

TESTING FOR TRUE STATE DEPENDENCE IN POVERTY DYNAMICS

Anna Giraldo, Enrico Rettore and Ugo Trivellato
Dipartimento di Scienze Statistiche, Università di Padova

This version: January 30, 2007

Very preliminary and incomplete, for comments only. Please, do not quote.

For additional information please contact:

Enrico Rettore
Dipartimento di Scienze Statistiche, Università di Padova
Via Cesare Battisti, 241 - 35121 Padova – Italy
E-mail: enrico.rettore@stat.unipd.it
Fax: ++39/0498274170
Phone: ++39/0498274164

TESTING FOR TRUE STATE DEPENDENCE IN POVERTY DYNAMICS

Anna Giraldo, Enrico Rettore and Ugo Trivellato *
Dipartimento di Scienze Statistiche, Università di Padova

Abstract

Evidence from several countries is that any household experiencing poverty today is much more likely to experience it again, which may be due to both true state dependence (TSD) and unobserved heterogeneity (UH). We deal with UH by specifying two sources of it: (i) the household's ability to obtain income in a specific time period, and (ii) the way in which this ability evolves from that time period onwards. We test for TSD using the panel component of the Italian Survey on Household Income and Wealth (SHIW). After testing for the ignorability of the massive attrition plaguing the panel and accepting it, we do not find any sign of TSD.

Keywords: Attrition ignorability, Discrete response panel data models, Poverty dynamics.

JEL-code: I32, C23, C25

1. Introduction

There is evidence from several countries that any household experiencing a poverty spell today is much more likely to experience it again in the future (for comparative cross-country analyses, see Duncan *et al.*, 1993; Oxley, Dang and Antolín, 2000; Mejer and Linden, 2000; OECD, 2001).

As an example referring to Italy, using data from the panel component of the Italian Survey on Household Income and Wealth (SHIW), late 1980s/early 1990s, Trivellato (1998) obtained the following figures:

$$\Pr(I_{it}=1 | I_{it-1}=0) \cong .05$$

$$\Pr(I_{it}=1 | I_{it-1}=1) \cong .50,$$

where $I_{it}=1$ if the i -th household disposable income is below the poverty threshold at time t , and $I_{it}=0$ otherwise. Analogous calculations, still on the panel component of SHIW, for the late 1990s and early 2000s confirm these orders of magnitude. Plainly, this evidence documents a high degree of persistence of poverty, as measured on disposable income.

There are two logically distinct, albeit possibly concomitant, processes which may generate such a serial persistence of poverty. It might be that the fact of experiencing poverty in a specific time period, in itself, increases the probability of undergoing poverty in subsequent periods (through human capital deterioration, decreasing self-esteem, *etc.*).

* Research for this paper was supported by grants No. xxxx and yyyy from the Italian Ministry of Education, University and Research under the PRIN Programme.

Since the work of Heckman (1978), such a process is said to exhibit *true state dependence* (TSD).

On the other hand, households might be heterogeneous with respect to characteristics which are relevant for the chance of falling into poverty and persistent over time. If that is the case, any household likely to experience poverty at time t because of (possibly unobserved) adverse characteristics will also be likely to experience poverty in any other period because of the very same adverse characteristics. [We refer to this process as steered by unobserved heterogeneity \(UH\).](#)

Distinguishing between the two processes is crucial, since the policy implications are very different. If the persistence of poverty is (at least partly) due to TSD, then a policy aimed at breaking the “vicious circle” via monetary transfers makes sense. Forcing households out of poverty at time t will reduce their chance of experiencing poverty in the future. But if the persistence of poverty is due only to UH, any monetary transfers policy aimed at breaking the “vicious circle” is pointless. Forcing households out of poverty today does not affect their adverse characteristics, hence does not reduce their chance of experiencing poverty spells in subsequent periods¹.

Eliminato: unobserved heterogeneity

It is worth noting that a not negligible part of the empirical literature only descriptively juxtaposes the two potential sources of poverty persistence, without trying to ascertain whether, after accounting for UH, there is also TSD, and without assessing their respective degrees of importance².

However, a growing strand of literature focuses on the issue of heterogeneity, both observed and unobserved, and TSD in poverty dynamics. Noticeable contributions include Stevens (1999), Cappellari and Jenkins (2002), Devicienti (2002). A similar route is taken by Chay, Hoynes and Hyslop (2001), whose focus is on welfare dynamics rather than on poverty dynamics³.

Broadly speaking, these papers conclude in favour of the presence of TSD. Two comments, and caveats, are in order. First, while the models and estimations methods used by the various authors vary considerably, they share some common traits. They rely on parametric specifications and distribution assumptions, which are partly motivated by computational convenience. In particular, all models assume a first-order stationary Markov chain for state dependence, and combine it with individual fixed-effect or random-effect models to deal with UH. The question arises whether results depend on the imposed functional forms and distribution assumptions.

¹ The argument is relevant for programmes of income support viewed (partly) as an active policy, inspired by an efficiency rationale – helping the poor to exit from poverty in order to stay out of it definitely. Of course, it does not call into question the reasonableness of such programmes simply as a passive welfare policy, resting on equity grounds. Clearly, it also leaves room for other active anti-poverty programmes, aimed at contrasting the adverse personal and/or household characteristics and permanent shocks responsible for the chance of persistently experiencing poverty.

² For instance, Oxley, Dang and Antolín (2000, p. 6) summarize the key results of their study across six OECD countries in the following terms: “(ii) *The probability of exiting poverty falls with previous experiences in poverty. At the same time, there is a high probability of falling back into poverty. Thus, for the longer-term poor, low probability of exit and high probability of re-entry tend to reinforce each other. [...]* (iv) *The characteristics of households experiencing shorter spells in poverty tend to be different from those of the longer-term poor. A large share of the longer-term poor would appear to be women, lone parents and elderly single individuals. A significant share of the longer-term poor are in paid work.*”

³ Related models have been applied to studies of income mobility. See, among others, Stewart and Swaffield (1999), Cappellari (2001), Cappellari and Jenkins (2004).

Second, these studies refer to countries where some general, nationwide measure(s) on income maintenance is/are in operation. In such contexts, it is problematic to identify the possible TSD in poverty dynamics from welfare dependence, that is from the “welfare trap” induced by income maintenance measure(s)⁴.

In this paper, we test for TSD in poverty/non poverty sequences while allowing for a flexible specification of UH. We use a panel sample from SHIW, a split-panel survey carried out on a two-year basis, over the period 1989-2004.

Since the works of Heckman (1978, 1981a), it is well-known that panel data allow the issue to be tackled. By studying the pattern of the sequences $\{I_{i1}, I_{i2}, \dots, I_{iT}\}$, we can identify whether TSD is at work. In order to test properly for TSD in poverty/non-poverty sequences based on disposable income, we argue that it takes accounting for two sources of UH: (i) the household’s ability to obtain income at a specific, initial time period, and (ii) the way in which this ability evolves from that time period onwards. A consequence of this double source of UH is that simple models for TSD in the presence of UH (e.g., fixed-effect or random-effect models) may badly miss the point.

Eliminato: y_{i2}, \dots

As for the empirical analysis, the fact that it is carried out on a pooled panel dataset from Italy has, ironically, the advantage of being substantially free from the confounding problem of welfare dependence. In fact, Italy does not have a general measure of income maintenance. A demonstration of *Reddito minimo d’inserimento*, a sort of minimum income programme, was implemented in the years 1998-2002, but at a very small scale (Sacchi, 2007). Besides, the overall relevance of regional or local measures of income maintenance is modest. Thus, for Italy the assumption of lack of welfare programmes for the poor is a reasonable simplification. As a consequence, on finding that TSD does exist we could unambiguously conclude that it is due to the occurrence of poverty. On the other hand, finding that there is no TSD in Italy would provide an indirect argument to argue that the existence of TSD in a country providing income supports to the poors should be interpreted as welfare dependence.

The paper proceeds as follows. In section 2 we present a sort of textbook model, that captures the essential features of most specifications adopted to analyze poverty (or welfare) dynamics. In section 3, moving from the life-cycle permanent income hypothesis we take on a richer model, that allows for a flexible specification of UH and for more complex dynamics. In this context, we also assess the consequences of mistakenly testing for TSD in poverty/non-poverty sequences under the textbook model presented in the previous section. In section 4 we present our Instrumental Variable (IV) tests for TSD. As for the empirical analysis, since the SHIW panel is plagued by massive attrition, preliminarily we develop a test on whether such sample selection is ignorable for the purpose of testing for TSD. We conclude in favour of ignorability (section 5). The main results for the model of interest are presented in section 6. They provide no evidence of TSD. Some sensitivity analyses, aimed at ascertaining the robustness of results to changes in the poverty line, corroborate that conclusion. Interestingly enough, we show that, had we tested for TSD under (some of) the textbook model restrictions on unobservables, we would have (mistakenly) concluded the opposite way.

Section 7 outlines our conclusions, which can be summarized in two statements. Firstly, while it is apparent that the SHIW panel sample is biased by attrition, with households less likely to experience poverty surviving longer in the sample, we also find evidence that attrition is basically ignorable for the specific purpose of testing for TSD.

⁴ Contini and Negri (2005) offer an interesting presentation of the problem.

Secondly, after accounting for the two sources of heterogeneity, we do not find any sign of TSD.

2. Testing for TSD in the presence of UH: the textbook model

The textbook model to test for TSD in the presence of UH, as adapted to our problem, is the following (see, for instance, Hsiao, 1986):

$$I_{it} = \mathbf{1}(\alpha_i + \varphi \cdot I_{it-1} + \varepsilon_{it} < c), \quad t=1, T, \quad (1)$$

where:

- $\mathbf{1}(\cdot)$ is an indicator function equal to 1 if the condition in its argument is satisfied and 0 otherwise;
- the binary variable I_{it} is equal to 1 if disposable income y_{it} falls below c , a poverty threshold usually referred to as the poverty line, and 0 otherwise⁵;
- the model allows for UH through α_i , an unobserved characteristic which makes individuals heterogeneous in a time-invariant way: the lower α_i the higher the chance for the i -th individual to experience $I_{it}=1$ in each time period;
- $\{\varepsilon_{it}\}$ is a sequence of serially independent, zero mean, identically distributed random variables.

The value of φ determines whether the sequence $\{I_{it}\}$ features TSD. If $\varphi < 0$, then experiencing poverty at time $t-1$, $I_{it-1}=1$, *causes* a lower disposable income at time t , hence increases the chance to experience $I_{it}=1$:

$$\Pr(I_{it}=1 | I_{it-1}=1, \alpha_i) = \Pr(\varepsilon_{it} < -\alpha_i - \varphi) > \Pr(\varepsilon_{it} < -\alpha_i) = \Pr(I_{it}=1 | I_{it-1}=0, \alpha_i).$$

With reference to this framework, an adequate representation for UH is crucial for proper testing for TSD. A direct check on whether $\Pr(I_{it}=1 | I_{it-1}=1) > \Pr(I_{it}=1 | I_{it-1}=0)$ does not provide that, since in the presence of UH (*i.e.*, $\text{var}\{\alpha_i\} > 0$) we are bound to observe $\Pr(I_{it}=1 | I_{it-1}=1) > \Pr(I_{it}=1 | I_{it-1}=0)$ even if $\varphi = 0$.

Moving from model (1) – a sort of reference model for most empirical research on poverty dynamics, alternative strategies to implement the test for TSD include (i) conditioning on a sufficient statistic for α_i , and (ii) imposing some structure on the distribution of α_i (see Arellano and Honoré, 2001, and Arellano, 2003, for an up-to-date reviews).

The first strategy was pioneered by Chamberlain (1982, 1985). It works in those instances in which a sufficient statistic, SS_i say, exists for parameter α_i . Exploiting this sufficient statistic, $\Pr(I_{i1}, \dots, I_{iT} | SS_i; \varphi, \alpha_i)$ – the probability of observing a specific sequence on the i -th unit conditional on SS_i – turns out to be independent of α_i , thus allowing one to infer on φ .

As regards the second strategy, by assuming that UH is distributed in a specific way (*i.e.*, by interpreting (1) as a random-effect model and imposing a distribution assumption), we can obtain a likelihood function for φ by integrating out the unobserved α_i . There is an additional problem here with initial condition I_{i1} , because the analyst very often does not

Eliminato: to test

Eliminato: α_i ,

Eliminato: SS_i ,

⁵ We reverse the inequality with respect to the conventional practice, in order to be consistent with the notation used throughout the paper, where $I_{it}=1$ denotes that the i -th household at time t is poor.

know whether I_{i1} has been generated by the same model as the subsequent observations (see Heckman, 1981b).

In the analysis of poverty/non-poverty or welfare participation sequences, the first strategy is adopted, *e.g.*, by Chay, Hoynes and Hyslop (2001), who model welfare dynamics in California. A nice example of the second strategy is given by Cappellari and Jenkins (2002): they model poverty dynamics in Great Britain by adding parameter specifications and distribution assumptions on initial conditions and the panel attrition process.

3. How does income evolve over time? A flexible specification for UH

In this section, we move from the well-established literature on life-cycle permanent income hypothesis, and derive a model for poverty/non-poverty sequences with a flexible specification for UH. We will then show why the textbook model (1) does not provide an adequate representation of the features of poverty dynamics.

Following Hall (1978), let us represent disposable income y_{it} as:

$$y_{it} = y_{it-1}^p + S_{it} \quad t=1, T, \quad (2)$$

with $y_{it-1}^p \perp S_{it}$, where y_{it-1}^p represents the *expected income* for time t on the basis of the information available up to time $t-1$. In the absence of any unexpected event, current income at time t equals y_{it-1}^p . S_{it} represents unexpected departures of current income from y_{it-1}^p , as seen from time $t-1$. Being a prediction error, S_{it} is orthogonal to the predicted value y_{it-1}^p .

Moreover, let us represent S_{it} as:

$$S_{it} = u_{it} + v_{it}, \quad (3)$$

where u_{it} is the permanent component of the shock. It summarizes the impact of all new information that becomes available in period t about the household lifetime well-being. In this sense, it lastingly affects income *from time t onwards*. As for v_{it} , it is the transitory component of the shock, which affects income *only at time t* .

As a consequence, the sequence of expected income follows a random walk:

$$y_{it}^p = y_{it-1}^p + u_{it}, \quad (4)$$

possibly with a drift if $E\{u_{it}\} \neq 0$, while the sequence of first differences in current income follows a MA(1) process:

$$\Delta y_{it} = u_{it} + v_{it} - v_{it-1}. \quad (5)$$

Compared to model (1), in model (2)-(3) there are *two sources* of across-household heterogeneity here, $\sigma_I^2 = \text{var}\{y_{it}^p\}$ and $\sigma_{ut}^2 = \text{var}\{u_{it}\}$. Households differ with respect to their expected income at time $t=1$ and they also differ with respect to the way in which the sequence of permanent shocks u_{it} shapes the pattern of expected income from period $t=1$ onwards. Model (1) emerges as a special case of model (2)-(3) by setting $\text{Var}\{u_{it}\} = 0$, $t =$

1, T , in which case $y_{it}^p = y_{it-1}^p$ plays as the time-invariant unobserved characteristic α_i in (1), with respect to which households are heterogeneous.

In this set-up, TSD adds a further source of serial dependence:

$$y_{it} = y_{it-1}^p + u_{it} + \varphi I_{it-1} + v_{it}, \quad t=1, T. \quad (6)$$

The qualitative difference made by TSD ($\varphi < 0$) is the following. If $\varphi = 0$, then:

$$y_{it} \perp v_{is}, \quad \forall s \neq t, \quad (7)$$

i.e., the transitory shock affects only contemporary income. Instead, if $\varphi < 0$, then y_{it} is *not* independent of lagged values of the transitory shock v_{it} .

With reference to this specification, five comments are in order. First, the theory of consumer behaviour under the life-cycle permanent income hypothesis offers implications to discriminate between model (1) and model (2)-(3). Under suitable conditions, it states that, household consumption, C_{it} , equals permanent income: $C_{it} = y_{it}^p$ (see again Hall, 1978). And, most importantly for us, the empirical evidence show that ΔC_{it} is uncorrelated to the past history of consumption (see, *e.g.*, Blundell and Preston, 1998), *i.e.*, it behaves as the permanent shocks in the permanent income sequences. This suggests that just one formulation of economic theory, the one bringing to model (9), has implications which are consistent with the evidence. At least for the case of poverty defined with reference to household disposable income, model (8) is thus incompatible with known properties of the time pattern of the income-consumption couple.

Second, there is a strand in the literature that models the dynamics of disposable income, and then recovers from that model the implications for the dynamics of poverty (see Lillard and Willis, 1978, and Stevens, 1999, section VI, among others). As it is apparent, we take a different route. We model directly the poverty/non poverty sequence $\{I_{i1}, \dots, I_{iT}\}$, that is to say, transition probabilities into and out of the lower portion of the income distribution (as, *e.g.*, Stevens, 1999, sections III-V; the same route is taken by Stewart and Swaffield, 1999, in modelling low pay dynamics). The motivation for our choice is that the dynamics of high income is hardly relevant to the study of the dynamics of poverty. Thus, our analysis has the advantage of being unaffected by movements within the upper portion of the distribution of income. There is, of course, a price to pay for this approach, which consists of imposing an arbitrary cut-off – the poverty line. Clearly, results might be sensitive to it. To mitigate this arbitrariness, we will use three thresholds, reasonably spread over the range usually considered for poverty analyses, and check whether results change as the poverty line is modified.

Third, a word of caution needs to be added on how rejection of the null hypothesis should be interpreted. In principle, rejecting the no TSD hypothesis, *i.e.* finding evidence that I_{it} is *not* independent of lagged values of v_{it} , need not be due to TSD, in that serially correlated transitory shocks may also induce a departure from (7). Note, however, that if we accept (7) we unambiguously conclude against TSD.

Finally, it is worth noting that eq. (6) allows us to assess the consequences of mistakenly testing for TSD in poverty/non-poverty sequences within model (1), *i.e.*, omitting across-household heterogeneity due to the sequence of permanent shocks. To exemplify, consider the couple of observations (I_{i2}, I_{i3}) under the no TSD case. Conditional on y_{i1}^p they are *not* independent, since they are both affected by the permanent shock u_{i2} :

$$\Pr(I_{i3} = 1 | I_{i2} = 1, y_{i1}^p) > \Pr(I_{i3} = 1 | I_{i2} = 0, y_{i1}^p).$$

Eliminato: -interruzione pagina-

Eliminato: y_{i2}, \dots

Eliminato: ic

Since model (1) does not account for this (positive) dependence of I_{i3} on I_{i2} , this [dependence](#) is picked up by the TSD parameter. Once again, it looks like TSD but in fact it is only omitted heterogeneity.

4. An IV test for TSD

In this section we present the econometrics relevant to our test. Exploiting the notation introduced in the previous section, disposable income at time t is linked to permanent income at time $t-1$, to permanent and transitory shocks at time t , and, under the alternative hypothesis of TSD, to the experience of a spell of poverty at time $t-1$ by the following relationship:

$$y_{it} = y_{it-1}^p + \theta_t u_{it} + \sigma_t v_{it} + \varphi I_{it-1}, \quad (10.1)$$

where we have redefined u_{it} and v_{it} as unit variance shocks, so that the standard deviations of the permanent and transitory shock hitting disposable income at time t are equal to θ_t and σ_t , respectively,

Consequently, the binary variable I_{it} is:

$$I_{it} = \mathbf{1}(y_{it} < c). \quad (10.2)$$

Let F be the cumulative distribution of v_{it} , which we assume shape stationary. The probability to experience a poverty spell at time t conditional on past history and on the contemporary permanent shock is:

$$\Pr(y_{it} < c) = F[(c - y_{it-1}^p - \theta_t u_{it} - \varphi I_{it-1}) / \sigma_t] \quad (11)$$

(to simplify the notation here and in the following we leave implicit the conditioning variables).

We develop our analysis on a first order Taylor series expansion of (11) approximating it in a neighbourhood of $\theta_t/\sigma_t=0$, $\sigma_{t-1}/\sigma_t=1$ and $\varphi/\sigma_t=0^6$:

$$\Pr(y_{it} < c) = \Pr(y_{it-1} < c) - \varphi f_{it-1} (I_{it-1} - I_{it-2}) / \sigma_{t-1} - \theta_t f_{it-1} u_{it} / \sigma_{t-1} + (\sigma_{t-1} / \sigma_t - 1) f_{it-1} (c - y_{it-1}^p) / \sigma_{t-1}, \quad (12)$$

in which f_{it-1} is the probability density associated to F as evaluated at $(c - y_{it-1}^p) / \sigma_{t-1}$.

Eq. (12) displays the basic features of the dynamics of poverty.

- The probability at time t is obtained from the probability at time $t-1$ by adding the permanent shock u_{it} scaled by the individual specific parameter $\theta_t f_{it-1} / \sigma_{t-1}$.
- If $\sigma_{t-1} / \sigma_t < 1$ ($\sigma_{t-1} / \sigma_t > 1$), *i.e.*, if the variance of the transitory shock increases (decreases) from time $t-1$ to time t , subjects whose permanent income was below the poverty line at time $t-1$ decrease (increase) their probability to experience a poverty spell, while subjects whose permanent income was above the poverty line increase (decrease) it.

⁶ These conditions have an immediate substantive interpretation. The Taylor expansion is useful in the case when (i) the variance of the permanent shock is small with respect to the variance of the transitory shock, (ii) the variances of transitory shocks change slowly over time, so that in two consecutive times, $t-1$ and t , they are almost the same, and (iii) the state dependence parameter φ is small with respect to the standard deviation of the transitory shock.

The intuition is that if the random variability of income increases (decreases), then the probability to flow from one state to the other increases (decreases) as well.

- In the presence of TSD ($\varphi < 0$), subjects entering poverty at time $t-1$ ($I_{it-1}=1$ and $I_{it}=0$) at time t have a higher chance to experience a spell of poverty again. On the other hand, subjects leaving poverty at time $t-1$ ($I_{it-1}=0$ and $I_{it-2}=1$) have a lower chance to experience a spell of poverty at time $t-1$.
- Finally, note that the one-period variation of the probability to experience a spell of poverty, $\Pr(y_{it} < c) - \Pr(y_{it-1} < c)$, is scaled by the individual specific term f_{it-1} . Since this individual specific term increases as the permanent income at time $t-1$ gets close to the poverty line, then permanent shocks, TSD and variation over time of the transitory shock shape the dynamics of poverty only for those individuals whose permanent income is not too far from c . Individuals whose permanent income at time $t-1$ is far from c are unaffected by these forces.

To derive a test of the null hypothesis $\varphi=0$ against the alternative $\varphi < 0$ let us start with the simpler case in which *transitory shocks feature a stationary variance*. Then, eq. (12) simplifies to:

$$\Pr(y_{it} < c) = \Pr(y_{it-1} < c) - \varphi f_{it-1} (I_{it-1} - I_{it-2})/\sigma - \theta_t f_{it-1} u_{it} / \sigma. \quad (13)$$

Let ϵ_{it} be the deviation of I_{it} from its mean $\Pr(y_{it} < c)$ and rewrite (13) as

$$I_{it} = I_{it-1} - \varphi f_{it-1} (I_{it-1} - I_{it-2})/\sigma - \theta_t f_{it-1} u_{it} / \sigma + \epsilon_{it} - \epsilon_{it-1}. \quad (14)$$

This is a regression in which both regressors are correlated to the component $\epsilon_{it} - \epsilon_{it-1}$ of the disturbance term (but note they are *not* correlated to $\theta_t f_{it-1} u_{it} / \sigma$, since the permanent shock is orthogonal to past history). Note however that in the absence of TSD I_{it-2} is a valid instrument for the explanatory variable I_{it-1} (it is uncorrelated both with u_{it} and $\epsilon_{it} - \epsilon_{it-1}$). As a consequence, if the null hypothesis holds true the resulting IV estimate should depart from 1 only because of sampling variability.

If the null fails to hold, it is straightforward to show that the IV estimate is systematically smaller than 1. This is because, as a result of the persistence induced by TSD, the following inequality holds:

$$E(I_{it} - I_{it-1} | I_{it-2}=1) < E(I_{it} - I_{it-1} | I_{it-2}=0). \quad (15)$$

By rearranging it we end up with the inequality:

$$E(I_{it} | I_{it-2}=1) - E(I_{it} | I_{it-2}=0) / [E(I_{it-1} | I_{it-2}=1) - E(I_{it-1} | I_{it-2}=0)] < 1, \quad (16)$$

where the left hand side term is the IV estimand in the case we are considering.

Summing up, when the variance of the transitory shock is stationary, testing for TSD amounts to reject the null when the estimate of the coefficient of I_{it} on I_{it-1} , using I_{it-2} as an IV, is smaller than 1 more than it is expected due to sampling variability.

When we let the *variance of v_{it} to change over time*, I_{it-2} is no longer a valid instrument since it is correlated to $(\sigma_{t-1}/\sigma_t - 1) f_{it-1} (c - y_{it-1}^p)/\sigma_{t-1}$ (negatively if $\sigma_{t-1}/\sigma_t < 1$, positively otherwise). In the Appendix we prove that at the first order of approximation ($I_{it-2} - I_{it-3}$) is a valid instrument in the first-differenced model:

$$I_{it} - I_{it-1} = -\varphi f_{it-1} (I_{it-1} - I_{it-2})/\sigma_{t-1} - \theta_t f_{it-1} u_{it}/\sigma_{t-1} + (\sigma_{t-1}/\sigma_t - 1) f_{it-1} (c - y_{it-1}^p)/\sigma_{t-1} + \epsilon_{it} - \epsilon_{it-1}. \quad (17)$$

Note however that if the null hypothesis holds true, *i.e.* $\varphi=0$, this IV is uncorrelated with the endogenous explanatory variable in (17), implying that the variance of the resulting estimate is infinitely large. To the specific purpose of testing for TSD – *not* to the purpose of reliably estimating the TSD parameter, a way out is to estimate the reduced form parameter by regressing $(I_{it} - I_{it-1})$ on the instrument $(I_{it-2} - I_{it-3})$. This is equal to zero if and only if $\varphi=0$.

5. Is attrition in the SHIW panel ignorable for the purpose of testing for TSD?

We move now to the empirical analysis. As anticipated, we use information on household disposable income from the SHIW, by far the most reliable survey on income and wealth in Italy. Since the late 1980s SHIW is carried out on a bi-annual⁷ basis according to a split-panel design (Banca d'Italia, various years, and Brandolini, 1999). Specifically, we exploit the panel component of the survey, available since 1989 up to 2004.

A major problem with this panel is a severe attrition (here and in the following, we use the term in a broad sense, *i.e.*, including the effect of the design). Table 1 shows how the number of households still in the survey among those entering the survey in a specific calendar year sharply decreases over time. While some of the decrease is by design – the split-panel, part of it is due to lack of survey manager over the following rules⁸.

Table 1 about here

To document the bias resulting from attrition, we partition the panel by grouping together households according to the years in which they entered and left the panel, respectively. As an example, out of the 8,274 households entering the 1989 sample, we get a two-wave sub-panel made up of 1,137 (=2,187–1,050) households who left the survey in 1991; a three-wave sub-panel made up of 223 (=1,050–827) households who left the survey in 1993 and so on, up to the eight-wave sub-panel made up of 230 households still in the survey in 2004. This way we get 28 mutually exclusive sub-panels of different length.

We computed the poverty head-count ratio on selected sub-panels. From Figure 1 it is apparent that time-in-survey is correlated to the probability to experience a spell of poverty: households staying longer in the survey are less likely to experience poverty throughout the whole time window we consider.

Figure 1 about here

Note however that the bias induced by attrition need not be a problem to our test for TSD. This is because our test for TSD depends on the micro-data only through the probabilities of transition between states, not through the head-count ratios. If the attrition

⁷ An exception is the 1998 wave, which took place *three* years after the previous one.

⁸ An analysis of the attrition process for the period 1989-1995 is in Giraldo, Rettore and Trivellato (2001).

process does not bias the inference on such transition probabilities, then we may say that it is ignorable to our test for TSD.

To get evidence on whether the attrition process biases inference on the transition probabilities we estimate them separately on each of the 28 mutually exclusive sub-panels. As an example, for the transition matrix 1989-91 we get seven independent estimates from the two-wave sub-panel 1989-91, from the first two waves of the three-wave sub-panel 1989-91-93, up to the first two waves of the eight-wave sub-panel 1989-91-...-02-04.

Table 2 about here

If such independent estimates were equal up to sampling variability we could confidently conclude that the attrition process is ignorable to our purpose of testing for TSD⁹. In Table 2 we report the usual likelihood-ratio statistics (and the corresponding degrees of freedom) separately for each transition matrix. The null hypothesis is not rejected in most cases. Just in a few cases there is some evidence that the estimates are not equal.

Overall, we conclude that even if the attrition process in SHIW badly biases the estimation of the head-count ratio, it is ignorable, or nearly so, for the estimation of the transition probabilities, hence for our test for TSD.

6. Testing for TSD: results

As for the empirical analysis of poverty dynamics and TSD, we go along the criteria suggested by Eurostat (xxxx) and keep to the following operational criteria in order to identify poor households:

- (a) We make use of the OECD modified equivalence scale, that weights 1 the first adult, 0.5 each additional household member at least 14 years old, 0.3 each member younger than 14.
- (b) We set the poverty line in the calendar year 1995 at 60% of the median equivalent income in that year;
- (c) As for the variability of the poverty line over time, we simply derive the poverty line for the other years deflating/inflating the 1995 poverty line by means of the consumer price index¹⁰.

As a robustness check, we replicate our analysis with two alternative poverty lines set at 50% and 70% of the median equivalent income, respectively.

Main results are summarized in Table 3. Our IV test robust to heteroschedastic transitory shocks as well as to the presence of permanent shocks does not provide any evidence pointing to the existence of TSD.

⁹ In fact, for the ignorability of attrition to our test it would be enough to show that time-in-survey is ignorable for the TSD parameter. To keep the exposition simple we omit this discussion.

¹⁰ When confronted to the (modest) real growth of equivalent household disposable income – which is the typical pattern in our sample period, except for the strong recession in 1993 (Miniaci e Weber 1999)–, that choice leads to an estimate of poverty persistence (and, similarly, of the poverty head-count ratio) slightly lower with respect to the one implied by a strictly relative threshold, *i.e.*, by computing a poverty line specific to every year.

Table 3 about here

With reference to eq. (17), using the base poverty line we get a largely insignificant t -statistic for the estimated φ : it is as large as 0,2869. Using the alternative poverty lines results do not change (the t -statistics are as large as 2,1633 and 1,1843 in the 50% and in the 70% case, respectively). Note that in the case of the lower poverty line the t -statistic is statistically significant, but the sign is the opposite of the one expected under TSD.

Overall we conclude that the dynamics of poverty we observe in Italy over the period 1995-2004¹¹ does not feature any TSD.

A final issue worth paying attention is whether by testing for TSD in a conventional manner, *i.e.*, by imposing the textbook model restrictions on unobservables, our conclusions would have been different. Exploiting results presented in section 4, this question is easily answered to at least for the case in which the variance of the transitory shock is restricted to be stationary. Following the discussion on eq. (14), in such a case the estimate of the coefficient of I_{it} on I_{it-1} , using I_{it-2} as IV, under the null hypothesis departs from 1 only because of sampling variability whereas under the alternative hypothesis is systematically lower than 1. Our evidence in that it is by far significantly lower than 1, no matter for the adopted poverty line. That is, by missing to allow for a time-varying variance of the transitory shock we would mistakenly conclude that poverty dynamics features TSD.

7. Concluding remarks

We summarize the results of our analysis of the dynamics of poverty in Italy, 1989-2004, in three statements.

As for the SHIW panel we use, it is plagued by a severe attrition. There is a clear-cut evidence that the measure of poverty head-count ratio is severely biased by attrition, since the longer a household survives in the survey the lower its probability to experience a poverty spell. On the other hand – and crucial to our analysis of true state dependence, attrition does not appreciably bias the estimation of transition probabilities. Hence it is ignorable with respect to the purpose of testing for true state dependence.

As for the issue of interest, *i.e.*, whether after controlling for unobserved heterogeneity there is evidence of true state dependence, moving from the literature on the dynamics of income and consumption we have argued that the dynamics of poverty/non-poverty states based on household disposable income is driven by two sources of across household unobserved heterogeneity: (i) the household permanent income at a given, initial point in time, and (ii) the way in which such permanent income evolves over time as shaped by permanent shocks. Once these two sources of unobserved heterogeneity have been properly accounted for, we do not find any sign of true state dependence.

On the contrary, had we tested for true state dependence by imposing the stationarity of the transitory shock variance – a typical restriction imposed by textbook-

¹¹ The first three observations are lost since we work with a first-differenced model in which the explanatory variable is the dependent variable lagged once and the IV is the dependent variable lagged twice.

type models, we would mistakenly conclude that the dynamics of poverty is driven also by true state dependence.

This result, which turns out robust to alternative definitions of the poverty line, bears implication for the design of anti-poverty policies. Any policy providing income support to households falling below a specified poverty line is *not* called into question by our results, if it is intended as a *passive* policy aimed at reducing inequalities. Instead, our results imply that it cannot be used as an *active* policy: an income transfer *today* to a household below the poverty line *today* does not improve that household's chance to exit poverty *tomorrow*.

Of course this is not to mean that there is no room for anti-poverty active policies. The point is that such policies should be targeted to remove the individual characteristics (lack of education and skill, poor health, lack of social networks, say) – affecting the individuals from the outset or induced by permanent shocks – relevant for the household risk to persist into poverty.

Appendix

Rewrite equation (17) for time $t-2$:

$$I_{it-2} - I_{it-3} = -\varphi f_{it-3} (I_{it-3} - I_{it-4})/\sigma_{t-3} - \theta_{t-2} f_{it-3} u_{it-2}/\sigma_{t-3} + (\sigma_{t-3}/\sigma_{t-2} - 1) f_{it-3} (c - y_{it-3}^p)/\sigma_{t-3} + \epsilon_{it-2} - \epsilon_{it-3}. \quad (18)$$

The covariance between $(I_{it-2} - I_{it-3})$ and the unobservables in (17) does not vanish only because of the term:

$$\text{Cov}[I_{it-2} - I_{it-3}, (\sigma_{t-3}/\sigma_{t-2} - 1) f_{it-3} (c - y_{it-3}^p)/\sigma_{t-3}],$$

since u_{it} and $\epsilon_{it} - \epsilon_{it-1}$ are uncorrelated with past history. Working out this covariance we get:

$$\begin{aligned} & (-\varphi/\sigma_{t-3}) (\sigma_{t-3}/\sigma_{t-2} - 1) \text{cov}(f_{it-3} (I_{it-3} - I_{it-4}), f_{it-3} (c - y_{it-3}^p)/\sigma_{t-3}) + \\ & (-\theta_t/\sigma_{t-1}) (\sigma_{t-3}/\sigma_{t-2} - 1) \text{cov}(f_{it-1} u_{it}, f_{it-3} (c - y_{it-3}^p)/\sigma_{t-3}) + \\ & (\sigma_{t-1}/\sigma_{t-1}) (\sigma_{t-3}/\sigma_{t-2} - 1) \text{cov}(f_{it-1} (PL - y_{it-1}^p)/\sigma_{t-1}, f_{it-3} (c - y_{it-3}^p)/\sigma_{t-3}). \end{aligned}$$

Apparently, in a neighbourhood of the point we chose for our Taylor series expansion the coefficients in each of the three terms are negligible, allowing us to conclude that $(I_{it-2} - I_{it-3})$ is a valid instrument in eq. (17).

References

- Arellano, M. (2003), *Panel data econometrics*, Oxford University Press, Oxford.
- Arellano, M., and B. Honoré (2001), “Panel data models: some recent developments”, in J.J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics, Vol. V*, Elsevier, Amsterdam.
- Banca d'Italia (various years), *I bilanci delle famiglie italiane nell'anno* [da 1989 a 2004], *Supplemento al Bollettino Statistico* (nuova serie), Banca d'Italia, Roma, various years.
- Blundell, R., and I. Preston (1998), “Consumption inequality and income uncertainty”, *The Quarterly Journal of Economics*, 113, 603-640.

- Brandolini, A. (1999), “The distribution of personal income in post-war Italy: source description, data quality, and the time pattern of income inequality”, *Giornale degli Economisti e Annali di Economia*, 58, 183-239.
- Cappellari, L. (2001), “Earnings mobility among Italian low paid workers”, ISER Working Paper 2001-13, University of Essex, Colchester.
- Cappellari, L., and S.P. Jenkins (2002), ‘Who stays poor? Who becomes poor? Evidence from the British Household Panel Survey’, *Economic Journal*, 112, C60-C67.
- Cappellari L., S.P. Jenkins (2004) ‘Modelling low income transitions’, *Journal of Applied Econometrics*, 19, 593-610.
- Chamberlain, G. (1982), “Panel data”, in Z. Griliches and M. Intriligator (Eds.), *Handbook of Econometrics, Vol. II*, North-Holland, Amsterdam.
- Chamberlain, G. (1985), “Heterogeneity, omitted variable bias, and duration dependence”, in J. Heckman and B. Singer (Eds.), *Longitudinal analysis of labor market data*, Cambridge University Press, Cambridge.
- Chay, K.Y., H. Hoynes and D. Hyslop (2001), “A non-experimental analysis of true state dependence in monthly welfare participation sequences”, Department of Economics, UC-Berkeley, Berkeley, CA (mimeo.).
- Contini, D., and N. Negri (2005), “Would declining exit rates from welfare provide evidence of welfare dependence in homogeneous environments?”, Progetto MIUR «Metodi e studi di valutazione degli effetti di politiche del lavoro, di aiuto alle imprese e di welfare», Working paper n. 68, Dipartimento di Scienze Statistiche, Università di Padova.
- Devicienti, F. (2002), “Poverty persistence in Britain: a multivariate analysis using the BHPS, 1991-1997”, *Journal of Economics*, Suppl. 9, pp. 1-34.
- Duncan, G.J., B. Gustafsson, R. Hauser, G. Schmauss, H. Messinger, R. Muffels, B. Nolan, and J.C. Ray (1993), “Poverty dynamics in eight countries”, *Journal of Population Economics*, 6, 215-234.
- Eurostat (xxxx),
- Giraldo, A., E. Rettore and U. Trivellato (2001), “Attrition bias in the Bank of Italy’s Survey of Household Income and Wealth”, in *Proceedings of the International Conference on Quality in Official Statistics*, Stockholm, May 14-15, 2001 (CD-Rom).
- Hall, R. (1978), “Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence”, *The Journal of Political Economy*, 86, 971-987.
- Heckman, J.J. (1978), “Simple statistical models for discrete panel data developed and applied to test the hypothesis of true state dependence against the hypothesis of spurious state dependence”, *Annales de l’INSEE*, 30-31, 227-269.
- Heckman, J.J. (1981a), “Statistical models for discrete panel data”, in C. Manski and D. McFadden (eds.), *Structural analysis of discrete data with econometric applications*, The MIT Press, Cambridge, MA.
- Heckman, J.J. (1981b), “The incidental parameters problem and the problem of initial conditions in estimating a discrete time – discrete data stochastic process”, in C. Manski and D. McFadden (eds.), *Structural analysis of discrete data with econometric applications*, The MIT Press, Cambridge (MA).

- Hsiao, C. (1986), *Analysis of panel data*, Cambridge University Press, Cambridge.
- Lillard, L.A., and R.J. Willis (1978), "Dynamic aspects of earning mobility", *Econometrica*, 46, 985-1012.
- Mejer, L., and G. Linden (2000), "Persistent income poverty and social exclusion in the European Union", *Statistics in Focus*, Population and Social Conditions, Theme 3, 13/2000.
- Miniaci, R., and G. Weber (1999), "The Italian recession of 1993: aggregate implications of microeconomic evidence", *Review of Economics and Statistics*, 81, 237-249.
- OECD (2001), *Employment Outlook*, Chapter 2, OECD, Paris.
- Oxley H., T.T. Dang, and P. Antolín (2000), 'Poverty dynamics in six OECD countries', *OECD Economic Studies*, No. 30, 2000/I, 7-52.
- Sacchi, x., (2007),
- Stevens A.H. (1999), 'Climbing out of poverty, falling back in. Measuring the persistence of poverty over multiple spells', *Journal of Human Resources*, 34, 557-588.
- Stewart, M.B., and J.K. Swaffield (1999), 'Low pay dynamics and transition probabilities', *Economica*, 66 (261), 23-42.
- Trivellato U. (1998), 'Il monitoraggio della povertà e della sua dinamica: questioni di misura e evidenze empiriche', *Statistica*, 58, 549-574.

Table 1. *SHIW's sample size by year of first interview and year of interview 1989-2004*

Year of first interview	Year of interview							
	1989	1991	1993	1995	1998	2000	2002	2004
1989	8,274	2,187	1,050	827	544	404	307	230
1991		6,001	2,420	1,752	1,169	832	613	464
1993			4,619	1,066	583	399	270	199
1995				4,490	373	245	177	117
1998					4,478	1,993	1,224	845
2000						4,128	1,014	667
2002							4,406	1,082
2004								4,408
	8,274	8,188	8,089	8,135	7,147	8,001	8,011	8,012

Source: Banca d'Italia (various years).

Table 2. *Test for attrition ignorability in the estimation of transition probabilities, order of transition one to seven, 1989-2004*

	<i>1989-91</i>	<i>1991-93</i>	<i>1993-95</i>	<i>1995-98</i>	<i>1998-00</i>	<i>2000-02</i>	<i>2002-04</i>
χ^2 stat.	14.76	25.37	.	41.15*	44.02**	.	20.60*
Dof	12	22	28	30	28	22	12
	<i>1989-93</i>	<i>1991-95</i>	<i>1993-98</i>	<i>1995-00</i>	<i>1998-02</i>	<i>2000-04</i>	
χ^2 stat.	9.18	13.74	31.20*	27.27	.	10.88	
Dof	10	18	22	22	18	10	
	<i>1989-95</i>	<i>1991-98</i>	<i>1993-00</i>	<i>1995-02</i>	<i>1998-04</i>		
χ^2 stat.	7.40	7.98	19.42	35.12**	10.66		
Dof	8	14	16	14	8		
	<i>1989-98</i>	<i>1991-00</i>	<i>1993-02</i>	<i>1995-04</i>			
χ^2 stat.	3.06	11.21	12.94	4.81			
Dof	6	10	10	6			
	<i>1989-00</i>	<i>1991-02</i>	<i>1993-04</i>				
χ^2 stat.	3.56	4.72	1.374				
Dof	4	6	4				
	<i>1989-02</i>	<i>1991-04</i>					
χ^2 stat.	3.23	0.17					
Dof	2	2					

** significant at level .05; * significant at level .10.

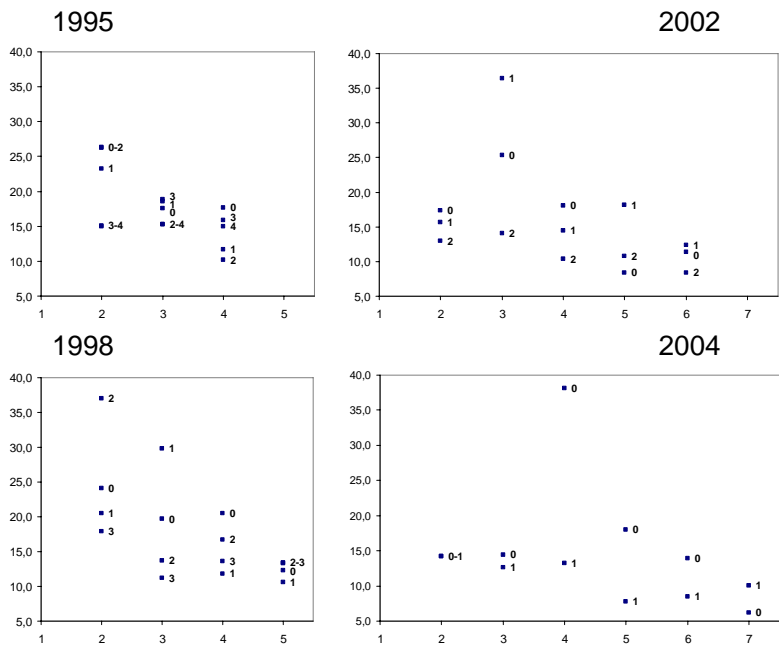
Table 3. *Testing the hypothesis of no true state dependence under models (17) and (14), respectively, and under alternative Poverty lines (number of households/years: 7,397)^a*

	Poverty line set at 60% of the median	Poverty line set at 50% of the median	Poverty line set at 70% of the median
Model (17) (left-tail t test) ^b	$t = 0.2869$	$t = 2.1633$	$t = 1.1843$
Model (14) (left-tail t test) ^b	$t = -6.2611$ ***	$t = -6.6376$ ***	$t = -6.8000$ ***

^a The test statistic has been obtained by weighting the contribution of each household by the number of its members (at the time period to which the dependent variable refers to).

^b: *** significant at level .01.

Fig. 1. Head-count ratios from SHIW in selected calendar years by number of years in the survey, 1989-2004



Note: Each panel refers to a specific calendar year. Number of interviews foregone by the household in the year to which the panel refers to is along the horizontal axis. Number of interviews the household experienced after the year to which the panel refers to are in the body of each panel attached to the point they refer to. As an example, the four points in the 1998 panel on the left-most column refer to households experiencing their second interview in 1998, *i.e.* these households entered the survey in 1995 and did not leave it yet by 1998. The lowest point refers to households who experienced three further interviews after the 1998 one, *i.e.* they were still in the survey in 2004; the second lowest point refers to households who experienced one further interview after the 1998 one, *i.e.* they left the survey after the 2000 interview.