Modelling poverty transitions in Spain: Do attrition and initial conditions really matter?

Sara Ayllón
Department of Applied Economics
Universitat Autònoma de Barcelona (UAB)

PRELIMINARY DRAFT. DO NOT CITE OR QUOTE.

Abstract

The availability of panel data has allowed a comprehensive description of poverty exits and entries in Spain. However, most of the literature, so far, has ignored or not explicitly modelled the process of sample attrition and/or the initial conditions problem we face when studying poverty dynamics with survey data. The main objective of this work is to assess whether attrition and the poverty status in the base year are endogenous processes to poverty transitions in the Spanish case. Our estimation strictly follows the model recently proposed by Cappellari and Jenkins (2004a). Data used is from the European Community Household Panel and refers to poverty transitions that take place between 1994 and 2000 in Spain. Results show that unobservables affecting initial conditions are correlated with unobservables related with poverty transitions and therefore should necessarily be modelled simultaneously when studying poverty dynamics. Instead we do not find evidence of unobservables affecting retention that also determine poverty transience.

JEL Classification: I32, D31

Keywords: poverty transitions, trivariate multinomial probit, attrition, initial conditions

1. Introduction

Nowadays we have a fairly good description of poverty transitions and poverty persistence in Spain. Works by Cantó (2002, 2003), Cantó, Del Río and Gradín (2003, 2006), Bárcena and Cowell (2006) or Arranz and Cantó (2007), among others, have allowed a comprehensive knowledge of poverty dynamics in the Spanish case. We have learnt about the most important labour market and demographic events associated with poverty transitions, the importance of spell duration and recurrence, the sensitivity of poverty transitions to the income accounting period used, etc. However, most of the literature that we know ignores or does not explicitly model sample retention and initial poverty status along with the analysis of poverty transitions.

Our contribution in this work is to apply to Spanish data a model recently proposed in the literature by Cappellari and Jenkins (2004a) that estimates poverty transience and deals with the 'initial conditions' problem and the possibility of non-random attrition of the sample. Accounting for initial conditions is important because individuals observed poor in the base year could have been in this very same status for long time before we observe them. Further, poverty dynamics estimates should control for the fact that transitions are only observed for those individuals in the survey at $t-1$ and at t who could be a *selected* group of the original sample. We assess the endogeneity of both processes to poverty transitions by freely estimating the correlations between unobservables affecting each process. Data used is from the eight waves of the European Community Household Panel (ECHP) and poverty transitions refer to the period between 1994 and 2000.

Our main findings show that poverty status in the base year is endogenous to poverty transitions and that individuals who were more likely to be initially poor are less likely to remain poor compared to the non-poor. Failure to account for this endogeneity would underestimate our poverty persistence. Unobservables affecting attrition prove to be exogenous to poverty transitions: retained individuals are neither less or more likely to remain poor or fall into poverty in comparison to individuals more likely to attrit. As far as we know, similar estimates do not exist for Spain. Neither, this type of model has been applied to ECHP data.

The paper is structured as follows. Section 2, after this introduction, revises briefly the existing literature on poverty dynamics devoted to the Spanish case. Section 3 presents the data and some methodological choices while Section 4 shows poverty transitions at a descriptive level. Section 5 presents the model by Cappellari and Jenkins (2004a) used in this work and Section 6 discusses the empirical results obtained. Section 7 concludes.

2. Review

In this section, we briefly revise some of the existing literature on poverty dynamics in Spain, either dedicated to its description and characterization or to methodological questions.[1]

Part of the literature on poverty transience in the Spanish case focuses on the description of trends. Cantó, Del Río, Gradín (2003) using the *Encuesta Contínua de Presupuestos Familiares* for the period 1985-1995 find that the decline of poverty that took place during late eighties can be associated with high exit rates rather than a great improvement of the economic vulnerability faced by households at risk of falling into it.

---

[1] This is not a review about the extensive literature on the different strategies used to model poverty transitions. See Jenkins (2000) for this purpose.

Instead the increase of poverty risk experienced by Spaniards at the beginning of the nineties was due both to the increase of poverty entries and, especially, the reduction of poverty exits. Bárcena and Cowell (2006) update these results by running a similar descriptive analysis for the period 1993-2000 with data from the ECHP.[2] They observe that the increase of poverty risk between 1993 and 1996 is associated with a high percentage of poverty entries and exits as opposed to the time between 1997 and 2001 when income grows steadily.

As for the characterization of poverty dynamics, Cantó (2003) assesses the importance of demographic and socio-economic characteristics on poverty exit probabilities. She finds that less than 10% of the transitions out of poverty are linked to demographic events and the rest to labour market changes or social assistance benefits reception. Poor households are highly likely to receive pension benefits and other regular transfers which help them out of poverty. However, it is new unemployment benefit which has the largest effect, even though received more infrequently. Also, the employment of other adult members is a key factor explaining exits from poverty in families with children (see Bane and Ellwood, 1986, for a similar result).

In relation to methodological questions, Cantó, Del Río and Gradín (2006) show that the choice of quarterly income or annual income has important consequences for the poverty estimates. Exit rate is fairly similar in both income definitions but entry rate is higher with quarterly income. Further, they show that only half of those classified as leavers by one income definition are not equally classified by the other one and the misclassification is even stronger when considering poverty entries.

Spanish researchers have also focus on the importance of duration dependence in their work. Cantó (2002) argues about the importance of time in poverty on the measurement of entries and exits. She proposes a discrete time duration-dependent *n*-order Markov process with heterogeneity jointly estimating exits and (re)entries. In the model, duration dependence is taken into account by including dummies for the distance of household equivalent income to the poverty line and also for time spent in or out of poverty just before or after an exit. Results show that one third of households that escape poverty shortly return to it while if they manage to be one year out of poverty, the chances to fall back into it strongly decrease. Similarly, Bárcena *et al.* (2004) also study the influence that time spent in poverty has on poverty entries and exits. They argue about the need to consider, not only the poverty status at t conditional on the poverty status at $t-1$, but also, the time spent in the same poverty status than $t-1$, Poverty transitions probabilities prove to be smaller when accounting for time inertia. More recently, Arranz and Cantó (2007) propose a multi-state multiple transition discrete hazard regression model that controls not only for observed and unobserved heterogeneity but also for the length of current poverty spell, time between spells, the occurrence of multiple spells and the accumulation of poverty spells. They find evidence of negative duration dependence -longer poverty spells reduce the exit probability and increase the re-entry hazard. Thus, second and third poverty spells are shorter than the first one. Among other results, they also show that non-poverty spells are of longer duration than poverty spells.

Amuedo-Dorantes and Serrano-Padial (2006), on the other hand, focus on the difficulties of building an econometric model for poverty transitions. They are concerned about feed-back effects and examine the poverty implications for past and current temporary employment in Spain. They find that holding a temporary contract increases not only the probability of current poverty but also of future poverty via an

[2] The first estimates of poverty entries and exits in Spain using data from the ECHP can be consulted in García and Toharia (1998).

indirect effect that increases the chances of holding a type of contract in the future with higher poverty risk (but not a direct effect of past temporary contract on poverty).

However, in none of the literature that we know, initial conditions are explicitly taken into account when modelling poverty transitions. As for attrition, Cantó, Del Río and Gradín (2003) and Cantó (2003) do take into account potential non-randomness of sample reduction by the construction of own sample weights (see also, Ayala, Navarro and Sastre, 2006, for a similar strategy when analysing income mobility). Nevertheless, we are left with the question whether there is a correlation between unobservables affecting poverty status at base year and/or attrition that also affect poverty transience.

3. Data and definitions

Data

The dataset used in the analysis is the Spanish component of the European Community Household Panel (ECHP) which is a harmonised cross-national longitudinal survey collected across all members of the (former) European Union-15. The panel runs from 1994 to 2001 except for Austria and Finland that joined the project in 1995 and 1996 respectively. Data is based on a standardised questionnaire that is answered each year by a representative sample of individuals and households on questions related to income, education, employment, household structure, housing, health, social relations and individual satisfaction, to name the most important sections. Questionnaires are generally answered in a face-to-face personal interview but other methods have been also used (self-completed interview, telephone, proxy, etc.). The target population consists of all private households throughout the national territory in every country.

Unit of analysis and sample size

Although the household is the unit of measurement for income, we examine poverty dynamics for individuals. As argued in OECD (2001), this methodological choice has the advantage of giving greater weight to larger families and allows tracing the poverty status of individuals when family structure changes (e.g. divorce, marriage, youth emancipation, etc.). Further, and following previous literature, we restrict our analysis to the population between 25 and 64 years old (see for instance, Cappellari and Jenkins, 2004a; Van Kerm, 2004). As pointed out by Arranz and Cantó (2007), it is among the Spanish working-age population that transitory or short-term poverty mainly takes place –and this is explicitly what we model in this study. See also, OECD (2001) for similar evidence.

Table 1 shows the number of sample observations used for the empirical analysis. Notice that apart from the age of the individuals, no other restrictions are imposed to our working sample.[3] We allow individuals to enter the panel even if we know their poverty status for one single year and they transit to missing in the following one. Thus, our panel is unbalanced and maximizes the use of the information available in the survey.

[3] We see this feature of our analysis as a great advantage compare to the existing literature. For instance, Arranz and Cantó (2007) need to limit their analysis to individuals present in the survey in 1994. Bárcena and Cowell (2006) restrict their analysis to adults participating during the eight waves of the panel which cuts considerably their working sample. Also Cantó, Del Río and Gradín (2003) with data from the ECPF limit the sample to those individuals answering at least 5 quarterly questionnaires. Poggi (2007) also uses a balanced panel when estimating persistence of social exclusion in Spain.

[TABLE 1 AROUND HERE]

In our descriptive analysis, we use longitudinal weights as recommended by Eurostat (2003) (except for table A.1. in the Appendix that we utilize cross-sectional). We do not use weights in the econometric analysis as it is more efficient not to do so (see Poggi, 2007).

Poverty

A person is considered poor if the equivalent income of the household where s/he lives is below the poverty line defined as 60% of the median of that distribution. The threshold is relative to time so there is a poverty line for each of the years analysed. We use the modified OECD equivalent scale as the scaling factor that takes into account the economies of scale within the household by giving a weight of 1 to the first adult, 0.5 to the rest of adult members in the household and 0.3 to children under 14 years of age.[4] As in other poverty studies, we accept that all incomes are pooled together and shared equally among household members.

Income in the ECHP

Alike in other surveys, all the annual income variables are collected retrospectively in the ECHP. For instance, in wave 1, run in 1994, annual income variables refer to incomes obtained by household members in 1993. Neglecting this time lag between the period to which household income refers (year $t-1$) and the period to which household composition and other variables of interest relate (year t) would introduce some bias (Debels and Vandecasteele, 2005). Therefore, net household income in year $t-1$ is finally constructed as the sum of the net personal income reported at t of the individuals that were present in the household at $t-1$ (see Debels and Vandecasteele, 2005; Arranz and Cantó, 2007).[5] This approach allows building household equivalent income at each year with the household composition (and equivalence scale) referring to the same year.

Note, however, that the choice of this income distribution implies, on one hand, that only seven waves of the panel can be used in our analysis, and, on the other, that a certain number of missing values arise when one of the members of the household does not inform of his/her income at t, either because of attrition or because this person refuses to collaborate in this part of the questionnaire. In few cases where there is missing income information only for one of the members of the household, we were able to impute to that individual the income information given in the survey for within household non-response (0,37% of the individuals_waves sample). Finally, note that household income for year $t-1$ cannot be computed in those households where one of the household members dies at t. In these cases, we proxied his/her personal income at $t-1$ with the one reported at $t-1$ (0,30% of the individuals_wave sample). In table A.1. of the Appendix you can check how differences in the population headcount ratio are very small when we either use the household equivalent income with time lag or without respect to household composition. Table A.2., on the other hand, shows how the use of the corrected household income definition increases about 2,9% the number of

[4] These methodological options follow the recommendations of the European Commission for the analysis of poverty and social exclusion in the European Union. You can check them at: http://europa.eu.int/comm/employment_social/social_protection_commitee/spc_indic_en.htm

[5] Debels and Vandecastele (2008) propose a more accurate measure that accounts for household composition change within waves.

transitions to missing. Note, however, that we use an estimation technique that explicitly accounts for sample attrition.

4. Poverty dynamics in Spain: a description

In this section we want to briefly describe poverty transitions in Spain during the analysed period for the proposed sample. Table 2 shows poverty status of Spanish individuals aged 25 to 64 at time t conditional on their status at $t-1$.[6] The first panel shows the results when missing income information is not taken into account and the second one when we do so.

[TABLE 2 AROUND HERE]

First of all, it is worth noticing, the important difference on the probability of being poor at time t depending on the poverty status at $t-1$. The chances of an individual to be poor at t were 58,37% if s/he was already poor at $t-1$ while only 8,16% if not.[7] One of the objective of this study is to address the possibility that an endogenous selection mechanism occurs and individuals observed to remain poor at t might have been over (or under) represented at $t-1$.

In the second place, it is also interesting to observe the pattern of transitions to missing. Results show that 14,5% of individuals not poor at $t-1$ are not observed any longer at t. Among those observed poor, the percentage is 14,13. At first sight, it seems that sample retention is exogenous to poverty status at $t-1$. In the next section, however, we explicitly address the question of potential non-random selection of the sample.[8]

Table 3 sheds additional light on the different transition probabilities year by year. As can be observed, entry rates fluctuate between 6,13% and 8,18% while exit rates do so between 31,13% and 41,15%. However, as similarly pointed out by Arranz and Cantó (2007), there is not a really clear pattern along the period. As for the income missing data, one can observe a smaller number of transitions out of the sample in years 1998 and 2000. Nonetheless, it is difficult at a descriptive level to distinguish a different sample retention process for those poor and not poor at $t-1$.[9]

[TABLE 3 AROUND HERE]

[6] We are aware that the elements in the diagonal may be under-estimated (*lumping of the diagonals*) as usually in first-order Markov matrices (see Bárcena *et al.*, 2004; Cantó, 2002) given poverty recall is not taken into account.

[7] These results are fairly similar to those obtained by Bárcena and Cowell (2006) with the same dataset as they estimate an entry rate of 8.07% and an exit rate of 39.80%. Note, however, that they base their estimates on a balanced panel for individuals that are in the panel eight consecutive waves and are aged 16 or above.

[8] In OECD (2001) it is argued that "Attrition bias may be particularly acute for the ECHP since attrition rates are quite high for some of the participating countries (…) and the poverty population appears to drop out of the sample at a disproportionate rate in most of these countries" (OECD, 2001, p.43). However, from our descriptive analysis we do not find such a clear pattern in the Spanish case.

[9] Possibly, the exception is for transitions between 1999 and 2000, with higher transitions out of the sample among those initially not poor.

5. The model

In this work we estimate poverty transitions by strictly following the proposed model by Cappellari and Jenkins (2004a) which allows the characterization of poverty persistence and poverty entries.[10] We build on a system of simultaneous equations that includes a first-order poverty transitions equation, the poverty status at $t-1$ (in order to account for the initial conditions problem) and the retention equation (to consider potential non-random attrition) plus the correlations between the three equations that are allowed to be estimated freely.[11] We refer the interested reader to the original article for a full econometric illustration of the model.

Let's define first the poverty transitions equation. We assume that in period t individuals can be characterized by a latent poverty propensity p_{it}^{*} that takes the form:

$$p_{it}^{*} = [(P_{it-1})\varepsilon_1{}' + (1-P_{it-1})\varepsilon_2{}']z_{it-1} + \omega_{it-1}$$
$$P_{it} = I(p_{it}^{*} > 0)$$

where $i = 1, 2, ...N$ refers to individuals, $\varepsilon_1{}'$ and $\varepsilon_2{}'$ are column vectors, z_{it-1} is the vector of explanatory variables and ω_{it-1} the error term. Note that the vector of explanatory variables always refers to $t-1$ in order to reduce endogeneity / simultaneity problems with the poverty transitions (see Jenkins, 2000; Cantó, 2003).[12] Further, ω_{it-1} can be written as $\omega_{it} = \zeta_i + \tau_{it}$ where the first term is the individual specific effect that stands for all unobserved determinants of conditional poverty that are time-invariant for a given individual (e.g., ability, motivation, etc.).[13] The second term is the usual white noise error assumed to follow a standard normal distribution. Both parts of the error term are assumed to follow a normal distribution. $I(p_{it-1}^{*} > 0)$ is a binary indicator function equal to one if the latent conditional poverty propensity is positive and equal to zero otherwise.

Our second equation accounts for (possibly) non-random attrition of the sample. It studies the probability to observe the poverty status of an individual both at $t-1$ and at t and it can be defined as follows,

$$r_{it}^{*} = \varphi' w_{it-1} + \mu_{it}$$
$$R_{it} = I(r_{it}^{*} > 0)$$

where r_{it}^{*} is the latent probability of consecutive participation (retention) and of knowledge of poverty status, w_{it-1} is the vector of explanatory variables, φ is the vector of parameters and μ_{it-1} is the error term. Again, μ_{it} can be written as $\mu_{it} = \alpha_i + \lambda_{it}$, being

[10] See Cappellari and Jenkins (2002) for a less technical version of the paper. In Cappellari and Jenkins (2004b) a distinction is made between attrition due to drop-out of the sample and economic item non-response when modelling low pay transitions among British men aged 18-64. However, distinguishing between the two attrition processes had little impact on the estimates.

[11] Note that the inclusion of a retention equation allows us the use of an unbalanced panel and therefore to draw on all the information available in the panel.

[12] We do not include as explanatory variables events taking place between $t-1$ and t in order to avoid contemporaneously correlated regressors.

[13] Notice that we suppose that the unobserved heterogeneity of conditional poverty is common in the case of poverty entry and poverty persistence. Van Kerm (2004) instead builds on a system of equations where the poverty entry rate is estimated separately from the persistence and the individual-specific effects are different in each case. Conceptually we find it easier to think of a shared individual unobserved heterogeneity for exits and (re)entries given it is a time-invariant concept.

the first term the individual specific effect and the second the usual white noise error. Both parts of the error term are assumed to follow a normal distribution. $I(r_{it}^{*} > 0)$ is a binary indicator function equal to one if the latent retention propensity is positive and equal to zero otherwise.

Finally, a third equation allows us to account for the initial conditions problem which arises because the start of the observation window may not be the same than the start of the poverty experience. In our case, individuals observed poor at $t-1$ could have been in this very same status for long time before we observe them. Thus, we assume that in period $t-1$, individuals can be characterized by a latent poverty propensity which is defined as,

$$p_{it-1}^{*} = \beta' x_{it-1} + u_{it-1}$$
$$P_{it-1} = I(p_{it-1}^{*} > 0)$$

where x_{it-1} to the vector of explanatory variables that both describe the individual and his/her household, β to the vector of parameters and u_{it-1} to the error term. Further, u_{it-1} can be written as $u_{it-1} = \eta_i + \delta_{it-1}$, where the first term is the individual specific effect and the second term is the usual white noise error. Both η_i and δ_{it-1} are assumed to be normally distributed. And, $I(p_{it-1}^{*} > 0)$ is a binary indicator function equal to one if the latent base year poverty propensity is positive and equal to zero otherwise.

Further, the model allows the three random effects to be freely correlated. If it turns that they are correlated, both sample selection processes are endogenous to poverty transitions and the estimation of a univariate probit would lead to inconsistent estimators of the parameters of interest (see Cappellari and Jenkins, 2006). If all the correlations are zero, we could actually estimate the equations separately.

Thus, we define,

$$\rho_1 = \mathrm{cov}(\zeta_i, \alpha_i) = \mathrm{corr}(\omega_{it-1}, \mu_{it})$$
$$\rho_2 = \mathrm{cov}(\alpha_i, \eta_i) = \mathrm{corr}(\mu_{it}, u_{it-1})$$
$$\rho_3 = \mathrm{cov}(\zeta_i, \eta_i) = \mathrm{corr}(\omega_{it-1}, u_{it-1})$$

where ρ_1 summarises the association between unobservable individual specific factors determining the poverty transitions and the unobservable individual specific factors determining the consecutive participation of a certain individual and the knowledge of his/her poverty status. If ρ_1 is positive (negative) it means that those individuals more likely to be consecutively poor -or fall into poverty- are also more (less) likely to be income retained in the sample. Similarly, ρ_2 summarises the association between the unobservable individual specific factors related to the consecutive participation of a certain individual in the sample and the unobservable individual specific factors determining the base year poverty status. If ρ_2 is positive (negative) it means that the individuals that are more likely to be consecutively participating in the sample and their household informs of their income are also more (less) likely to be initially poor. Finally, ρ_3 summarises the association between the unobservable individual specific factors determining poverty transitions and those determining the base year poverty status. If ρ_3 is positive (negative) those individuals more likely to be observe to fall into poverty, or persist in it, are also more (less) likely be observed initially poor.

Once the model is estimated, we can derive the transition probabilities. The (re)entry probability is given by,

$$s_{it}(z_{it-1}, x_{it-1}) \equiv \Pr(P_{it} = 1 \mid P_{it-1} = 1) = \frac{\Phi_2(\varepsilon_1' z_{it-1}, \beta' x_{it-1}, \rho_3)}{\Phi(\beta' x_{it-1})}$$

And, the entry probability by,

$$e_{it}(z_{it-1}, x_{it-1}) \equiv \Pr(P_{it} = 1 \mid P_{it-1} = 0) = \frac{\Phi_2(\varepsilon_1' z_{it-1}, -\beta' x_{it-1}, -\rho_3)}{\Phi(-\beta' x_{it-1})}$$

where $\Phi(\cdot)$ and $\Phi_2(\cdot)$ are the cumulative density functions of the univariate and bivariate standard normal distributions. Note that our estimation technique allows us to predict what would have been the conditional poverty probability of those individuals that actually attrit.

The model is estimated by Maximum Simulated Likelihood (MSL) using a Geweke-Hajivassiliou-Keane (GHK) simulator.[14] Differently from the original article, we have used Halton draws instead of pseudo-random ones.[15] Antithetics is also applied.[16] As argued by Cappellari and Jenkins (2006), Halton draws are more effective than pseudo-random ones because they provide the same accuracy with a smaller number of draws which saves computer time.[17] We do take into account for intra-household correlation in poverty status by using robust standard estimates and by clustering with the household identification number.

Further, and as pointed by Wooldridge (2002), in order to identify the model we need to use *exclusion restrictions* that is, variables that influence the probability of sample retention and poverty status at $t-1$ but have no effect on the probability of poverty transition. In the case of the retention equation, we have used a dummy variable that identifies original sample members –as suggested by Cappellari and Jenkins (2004a)– together with a set of dummies describing the mode of interviewing. The idea being that those individuals answering a self-administered questionnaire, a telephone

---

[14] The model has been estimated using the `ml` commands available in Stata® and following the instructions in Cappellari and Jenkins (2003, 2006). Notice the difficulty of estimating a simultaneous model where some observations contain missing values (individuals transiting from a known poverty status to a missing one). The maximization technique used is the modified Newton-Raphson algorithm, default in Stata®. See also Gould and Sribney (1996) for more information on `ml` use. We intended to integrate out individual-specific effects across the pooled transitions in order to assure consistent and efficient estimates by using the software package aML and its Gauss-Hermite quadrature, however, it turned computationally infeasible.

[15]. In simulation, draws from a density are used to calculate the average of a statistic over that density. As proven by Train (2003), MSL is consistent, asymptotically normal, efficient and asymptotically equivalent to Maximum Likelihood (ML) as long as the number of draws used in the simulation (R) rises faster than the square root of the sample size ($\sqrt{N}$). In other words, the bias between MSL and ML diminishes as more draws are used in the simulation. See also Chapter 5 in Greene (2000). Also, we make use of the programme `mdraws` made available by Cappellari and Jenkins (2006) for Stata® users. The primes used for the creation of the draws are 2, 3 and 5. As explained in Train (2003), given that the simulated log-likelihood function is a sum over observations of the log of simulated probabilities, if the draws are taken in such a way that negative correlation over observations is created, then the variance of the sum is lower. And this is precisely what Halton sequences do: they induce a negative correlation over observations. Halton draws are created to fill the unit interval evenly with elements placed equidistantly apart. Each cycle covers the areas not covered by previous cycles. Further, because Halton sequences in our study are created over 3 dimensions, we use the option `burn` to eliminate the initial part of the series, as it is customary (see Train, 2003, p.230).

[16] Antithetic draws are obtained by creating various types of mirror images of every draw (see Train, 2003, p.219). That way, the 100 Halton draws created per each of the 3 dimensions are finally 200 per equation. Train (2003) argues that the antithetics substantially improve the estimation of probit models.

[17] Cappellari and Jenkins (2006) show that calculations based on 1000 pseudo-random draws get very close to those derived using directly the bivariate normal probability distribution function as opposed to 50 pseudo-random draws. However, they also show that 100 Halton draws get even closer.

interview or a proxy interview as opposed to a face-to-face personal interview might be less interested with the survey project or have less time for it and therefore may have a smaller probability of being retained in the future.[18] The instruments for the initial conditions are more difficult to find, especially because the ECHP does not collect family background information (as used in Cappellari and Jenkins, 2004a or Van Kerm, 2004). Conceptually is neither an easy exercise: we need to find a variable that influences the chances to be poor at a given point in time but that does not affect the probability of falling into or reentering poverty. After, trying several sets of instruments, our final estimation includes a dummy in the initial conditions equation that identifies whether the household head suffers or not from a chronic disease.[19] The validity of these instruments is discussed in the next section.

As noted by Biewen (2004), one of the main advantages of this pooled estimation strategy is that it circumvents the strict exogeneity assumption by which conditional on past poverty and unobserved individual characteristics, current poverty must not be related to the value of the covariates in past or future periods. In the model just presented, by definition, changes in employment status or family structure are not allowed to affect poverty until next period.

On the contrary, we are aware that possibly the most important drawback of the model is the impossibility to control for duration dependence in the poverty status. Arranz and Cantó (2007), to mention the most recent work on the Spanish case, have shown that the length of current poverty spell, the time between spells, the occurrence of multiple spells and the accumulation of poverty spells do have an effect on poverty entries and exits (see also Devicienti and Gualtieri, 2007). The model we apply in this study does not substitute the key findings obtained from duration dependence models. Rather, we believe it can complement some of their results.

6. Empirical results

Table 4 shows the obtained results related to correlations, the validity of the used instruments and a test for state dependence. Table 5, on the other hand, presents the impacts of the explanatory variables on transitions probabilities and Table 6 those of the retention and the base year poverty status equations.

Let's discuss first the results relative to the correlations. As can be seen in Table 4, ρ_1 that accounts for the correlations between unobservables affecting poverty transitions and sample retention is negative but not significant. This result indicates that those individuals more likely to be retained are neither less nor more likely to remain poor or fall into poverty. Similarly, ρ_2 that measures the correlation between unobservables affecting the base year poverty status and the retention equation was not statistically different from zero. Nothing can be said, therefore, about the retention propensity among those observed initially poor compared to the non-poor.[20]

The correlation between unobservables affecting poverty transitions and poverty status at the base year (ρ_3) is negative and statistically significant at 99% which

---

[18] Coefficients in Table 6 show this is actually the case.

[19] We tried with a set of dummies that informed of the time that the household head spent between finishing his/her education and becoming employed (if ever). However, in more than 20% of the cases the information on the age when the highest level of education was completed was missing and the instrument turned to be invalid. Also, we tried with a dummy that informed whether the person had been unemployed for one month or longer before first job or business. Similarly, this instrument proved null.

[20] Arranz and Cantó (2007) with data from the ECHP also argue that "[…], the probability of attrition does not appear to be determined by the individual poverty situation" (Arranz and Cantó, 2007, p.13).

indicates that individuals more likely to be observed poor at the base year are less likely to remain poor compared to the non-poor group. Ignoring this initial condition endogeneity would lead to an underestimation of poverty persistence..

[TABLE 4 AROUND HERE]

Table 4 shows also the results of the exogeneity tests of the two selection processes considered. First we test whether the initial conditions are exogenous to poverty transitions. This hypothesis is strongly rejected by a Wald test with a *p*-value of 0,0000. The second hypothesis testing the exogeneity of retention to poverty transience could not be rejected.[21] Results show initial conditions is a mechanism of selection endogenous to the estimation of low income transitions and this fact needs to be taken into account when modelling poverty transitions. In turn, unobservables affecting attrition can be neglected when modelling poverty transience in Spain with data from the ECHP.

As commented, finding the appropriate instruments to identify the model was not an easy task (especially in the base year poverty status equation). Still, the instruments we found prove to be valid in our model. Tests indicate that the fact that the household head suffers from a chronic disease could be excluded from the transitions equation, and the same for the original sample member dummy and the type of questionnaire answered. On the contrary, both sets of instruments increase the precision of the initial conditions equations and the retention equation, respectively. Note in Table 6 how each coefficient is also statistically significant.

Our final test checks whether ε_1' is equal to ε_2'. If that is the case, poverty status at t would not depend on poverty status at $t-1$ since the overall effect of poverty entry would be the same than those of poverty persistence and no sign of state dependence would be present in the Spanish case. As can be seen in the last row of Table 4, this is not the case. The null hypothesis of no state dependence is rejected with a p-value of 0.0000.

Let's next turn to the impacts of explanatory variables on the probability of entry or persist in the poverty status (Table 5) and those on retention and base year poverty status (Table 6). Let's next turn to the impacts of explanatory variables on the probability of entry or persist in the poverty status. Age measured at the individual level does not seem to have any effect on the probability of entry or persist but it does when it refers to the household head.[22] An increase in the household head's age increases the probability of entry but this phenomenon reverses when becoming older.[23] And, the contrary is true for poverty persistence. Note also the different sign of the coefficients in the base year poverty status when referred to household members' age and household head age. On the other hand, households headed by a woman are more likely to be poor at a given point in time, but this characteristic is not relevant on the chances of poverty

---

[21] This result reinforces Perez-Mayo (2004)'s findings in his analysis of poverty transitions in Spain with data from the ECHP. After carrying a grouped analysis, he argues that consistent poverty (income poor with low living standards) is independent to definitive or temporarily attrition.

[22] In our analysis, the household head is the member with higher personal earnings. Again, we chose the household head at $t-1$ with the personal income asked at t that refers to $t-1$. We find this criterion more objective than taking the reference person, the person responsible for accommodation or the respondent to the household interview.

[23] Recall that our analysis includes only those aged between 25 and 64 years of age. This result probably reflects upward mobility of the household head labour market career.

transition. None of the mentioned characteristics is neither precisely estimated in the retention equation.

[TABLE 5 AND 6 AROUND HERE]

Education level of the household head is not strongly associated with poverty entry but it is with poverty persistence and poverty status at the base year. We believe these results are consistent with the fact that household head's incapacity to read and write is a fairly time-invariant characteristic and therefore it may determine long periods of poverty.

When the household head is at work, the chances of his/her household to be found in poverty at a given point in time are smaller than among inactive or unemployed. However, economic status of the household head is not relevant in the entry probability. Instead, household headed by a worker or (especially) an unemployed have greater chances to persist in poverty than inactive heads. Despite more research needs to be done on this finding, the results may be picking up persistence in low pay among precarious workers. Also, note that if the household head works, its members are more likely to be retained. Number of workers in the household, as expected, is negatively related to base poverty status, poverty entry and poverty persistence. [24]

As for family structure, neither living with a household head that has a partner or in a lone parent household has an impact on the poverty entry and poverty persistence probability. However, note that both characteristics are significant in the case of base year poverty status. Having a partner is also significant in the retention equation – probably picking up greater family stability.

The number of children in the household is another key characteristic related to the probability of falling into poverty. Notice in the case of entry how the chances to fall into poverty increase as the age of children also does. This effect adds to the greater chances of being poor that families with children have at a given point in time. Chances to persist in poverty are also higher for families with children but it does not seem to be a matter of children's age. The number of elderly people (above 65) reduces the chances to be found in poverty and also to be permanently poor but their effect is null in the case of entries. Differently, young people increase the chances of their families to be poor but have no effect on poverty transience.[25] Children and youth increase the chances of a family to be retained while adults over 75 decrease them.

Finally, being an immigrant increases the chances of entry, persist and being found poor. Living in own property and paying mortgage reduces the chances of becoming poor and being permanently poor compared to those owners without any mortgage payments left. Note also the use of year dummies in order to control for labour market conditions and the business cycle.

7. Conclusions

Our main objective in this study has been to assess whether the processes of sample retention and initial conditions are endogenous selection processes to poverty transitions not only via observed heterogeneity but also through unobserved one.

[24] Bane and Ellwood (1986) in their seminal work already underlined the importance of the earnings of others than the head and the spouse in order to move a household out of poverty.

[25] Note that the help-effect of working young people cohabiting with their parents will be picked up by the 'number of workers' variable.

Following the model recently proposed by Cappellari and Jenkins (2004a), we find that the unobservables affecting poverty status at base year are negatively correlated with unobservables affecting poverty transition. As for attrition, we do not find evidence of unobservables associated with retention that would also have an effect of poverty transience. That is, a control for observed heterogeneity in the case of sample retention may be sufficient when modellling poverty transience in Spain with data from the ECHP.

Much research needs yet to be done. To start with one should consider, for instance, how to adjust our findings to duration dependence models that have also proven to be relevant in the description of individual and household characteristics associated with poverty dynamics.

Acknowledgments

I am grateful for the hospitality and the IRISS grant received while visiting CEPS/INSTEAD in Luxembourg where I wrote most of this paper. I owe special thanks there to Philippe Van Kerm for helpful advice. All the errors in this paper are my own. Financial support is also greatly acknowledged from the Spanish project SEJ2004-07373-C03-01/ECON.

References

Amuedo-Dorantes, C.; Serrano-Padial, R. (2006). 'Labour market flexibility and poverty dynamics: Evidence from Spain', unpublished working paper.

Arranz, J. M.; Cantó, O. (2007). 'Measuring the effect of spell recurrence on poverty dynamics: Evidence from Spain', paper presented at *XV Encuentro de Economía Pública*, Salamanca, Spain, 7-8 February 2008.

Ayala, L.; Navarro, C.; Sastre, M. (2006). 'Cross-country income mobility comparisons under panel attrition: the relevante of weighting schemes'. ECINEQ Working Paper Series, 2006-47.

Bane, M.J.; Ellwood, D.T. (1986). 'Slipping into and out of poverty: the dynamics of spells', *Journal of Human Resources*, 21(1): 1-23.

Bárcena Martín, E.; Fernández Morales, A.; Lacomba Arias, B.; Martín Reyes, G. (2004). 'Dinámica de la pobreza a corto plazo en España y Reino Unido a través de los datos del Panel de Hogares Europeo', *Estadística Española*, 46(157): 461-481.

Bárcena Martín, E.; Cowell, F. A. (2006) 'Static and dynamic poverty in Spain: 1993-2000', *Hacienda Pública Española / Revista de Economía Pública*, 179-(4/2006): 51-77.

Biewen, M. (2004). 'Measuring state dependence in individual poverty status: are there feedback effects to employment decisions and household composition?, IZA Discussion Papers, n. 1138.

Cantó, O. (2002). 'Climging out of poverty, falling back in : low income stability in Spain', *Applied Economics*, 34: 1903-1916.

Cantó, O. (2003). ´Finding out the routes to escape poverty: the relevance of demographic vs. labor market events in Spain', *Review of Income and Wealth*, 49(4): 569-588.

Cantó, O.; Del Río, C.; Gradín, C. (2003). 'La evolución de la pobreza estática y dinámica en España en el período 1985-1995', *Hacienda Pública Española / Revista de Economía Pública*, 167-(4/2003): 87-119.

Cantó, O.; Del Río, C.; Gradín, C. (2006). 'Poverty statics and dynamics: Does the accounting period matter?, *International Journal of Social Welfare*, 15(3): 209-218.

Cappellari, L.; Jenkins, S. P. (2002). 'Who stays poor? Who becomes poor? Evidence from the British Household Panel Survey', *The Economic Journal*, 112(478): C60-C67.

Cappellari, L.; Jenkins, S. P. (2003). 'Multivariate probit regression using simulated maximum likelihood', *The Stata Journal*, 3(3): 278-294.

Cappellari, L.; Jenkins, S. P. (2004a). 'Modelling low income transitions', *Journal of Applied Econometrics*, 19: 593-610.

Cappellari, L.; Jenkins, S. P. (2004b). 'Modelling low pay transition probabilities, accounting for panel attrition, non-response and initial conditions', ISER Working Paper, n. 2004-08.

Cappellari, L.; Jenkins, S. P. (2006). 'Calculation of multivariate normal probabilities by simulation, with applications to maximum simulated likelihood estimation', *The Stata Journal*, 6(2): 156-189.

Debels, A.; Vandecasteele, L. (2005). 'Correcting the income time lag in panel data: the impact on income and poverty dynamics', unpublished Working Paper.

Debels, A.; Vandecasteele, L. (2008). 'The time lag in annual household-based income measures: assessing and correcting the bias', *Review of Income and Wealth*, 54(1): 71-88.

Devicienti, F.; Gualtieri, V. (2007). 'The dynamics and persistence of poverty: evidence from Italy', Laboratorio Riccardo Revelli, Working Paper n. 63.

Eurostat (2003). 'ECHP UDB Manual, European Community Household Panel, Longitudinal User's Database', DOC. PAN 168/2003-12.

García Mainar, I.: Toharia, L. (1998). 'Paro, pobreza y desigualdad en España: análisis transversal y longitudinal', *Ekonomiaz*, 40: 135-165.

Gould, W.; Sribney, W. (1999). *Maximum Likelihood estimation using Stata®*, College Station: Stata Press.

Greene, W. H. (2000). *Análisis econométrico*, Prentice Hall, 3rd. edition.

Jenkins, S.P. (2000). ‘Modelling household income dynamics’, *Journal of Population Economics*, 13: 529-567.

OECD (2001). *Employment Outlook,* Chapter 2: ‘When money is tight: poverty dynamics in OECD countries’. [http://www.oecd.org/dataoecd/29/55/2079296.pdf]

Pérez-Mayo, J. (2004). ‘Consistent poverty dynamics in Spain’, IRISS Working Paper Series, no. 2004-19.

Poggi, A. (2007). ‘Does persistence of social exclusion exist in Spain?’, *Journal of Economic Inequality*, 5: 53-72.

Train, K.E. (2003). *Discrete Choice Methods with Simulation*. Cambridge University Press [pre-print available at http://emlab.berkeley.edu/users/train/distant.html].

Van Kerm, P. (2004). ‘Une evaluation econometrique des flux vers et hors de la pauvreté en Belgique’, IRISS Working Paper Series, no. 2004-04.

Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*, Cambridge MA: MIT Press.

Table 1. Number of observations in the sample. Individuals between 25 and 64 years old (included)

Wave	Individuals	Total Individuals_Wave
1 (1994)	9.443	9.433
2 (1995)	8.914	18.347
3 (1996)	8.243	26.590
4 (1997)	7.691	34.281
5 (1998)	7.319	41.600
6 (1999)	6.926	48.526
7 (2000)	6.791	54.462
Total	**54.462**	

Source: Own construction using the ECHP, 1994-2001. Note that the last poverty transitions take place between 1999 and 2000 and therefore the total number of transitions observed is 48.526.

Table 2. Poverty status at t conditional of poverty status at $t-1$ in Spain without and with income missing data, 1994-2000

Year		t		
		Not poor	**Poor**	**Missing**
	Not poor	91,84	8,16	-
$t-1$	**Poor**	41,63	58,37	-
	Total	*83,27*	*16,73*	-
	Not poor	78,52	6,98	14,50
$t-1$	**Poor**	35,75	50,12	14,13
	Total	*71,25*	*14,31*	*14,43*

Source: Own construction using the ECHP, 1994-2001. Individuals aged 25 to 64. *N*=48.526 observations.

Table 3. Poverty status at t conditional of poverty status at $t-1$ in Spain with income missing data year by year, 1994-2000

Year		**Not poor**	**Poor**	**Missing**
			1995	
1994	**Not poor**	79,43	7,41	13,17
	Poor	41,15	43,88	14,97
			1996	
1995	**Not poor**	75,95	8,18	15,87
	Poor	31,13	54,11	14,75
			1997	
1996	**Not poor**	79,39	6,13	14,47
	Poor	34,59	50,06	15,35
			1998	
1997	**Not poor**	78,73	6,67	14,60
	Poor	36,52	50,58	12,90
			1999	
1998	**Not poor**	77,00	6,55	16,45
	Poor	35,45	47,73	16,82
			2000	
1999	**Not poor**	80,97	6,65	12,38
	Poor	34,91	56,35	8,74

Source: Own construction using the ECHP, 1994-2001. Individuals aged 25 to 64. *N*=48.526 observations.

Table 4. Estimates of cross-equation correlations, test of exogeneity of selection processes, validity of instruments and state dependence test

	Estimate	*t*-ratio
Correlations:		
ρ_1 [Poverty transitions, Retention]	-0.094	-1.00
ρ_2 [Retention, Base year poverty status]	0.008	0.47
ρ_3 [Poverty transitions, Base year poverty status]	-0.292***	-2.93
	Test	*p*-value
Exogeneity of:		
Initial conditions: $\rho_2 = \rho_3$	8.87	0.0118
Sample retention: $\rho_1 = \rho_2$	1.13	0.5672
Joint exogeneity: $\rho_1 = \rho_2 = \rho_3$	9.92	0.0193
Validity of instruments:		
Exclusion of hh chronic disease from transitions eq. (d.f.=2)	4.33	0.1146
Exclusion of sample membership from transitions eq. (d.f.=8)	8.33	0.4017
@ Exclusion of both instruments sets from transitions equation		
Inclusion of hh chronic disease in IC equation (d.f.=1)	38.35	0.0000
Inclusion of sample membership status in retention eq. (d.f.=4)	98.62	0.0000
@State dependence		

Source: Own construction using the ECHP, 1994-2001.

Table 5. Poverty entry and poverty exit coefficients in Spain, 1994-2000

	Poverty Entry			Poverty Persistence		
	Coefficient		*[t-ratio]*	Coefficient		[*t*-ratio]
Individual characteristics						
Ref. Male						
Female	0.0434	**	*(2.89)*	0.0365	***	*(3.53)*
Age	-0.0103		*(-0.83)*	-0.0125		*(-1.23)*
Age^2	0.000218		*(1.54)*	0.000161		*(1.38)*
Household head characteristics						
Demographic characteristics						
Ref. Male						
Female	0.0356		*(0.50)*	0.0231		*(0.50)*
Age	0.0272	*	*(2.19)*	-0.0419	***	*(-4.84)*
Age^2	-0.00026		*(-1.95)*	0.000378	***	*(3.95)*
Ref. Spanish						
Immigrant	0.452	*	*(2.24)*	0.507	***	*(3.67)*
Num. of children aged 0-2	0.166	*	*(2.31)*	0.0403		*(0.82)*
Num. of children aged 3-5	0.174	*	*(2.29)*	0.144	**	*(2.94)*
Num. of children aged 6-11	0.186	**	*(2.92)*	0.149	***	*(3.87)*
Num. of children aged 12-15	0.241	***	*(3.63)*	0.152	***	*(3.71)*
Num. of children aged 16-18	0.0791		*(1.15)*	0.146	***	*(3.33)*
Youth aged 16-24	-0.0354		*(-0.62)*	0.0357		*(0.97)*
Older adults aged 65-74	-0.0667		*(-0.78)*	-0.158	***	*(-3.43)*
Older adults aged +75	-0.0221		*(-0.17)*	-0.138	*	*(-2.08)*
Household size	-0.0233		*(-0.58)*	0.0393		*(1.55)*
Education						
Ref. Completed primary school						
Can't read or write	0.162	*	*(2.20)*	0.195	***	*(3.42)*
Secondary school	-0.0566		*(-0.71)*	-0.385	***	*(-8.73)*
University degree	-0.177		*(-0.94)*	-0.812	***	*(-12.08)*
Labour market characteristics						
Ref. Inactive						
Employed	0.0497		*(0.58)*	0.146	*	*(2.27)*
Unemployed	-0.122		*(-1.46)*	0.396	***	*(4.79)*
Number of workers in the household	-0.230	***	*(-4.13)*	-0.201	***	*(-6.36)*
Family structure						
Ref. No partner						
With partner	0.0451		*(0.59)*	-0.00929		*(-0.19)*
Ref. No lone family with dependent children						
Lone family with dependent children	0.0400		*(0.24)*	0.197		*(1.05)*
Housing						
Ref. Own housing (no mortgage)						
Own housing, mortgage	-0.216	**	*(-3.26)*	-0.154	***	*(-3.70)*
Rent	0.0710		*(0.98)*	-0.0719		*(-1.35)*
Subsidized or rent free	0.0200		*(0.26)*	-0.00434		*(-0.07)*
Time						
Ref. Wave 1						
Wave 2	0.241	***	*(3.33)*	0.0557		*(1.14)*
Wave 3	0.108		*(1.56)*	-0.0679		*(-1.39)*
Wave 4	0.169	*	*(2.29)*	-0.00239		*(-0.05)*
Wave 5	0.134		*(1.74)*	0.0386		*(0.77)*
Wave 6	0.156	*	*(2.03)*	0.123	*	*(2.46)*
Constant						
Log-pseudolikelihood						

Model chi-square			
Obs. individual_wave	47.959		
Clusters			

Source: Own construction using the ECHP, 1994-2001. Individuals aged 25 to 64. The reference person is a Spanish male aged around 45 that leaves alone, is inactive, completed primary school, lives in his own property, does not have a chronic disease, is an original sample member and replied to a face-to-face interview. Significance level; * `p<-05, ** p<.01, ***p<-001

Table 6. Retention and poverty status at base year coefficients for Spain 1994-2000

	Retention			Poverty at $t-1$		
Individual characteristics	Coefficient		*(t-ratio)*	Coefficient		*(t-ratio)*
Ref. Male						
Female	0.00705		*(0.81)*	-0.00506		*(-0.48)*
Age	0.00134		*(0.19)*	0.0194		*(1.74)*
Age^2	0.0000001		*(0.00)*	-0.00024		*(-1.87)*
Household head characteristics						
Demographic characteristics						
Ref. Male						
Female	-0.00827		*(-0.27)*	0.318	***	*(7.58)*
Age	-0.00129		*(-0.24)*	-0.0742	***	*(-9.13)*
Age^2	0.0000142		*(0.25)*	0.00080	***	*(9.23)*
Ref. Spanish						
Immigrant	-0.246	**	*(-2.87)*	0.297	*	*(2.06)*
Num. of children aged 0-2	0.109	**	*(3.02)*	0.195	***	*(4.39)*
Num. of children aged 3-5	0.136	***	*(3.52)*	0.265	***	*(6.12)*
Num. of children aged 6-11	0.118	***	*(4.10)*	0.332	***	*(8.90)*
Num. of children aged 12-15	0.118	***	*(3.91)*	0.358	***	*(9.05)*
Num. of children aged 16-18	0.119	***	*(3.82)*	0.371	***	*(9.22)*
Youth aged 16-24	0.0448		*(1.80)*	0.181	***	*(5.14)*
Older adults aged 65-74	0.0135		*(0.48)*	-0.398	***	*(-9.08)*
Older adults aged +75	-0.153	***	*(-4.05)*	-0.624	***	*(-8.34)*
Household size	-0.121	***	*(-6.84)*	0.0117		*(0.44)*
Education						
Ref. Completed primary school						
Can't read or write	0.0295		*(0.77)*	0.540	***	*(11.03)*
Secondary school	0.0521		*(1.75)*	-0.424	***	*(-9.77)*
University degree	0.0306		*(0.92)*	-1.029	***	*(-16.29)*
Labour market characteristics						
Ref. Inactive						
Employed	0.0986	*	*(2.37)*	-0.219	***	*(-4.10)*
Unemployed	-0.0229		*(-0.41)*	0.435	***	*(7.19)*
Number of workers in the household	-0.0218		*(-1.20)*	-0.354	***	*(-12.82)*
Family structure						
Ref. No partner						
With partner	0.170	***	*(5.27)*	0.0842		*(1.85)*
Ref. No lone family with dependent children						
Lone family with dependent children	-0.116		*(-1.05)*	0.489	***	*(3.85)*
Housing						
Ref. Own housing (no mortgage)						
Own housing, mortgage	-0.0326		*(-1.12)*	-0.178	***	*(-4.54)*
Rent	-0.169	***	*(-4.88)*	0.141	**	*(2.74)*
Subsidized or rent free	0.00722		*(0.15)*	0.329	***	*(6.06)*
Time						
Ref. Wave 1						
Wave 2	-0.0987	**	*(-2.88)*	-0.00967		*(-0.31)*
Wave 3	-0.0852	*	*(-2.44)*	0.0847	**	*(2.58)*

Wave 4	0.00726		*(0.20)*	0.0907	*	*(2.57)*
Wave 5	-0.0808	*	*(-2.20)*	0.117	***	*(3.31)*
Wave 6	0.0975	*	*(2.52)*	0.133	***	*(3.55)*
Instruments						
Ref. Participates 1st. time after w1						
Original sample member	0.241	***	*(6.21)*			
Ref. Face-to-face interview						
Self-administered interview	-0.171	**	*(-2.99)*			
Telephone interview	-0.193	**	*(-2.87)*			
Proxy interview	-0.0778	***	*(-3.49)*			
Ref. HH does not have a chronic disease						
HH has a chronic disease				0.232	***	*(3.89)*
Constant	1.040	***	*(5.82)*	0.273		*(1.04)*

Source: Own construction using the ECHP, 1994-2001. Individuals aged 25 to 64. The reference person is a Spanish male aged around 45 that leaves alone, is inactive, completed primary school, lives in his own property, does not have a chronic disease, is an original sample member and replied to a face-to-face interview. Significance level; * `p<-05, ** p<.01, ***p<-001

Appendix

Table A.1. Headcount ratio (whole population) and number of individual_wave observations by income definition.

	Income refers to	Income definition [1]		Income definition [2]	
Wave	$t-1$	Poor	Obs	Poor	Obs.
1 (1994)	1993	19,59	22.837	-	
2 (1995)	1994	18,98	20.458	18,66	18.677
3 (1996)	1995	17,97	19.278	18,21	17.488
4 (1997)	1996	20,34	17.916	20,63	16.183
5 (1998)	1997	18,18	16.598	18,31	15.026
6 (1999)	1998	18,89	15.863	18,89	14.168
7 (2000)	1999	18,02	14.784	18,48	13.349
8 (2001)	2000	18,82	14.270	18,86	12.935

Source: Own construction using the ECHP, 1994-2001. [1] Household income refers to $t-1$ but the equivalence scale refers to t. [2] Household income and equivalence scale refer to t. Notice that the headcount ratio for 1993 cannot be obtained as we do not know household composition for this year. Cross-sectional individual weights used.

Taula A.2. Poverty status at t conditional of poverty status at $t-1$ in Spain with missing income information and by income definition (whole population)

Year		t		
		Not poor	**Poor**	**Missing**
[1] $t-1$	**Not poor**	80,89	7,75	11,36
	Poor	34,56	53,85	11,59
	Total	*72,13*	*16,46*	*11,40*
[2] $t-1$	**Not poor**	78,22	7,36	14,42
	Poor	32,83	53,07	14,10
	Total	*69,66*	*15,98*	*14,36*

Source: Own construction using the ECHP, 1994-2001. [1] Household income refers to $t-1$ but the equivalence scale refers to t (1993-2000). [2] Household income and equivalence scale refer to t (1994-2000).[26] Longitudinal individual weights used.[27]

[26] Results are very similar if we restrict the transition matrix to the period 1994-2000. Entry rate is 7,70 and exit rate is 34,79.

[27] We are aware of Eurostat's (2003) recommendation to use longitudinal weights for the latest wave available. However, in our case, weights refer to t -1 given that transitions to missing would otherwise be erased.