TESTING FOR TRUE STATE DEPENDENCE IN POVERTY DYNAMICS

Anna Giraldo*, Enrico Rettore[+] and Ugo Trivellato[§]

* *Dipartimento di Scienze Statistiche, Università di Padova*

[+] *Dipartimento di Scienze Statistiche, Università di Padova and IRVAPP*

[§] *Dipartimento di Scienze Statistiche, Università di Padova. IRVAPP and IZA*

This version: November 18, 2010

Work in progress. For comments only. Please do not cite.

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, initial 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. After ascertaining the ignorability of the process that generates the massive attrition plaguing the panel, we do not find any sign of TSD.

Keywords: Poverty dynamics, Dynamic discrete-response models, Attrition ignorability

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 and for Italy, see Duncan *et al.*, 1993; Trivellato, 1998; Oxley, Dang and Antolín, 2000; Mejer and Linden, 2000; OECD, 2001).

This is commonly interpreted as proof of the existence of *true state dependence* (TSD): experiencing poverty in a specific time period *by itself* increases the probability of undergoing poverty in subsequent periods (through human capital deterioration, decreasing self-esteem, *etc*.: see Bane and Ellwood, 1994). However, this outcome might stem from the existence of household characteristics which are *both* relevant for the chance of falling into poverty *and* persistent over time. These adverse characteristics may induce repeated or prolonged poverty spells without the need to assume that poverty itself raises the probability of future poverty episodes. Many characteristics are usually controlled for in econometric analysis, but we cannot exclude that unobserved ones play a role. Hence, we label this possibility *unobserved heterogeneity* (UH).

The literature has so far failed to neatly distinguish these two potential sources of poverty persistence[1]. The aim of this paper is to provide a test to disentangle TSD from UH. Noticeable contributions (*e.g.*, Stevens, 1999; Cappellari and Jenkins, 2002; Devicienti, 2002; Biewen, 2009) use different models and estimation methods, but all tend to rely on parametric specifications and distributional assumptions. In particular, all models assume a first-order stationary Markov chain for state dependence, and combine it with *time-invariant* UH.

We depart from this approach by explicitly referring to the life-cycle permanent income hypothesis. It takes accounting for two sources of UH: (i) the household's permanent income at a specific, initial time period, and (ii) the way in which it evolves from that time period onwards as a result of permanent shocks. A crucial consequence of this double source of UH is that simple models for TSD in the presence of UH (*e.g.*, time-invariant UH models) may badly miss the point.

A second innovation relative to existing literature is our explicit recognition that "being on welfare" is by and large observationally equivalent to "being poor" according to a conventionally established poverty line (at least to the extent that lines for welfare measures are not too different from lines for statistical purposes). This observation raises the issue of whether the evidence of TSD in the dynamics of welfare participation (as in Hoynes, Chay and Hyslop, 2004) is induced by the welfare system or by poverty itself (see the discussion in Contini and Negri, 2007).

[1] For instance, Oxley, Dang and Antolín (2000, p. 6) summarise 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*.»

Our empirical test does not suffer from this ambiguity since it is applied to Italian data. In Italy up to 2004 there was no countrywide intervention providing income support to the poor[2]. On finding that TSD exists we could unambiguously conclude that it is due to the occurrence of poverty.

Our test is crucially important for policy design. If the persistence of poverty is (at least partly) due to TSD, then welfare policies providing income transfers to the poor to lift them to a minimum standard-of-living threshold would also include an "activation" component: forcing households out of poverty at time *t* would reduce their chance of experiencing poverty in the future. Instead, if the persistence of poverty is due only to UH, any monetary transfers policy would still be a relief for the poor but would not act on the determinants of their status. Different policies would be called for this latter purpose.

The paper proceeds as follows. In section 2 we present a 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 analytically derive the consequences of mistakenly testing for TSD in poverty dynamics under the textbook model presented in the previous section. In section 4 we present our Instrumental Variable (IV) test for TSD. In the empirical analysis 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 SHIW panel is plagued by massive attrition, preliminarily we develop a test for 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.

Section 7 outlines our conclusions, which can be summarised in two statements. First, while it is apparent that the SHIW panel sample is biased by attrition, with households less likely to experience poverty surviving longer in the panel, we also find evidence that attrition is basically ignorable for the specific purpose of testing for TSD. Secondly, after accounting for the two sources of heterogeneity, we do not find any sign of TSD. Some sensitivity analyses, aimed at ascertaining the robustness of results to changes in the poverty line, corroborate that conclusion.

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):

$$I_{it} = \mathbf{1}(\alpha_i + \varphi \cdot I_{it-1} + \varepsilon_{it} < c), \qquad t=2\ldots T \tag{1}$$

[2] A demonstration of *Reddito minimo d'inserimento* – a minimum income programme – was implemented in the years 1998-2002, but at a very small scale (Sacchi, 2007). Besides, income transfers are not universal; rather, they are limited to particular categories, *i.e.*, low income families with three or more children, and in general are quite modest. Local measures of income maintenance are limited. Starting from 2005, some measure of minimum income was introduced by a few Regions. This is why we do not exploit data from SHIW after 2004.

where:

- 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 otherwise[3];
- **1**(.) is an indicator function equal to 1 if the condition in its argument is satisfied and 0 otherwise;
- 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 sequences $\{I_{it}\}$ feature 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, conditioning on 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 the right answer, since in the presence of UH (*i.e.*, $var\{\alpha_i\}>0$) we are bound to observe $\Pr(I_{it}=1|\, I_{t-1}=1) > \Pr(I_{it}=1|\, I_{it-1}=0)$ 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 review).

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, $\Pr(I_{i1},...,I_{iT}\,|SS_i;\, \varphi,\, \alpha_i)$ – 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-effects model and imposing a distributional assumption), we can obtain a likelihood function for φ by integrating out the unobserved α_i. There is an additional problem here with the initial condition I_{i1}, because the analyst very often does not know whether I_{i1} has been generated by the same model as the subsequent observations (Heckman, 1981b).

In the analysis of poverty/non-poverty or welfare participation sequences, the first strategy is adopted, *e.g.*, by Hoynes, Chay and Hyslop (2004), who model welfare dynamics in California. A good example of the second strategy is given by Cappellari and

[3] 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.

Jenkins (2002): they model poverty dynamics in Great Britain by adding parameter specifications and distributional assumptions on initial conditions and the panel attrition process.

In the next section we argue that model (1) is *not* an appropriate framework to test for TSD in poverty dynamics. The point is that any test for TSD is designed to detect an *abnormal* persistence of poverty over time. Hence, to begin with the analyst must carefully specify what a *normal* persistence should be, *i.e.*, the kind of persistence one expects to observe under the no TSD hypothesis. Model (1) takes as normal the persistence induced by the following autocovariance of the unobservables:

$$cov\ \{\alpha_i + \varepsilon_{it}, \alpha_i + \varepsilon_{is}\} = \sigma_\alpha^2, \qquad s \neq t.$$

Any observed departure from the pattern of poverty persistence induced by that autocovariance is taken as evidence of TSD. Our argument is that when the poverty indicator I_{it} is based on *disposable income* the kind of persistence implied by model (1) is inconsistent with known facts about the pattern of income and consumption over the life-cycle. As a consequence, moving from model (1) the null hypothesis is bound to be rejected even in the absence of TSD.

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

In this section, we move from the empirical literature on the 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:

$$y_{it} = y_{it-1}^P + S_{it}, \qquad t=2\ldots T \tag{2}$$

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, and S_{it} represents unexpected (as seen from time t–1) departures of time t income from y_{it-1}^P. Being a prediction error, S_{it} is orthogonal to the predicted value y_{it-1}^P.

Moreover, let us represent S_{it} as:

$$S_{it} = u_{it} + v_{it}, \tag{3}$$

where u_{it} is the permanent component of the shock. It summarises the impact of all new information that becomes available in period t relevant for 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:

$$y_{it}^P = y_{it-1}^P + u_{it}, \tag{4}$$

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

$$\Delta y_{it} = u_{it} + v_{it} - v_{it-1} \,. \tag{5}$$

Compared to model (1), in model (2)-(3) there are *two sources* of across-household heterogeneity. Households are heterogeneous with respect to their expected income as evaluated at time t=1 as well as 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 $var\{u_{it}\}= 0, \forall\, t,$ in which case $y_{it}^P = y_{it-1}^P = y_i^P$ plays as the time-invariant unobserved characteristic α_i in (1). Let $\sigma_I^2 = \mathrm{var}\{y_{i1}^P\}$ and $\sigma_{ut}^2 = \mathrm{var}\{u_{it}\}$.

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

$$y_{it}= y_{i1}^P + u_{i1} + \ldots + u_{it}+ \varphi\, I_{it\text{-}1} + v_{it}, \qquad t=2\ldots T. \tag{6}$$

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

$$y_{it} \perp v_{is}, \qquad \forall\, s\neq t, \tag{7}$$

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

With reference to this specification, four comments are in order. First, the theory of consumer behaviour under the permanent income hypothesis offers testable 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), which implies that $var\{\Delta C_{it}\}=cov\{\Delta C_{it}\Delta y_{it}\}$ and that ΔC_{it} is uncorrelated to the past history of consumption, *i.e.*, it behaves as the permanent shocks in the permanent income sequences. These implications have been the focus of several empirical papers, and the resulting evidence supports the permanent income hypothesis (see, *e.g.*, Blundell and Preston, 1998). This suggests that model (1) assuming one source of across-household heterogeneity is inconsistent with the evidence. When poverty is defined with reference to household disposable income a model with two sources of heterogeneity should be preferred to the textbook one.

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 sequences $\{I_{i1, \ldots, } I_{iT}\}$, 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; Heckman, 1978 and 1981a, offers pioneering contributions to model the pattern of such sequences and identify whether TSD is at work.). 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. Of course, there is a price to pay for this approach, which consists of imposing an arbitrary cut-off – the poverty line. Results might be sensitive to it. To mitigate this arbitrariness, we will use three alternative 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 triple of observations (I_{i1}, I_{i2}, I_{i3}) under the no TSD case and in the presence of permanent shocks. The time-invariant individual-specific component of the triple is y_{i1}^P, the permanent income at time 1. Conditional on that fixed-effect, I_{i2} and I_{i3} are *not* independent, since they are both affected by the permanent shock u_{i2}:

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

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

4.1. The test under the mis-specified textbook model

We present now the econometrics relevant to our test. We start by showing how our instrumental variables (IV) test works in the standard case in which there is *no permanent shock* ($var\{u_{it}\}=0$), so that we can drop the t subscript from the permanent income and write y_i^P, since permanent income is not allowed to vary over time. An equivalent representation of the fixed-effect model (1) is[4]:

$$I_{it}=F(y_i^P) + \beta_i\, I_{it-1}+\epsilon_{it} \tag{8}$$

where:

$F(y_i^P)=Pr(y_i^P + v_{it} < c)$, with $F(\cdot)$ the distribution function of the transitory shock

[4] Model can be written as $I_t = y_i^P + \varphi I_{t-1} + v_{it}$. Thus, the probability for household i-th at time t to be poor, *not* having been poor at time t-1, is equal to the probability that $v_{it} < (c - y_i^P)$. We denote that probability $F(y_i^P)$.

and

$\beta_i = F(y_i^P + \varphi) - F(y_i^P)$

is the causal effect of experiencing a poverty spell at time t-1 on the probability of experiencing it at time t for a subject with permanent income y_i^P.

By applying standard panel data econometrics, to get rid of the individual specific fixed-effect $F(y_i^P)$ it suffices to first-difference eq. (8):

$$\Delta I_{it} = \beta_i \, \Delta I_{it\text{-}1} + \Delta\epsilon_{it} \tag{9}$$

Then we can estimate it by using $\Delta I_{it\text{-}2}$ as an IV for $\Delta I_{it\text{-}1}$. This way we get an estimate of the average causal effect $E\{\beta_i | \Delta I_{it\text{-}2} \, \Delta I_{it\text{-}1} = -1\}$. This is because, disregarding asymptotically vanishing terms, the IV estimator:

$$\sum_i \Delta I_{it\text{-}2} \, \Delta I_{it} / \sum_i \Delta I_{it\text{-}2} \, \Delta I_{it\text{-}1} \tag{10}$$

is equal to the weighted average $\sum_i w_i \beta_i$ with:

$$w_i = \Delta I_{it\text{-}2} \, \Delta I_{it\text{-}1} / \sum_i \Delta I_{it\text{-}2} \, \Delta I_{it\text{-}1} \, . \tag{11}$$

It is straightforward to see that *only* individuals experiencing either the sequence (0, 1, 0, I_{i4}) or the sequence (1, 0, 1, I_{i4}) contribute to the estimator, otherwise the weight w_{it} is zero. As usual (see Chamberlain, 1982), *at least four observations* on each individual are needed. It is clear from the definition of β_i that the IV estimand $E\{\beta_i | \Delta I_{it\text{-}2} \, \Delta I_{it\text{-}1} = -1\}$ is equal to zero if and only if φ=0. As a consequence, we can directly use the IV estimator to test for TSD.

4.2. The test under the model with flexible UH

We deal with the *permanent shock* u_{it} by considering a small-sigma approximation to the probability to fall in poverty at time t as a function of past history and current permanent shock:

$$\Delta I_{it} \approx -\varphi \, f_{it\text{-}1} \, \Delta I_{it\text{-}1} - \theta \, f_{it\text{-}1} \, u_{it} + \Delta\epsilon_{it} \, . \tag{12}$$

It holds when φ is small and the variance of the permanent shock is small with respect to the variance of the transitory one ($f_{it\text{-}1}$ is the probability density of the transitory shock at ($c - y_{it-1}^P$)). Provided that the permanent shock is zero mean, $\Delta I_{it\text{-}2}$ is still a valid instrument for $\Delta I_{it\text{-}1}$ to estimate the average causal effect $E\{\beta_i | \Delta I_{it\text{-}2} \, \Delta I_{it\text{-}1} = -1\}$, since it is uncorrelated to the unobservables in eq. (12) ($f_{it\text{-}1} \, u_{it} + \Delta\epsilon_{it}$). As a consequence, the IV estimator (10) is consistent for the estimand $E\{\beta_i | \Delta I_{it\text{-}2} \, \Delta I_{it\text{-}1} = -1\}$ even in the presence of permanent shocks. Hence, it can be used to test for TSD.

In the Appendix we provide the details of the small-sigma approximation. We also prove that at the first order of approximation ($I_{it\text{-}2} - I_{it\text{-}3}$) is a valid instrument in the first-differenced model even when the *variance of* v_{it} *changes over time*. So our test is robust to the presence of longitudinally heteroskedastic transitory shocks.

4.3. The Chamberlain conditional likelihood test is not robust to the presence of permanent shocks

Before moving on to the empirical analysis we exploit our small-sigma approximation to show that the Chamberlain conditional likelihood strategy to the identification of the TSD parameter does not work in the presence of permanent shocks.

Consider the case T=4. It is known that in the absence of permanent shocks and if the transitory shock is Logistic distributed, conditioning on the sufficient statistic $(I_{i1}, I_{i2}+I_{i3}, I_{i4})$ provides a likelihood function for the TSD parameter free of the time-invariant fixed-effects (see Chamberlain, 1982). The procedure is computationally simple, and amounts to estimate the logistic regression of $I_{i3}-I_{i2}$ on $I_{i4}-I_{i1}$ exploiting the sub-sample of units satisfying the condition ($I_{i3}+I_{i2}=1$, $I_{i4}+I_{i1}=1$).

Moving from eq. (12) one immediately gets that *under the null hypothesis* of no TSD our small-sigma approximation yields:

$$I_{i3}-I_{i2}= -\theta f_{i2} u_{i3}+\epsilon_{i3}-\epsilon_{i2}$$

$$I_{i4}-I_{i1}= -\theta f_{i1} (u_{i2}+u_{i3}+u_{i4})+\epsilon_{i4}-\epsilon_{i1}$$

with $f(.)$ the Logistic density function.

In the absence of permanent shocks, *i.e.*, θ=0, the regression of $I_{i3}-I_{i2}$ on $I_{i4}-I_{i1}$ features a zero coefficient because the transitory shocks ϵ_{it} are uncorrelated. Instead, when $\theta\neq 0$ that coefficient is no longer zero, leading to reject the null even when it is true.

5. Is attrition in the SHIW panel ignorable to 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-annual 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 this happens partly by design – the split-panel, part of it is due to lack of survey management over the following rules[5].

Table 1 about here

---

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

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 ratios 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 needs 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 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-…-2004.

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 TSD. 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 ratios, 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 (2000) and keep to the following operational criteria in order to identify poor households:

a) we make use of the OECD modified equivalence scale[6];
b) we set the poverty line in the calendar year 1995 at 60% of the median equivalent income in that year;
c) we derive the poverty line for the other years deflating/inflating the 1995 poverty line by means of the consumer price index[7].

As a robustness check, we replicate our analysis with two alternative poverty lines set at 80% and 120%, respectively, of the poverty line as defined above.

The main results are summarised in Table 3. The first stage, namely the OLS regression of the first-differenced poverty status lagged one on the first-differenced poverty status lagged two, is well determined with all the poverty lines we consider. Our IV test robust to the presence of permanent shocks (as well as to longitudinally heteroskedastic transitory shocks: see Appendix) does not provide any evidence pointing to the existence of TSD.

Table 3 about here

With reference to eq. (15), using the base poverty line we get a largely insignificant estimate of φ, the state dependence parameter, its .95 confidence interval being (-.07599, .06818). Using the alternative poverty lines results do not change (the .95 confidence intervals are (-.09706, .07374) and (-.06411, .07876) in the 80% and in the 120% case, respectively). Overall we conclude that the dynamics of poverty we observe in Italy over the period 1995-2004[8] does not feature any TSD.

7. Concluding remarks

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

---

[6] It weights 1 the first adult, 0.5 each additional household member at least 14 years old, 0.3 each member younger than 14.

[7] 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 ratios) slightly lower with respect to the one implied by a strictly relative threshold, *i.e.*, by computing a poverty line specific to every year.

[8] 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.

As for the SHIW panel we use, it is plagued by a severe attrition. There is a clear-cut evidence that the 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 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.

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 also 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 active anti-poverty policies. The point is that such policies should be targeted to contrast the adverse characteristics (lack of education and skills, poor health, lack of social networks, say) – affecting individuals and households from the outset or induced by permanent shocks – relevant for the household's risk to persist into poverty.

Appendix

A (i). The derivation of the small-sigma approximation

Introducing the permanent shocks and exploiting the notation introduced in the main text, disposable income at time t is linked to permanent income at time t–1, to permanent and to 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:

$$y_{it} = y^{P}_{it-1} + \theta\, u_{it} + \sigma\, v_{it} + \varphi\, I_{it\text{-}1}\,, \qquad \text{(A.1)}$$

where we redefine the permanent and transitory shock u_{it} and v_{it} as unit variance and the standard deviations of the shocks hitting income are equal to θ and σ, respectively. Consequently, the binary variable I_{it} is:

$$I_{it} = \mathbf{1}(y_{it} < c). \qquad \text{(A.2)}$$

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

$$\Pr(y_{it} < c) = F[(c - y^{P}_{it-1} - \theta\, u_{it} - \varphi\, I_{it-1})/\,\sigma\,] \qquad \text{(A.3)}$$

(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 (A.3) approximating it in a neighbourhood of $\theta/\sigma=0$ and $\varphi/\sigma=0$[9]:

$$\Pr(y_{it} < c) = \Pr(y_{it-1} < c) - \varphi\, f_{it-1}\,(I_{it-1} - I_{it-2})/\sigma - \theta\, f_{it-1}\, u_{it}\,/\sigma\,, \qquad \text{(A.4)}$$

in which f_{it-1} is the probability density associated to F as evaluated at $(c - y^{P}_{it-1})/\sigma$.

Eq. (A.4) displays the basic features of the dynamics of poverty:

- The probability to fall in poverty at time t is obtained from the corresponding probability at time t–1 by adding the permanent shock u_{it} scaled by the individual specific parameter $\theta\, f_{it-1}/\sigma$.
- In the presence of TSD (φ <0), subjects entering poverty at time t–1 (I_{it-1}=1 and I_{it-2}=0) have a higher chance to experience a spell of poverty again at time t. 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.
- 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 both permanent shocks and TSD 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 φ=0 against the alternative φ <0 let ϵ_{it} be the deviation of I_{it} from its mean $\Pr(y_{it} < c)$ and rewrite (A.4) in first-differences as:

$$I_{it}\text{-}I_{it-1} = -\,\varphi\, f_{it-1}\,(I_{it-1} - I_{it-2})/\sigma - \theta\, f_{it-1}\, u_{it}\,/\sigma + \epsilon_{it} - \epsilon_{it-1}. \qquad \text{(A.5)}$$

Since $(I_{it-2} - I_{it-3})$ is orthogonal to the unobservables $\theta_t\, f_{it-1}\, u_{it}\,/\sigma$:

$$E\{f_{it-1}\, u_{it} \mid \Delta I_{it-2}\} = E\{\, f_{it-1} \mid \Delta I_{it-2}\}\, E\{u_{it} \mid \Delta I_{it-2}\} = E\{\, f_{it-1} \mid \Delta I_{it-2}\}\, E\{u_{it}\} = 0^{10}$$

it can be used as an instrument for $(I_{it-1} - I_{it-2})$.

A (ii). The robustness of the IV test for TSD to longitudinally heteroskedastic shocks

Consider the first differenced model as in (A.5) but now both θ and σ are allowed to vary over time:

$$I_{it} - I_{it-1} = -\,\varphi\, f_{it-1}\,(I_{it-1} - I_{it-2})/\sigma_{t-1} - \theta_t\, f_{it-1}\, u_{it}/\sigma_{t-1} + (\sigma_{t-1}/\sigma_t - 1)\, f_{it-1}\,(c - y^{P}_{it-1})/\sigma_{t-1} + \epsilon_{it} - \epsilon_{it-1} \qquad \text{(A.6)}$$

[9] These conditions have an immediate substantive interpretation. The Taylor expansion is useful in the case in which (i) the variance of the permanent shock is small with respect to the variance of the transitory shock, (ii) the state dependence parameter φ is small with respect to the standard deviation of the transitory shock.

[10] Full independence between y^{P}_{it-1} and u_{it} is required for the first equality to hold. $E\{u_{it}\}$=0 is required for the last equality to hold.

Write it with reference to time $t-2$:

$$I_{it\text{-}2} - I_{it\text{-}3} = -\varphi\, f_{it\text{-}3}\,(I_{it\text{-}3} - I_{it\text{-}4})/\sigma_{t\text{-}3} - \theta_{t\text{-}2}\, f_{it\text{-}3}\, u_{it\text{-}2}/\sigma_{t\text{-}3} + (\sigma_{t\text{-}3}/\sigma_{t\text{-}2} - 1) f_{it\text{-}3}\,(c - y^P_{it-3})/\sigma_{t\text{-}3} + \epsilon_{it\text{-}2} - \epsilon_{it\text{-}3}.$$

The covariance between $(I_{it\text{-}2} - I_{it\text{-}3})$ and the unobservables in (A.6) does not vanish only because of the term:

$$cov\{I_{it\text{-}2} - I_{it\text{-}3}, (\sigma_{t\text{-}3}/\sigma_{t\text{-}2} - 1) f_{it\text{-}3}\,(c - y^P_{it-3})/\sigma_{t\text{-}3}\},$$

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

$$(-\varphi/\sigma_{t\text{-}3})\,(\sigma_{t\text{-}3}/\sigma_{t\text{-}2} - 1)\,\mathrm{cov}(f_{it\text{-}3}\,(I_{it\text{-}3} - I_{it\text{-}4}), f_{it\text{-}3}\,(c - y^P_{it-3})/\sigma_{t\text{-}3}) +$$

$$(-\theta_t/\sigma_{t\text{-}1})\,(\sigma_{t\text{-}3}/\sigma_{t\text{-}2} - 1)\,\mathrm{cov}(f_{it\text{-}1}\, u_{it}, f_{it\text{-}3}\,(c - y^P_{it-3})/\sigma_{t\text{-}3}) +$$

$$(\sigma_{t\text{-}1}/\sigma_t - 1)\,(\sigma_{t\text{-}3}/\sigma_{t\text{-}2} - 1)\,\mathrm{cov}(f_{it\text{-}1}\,(c - y^P_{it-1})/\sigma_{t\text{-}1}, f_{it\text{-}3}\,(c - y^P_{it-3})/\sigma_{t\text{-}3}).$$

Apparently, in a neighbourhood of the point we choose for our Taylor series expansion the coefficients in each of the three terms are negligible, allowing us to conclude that $(I_{it\text{-}2} - I_{it\text{-}3})$ is a valid instrument in (A.6).

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* [from 1989 to 2004], Supplemento al Bollettino Statistico (nuova serie), Banca d'Italia, Roma.

Bane M.J., Ellwood D.T. (1994) *Welfare realities. From rhetoric to reform*, Harvard University Press, Harvard.

Biewen M. (2009), "Measuring state dependence in individual poverty histories whem there is feedback to employment status and household composition, *Journal of Econometrics*, 24, 1095-1116.

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

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.J. Heckman and B. Singer (Eds.), *Longitudinal analysis of labor market data*, Cambridge University Press, Cambridge.

Hoynes H.W., K.Y. Chay and and D. Hyslop (2004), "True state dependence in monthly welfare participation: a nonexperimental analysis", Working Papers 05-33, Department of Economics, University of California Davis, Davis, CA.

Contini, D., and N. Negri (2007), "Would declining exit rates from welfare provide evidence of welfare dependence in homogeneous environments?", *European Sociological Review*, 23, 21-33.

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 (2000), "Social exclusion in the EU Member States", *Statistics in Focus*, Population and Social Conditions, Theme 3, 11/2000.

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.

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, S. (2007), “L’esperienza del Reddito minimo di inserimento”, in A. Brandolini e C. Saraceno (eds.), *Povertà e benessere. Una geografia delle disuguaglianze in Italia*, il Mulino, Bologna, 423-448.

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*

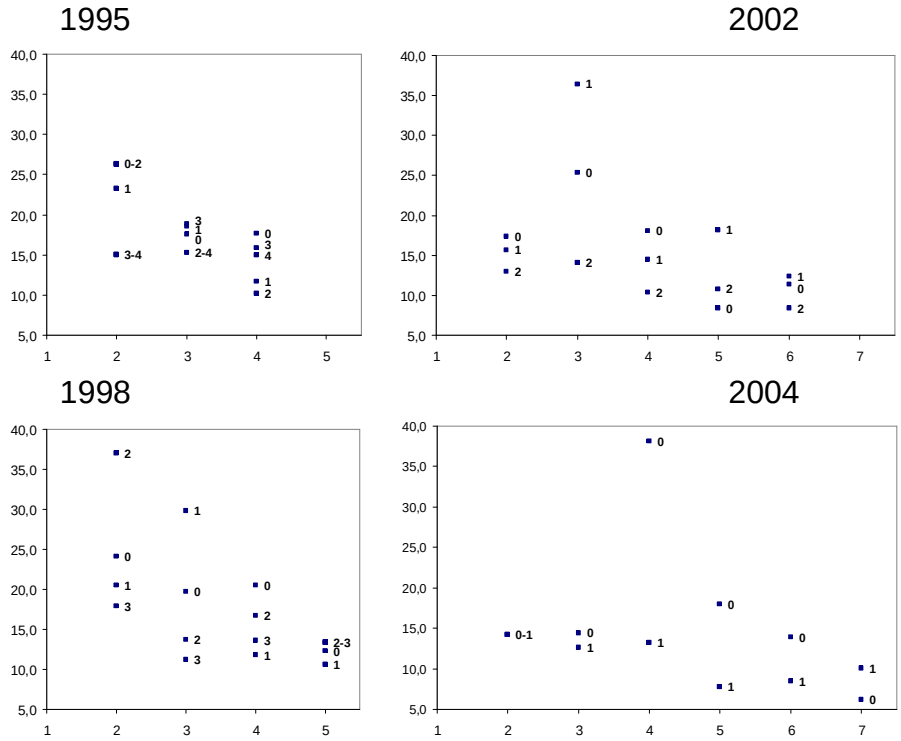
	1989-91	*1991-93*	*1993-95*	*1995-98*	*1998-00*	*2000-02*	*2002-04*
χ^2-stat.	14.76	25.37	...	41.15*	44.02**	...	20.60*
D.of f.	12	22	28	30	28	22	12
	1989-93	*1991-95*	*1993-98*	*1995-00*	*1998-02*	*2000-04*	
χ^2-stat.	9.18	13.74	31.20*	27.27	...	10.88	
D. of f.	10	18	22	22	18	10	
	1989-95	*1991-98*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

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

Table 3. *Testing the hypothesis of no True State Dependence under alternative poverty lines (number of households/years: 7,397). IV estimates of the first-differenced model (standard errors in parentheses).* .

	Poverty line set at 60% of the median (*PL*)	Poverty line set at 80% of *PL*	Poverty line set at 120% of *PL*
First stage	- .4190 (.01996)	- .4108 (.02408)	- .4121 (.01912)
IV estimate	- .003906 (.03678)	-.01166 (.04357)	.007326 (.03645)

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



Note: Each graph refers to a specific calendar year. Number of interviews foregone by the household in the calendar year to which the graph refers to is along the horizontal axis. Number of interviews the household experienced after the year to which the graph refers to are in the body of each graph, attached to the point they refer to. As an example, the four points in the 1998 graph 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; and so on.