 8 percentage points depending upon the case considered, while the underlying estimated coefficient for the female dummy is statistically significant at usual confidence levels. Increasing labour market experience from 20 to 30 years, on the other hand, reduces conditional low pay probabilities by a lesser extent and the underlying coefficients do not always appear to be precisely estimated. Holding a high school degree reduces the probability of persistence below the lowest threshold by 14 percentage points, while the effect is smaller in size, but with underlying coefficients still precisely estimated, when the higher cut-off point or drops into low pay are taken into account. Marginal effects for occupation reveal an asymmetric impact on conditional low pay probabilities: while for the initially high paid in non-manual jobs the probability of falling into low pay is some 8 to 13 percentage points lower compared to high paid manual workers, for employees below the low pay threshold no statistically significant association can be detected. For high-level non-manual workers this finding might reflect a small cell size problem. For low-level non-manual workers and teachers, on the other hand, this result suggests that factors which keep employees out of low pay may lose their effectiveness once low pay has been experienced. The public sector dummy displays the same kind of asymmetric effect noted above for occupation dummies, but only for the lowest threshold. Marginal effects for private sector industrial affiliation, on the other hand, do not reveal any clear pattern. Conditional low pay

[16] I generalise the procedure proposed by Stewart and Swaffield (1999) for the bivariate probit case.

probabilities are significantly lower for employees in medium sized private sector firms compared to small firms. Conversely, when large sized firm are taken into account, the kind of asymmetric impact characterising occupation dummies applies. An asymmetric impact of observed factors on conditional low pay probabilities applies also for regional dummies, but this time in the opposite direction. For example, north-western employees have a probability of low pay persistence that is 10 to 22 percentage points lower than that of workers from the South or Islands, while no significant differential characterises the probability of falling into low pay from higher pay. Finally, we can see that conditional low pay probabilities do not vary significantly over transitions. Since the latter transition occurs over a wider interval, this evidence points towards increasing distributional rigidity in the second half of the 1990s.

Results about differences in the impact of personal attributes on conditional low pay probabilities between workers above and below the low pay threshold are consistent with the existence of GSD effects. A formal test for the absence of GSD (formulated as equality of parameter vectors in the two conditional low pay equations) is reported at the bottom of Table 3. For both low pay thresholds the null hypothesis H_0:η_1=η_2 is overwhelmingly rejected. Measures of ASD and GSD computed according to (9) and (10) are also reported. GSD constitutes a substantial share of aggregate figures, the ratio GSD/ASD being approximately 53% for both thresholds. These figures are in line with the ones reported by Stewart and Swaffield (1999) for Britain. The test and measures of state dependence thence indicate that a relevant share of low pay persistence may be ascribed to past low pay experiences, which modify individual tastes or constraints and make more difficult for individual to move onto the higher part of the distribution, irrespective of personal attributes.

V. Taking a longer run view

Results presented so far refer to low pay probabilities conditional on both retention and one period lagged low pay states. This Section proposes an extension of the analytical framework aimed at assessing the features of longer term low pay persistence. In particular, I will look at the three SHIW waves simultaneously and will estimate 1998 low pay probabilities conditional on earnings attrition and pay states in 1993 and 1995. The model presented in this Section has never been applied before to the analysis of earnings mobility – at least to my knowledge – and represents an intermediate analytical perspective between models of first order transitions like the one of Section III and low income spells analyses like the ones discussed in Jenkins (2000). As such, it allows studying the covariates of low pay persistence distinguishing between different sequences of previous low pay while controlling for the endogeneity issues outlined in

Section III. In addition, parameter estimates can be used to assess state dependence over the longer run.

[Table 5 around here]

Table 5 provides an illustration of transition patterns considering the three available waves simultaneously and using the first quintile as low pay threshold, with the relevant sample now being given by 1993 employees with valid earnings. The first row of the Table shows that a different treatment of earnings attrition is needed when modelling the three-years transition compared to the two-years case. The latent retention propensity needs now to cross two thresholds in order for employees to be in balanced panel of earnings recipients, i.e. being in sample between 1993 and 1995 and between 1995 and 1998. Treating retention as binary, i.e. distinguishing only between balanced panel versus non-balanced panel observations would imply a loss of information in estimating the 1995 earnings distribution. The Table also show that the probability of being in the balanced sample is much larger for those who start the transition from above rather then below the low pay threshold, again pointing towards the importance of jointly modelling earnings attrition and transitions. The second row of the Table looks at the 1993-98 balanced sample and reports earnings transition probabilities from 1993 to 1998. It can observed that while the probability of falling into low pay from higher pay is approximately the same as for the shorter term transitions of Table 3, the probability of low pay persistence is some 6 to 8 percentage points lower, as can be expected by the fact that a wider time windows is taken into account. Finally, the bottom line of the Table provides the probabilities of 1998 pay states conditional on 1993 and 1995 pay states. Employees who have been low paid in both 1993 and 1995 face a probability of being low paid in 1998 in the order of 68%, larger than the ones characterising two years transitions. Having entered low pay after an initial high pay experience is also associated with considerable low pay persistence, in the order of 45%; comparing this figure with the one for employee who have always been low paid suggests the existence of positive duration dependence at the aggregate level. Climbing out of low pay and falling back into it is a less likely but still relevant phenomenon, with an associated probability of 31%. Finally those who have never been low paid before 1998 have conditional low pay probability below 5%.

Modelling earnings mobility and attrition over three waves

Since it is no longer appropriate to treat retention into the earnings distribution as binary, I model retention outcome as a multiple response discrete ordered variable:

$$
(11) \qquad \begin{aligned} & r^*_{it} = \boldsymbol{\psi}'\mathbf{w}_{it-5} + v_{it} \quad ; \quad R_{it} = t(r^*_{it}) \\ & t(r^*_{it}) = \begin{cases} -1 & \text{if} \quad r^*_{it} \in (-\infty, \tau] \\ 0 & \text{if} \quad r^*_{it} \in (\tau, 0] \\ 1 & \text{if} \quad r^*_{it} \in (0, +\infty) \end{cases} \\ & i = 1, \ldots, N_{93}; \quad \tau < 0 \end{aligned}
$$

Expression in (11) uses the same notation as in (3), but there are differences to be stressed. First of all, the sample is restricted to 1993 employees with valid earnings since we are now studying a single three-year transition. Secondly, there is now an additional intermediate outcome of the earnings attrition process, i.e. having valid earnings in the first two years of the transition but not in the third.[17] Accordingly, I introduce an additional threshold in the support of r^*_{it}, τ, while holding the threshold for being a (three-year) balanced panel member normalised to 0. The mapping $t(.)$ transforms r^*_{it} into a discrete ordered variable R_{it}. Explanatory variables are measured at the beginning of the transition.

The 1993 earnings distribution is specified analogously to equation (1), the analogy being that 1993 is the starting year of the transition, with the dummy variable $L_{it\text{-}5}$ indicating the occurrence of low earnings at the start of the transition. The 1995 earnings distribution, on the other hand, can only be observed conditionally on staying in the sample of earnings recipients:

$$
(12) \qquad \begin{aligned} & g_{t-3}(y_{it-3}) = \begin{cases} \boldsymbol{\omega}'\mathbf{x}_{it-5} + e_{it-3} & \text{if} \quad R_{it} \geq 0 \\ \textit{unobserved} & \textit{otherwise} \end{cases} \\ & L_{it-3} = I(y_{it-3} \leq \lambda_{t-3}) \end{aligned}
$$

Finally, I study 1998 earnings conditionally on past pay states and retention:

$$
(13) \qquad \begin{aligned} & h(y_{it}) = \begin{cases} \boldsymbol{\gamma_1}'\mathbf{z}_{it-5} + \varepsilon_{it} & \text{if} \quad L_{it\text{-}5} = 1 \quad \text{and} \quad L_{it-3} = 1 \quad \text{and} \quad R_{it} = 1 \\ \boldsymbol{\gamma_2}'\mathbf{z}_{it-5} + \varepsilon_{it} & \text{if} \quad L_{it\text{-}5} = 0 \quad \text{and} \quad L_{it-3} = 1 \quad \text{and} \quad R_{it} = 1 \\ \boldsymbol{\gamma_3}'\mathbf{z}_{it-5} + \varepsilon_{it} & \text{if} \quad L_{it\text{-}5} = 1 \quad \text{and} \quad L_{it-3} = 0 \quad \text{and} \quad R_{it} = 1 \\ \boldsymbol{\gamma_4}'\mathbf{z}_{it-5} + \varepsilon_{it} & \text{if} \quad L_{it\text{-}5} = 0 \quad \text{and} \quad L_{it-3} = 0 \quad \text{and} \quad R_{it} = 1 \end{cases} \\ & L_{it} = I(y_{it} \leq \lambda_t) \end{aligned}
$$

Equations (1) (applied to the unique starting year, 1993), (11), (12) and (13) provide a framework for the analysis of three-years transition controlling for multiple responses in attrition. As before, errors are assumed to be jointly normally distributed:

$$
(14) \qquad (u_{it-5} \quad v_{it} \quad e_{it-3} \quad \varepsilon_{it}) \sim \mathrm{N}_4[\mathbf{0}, \boldsymbol{\Omega}]
$$

[17] There are few cases (42 observations) of “re-joiners”, i.e. employees who re-enter into the earnings distribution in 1998 after having left in 1995. As explained above (see Section III) I treat earnings attrition as an absorbing state (I borrow this definition from Zabel, 1998); consequently I ignore re-entries into the distribution and consider these cases as “attritors” also in estimation of the 1998 distribution.

where **0** is a row vector of zeros and the **Ω** matrix has diagonal elements equal to unity and extra-diagonal elements equal to cross-equations correlation coefficients, while N_4 is the four-variate normal density. After applying a change in parameters similar to the one applied in the case of the two-year transition to equations (12) and (13) we obtain vectors $\boldsymbol{\kappa}$ and $\boldsymbol{\eta_1}$ to $\boldsymbol{\eta_4}$, respectively, which index the probability of having earnings below the low pay threshold. Likelihood contributions for this model are given in the Appendix.

Additional identifying restrictions would be required for this model compared to the one in Section III. In fact, the 1995 earnings distribution is observed conditionally on $R_{it}>-1$, i.e. if individuals survive in the sample of income recipients after 1993. In turn, 1995 pay states enter the conditioning set for 1998 low pay probabilities. Thence, assuming, as we did in Section III, that parental background variables (plus the square of experience) enter earnings levels but not earnings changes, we would need to include these variables into the 1995 earnings equations, implying that additional instruments should be included into the retention equation. However, as pointed out earlier, there are no variables in the SHIW that could be used for this purpose. Thence, I base identification of the relationship between retention and 1995 earnings levels on functional form, and the vector of regressors will be the same in all the three conditioning equations.

[Table 6 around here]

Results

Results from simulated estimation of the four-variate probit model are reported in Table 6. The level of aggregation of explanatory variables is higher compared to the two-year model of Section III since, as seen in Table 5, cell size is now tiny. By first considering the estimates of the cross-equations correlation coefficients reported at the bottom of the Table, it can be observed that all the patterns emerged from the two-year model are confirmed. Those who earn below the low pay threshold of a given year are more likely to drop out of the earnings distribution during the transition compared to the higher paid, as indicated by the negative estimates of ρ_1 and ρ_4, although in the second case the estimate precision is lower. The coefficient ρ_2 measures reduced form correlation between low pay probabilities in 1993 and 1995 and it is positive and precisely estimated. The correlation between initial conditions and low pay transition measured by ρ_3 and ρ_6 is negative (as it was in the case of the two-year model), indicating the presence of Galtonian regression effects. The correlation between unobservables of low pay transition and 1995 low pay probabilities is not precisely estimated: it may well be that this effect is absorbed by the simultaneous control for correlation between

1993 initial conditions and transition and reduced form low pay correlation (i.e. ρ_3 and ρ_2, respectively). Finally, the correlation between retention and transition probabilities measured by ρ_5 is positive and precisely estimated. Comments analogous to those made when the result was found for the two years case also apply now.

Among those who experienced low pay in 1993 and 1995, the probability of experiencing low pay in 1998 is lower for non-manual workers compared to employees in blue collar occupations, as well as for northern workers compared to workers living in the rest of the country.[18] This latter effect can also be observed among those who entered low pay in 1995 after having been in the high pay area of the distribution in 1993, whereas the remaining coefficients estimated for this group are not statistically significative at usual confidence levels. For the groups of employees who climb out of low pay in 1995 and fall back into it in 1998, no clear association can be detected between 1998 conditional low pay probabilities and personal attributes. Finally, for employees who did not experience low pay in 1993 and 1995, estimated coefficients indicate quite clearly that the probability of falling into low pay in 1998 is higher for female workers, less educated employees, blue collar workers and workers in the public sector.

Estimated coefficients from the four-variate probit can be used to investigate the extent of GSD in longer term transition in a fashion similar to the one adopted in Sections III and IV. The bottom panel of the Table reports 1998 conditional low pay probabilities estimated for each of the four sequences of past low pay; again, we can note how model predictions replicate the aggregate figures of Table 5. The Table also reports measures of ASD and in particular, the 1998 conditional low pay probability of those who have always been low paid is contrasted with the ones from the other sequences of past low pay. The next row in the Table reports a test for GSD, the null hypothesis given by the equality of coefficients vectors across the four low pay sequences (H_0: $\eta_1=\eta_2=\eta_3=\eta_4$). The null hypothesis is clearly not rejected, an outcome which is opposite to the findings of the previous Section. However, it might be that the test is biased by the tiny cell size, which drives estimated coefficients towards zero and thence towards equality across low pay sequences. Finally, measures of GSD analogous to those of the previous Section are reported. For each balanced panel observation 1998 conditional low pay probabilities have been computed for each sequence of past low pay. Next the differences between the probability conditional on having always been low paid and each of the three other sequences have been

[18] Estimates precision is not particularly high. Besides the aforementioned cell size problem, estimates imprecision could also be due to the use of regressors measured at the start of the transition, i.e. five years before the outcome of interest.

computed. Reported figures are the average of these differences over the balanced sample. Again, we should expect tiny cell size to bias the measurement of GSD downward, a *caveat* to bear in mind when interpreting results. What emerges from these calculations is that the ratio GSD/ASD is approximately 25% when columns (1) and (2) are compared. Thence, when comparing conditional low pay probabilities between employees who entered low pay in the previous period and those who had always been observed in low pay, the incidence of GSD is much lower than the one emerged in the two years model. This finding suggests that whatever the causes of GSD, they produce their effect as soon as individuals are touched by low pay, while subsequent low pay experiences contribute to a lesser extent. The GSD/ASD ratio is instead higher at approximately 68% for the two other sequences of past low pay. Interestingly, the low pay experience in the first year of observation does not seem to matter here, i.e. the relevance of GSD is the same for individuals who have never been low paid and for those who managed to escape from low pay during the 1993-95 transition.

VI. Concluding remarks

This paper has used data from the Survey on Household Income and Wealth to analyse the earnings mobility of low paid Italians. In particular models of low pay transition probabilities have been estimated while controlling for endogenous initial conditions and endogenous attrition from the earnings distribution. With this aim, Simulated Maximum Likelihood techniques have been used.

Results from models of wave-to-wave transitions indicate that both initial conditions and attrition are endogenous and should be properly controlled for. In particular, results on earnings attrition suggest that employees below the low pay threshold of a given year are less likely to survive into the earnings distribution of the next observation period compared to higher paid individual, a symptom of higher instability of low paid employment. The analysis of the relationship between personal attributes and low pay transition probabilities has shown that employees with low educational qualifications, female employees and southern workers have higher risks of being trapped into low pay. The probability of dropping into low pay, on the other hand, appears to be associated with manual jobs and with jobs in the metal-manufacturing industry and in small firms.

The analysis also indicates that state dependence effects play a relevant role in creating low pay traps: it is the experience of low pay which modifies the economic environment faced by individuals, increasing the probability of future low pay experiences irrespective of personal attributes. While the paper does not investigate the causes of these effects, these results points towards the need of policies targeted on the whole pool of low paid employees, rather than on specific "problem groups" within it.

I also studied transition probabilities allowing for second order dynamics. Results show that longer term low pay traps tend to occur among manual and southern workers. On the other hand, female employees and employees with low educational attainment are more likely than otherwise similar individuals to drop into low pay after having stayed out of it in the two periods prior to observation. Investigation of state dependence effects show that the bulk of it occurs at the beginning of a low pay spell, while the contribution of subsequent low pay experiences is less pronounced. Caution has to be exerted when considering results from this latter model due to tiny sample sizes which prompts for future applications on richer data.

Appendix : Likelihood contributions for the 3-year transition model

This Appendix reports likelihood contributions for the model of Section V.

If R_{it}=1, i.e. for employees with valid earnings in 1993, 1995 and 1998, likelihood contributions are given by:

(A.1)
$$\begin{aligned}\ell_i &= \Phi_4(k_i \boldsymbol{\eta}_{\boldsymbol{\theta}}'\mathbf{z}_{it-5}, p_i\boldsymbol{\kappa}'\mathbf{x}_{it-5}, \boldsymbol{\psi}'\mathbf{w}_{it-5}, q_i\boldsymbol{\beta}'\mathbf{x}_{it-5}; \\ &\qquad k_i p_i \rho_6, k_i\rho_5, p_i\rho_4, k_i q_i\rho_3, p_i q_i\rho_2, q_i\rho_1) \\ k_i &= 2L_{it}-1; \quad p_i = 2L_{it\text{-}3}-1; \quad q_i = 2L_{it-5}-1\end{aligned}$$

with $\boldsymbol{\theta}$=1 if $L_{it\text{-}3}$=$L_{it\text{-}5}$=1, $\boldsymbol{\theta}$=2 if $L_{it\text{-}3}$=1 and $L_{it\text{-}5}$=0, $\boldsymbol{\theta}$=3 if $L_{it\text{-}3}$=0 and $L_{it\text{-}5}$=1 and $\boldsymbol{\theta}$=4 if $L_{it\text{-}3}$=$L_{it\text{-}5}$=1. If R_{it}=0, i.e. when individuals exit from the earnings distribution in 1995, likelihood contributions are given by:

(A.2)
$$\begin{aligned}\ell_i &= \Phi_3(p_i\boldsymbol{\kappa}'\mathbf{x}_{it-5}, -\boldsymbol{\psi}'\mathbf{w}_{it-5}, q_i\boldsymbol{\beta}'\mathbf{x}_{it-5}; \text{-}p_i\rho_4, p_i q_i\rho_2, \text{-}q_i\rho_1) - \\ &\quad \Phi_3(p_i\boldsymbol{\kappa}'\mathbf{x}_{it-5}, \tau - \boldsymbol{\psi}'\mathbf{w}_{it-5}, q_i\boldsymbol{\beta}'\mathbf{x}_{it-5}; \text{-}p_i\rho_4, p_i q_i\rho_2, \text{-}q_i\rho_1)\end{aligned}$$

Finally, if R_{it}=-1, i.e. for observations with valid earnings only at the start of the transition, likelihood contributions take the following form:

(A.3)
$$\ell_i = \Phi_2(-\boldsymbol{\psi}'\mathbf{w}_{it-5}, q_i\boldsymbol{\beta}'\mathbf{x}_{it-5}; \text{-}q_i\rho_1)$$

Multivariate normal c.d.f. 's of order 3 and 4 are computed via simulation applying the GHK simulator.

The cross-equations correlation coefficients have the following meaning:

- ρ_1=correlation between 1993 unconditional low pay probability and retention
- ρ_2=reduced form correlation between low pay probabilities in 1993 and 1995 (1995 conditional on retention)

- ρ_3=correlation between 1993 unconditional low pay probability and conditional 1998 low pay probability
- ρ_4=correlation between retention and 1995 low pay probability (conditional on retention)
- ρ_5=correlation between retention and 1998 conditional low pay probability
- ρ_6=correlation between 1998 conditional low pay probability and 1995 low pay probability (conditional on retention)

References

Arulampalam, W., Booth, A.L. and Taylor, M.P. (2000): Unemployment Persistence. *Oxford Economic Papers*, 52, pp 24-50.

Bingley, P., Bjørn, J. and Westergård-Nielsen, N. (1995): Determinants of Wage Mobility in Denmark 1980-1990. *CLS* working paper, University of Aarhus, 10-95.

Cappellari, L. (2000): Low Wage Mobility in the Italian Labour Market. *International Journal of Manpower*, 21, pp. 264-90.

D'Alessio, G. and Faiella, I. (2000): I Bilanci delle Famiglie Italiane nell'Anno 1998. *Supplementi al Bollettino Statistico*, 22, Banca d'Italia.

Dickens, R. (2000): Caught in a trap? Wage Mobility in Great Britain: 1975-1994. *Economica*, 67, pp. 477-97.

Freeman, R.B. (1996): The Minimum Wage as a Redistributive Tool. *The Economic Journal*, 106; pp. 639-49.

Gregory, M. and Elias, P. (1994): Earnings Transitions of the Low Paid in Britain, 1976-91: a Longitudinal Study. *International Journal of Manpower*, 15; pp. 170-88.

Gourieroux, C. and Monfort, A. (1996): *Simulation-Based Econometric Methods*. Oxford: Oxford University Press.

Hajivassiliou, V. and Ruud, P. (1994): Classical Estimation Methods for LDV Models Using Simulation. in Engle R. and McFadden D. (eds.), *Handbook of Econometrics, vol. IV*, Amsterdam: North Holland.

Heckman, J.J. (1981a): Heterogeneity and State Dependence. in Rosen S. (eds), *Studies in Labour Markets*, Chicago: Chicago University Press.

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. and McFadden D. (eds.), *Structural Analysis of Discrete Data with Economic Applications*, Cambridge: MIT Press.

Jenkins, S.P. (2000): Modelling Household Income Dynamics. *Journal of Population Economics*, 13, pp 529-67.

Lillard, L.A. and Willis, R.J. (1978): Dynamic Aspects of Earnings Mobility. *Econometrica*, pp. 985-1012.

Lillard, L.A. and Panis, C.W.A (1998): Panel Attrition from the PSID. *Journal of Human Resources*, 33, pp.437-57.

Lucifora, C. (1998): Working Poors? An Analysis of Low Wage Employment in Italy. in *Low Pay and Earnings Mobility in Europe*, Asplund, R., Sloane, P. and Thedossiou, I. (eds.), Cheltenham: Edward Elgar.

OECD (1993): Earnings Inequality: Changes in the 1980s. *Employment Outlook*, pp. 157-184.

OECD (1996): Earnings Inequality, Low-paid Employment and Earnings Mobility. *Employment Outlook*, pp. 59-108.

Smith, R.E. and Vavricheck, B. (1992): The Wage Mobility of Minimum Wage Workers. *Industrial and Labour Relations Review*, 46; pp. 82-8.

Stewart, M.B. (1999): Low Pay, No Pay Dynamics. in *Persistent Poverty and Lifetime Inequality: The Evidence*, Centre for Analysis of Social Exclusion, London School of Economics, Report no.5.

Stewart, M.B and Swaffield, J.K. (1999): Low Pay Dynamics and Transition Probabilities. *Economica*, 66; pp. 23-42.

Zabel, J.E. (1998): An Analysis of Attrition in the PSID and the Survey of Income and Program Participation. *Journal of Human Resources*, 33, pp.479-506.

Table 1: Sample means (col. 1-3) and probability of being in the balanced panel (col. 4)

	(1)	(2)	(3)	(4)
Female	0.36	0.36	0.35	0.66
Male	0.64	0.64	0.65	0.68
Potential labour market experience	19.31	19.11	19.23	0.67[a]
				0.57[a]
Education<High school	0.46	0.48	0.43	0.63
Education≥ High school	0.54	0.52	0.57	0.70
Blue collar	0.43	0.45	0.40	0.62
White collar (low level) - Teacher	0.48	0.46	0.50	0.71
White collar (high level) – Manager – Magistrate - Professor	0.10	0.09	0.10	0.71
Manufacturing	0.28	0.30	0.28	0.66
Agriculture	0.02	0.03	0.02	0.51
Construction	0.05	0.06	0.05	0.57
Retail trade	0.09	0.09	0.08	0.59
Transport and Communication	0.03	0.03	0.03	0.61
Financial and related services	0.07	0.06	0.06	0.65
Personal and household services	0.03	0.03	0.02	0.56
Public sector	0.42	0.39	0.46	0.73
Firm size ≤19 [b]	0.39	0.41	0.35	0.57
20≤ Firm size≤99	0.23	0.24	0.23	0.64
100≤ Firm size≤499	0.15	0.15	0.17	0.71
Firm size ≥500	0.22	0.21	0.24	0.70
North-west	0.23	0.25	0.22	0.65
North-east	0.23	0.22	0.24	0.68
Centre	0.20	0.22	0.20	0.69
South and Islands	0.34	0.31	0.34	0.66
1993	0.49	0.51	0.59	0.81
1995	0.51	0.49	0.41	0.54
Number of observations	5581	11282	3730	5581

Notes:

Pooled SHIW data for 1993 and 1995.

Full time employees aged 18-58 if female and 18-60 if male.

Column (1) considers only employees from panel households

Column (2) considers the whole SHIW sample (employees in both panel and non-panel households).

Column (3) considers only employees from panel households and with valid earnings in two consecutive waves (balanced panel)

Column (4) provides the proportion of observations in the balanced panel conditional on the indicated personal characteristic

(a) Figures computed on the samples with less than 20 years and more than 30 years of experience

(b) Estimates by firm size are conditional on being a private sector employee

Table 2: Descriptive statistics of the distribution of nominal hourly earnings

	1993	**1995**	**1998**
p10	6.73	7.05	7.97
p20	8.01	8.55	9.44
p30	8.93	9.62	10.58
p50	10.81	11.54	12.50
p90	19.42	20.03	21.88
2/3 median	7.21	7.69	8.33
Mean	12.32	12.83	14.22
Std. Dev.	7.53	6.69	8.29
Log(p75/p25)	0.22	0.22	0.22
Log(p90/p10)	0.46	0.45	0.44
Log(p99/p1)	1.04	0.99	1.00
CPI (1993=100)[(a)]		108.70	116.90
Number of observations	5686	5554	4934

Notes:
SHIW cross-sections - Full time employees aged 18-58 if female and 18-60 if male. Thousands of lire. (a) source ISTAT

Table 3: Low pay transition probabilities

	Sample and low pay definition	destination state	high pay	low pay	out of the earnings distribution of full time employees								n.obs.
					partime	self-empl	unempl.	housewife	retired	student	other	attrited	
		origin state											
(1)	transition 1993-95	high pay	93.8	6.2									1871
	1st quintile	low pay	43.79	56.21									338
(2)	transition 1995-98	high pay	93.12	6.88									1293
	1st quintile	low pay	42.11	57.89									228
(3)	pooled transitions	high pay	93.52	6.48									3164
	1st quintile	low pay	43.11	56.89									566
(4)	pooled transitions	high pay	90.93	9.07									2834
	3rd decile	low pay	32.25	67.75									896
(5)	pooled transitions	high pay	65.31	4.52	30.17								4531
	1st quintile	low pay	23.67	31.23	45.1								1031
(6)	pooled transitions	high pay	65.31	4.52	1.32	1.77	1.88	0.35	5.32	0.09	0.35	19.09	4531
	1st quintile	low pay	23.67	31.23	3.49	3.01	8.63	1.94	2.13	0.19	0.97	24.73	1031

Notes:
SHIW data - Full time employees from panel households aged 18-58 if female and 18-60 if male.
Rows (1) to (4) consider balanced panels
Row (5) includes employees with valid earnings in the starting year but not in the arrival one
Row (6) distinguishes between exits from the earnings distribution and attritors

Table 4: Results[(a)] from SML estimation[(b)] of trivariate probit models for conditional low pay probabilities

Low pay threshold	**First quintile**				**Third decile**			
Initial pay state	**Low pay**		**High pay**		**Low pay**		**High pay**	
Average prediction	0.57		0.06		0.68		0.09	
Base category[(c)]	0.74		0.17		0.81		0.21	
Deviations from base category								
Female	0.07	(1.67)	0.08	(3.06)	0.05	(1.74)	0.08	(2.97)
30 years of potential labour market experience	-0.02	(1.05)	-0.02	(2.12)	-0.03	(2.15)	-0.01	(1.00)
Education≥ High school	-0.14	(2.37)	-0.06	(2.64)	-0.08	(1.92)	-0.06	(2.26)
White collar (low level) - Teacher	-0.05	(0.83)	-0.09	(3.74)	-0.06	(1.18)	-0.10	(3.89)
White collar (high level) – Manager – Magistrate - Professor	0.04	(0.29)	-0.08	(2.20)	0.05	(0.52)	-0.15	(4.01)
Public sector	0.04	(0.52)	-0.11	(4.20)	-0.15	(2.11)	-0.11	(3.34)
Agriculture	-0.10	(1.01)	0.03	(0.52)	-0.13	(1.77)	0.0002	(0.002)
Construction	0.09	(1.47)	0.00	(0.10)	0.01	(0.11)	0.05	(1.00)
Retail trade	-0.06	(1.03)	0.05	(1.26)	0.02	(0.52)	0.05	(1.16)
Transport and Communication	0.02	(0.14)	-0.08	(1.37)	-0.07	(0.81)	-0.03	(0.45)
Financial and related services	0.06	(0.69)	0.00	(0.02)	0.08	(1.33)	-0.07	(1.28)
Personal and household services	-0.11	(1.14)	0.04	(0.58)	-0.02	(0.30)	-0.05	(0.69)
20≤ Firm size≤99	-0.10	(1.68)	-0.06	(2.20)	-0.07	(1.83)	-0.05	(1.51)
100≤ Firm size≤499	-0.17	(1.75)	-0.06	(1.83)	-0.08	(1.55)	-0.06	(1.73)
Firm size ≥500	0.02	(0.24)	-0.10	(3.75)	-0.03	(0.50)	-0.11	(3.22)
North-west	-0.22	(3.17)	-0.03	(1.12)	-0.11	(2.43)	0.01	(0.32)
North-east	-0.16	(2.69)	-0.04	(1.47)	-0.04	(0.99)	-0.02	(0.70)
Centre	-0.10	(1.64)	0.01	(0.32)	0.02	(0.65)	0.03	(1.03)
transition 1995-98	-0.04	(0.83)	0.003	(0.12)	-0.10	(2.43)	0.0001	(0.002)

Table 4 (continued)

Low pay threshold	First quintile		Third decile	
ρ1 (initial conditions-retention)	-0.069	(2.18)	-0.092	(3.17)
ρ2 (initial conditions-transition)	-0.427	(3.32)	-0.351	(2.81)
ρ3 (retention-transition)	0.299	(1.88)	0.343	(2.16)
test1 (df=20) [d]	20.75	*0.4119*	19.15	*0.5122*
test2 (df=20) [e]	82.95	*0.0000*	91.18	*0.0000*
test 3 (df=20) [f]	62.65	*0.0000*	82.73	*0.0000*
State dependence [g]	0.50	0.26	0.59	0.32
Model chi2(df=98)	2182.86	*0.0000*	2541.48	*0.0000*
Log likelihood	-5696.67		-6350.99	
Number of observations	5535		5535	

Notes:SHIW data, pooled transitions 1993-95 and 1995-98 - Full time employees aged 18-58 if female and 18-60 if male

(a) Marginal effects, see text for computation. Absolute t-ratios in parentheses refer to SML coefficients

(b) GHK simulator with 75 random draws

(c) Male, 20 years of potential experience, blue collar worker, manufacturing, firm size <20, lives in the South or Islands, 1993-95 transition

(d) LR test of exclusion of parental background dummies from conditional low pay equations, p-value in Italic

(e) Wald test of exclusion of parental background dummies from selection equations, p-value in Italic

(f) Wald test of equality of coefficients in conditional low pay equations, p-value in Italic

(g) Aggregate state dependence (left) and genuine state dependence (right), see text for computation

Table 5: Retention and low pay conditional probabilities – Three years transition

(1)	Observed in	1993 only	1993 and 1995	1993,1995 and 1998	n. obs
	1993 pay state				
	High pay	16.17	37.54	46.28	2232
	Low pay	30.88	40.70	28.43	489
(2)	**1998 pay state**	High pay	Low pay		
	1993 pay state				
	High pay	93.22	6.78		1033
	Low pay	49.64	50.36		139
(3)	**1998 pay state**	High pay	Low pay		
	1993 and 1995 pay states				
	L_{93},L_{95}	32.43	67.57		74
	H_{93},L_{95}	55.36	44.64		56
	L_{93},H_{95}	69.23	30.77		65
	H_{93},H_{95}	95.39	4.61		977

Notes:
SHIW data, 1993-98 transition - Full time employees from panel households aged 18-58 if female and 18-60 if male. Low pay defined as first quintile of the hourly earnings distribution. L_t=low pay in year t; H_t=high pay in year t.

Table 6: Results[a] from estimation of four-variate probit model for 1998 conditional low pay probabilities (absolute t-ratios, p-values)

1993 and 1995 pay states	L93 L95		H93 L95		L93 H95		H93 H95	
Female	0.410	(1.07)	0.202	(0.54)	0.241	(0.63)	0.368	(2.17)
Potential labour market experience/10	-0.006	(0.04)	-0.010	(0.05)	0.194	(1.11)	0.048	(0.55)
Education≥ High school	0.958	(1.35)	-0.332	(0.51)	-0.308	(0.71)	-0.345	(1.67)
Non-manual occupation	-1.342	(1.91)	0.504	(0.79)	-0.542	(0.95)	-0.283	(1.39)
Other manufacturing, Retail trade, Financial and other services	0.263	(0.75)	0.006	(0.02)	0.178	(0.50)	0.169	(0.86)
Public sector	0.553	(0.93)	0.174	(0.34)	0.446	(0.70)	-0.332	(1.65)
North	-0.484	(1.50)	-0.485	(1.41)	0.126	(0.38)	-0.167	(1.08)
Constant	0.401	(0.71)	-0.516	(0.99)	-0.568	(1.08)	-1.695	(5.70)
τ			-1.156			(37.32)		
ρ1(initial conditions 1993-retention)			-0.110			(3.02)		
ρ2(initial conditions 1993- 1995)			0.589			(14.39)		
ρ3(initial conditions 1993 – transition)			-0.413			(1.65)		
ρ4(initial conditions 1995- retention)			-0.071			(1.16)		
ρ5(retention-transition)			0.359			(1.57)		
ρ6(initial conditions 1995 – transition)			-0.154			(0.56)		
Average prediction	0.68		0.44		0.3		0.04	
Aggregate State Dependence[b]			0.24		0.38		0.64	
Test(df=24, H0: $\eta1=\eta2=\eta3=\eta4$)			26.11			(0.3475)		
Genuine State Dependence[b]			0.06		0.26		0.43	
Log likelihood					-4351.957			
chi2(df=82)					1074.94 (0.0000)			
Number of observations					2703			

Notes: SHIW data, transition 1993-98 - Full time employees aged 18-58 if female and 18-60 if male. Reference category for dummy variables: male, blue collar worker, manufacturing lives in the Centre, South or Islands. Lt=low pay year t; Ht= high pay year t. Low pay defined as first quintile of hourly earnings.

(a) SML coefficients. GHK simulator with 50 random draws

(b) Deviations of 1998 conditional low pay probabilities from the L93 L95 case.