Elena Giarda

Persistency of financial distress amongst Italian households: evidence from dynamic probit models

Quaderni di Dipartimento

Serie Ricerche 2010, n. 3
ISSN 1973-9346

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Dipartimento di Scienze Statistiche "Paolo Fortunati"

Persistency of financial distress amongst Italian households: Evidence from dynamic probit models

Elena Giarda*

This version: 5 December 2011

Abstract

This paper aims to build on existing studies on households' financial distress and provides new evidence on the determinants of financial hardship in Italy and its persistence over time. It suggests a quantitative definition of financial distress based on the distribution of net wealth and tests whether the probability of experiencing financial difficulties is persistent over time by means of dynamic probit models. The analysis exploits the longitudinal component of the Bank of Italy Survey on Household Income and Wealth for the period 1998-2006. Results provide evidence that, after accounting for unobserved heterogeneity, past values of the outcome variable play a large role in explaining the probability of experiencing financial distress. Additionally, the probability of financial vulnerability decreases with income and increases in areas where unemployment rates are higher.

JEL classification: D14; C23; C35; G11

Keywords: Household financial distress; Net wealth; Dynamic probit models; SHIW

* Address for correspondence: Elena Giarda, Prometeia Associazione per le Previsioni Econometriche, Via G. Marconi 43, 40122 Bologna, Italy. E-mail: elena.giarda@prometeia.com. The views expressed are mine and do not necessarily reflect those of Prometeia.

1. Introduction and literature review

The aim of this paper is to analyse Italian households' financial conditions and to test the hypothesis of whether the probability they may incur financial distress is persistent over time. Testing whether the probability of experiencing financial difficulties is persistent over time is important, as the presence of a number of households "trapped" in financial distress generates socially vulnerable groups of people, posing problems of social and economic cohesion and consequently calling for appropriate policies. Our study draws from the international literature on indebtedness problems and perceived financial hardship and proposes a quantitative measure of potential financial weakness of households to adverse shocks in the economy. It applies dynamic binary-choice models to the longitudinal component of the Bank of Italy Survey of Household Income and Wealth for the period 1998-2006, taking account of unobserved heterogeneity and true state dependence. By so doing, the estimation methodology advances the existing empirical literature that analyses the financial conditions of Italian households.

Households financial conditions have been raising concerns of scholars and policy makers since household borrowing started increasing considerably in the nineties, both in absolute terms and relative to household income in most OECD countries. Between 1995 and 1999 the debt to disposable income ratio averaged 108.7% in the UK, 96.7 in the US, 69.5 in France and 41.4 in Italy; between 2000 and 2007 it averaged 149.9% in the UK, 120.2 in the US, 85.2 in France and 64.9 in Italy (OECD, 2010). Annual growth rates between 1995 and 2007 were 4.7% in the UK, 3.3 in the US, 3.6 in France and 8.0 in Italy. In 2009 differences among countries still exist, with a ratio of 82.0% in Italy against 170.6, 127.5 and 106.6% in the UK, US and France respectively. The gap between Italy and the other countries is partly explainable by the lower number of indebted households, which, over the period 2002-2008, were around 25.6% of all households, against values of 85.8, 54.0 and 44.9% in the US, UK and Germany.

Despite the fact that Italian households seem to be better off than households in other countries (lower levels of debt and higher levels of wealth), analysing financial conditions of Italian households is relevant. They express much higher levels of perceived financial distress (Boeri and Brandolini, 2005; Brandolini *et al*., 2009) and, recently, concerns about financial conditions of Italian households have been expressed by the Ministry of Labour and Social Policies and the Italian Banking Association. At the end of 2009 they started monitoring the financial status of Italian households through indicators of indebtedness and financial vulnerability, with the aim to identify the risks potentially stemming from rising debt and financial exposure in the aftermath of the financial crisis of 2008-2009.

Indicators of financial distress at the macro level have their counterpart in micro data which allow to identify areas of malaise among households and its possible persistence over time. Use of microeconomic indicators was recommended in 2009 by the Stiglitz-Sen-Fitoussi Commission, which started the debate on the necessity to integrate information deriving from the GDP with other indicators, many of them extra-income, in order to obtain more complete a picture of households living conditions. Specifically recommendations 3 and 4 suggest the joint monitoring of financial variables - stock of debt and assets - as additional indicators of possible financial pressure. Moreover the literature stresses the role of wealth in affecting current well-being and the importance of monitoring wealth for social policy such as in the definition of eligibility for means-tested public benefits (e.g. Brandolini *et al.*, 2010). Another motivation for our study is related to remarks in the household debt-related literature about whether outstanding debt, particularly in countries such as the US and the UK, is considered excessive, where "excessive" means carrying the risk of default or financial hardship in the event that the household is exposed to unexpected adverse shocks. The literature highlights that the riskiness associated to debt holdings increases with the income gearing ratio, but is softened by the coexistence of real or financial assets in the household portfolio, implying that quantitative indicators of financial distress derive from a combination of both factors, assets and liabilities.

However, the international literature on financial problems at the household level mainly assesses indebtedness problems and related perceived hardship, by using both subjective and qualitative indicators of financial malaise. As regards subjective indicators, Boheim and Taylor (2000) and May and Tudela (2005) find that, amongst British households, the probability of experiencing financial difficulties, as expressed by problems paying for housing, is persistent over time. A static analysis on the probability of reporting high debt burden is performed by del Rio and Young (2008) again for the UK and find that unsecured debt raises the vulnerability of households to adverse shocks. Pudney (2008) models the dynamics of individuals' subjective assessments of their financial wellbeing, in a short panel of UK households. Brandolini *et al*. (2009) find wide differentials in perceived housing costs and consumer credit burdens in Europe, with Mediterranean countries such as Italy and Spain expressing higher degree of malaise. Discontent of Italian households is also analysed in Boeri and Brandolini (2005) and possible explanations are related to the poor performance of disposable income in the previous decade, the unprecedented rapid increase in Italian households' debt exposure and the high proportion mortgages with variable interest rates. Other empirical applications on arrears and burden of household debt can be found in Whitley *et al*. (2004), Jappelli *et al*. (2008) and Georgarakos *et al*. (2010). The literature using quantitative indicators is also concerned chiefly with a concept of financial distress related to indebted households. However we wish to take account of the whole household portfolio, including assets (financial and real) an liabilities. Cox *et al.* (2002) and May *et al.* (2004) suggest a series of indicators to identify households suffering financial distress, which include flow and stock financial variables, whereas Barwell *et al.* (2006) address the issue more explicitly by suggesting analysis of the net worth distribution, emphasizing that to examine households' financial conditions in more depth requires account to be taken of the levels and composition of assets more than analysis of liabilities alone. Brown and Taylor (2008) define households in distress as those with negative wealth holdings and analyse financial conditions of households in Great Britain, Germany and the US, whereas Christelis *et al*. (2009) use the net worth/income ratio as a measure of financial distress.

We take Brown and Taylor (2008) as our starting point and build our framework on the idea that low levels of net wealth can be taken as indicators of financial vulnerability. By looking at the distribution of household net wealth from the SHIW, we chose the level of wealth corresponding to the second decile of its distribution to split households between those who experience financial hardship and those who do not. The value of the resulting threshold amounts roughly to half median income of the sample, one of the measures proposed in the literature for identifying households in poverty; the resulting percentage of households in distress (15.3% of total households) is in line with ISTAT (2010) and Brandolini *et al.* (2010). The choice to include households with positive net wealth can also be justified by the fact that, on average, only 3% of Italian households have negative net wealth, a largely lower percentage than is observed in countries such as the US, the UK and France (Sierminska *et al*., 2008).

A dynamic probit model is then estimated on the longitudinal component of the SHIW to test for the presence of true state dependence in experiencing financial distress, by means of the Heckman method (1981a; 1981b). Comparisons with other estimation methods are then made, namely the pooled probit, the random effects probit with exogenous initial conditions and the Orme model (2001). The estimation of the Heckman model shows that, after controlling for unobserved heterogeneity, there is evidence of true state dependence, that is, the probability of experiencing financial distress at time t positively depends upon the probability of having experienced financial fragility at time (t – 1). The null hypothesis of non-significance of unobserved heterogeneity is rejected, with the fraction of total explained variance due to unobserved household-level characteristics being around one third of the total variance. The size of this parameter shows the importance of individual components in the analysis of household financial problems and the adequateness of using panel data. Together with income, regional unemployment rates play a prominent role in determining a situation of financial vulnerability: lower income households display a higher probability of financial distress, and the sign of the unemployment rate denotes a higher probability of financial distress among households in areas of high unemployment.

The remainder of the paper is organised as follows. Section 2 reviews the econometric approaches dealing with the estimation of dynamic, nonlinear, panel data models. Section 3 describes the data and discusses the choice of the dependent variable. The estimation results and robustness analysis are discussed in Section 4, while Section 5 concludes with a summary of the main findings.

2. Dynamic probit models with panel data

The main issue in estimating dynamic panel data models consists of distinguishing between *true state dependence* – the impact of the lagged dependent variable on the dependent variable, and *spurious state dependence* - caused by the presence of time-invariant unobserved heterogeneity. This requires resolution of the so-called *initial conditions problem,* which arises from the fact that the observed start of the stochastic process – period t_0, the first available observation – does not coincide with the true start of the process. It follows that the dependent variable at period t_0 cannot generally be considered to be an exogenous variable that gives rise to the process. The dynamic probit model can be written as:

$$y_{it} = \mathbf{1}[y_{it}^* = x_{it}'\beta + \gamma y_{it-1} + \varepsilon_{it} > 0] \qquad i = 1,..., N \quad t = 1,...,T \qquad (1)$$

where y_{it} is a binary variable taking the value of one if the household is experiencing financial distress at the time of the interview and zero otherwise, **1**(.) is the indicator variable, y_{it}^* the latent variable (the unobservable propensity to be in financial distress), x_{it} the explanatory variables, y_{it-1} the previous state of the endogenous variable and ε_{it} is the error term. β is the vector of parameters associated with x_{it} and γ is the coefficient expressing state dependence. In Eq. (1) the latent propensity to experience financial difficulties y_{it}^* depends upon the observed distress status of the previous period y_{it-1}: dynamic feedback on the current state of the latent variable involves y_{it-1}. The inclusion of y_{it-1} allows to test for the presence of state dependence, via estimation of the coefficient γ. Households are observed to be in financial distress when their unobserved propensity to be in distress is grater than zero. However controlling for appropriate unobserved individual characteristics is necessary to obtain unbiased estimates of the state dependence parameter, thus distinguishing *true* from *spurious* state dependence.[1]

2.1. Modelling unobserved heterogeneity

Assuming the unobservable individual-specific heterogeneity is time-invariant, the error term can be decomposed as follows:

$$\varepsilon_{it} = \alpha_i + u_{it}$$

where α_i is unobservable individual heterogeneity and $u_{it} \approx N(0,1)$ is a random error. As in any panel data model, assumptions are required about α_i. In a fixed effects specification individual effects α_i are allowed to be correlated with the explanatory variables. This setting does not require specification of a functional distribution of α_i, as they are treated as parameters to be estimated together with the vector θ. However, this approach suffers from the so-called "incidental parameter problem" which, with a fixed T, causes inconsistency in the estimators of θ (Wooldridge, 2005). Honoré (1993) and Honoré and Kyriazidou (2000) suggest semi-parametric models that do not require specification of the distribution of individual effects. However, this requires strongly exogenous explanatory variables to resolve the identification problem and consequently a random effects specification of the model is generally assumed (Wooldridge, 2005). The standard random effects specification implies $\alpha_i \approx iidN(0,\sigma_\alpha^2)$ and zero correlation between individual effects and the exogenous variables, that is $corr(\alpha_i, x_{it}) = 0$. Additionally, according to the mainstream literature, we assume zero serial correlation in the idiosyncratic term u_{it}, and equicorrelation of the composite error term ε_{it}, that is $\rho = corr(\varepsilon_{it}, \varepsilon_{is}) = \dfrac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_u^2}$ for t, s = 1,..., T; $t \neq s$.

2.2. The initial conditions problem

Inclusion of the previous state to allow for state dependence requires specific assumptions about the generation of the initial observations y_{i0}. The estimators proposed in the literature for estimating the lagged-variable coefficient γ differ in terms of how the initial conditions problem is dealt with.

The simplest case treats the initial observations as exogenous, that is the distribution of y_{i0} does not depend on α_i. The likelihood function thus consists of two independent terms,

[1] A review of the empirical literature on non-linear dynamic models can be found in Stewart (2007) and Arulampalam and Stewart (2009).

one relative to the initial period, the other to subsequent periods. It follows that the joint probability at t = 1,...,T is maximised independent of the probability at time t = 0. For more realistic cases of endogenous initial conditions, methods have been proposed to integrate out unobserved heterogeneity from the likelihood function.

Heckman (1981a; 1981b) was the first to take explicit account of the initial conditions problem, assuming endogenous variables with a probability distribution conditional on the exogenous variables and unobserved heterogeneity. Heckman's is a simultaneous two stage approach, in which the first stage, approximating the initial conditions with a reduced form equation in which the explanatory variables are a set of instrumental variables, is simultaneously estimated with the structural model. Recall Eq. (1), the dynamic random effects probit specification:

$$y_{it} = \mathbf{1}[y_{it}^* = x_{it}'\beta + \gamma y_{it-1} + \varepsilon_{it} > 0] \qquad i = 1,..., N \quad t = 1,...,T \qquad (2)$$

with $\varepsilon_{it} = \alpha_i + u_{it}$. In Heckman's terminology this is the "structural model".

Let the first period equation at $t = 0$ (the "reduced form equation") be, for i = 1,..., N:

$$y_{i0} = \mathbf{1}[y_{i0}^* = z_i'\pi + \varepsilon_{i0} = z_i'\pi + \vartheta\alpha_i + u_{i0} > 0] \qquad (3)$$

where z_i is a vector of exogenous variables, such as x_{i0}, and additional variables that can be regarded as instruments (Akay, 2009; Arulampalam and Stewart, 2009); ε_{i0} is correlated with α_i, but uncorrelated with u_{i0}; u_{i0} is independent of α_i and the distributions are respectively N(0,1) and N(0, σ_α^2). Individual effects are thus defined as $\eta_i = \vartheta\,\alpha_i + u_{i0}$. The model can be estimated by noting that the distribution of y_{it}^* conditional on α_i is independent normal; then the likelihood can be marginalised with respect to the α to obtain the appropriate likelihood function for the maximisation (Arulampalam *et al.*, 2000: 32) and a test of $\vartheta = 0$ provides a test for exogeneity of the initial condition. In the equi-correlated probit specification, simultaneous estimation of the parameters of Eqs. 2 and 3 (the structural and reduced equations respectively) can be achieved by maximising the log-likelihood function:

$$\ln L = \sum_{i=1}^{N} \ln \int_{-\infty}^{+\infty} \left\{ \Phi[(\pi' z_i + \vartheta\alpha_i)(2y_{i0} - 1)] \prod_{t=1}^{T} \Phi[(x_{it}'\beta + \gamma y_{it-1} + \alpha_i)(2y_{it} - 1)] \right\} g(\alpha_i) d\alpha_i \qquad (4)$$

where $g(\alpha)$ is the probability density of unobserved heterogeneity and Φ is the standard normal cumulative function. The main problem in the Heckman model is the computational burden of maximising the likelihood function, which requires simultaneous estimation of two composite functions, and this has led researchers searching for simplified solutions to its implementation (e.g. Orme, 2001; Wooldridge, 2005; Arulampalam and Stewart, 2009).[2]

Orme (2001) suggests a first computational simplification of the Heckman estimator, defined as a "two-step pseudo-ML estimator", which has been widely utilised in subsequent applications. He proposes a two-step procedure to address the initial condition problem that is locally valid when the correlation between y_{i0} and y_{it} (ρ) tends to zero and can perform well also when ρ is not small. Orme uses an approximation to substitute α_i with another unobservable component that is uncorrelated with the initial observation. By assuming bivariate normality of the composite error term ε_{i0} and unobserved heterogeneity α_i, that is $(\varepsilon_{i0},\alpha_i)\approx BVN(0,0,\sigma_\varepsilon,\sigma_\alpha,\rho)$, individual effects can be defined as: $\alpha_i=\rho\frac{\sigma_\alpha}{\sigma_\varepsilon}\varepsilon_{i0}+\sigma_\alpha\sqrt{(1-\rho^2)}w_i$, with $w_i\approx N(0,1)$, orthogonal to ε_{i0} by construction and distributed as $N(0,1)$. The structural model thus becomes:

$$y_{it}^*=\gamma y_{it-1}+x_i'\beta+\left[\rho\frac{\sigma_\alpha}{\sigma_\varepsilon}\varepsilon_{i0}+\sigma_\alpha\sqrt{(1-\rho^2)}w_i\right]+u_{it} \quad (5)$$

which encompasses two time-invariant components of unobserved heterogeneity, ε_{i0} and w_i. Orme suggests estimating Eq. (3) and computing its generalised residual:

$$e_i\equiv E(\varepsilon_{i0}\mid y_{i0})=(2y_{i0}-1)\sigma_\varepsilon\varphi(\pi^{*\prime}z_i/\sigma_\varepsilon)/\Phi[(2y_{i0}-1)\pi^{*\prime}z_i/\sigma_\varepsilon] \quad (6)$$

and then use it as an explanatory variable in (5), that is $\varepsilon_{i0}\equiv e_i$.

More recently Wooldridge (2005) has proposed a conditional maximum likelihood estimator as an alternative to the Heckman model, suggesting that the distribution of unobserved heterogeneity should be modelled conditional on the initial value and any exogenous explanatory variables, in order to integrate out individual effects α_i. The methodology consists of "finding the distribution conditional on the initial value and the observed history of strictly exogenous explanatory variables", rather than attempting "to obtain the joint distribution of all outcomes of the endogenous variables" as in the Heckman's approach (Wooldridge, 2005: 39). Wooldridge, however, is aware of the drawbacks in specifying a parametric auxiliary conditional distribution for the unobserved heterogeneity and states that "misspecification of this distribution generally results in inconsistent parameter estimates" (Wooldridge, 2005: 40).

[2] The shortcut implementation of Heckman's estimator proposed by Arulampalam and Stewart (2009) is equivalent to a standard random effects specification, where unobserved heterogeneity contains a heteroskedastic factor loading.

The model is a correlated random effects model (Mundlak, 1978; Chamberlain, 1984), which relaxes zero-correlation of the random effects model by assuming the following specification of unobserved individual effects:

$$\alpha_i = a_0 + \xi' \bar{x}_i + a_i \qquad (7)$$

where $a_i \approx iidN(0, \sigma_a^2)$ and is independent of x_{it} and u_{it} for each i and t, and where $\bar{x}_i$ are the means (group-means) of the time-varying covariates. A different specification for the individual effects α_i is suggested, which also includes the initial values of the endogenous variable y_{i0}:

$$\alpha_i = \xi_0 + \xi_1 y_{i0} + \xi' \bar{x}_i + a_i \qquad (8)$$

The so-called dynamic "correlated random effects" probit model can thus be written as:

$$y_{it} = \mathbf{1}(x_{it}'\beta + \gamma y_{it-1} + \xi_0 + \xi_1 y_{i0} + \xi' \bar{x}_i + a_i + u_{it} > 0) \qquad (9)$$

where a_i is the "new" unobserved heterogeneity and u_{it} is the idiosyncratic term. The model assumes the distribution of α given y_{i0} and x_i be $\alpha_i \mid y_{i0}, x_i \approx N(\xi_0 + \xi_1 y_{i0} + \xi' \bar{x}_i, \sigma_a^2)$ and the explanatory variables $z_{it} \equiv (x_{it}, y_{it-1}, y_{i0}, \bar{x}_i)$. It follows that the log-likelihood function is specified as:

$$\ln L = \sum_{i=1}^{N} \ln \int_{-\infty}^{+\infty} \left\{ \prod_{t=1}^{T} \Phi\left[(x_{it}'\beta + \gamma y_{it-1} + \xi_1 y_{i0} + \xi' \bar{x}_i + a_i)(2y_{it} - 1)\right] \right\} g^*(a_i) da_i \qquad (10)$$

where $g^*(a_i)$ is the normal density of the "new" unobserved heterogeneity a_i in Eq. (8). The likelihood function of Eq. (10) is equivalent to the likelihood function of a static random effects probit model where the explanatory variables are $z_{it} \equiv (x_{it}, y_{it-1}, y_{i0}, \bar{x}_i)$ and the maximum likelihood estimator can be obtained via standard random effects probit estimation.

Much of the evidence in the literature indicates that the Heckman, Orme, and Wooldridge methods produce comparable results. The results of Monte Carlo simulations show that when one or both longitudinal dimensions (T and N) are relatively large, $T \geq 6$ and $N \geq 800$ (Arulampalam and Stewart, 2009) or T longer than 5-8 periods (Akay, 2009), the bias is relatively small for all three estimators, whereas for smaller sample sizes, the bias increases although none of the estimators dominates.

3. Data and definition of financial distress

The dataset used in this paper is the Bank of Italy Survey of Household Income and Wealth (SHIW) for the period 1998-2006 and a total of five waves (1998, 2000, 2002, 2004 and 2006). The survey collects detailed data on demographics, household consumption, income and balance sheet items on a biannual basis, with the exception of a three-year gap between 1995 and 1998. The number of households interviewed in each wave is around 8000, providing a representative sample of the Italian resident population. Sampling is in two stages, municipalities in the first stage and households in the second stage, with municipalities divided into 15 strata defined by 17 regions and 3 classes of population (more than 40000, 20000 to 40000, and less than 20000 inhabitants) and households randomly selected from registry office archives.[3]

The sample of our empirical application consists of the longitudinal component drawn from the Historical Archive of the SHIW, namely an unbalanced panel with common entrance in 1998 and exits after at least three consecutive periods. The total number of observations is 8619, of which 1911 are observations relative to 637 households remaining in the sample for at least 3 waves, 1328 observations of 332 households remaining in the sample for at least 4 periods, and 5380 observations of 1076 households in the sample for 5 periods.

As anticipated in the Introduction, we take Brown and Taylor (2008), who suggest negative net wealth as a measure of financial hardship, as our starting point and build our framework on the idea that low levels of net wealth can be taken as indicators of financial vulnerability. However, two issues arise when trying to define households in "financial distress". The first is whether holding negative net wealth is a sufficient condition to identify a situation of financial stress. The second is whether positive or null net wealth values can be associated with financial vulnerability.

Fig. 1 Households in financial distress
(area below the line W = W_A)

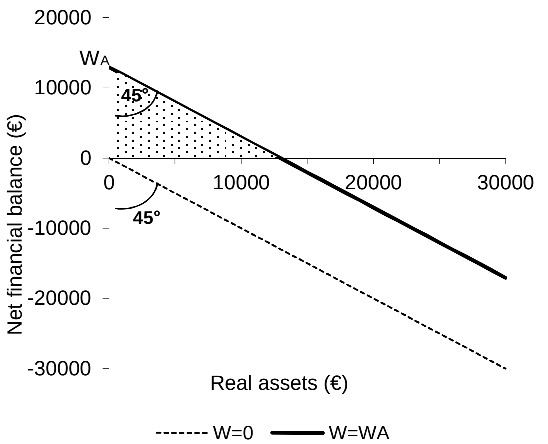
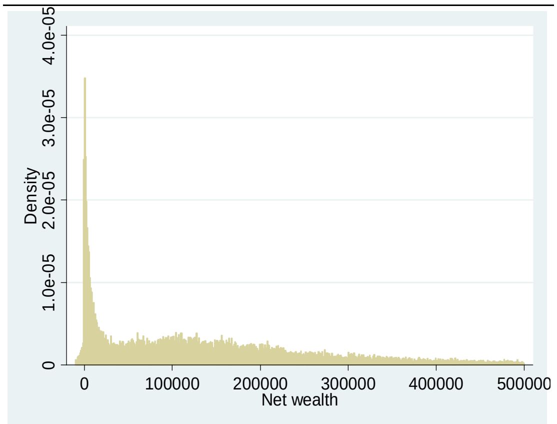


Fig. 1 depicts the possible combinations of real wealth holdings (positive *x*-axis) and net financial balance (*y*-axis). The negative *y*-axis describes a situation of negative net balance and null real wealth; the forty five degree line splitting the lower-right quadrant describes a situation on null wealth (*W* = 0); the area below the zero-wealth line encompasses all cases of negative net worth. Financially vulnerable households can be identified by including all individuals on the negative *y*-axis and in the area underneath the 45° line, and individuals with a "small" amount of positive net worth. It is clear that negative net wealth holdings define a situation of financial fragility, even though, on average, only 3%

[3] See Brandolini and Cannari (1994) and Faiella (2008) for a detailed description of the survey.

of Italian households have negative net wealth (as noted in Sierminska *et al*., 2008, a largely lower percentage than is observed in countries such as the US, the UK and France). Conversely, it requires some thought about whether a “small” entity of negative worth can determine a critical situation, without examining the associated real wealth value. It could be assumed that small amounts of negative worth are not critical if associated with high values of real wealth. Equally, individuals with moderate amounts of real wealth, but above the absolute value of the net financial balance and therefore with positive net worth, could be assumed to be in economic distress. Even if real wealth is high but is counterbalanced by a high value of debt, positive net worth holdings identify households with just small amounts of liquidity (the data show that households in financial distress are characterised, on average, by very low levels of real financial wealth). When considering positive net wealth holdings, financial distress is identified with the area below a parallel line to $W = 0$ and shifted upwards by a “certain amount”, so that it intersects the positive y-axis at the level W_A (line $W = W_A$ in Fig. 1): individuals in financial distress are also those with combinations of real wealth and net financial balance lying in the triangle defined by the line $W = W_A$ and the x and y axes (dotted area in Fig. 1).

Fig. 2 Net wealth distribution

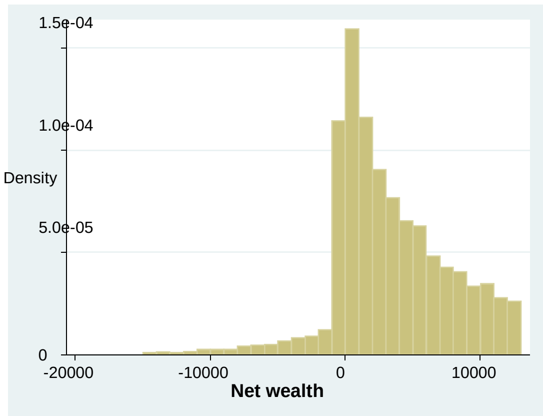


The choice of the threshold W_A was made with reference to the net wealth distribution of the SHIW, which is highly concentrated, with the majority of households owning low or null wealth (Fig. 2). We chose the threshold corresponding to the second decile of the wealth distribution, whose corresponding value is 12960 euro at 2006 prices, resulting in 15.3% of households being in financial distress (Table 1). First, the value corresponds roughly to half median income (26470 euro at 2006 prices), one of the measures proposed in the literature for identifying households in poverty. Second, the resulting percentage of households in distress is comparable with figures found in analogous contexts, such as the Eurostat deprivation index (15.4% of households in deprivation) published by ISTAT (2010), and percentages of poor households defined with asset-based criteria found by Brandolini *et al.* (2010), around 12% when considering financial assets and around 9% when considering net wealth.

Table 1	Distribution of the response variable (percentages)		
	In distress	Not in distress	Total
1998	17.2	82.8	100
2000	16.0	84.0	100
2002	15.1	84.9	100
2004	13.1	86.9	100
2006	13.3	86.7	100
Total	15.3	84.7	100

Fig. 3 depicts the distribution of net wealth of financially vulnerable households. It is worth stressing that the aim of the paper is not finding an exact measure of financial vulnerability, but contributing to the literature by suggesting the appropriateness of asset-based measures, in addition to income-based or perceived distress indicators, in assessing persistency of financial vulnerability of households.

Fig. 3 Net wealth distribution for values below threshold W_A



4. Empirical results

To test the hypothesis of true time persistency of financial distress among Italian households, we estimate the Heckman model (1981a; 1981b) defined by Eqs. (2), (3) and (4). The choice of the model covariates is based on the reduced form models for the determinants of household debt and financial assets and from the literature on household financial distress (see, amongst others, Cox and Jappelli, 1993; May and Tudela, 2005; Brown and Taylor, 2008; del Rio and Young, 2008). The probability of experiencing financial hardship is thus a function of the following variables (Table 2):

- The lagged dependent variable
- Income by quartile to capture distributive effects

- A set of household-level control variables: three age classes (below 40, 40-60, above 60); gender; education levels (primary, lower secondary, upper secondary, university); number of household components; ownership of risky assets, defined according to Guiso and Jappelli (2002); homeownership; whether the head of household is indebted
- Two aggregate variables: macro-area unemployment rates and the regional house price index.[4]

Table 2 Descriptive statistics

variables	mean	sd	min	max	N
distress (t-1)	0.1576	0.3644	0	1	6578
young	0.1186	0.3233	0	1	8627
old	0.4341	0.4957	0	1	8627
1st income quartile	0.2120	0.4088	0	1	8627
3rd income quartile	0.2547	0.4357	0	1	8627
4th income quartile	0.2918	0.4546	0	1	8627
education: primary	0.3388	0.4733	0	1	8627
education: lower secondary	0.2770	0.4476	0	1	8627
eduation: university	0.0873	0.2823	0	1	8627
female	0.2401	0.4271	0	1	8627
no. components	2.7978	1.2900	1	9	8627
risky portfolio	0.1602	0.3668	0	1	8627
homeowner	0.7443	0.4363	0	1	8627
indebted	0.2095	0.4069	0	1	8627
unemployment	9.6331	5.6565	3.8	19.6	8627
house prices	1.1063	0.1306	0.94	1.63	8627

The head of household is defined as the individual who declares herself responsible for the financial and economic choices of the household, a reasonable option as the focus of the study concerns financial decisions by the household (as in Bertocchi *et al*., 2011). In estimating the model, the reference household is assumed to have a head in the age class 40-60, with income belonging to the second quartile of the income distribution, male, with an upper secondary education level, owning a risky portfolio, homeowner, and indebted. Initial conditions (Eq. 3) are estimated with the explanatory variables set of the structural equation, adding the initial values x_{i0} of the structural equation and three supplementary dummy instrumental variables for if the household lives in the South of Italy, if the head of household is self-employed, and if the household resides in a municipality with less than 20000 inhabitants (similar instruments are used in Cox and Jappelli (1993) and related literature).

[4] Unemployment rates are taken from Italian Regional Accounts by ISTAT. The data source for house prices is Muzzicato *et al*. (2008).

Table 3 reports the estimates of the initial conditions equation and the structural equation for the Heckman model, where the main parameter of interest is the coefficient of the lagged dependent variable, γ of Eq. (4).[5] After controlling for unobserved heterogeneity, there is evidence of true state dependence, that is, the probability of experiencing financial distress at time *t* positively depends upon the probability of having experienced financial fragility at time *t*-1. The previous state parameter is equal to 0.563 and is statistically significant at the 95% level.

Table 3 Dynamic probit model estimation (Heckman method)

	Initial condition equation		Structural equation	
	Coef.	t-stat	Coef.	t-stat
constant	2.943	5.42	-0.917	-2.29
		Lagged response variable		
distress (t-1)			0.563	5.42
		Socio-economic explanatory variables		
young	0.426	2.43	0.092	0.72
old	-0.011	-0.06	0.107	0.96
1st income quartile	0.375	2.38	0.453	4.52
3rd income quartile	-0.466	-2.44	-0.322	-2.77
4th income quartile	-0.756	-3.19	-0.588	-4.20
education: primary	0.695	3.45	0.623	4.72
education: lower secondary	0.132	0.78	0.380	3.30
eduation: university	-0.787	-2.19	-0.447	-2.06
female	0.243	1.58	0.104	1.03
no. components	-0.002	-0.04	0.019	0.47
risky portfolio	-0.295	-1.31	-0.438	-3.30
homeowner	-3.377	-10.05	-3.258	-14.80
indebted	0.621	3.63	0.303	2.90
		Aggregate explanatory variables		
unemployment	0.110	1.47	0.03	3.440
house prices	-4.240	-1.32	0.12	0.390
		Other instruments for initial conditions		
south-isles			-0.719	-0.81
self-employed			-1.085	-4.57
small area			0.217	1.49
rho			0.244	3.30
theta			1.233	2.62
Log-likelihood			-1279.3	
LR test: rho=0			chi2(1) =	29.39
			p-value =	0.000
No. of observations	8619			

[5] The model is estimated in Stata using -redprob- (Stewart, 2006).

The results confirm the presence of unobserved individual effects, with a value of the LR test on ρ of 29.39 (p-value=0.000). According to Arulampalam (1999), the fraction of total explained variance due to unobserved individual characteristics can be derived from ρ as $\sigma_{\alpha}^{2} = \rho/(1-\rho)$. In our model 32.2% of the total variance is explained by unobserved household-level characteristics, in line with the 34% in May and Tudela (2005) (Boheim and Taylor, 2000, found a value around 19%). The size of this parameter shows the importance of individual components in the analysis of household financial problems, and the adequateness of using panel data. The *t*-test on coefficient ϑ (the parameter that defines the presence of individual effects η_i correlated with α_i in the initial conditions in Eq. 3 and, we recall it, defined as $\eta_i = \vartheta\,\alpha_i + u_{i0}$) rejects the null of non-exogeneity of initial conditions. [6]

As regards the explanatory power of our model, we observe that the age structure is not very strong: for the younger age group (household heads under 40) and the older age group (household heads over 60) the dummies are not significant.[7] However their coefficients are positive, which may suggest greater distress than in the reference age group (40-60 years old), as revealed by descriptive statistics where a relatively higher percentages of younger and older than middle-aged households are in distress.

A crucial variable in the model is income and it enters the equation in the form of quartiles, with the second one as the reference category. The first quartile coefficient is positive, while the third and fourth quartile coefficients are negative and increasing in absolute terms, meaning that lower income households display a higher probability of financial distress, whilst higher income households display lower probability to get into difficulties. The results are as expected: a low level of income if persistent over time, generates null or limited savings, and likely induces indebtedness to sustain household consumption. Low levels of savings are nearly always a sufficient condition for low wealth levels; it is therefore reasonable to associate lower income with higher probability of financial distress. In other of the models proposed in the literature financial fragility is associated with

[6] Estimates were also run with different thresholds of net wealth to define households in distress. Alternative thresholds to the second decile were identified by percentiles between the 17^{th} and the 22^{nd}, corresponding to values of net wealth in the interval 8311-17593 euro at 2006 prices. Results provide statistically significant coefficients of the lagged dependent variable between 0.484 and 0.673, which are justified by differences in the thresholds. As regards model specification, there is evidence of unobserved heterogeneity and coefficients of explanatory variables are of the same sign and order of magnitude as our main model.

[7] In an alternative specification (not shown) which excludes education level, the oldest age group dummy is significant.

the income gearing ratio and, therefore, implicitly expresses an inverse relationship with income (May and Tudela, 2005). Our result is coherent with Boheim and Taylor's (2000) model where income has a negative sign. In line with these results, the probability of financial distress displays a negative relationship with education levels.

The dummy for risky portfolio is significant with a negative coefficient. As risky portfolios are owned by higher income households, this variable can reinforce the role of income in defining the "ability" of the household to pay, and indicates lower exposure to financial fragility. The dummies for homeownership and being indebted have the expected signs: respectively negative and positive. There are weak signs of gender and composition effects: the female dummy and the number of household components show a positive but not significant coefficient. Living in a household with a female head (and with a greater number of members) can potentially expose households to higher financial distress. Gender effects can play a role in determining household's financial decisions, affecting preferences and the nature of background risk as "it is realistic to assume that women, being in a more vulnerable position in the labour market, tend to bear more labour risk and therefore more background risk" (Bertocchi *et al*., 2011: 2).

As suggested in the literature, our model specification includes aggregate variables such as regional unemployment rates and house prices (see e.g. May and Tudela, 2005).[8] Aggregate explanatory variables, being annual, can capture time effects, and being disaggregated at the territorial level, capture regional or wider area effects; unlike time dummies, they provide a direct identification of the macroeconomic events that affect the dependent variable. Ex-ante expectations are of lower probabilities of financial distress for households living in areas of lower unemployment. The overall effect of the house price index, instead, is not predetermined ex-ante: on the one hand, an increase can have a dampening effect on non-homeowners and make it more difficult to access the property market, on the other hand, it will positively affect house owners (as the value of their property increases). In our model, unemployment rate by geographical location is statistically significant and positive, denoting a higher probability of financial distress among households in areas of high unemployment; the regional house price index is not significant. Both results are in line with May and Tudela (2005).

[8] May and Tudela (2005) include in their model specification two additional variables, the income gearing ratio and the loan to value ratio. However they focus on indebted households and therefore both variables, if not missing in the survey, are available for each observation. Our study sample instead covers the whole survey sample and includes households without debt and/or without real wealth, generating a large number of missing values. For this reason these variables are not included in the analysis.

4.1. Comparisons with alternative estimation methods and robustness analysis

Results of the Heckman model are now compared to those obtained from the Orme (Eqs. 5 and 6), exogenous random effects and pooled models, as reported in Table 4.[9] In line with the literature, the Orme model gives equivalent results to the Heckman's, with an estimated coefficient of the lagged response variable of 0.563 and a significant coefficient of the generalised residual, confirming the validity of the model. The coefficient of the previous state γ is larger in the case of the random effects model with exogenous initial conditions than in the Heckman model: 0.790 compared to 0.563. This result is coherent with the literature and shows that the hypothesis of exogenous initial conditions tends to overestimate state persistence. The coefficients of the other variables are of the same magnitude, sign and significance as in the Heckman model. The final comparison is with the pooled model which produces a coefficient of the lagged variable of 0.881. However, the pooled and the random effects models involve different normalisations of the error term: normalisation for error variance in the pooled model is $\sigma_\varepsilon^2 = 1$, and in the random effects model it is $\sigma_u^2 = 1$. To allow comparison of the coefficients, those in the random effects model need be multiplied by $\sigma_u / \sigma_\varepsilon = \sqrt{1-\rho}$, where $\rho = \sigma_\alpha^2 / (\sigma_\alpha^2 + 1)$ is the constant cross-period error correlation (Arulampalam, 1999). Scaled coefficients of lagged financial distress are 0.736 in the exogenous random effects model, 0.490 in the Heckman model and 0.501 in the Orme model.

[9] The model was estimated also with the Wooldridge methodology (see Eq. 9), but when comparing the results with other estimation methods, substantial differences in the coefficient of the lagged variable were found. This is possibly explained by the fact that the model includes, in addition to explanatory variables, their means over time (which, together with the initial value of the dependent variable, define unobserved heterogeneity): low variation over time of explanatory variables can introduce collinearity problems in the estimation. As regards overall model specification, the lagged dependent variable is statistically significant and, when testing for joint significance in the group means and initial values, we do not reject the presence of unobserved individual effects.

Table 4 **Dynamic probit model estimation (other methods)**

	Pooled		Exogenous RE		Orme	
	Coef.	t-stat	Coef.	t-stat	Coef.	t-stat
constant	-0.905	-2.55	-0.912	-2.40	-0.913	-2.29
			Lagged response variable			
distress (t-1)	0.881	12.67	0.790	8.77	0.563	5.420
			Socio-economic explanatory variables			
young	0.073	0.70	0.094	0.81	0.086	0.68
old	0.116	1.25	0.116	1.13	0.106	0.95
1st income quartile	0.383	4.54	0.426	4.5	0.452	4.53
3rd income quartile	-0.258	-2.58	-0.284	-2.59	-0.323	-2.78
4th income quartile	-0.475	-4.06	-0.533	-4.04	-0.589	-4.21
education: primary	0.480	4.76	0.544	4.54	0.624	4.78
education: lower secondary	0.289	3.23	0.328	3.15	0.382	3.37
eduation: university	-0.374	-2.11	-0.399	-2	-0.439	-2.06
female	0.057	0.72	0.070	0.77	0.100	1.02
no. components	0.016	0.48	0.021	0.57	0.020	0.49
risky portfolio	-0.353	-3.15	-0.406	-3.23	-0.441	-3.33
homeowner	-2.681	-26.02	-2.940	-15.61	-3.221	-15.12
indebted	0.269	2.96	0.285	2.88	0.302	2.90
			Aggregate explanatory variables			
unemployment	0.023	3.35	0.024	3.16	0.029	3.43
house prices	0.076	0.28	0.081	0.28	0.120	0.4
			Generalised residual			
generalised res.					0.323	4.53
sigma_alpha			0.388	3.20	0.514	4.62
rho			0.131	1.84	0.209	2.92
Log-likelihood	-907.1		-905.3		-892.5	
LR test: rho=0	chi2(1) =		3.63		9.11	
	p-value =		0.028		0.001	
No. of observations: 8619						

As a robustness check, the model was also estimated on the balanced panel data structure with T = 5, and a total number of observations of 5380, for a total of 1076 households. The overall percentage of households in distress is 14.9%. The previous results of the pooled, exogenous initial conditions, Heckman and Orme models also hold with this alternative data structure, and the coefficients of the lagged dependent variable are of the same order of magnitude (despite slightly larger) in all the models as in the unbalanced sample: 0.970 in the pooled specification, 0.852 in the RE model with exogenous initial conditions, 0.644 with Orme and 0.636 with Heckman (Table 5). In the random effects models the null $\rho = 0$ is rejected at the standard significance levels; the null of joint non-significance of the instruments in the Heckman model is also rejected.

Table 5 Dynamic probit model estimation (balanced panel)

	Pooled		Exogenous RE		Orme		Heckman	
	Coef.	z	Coef.	z	Coef.	z	Coef.	z
constant	-1.140	-2.66	-1.154	-2.49	-1.181	-2.43	-1.185	-2.42
Lagged response variable								
distress (t-1)	0.970	11.02	0.852	6.87	0.644	4.78	0.636	4.70
Socio-economic explanatory variables								
young	0.251	1.86	0.314	1.97	0.329	1.95	0.337	1.97
old	0.119	1.03	0.110	0.84	0.116	0.83	0.117	0.83
1st income quartile	0.317	3.04	0.364	3.08	0.363	2.95	0.362	2.92
3rd income quartile	-0.187	-1.48	-0.207	-1.49	-0.240	-1.65	-0.239	-1.64
4th income quartile	-0.407	-2.76	-0.451	-2.74	-0.497	-2.88	-0.497	-2.87
education: primary	0.450	3.58	0.511	3.38	0.590	3.61	0.587	3.54
education: lower secondary	0.233	2.07	0.260	1.97	0.312	2.19	0.309	2.13
eduation: university	-0.299	-1.43	-0.327	-1.36	-0.356	-1.39	-0.369	-1.41
female	0.082	0.81	0.103	0.87	0.140	1.09	0.146	1.11
no. components	0.014	0.34	0.018	0.37	0.020	0.38	0.018	0.34
risky portfolio	-0.441	-3.10	-0.498	-3.10	-0.536	-3.20	-0.533	-3.19
homeowner	-2.589	-20.69	-2.873	-11.74	-3.123	-11.83	-3.167	-11.65
indebted	0.300	2.71	0.312	2.59	0.337	2.67	0.341	2.69
Aggregate explanatory variables								
unemployment	0.033	3.64	0.036	3.40	0.041	3.63	0.042	3.63
house prices	0.178	0.57	0.198	0.59	0.233	0.67	0.232	0.67
Generalised residual								
generalised res.					0.325	3.56		
sigma_alpha			0.407	2.59	0.514	3.62	0.571	
rho			0.142	1.51	0.209	2.29	0.246	2.64
theta							1.312	2.01
Log-likelihood	-585.0		-573.7		-565.7		-757.7	
LR test: rho=0	chi2(1) =		2.45		5.61		19.1	
	p-value =		0.059		0.009		0.000	
No. of obs.: 5380								

5. Conclusions

This paper analysed the determinants of financial distress amongst Italian households and its persistency over time, using the longitudinal component of the Bank of Italy Survey of Household Income and Wealth for the period 1998-2006. The study drew from the international literature on indebtedness problems and perceived financial hardship and proposed a quantitative measure of potential financial weakness of households to adverse shocks in the economy. By means of a dynamic probit model accounting for time-invariant unobserved heterogeneity and dealing with the initial condition problem, we were able to distinguish between true and spurious state dependence and show that Italian households experience financial fragility persistently over time. This means that the probability of experiencing financial distress at time t is positively related to the probability of having experienced distress at time (t – 1). By taking account of both true state dependence and unobserved heterogeneity, the estimation methodology advanced the existing empirical literature that uses data to analyse the financial conditions of households in Italy. Our definition of financial distress consists of a quantitative measure of financial distress based on combinations of assets and liabilities: it is the sign and dimension of net wealth rather than just debt levels that identify households experiencing financial fragility. The choice of the threshold to split households between those who experience financial hardship and those who do not was made with reference to the net wealth distribution. We chose the level of wealth corresponding to the second decile of its distribution, resulting in 15.3% of households being in financial distress. The value of the threshold corresponds roughly to half median income, one of the measures proposed in the literature for identifying households in poverty. Moreover, the resulting percentage of households in distress is comparable with figures found in analogous contexts, such as the Eurostat deprivation index, and percentages of poor households defined with asset-based criteria found in the literature.

Estimation results indicate that financial distress of Italian households is persistent over time: the Heckman model estimates a statistically significant coefficient of the lagged dependent variable of 0.563. The null hypothesis of non-significance of unobserved heterogeneity is rejected, with the fraction of variance explained by individual unobserved effects of value 32.2%. The hypotheses of non-exogeneity of initial conditions and joint non-significance of the initial values are also rejected. In terms of the other explanatory variables, "ability to pay" confirms our expectations: higher income, higher education and owning a risky portfolio lower the probability of experiencing financial distress. Age is not very relevant, with most dummies not significant. There is weak evidence also that households with a female head have a higher probability of incurring financial fragility. For aggregate variables, higher unemployment positively affects the probability of distress, whereas house prices are not significant. When comparing methods, the results of the pooled and random effects probit models are in line with the findings in the literature, with coefficients of the previous state showing higher values than the Heckman model, indicating that taking no account of unobserved heterogeneity or of exogenous initial conditions leads to overestimation of the coefficient of interest. The Orme model is equivalent to the Heckman's. Robustness checks, consisting of estimation of the same set of models on a balanced longitudinal data set, confirm our results.

Bibliography

Akay, A., 2009. The Wooldridge method for the initial values problem is simple: what about performance?. Discussion Paper 3943, IZA.

Arulampalam, W., 1999. A note on estimated effects in random effect probit models. Oxford Bulletin of Economics and Statistics 61, 597-602.

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

Arulampalam, W., Stewart, M.B., 2009. Simplified implementation of the Heckman estimator of the dynamic probit model and a comparison with alternative estimators. Oxford Bulletin of Economics and Statistics 71, 659-681.

Barwell, R., May, O., Pezzini, S., 2006. The distribution of assets, income and liabilities across UK households: results from the 2005 NMG Research survey. Bank of England Quarterly Bulletin, Spring.

Bertocchi, G., Brunetti, M., Torricelli, C., 2011. Marriage and other risky assets: a portfolio approach. Journal of Banking and Finance 35, 2902-2915.

Boeri, T., Brandolini, A., 2005. The age of discontent: Italian households at the beginning of the decade". IZA discussion paper 1530.

Boheim, R., Taylor, M. P., 2000. My home was my castle: evictions and repossessions in Britain. Journal of Housing Economics 9, 287-319.

Brandolini, A. and Cannari, L., 1994. Methodological appendix: The Bank of Italy's survey of household income and wealth. In: Ando, A., Guiso, L., Visco, I. (Eds.), Saving and the Accumulation of Wealth: Essays on Italian Households and Government Behavior. Cambridge University Press, Cambridge, pp. 369–386.

Brandolini, A., Magri, S., Smeeding, T.M., 2010. Asset-based measurement of poverty. Journal of Policy Analysis and Management 29, 267-284.

Brandolini, M., Giarda, E., Moriconi, C., 2009. The burden of housing costs and consumer credit in Europe: preliminary evidence of differences in perceptions among countries. Presented at the BHPS Conference, University of Essex, 9-11 July.

Brown, S., Taylor, K., 2008. Household debt and financial assets: evidence from Germany, Great Britain and the USA. Journal of the Royal Statistical Society, Series A 171, 615–643.

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

Christelis, D., Jappelli, T., Paccagnella, O., Weber, G., 2009. Income, wealth and financial fragility in Europe. Journal of European Social Policy 19, 359-376.

Cox, D., Jappelli, T., 1993. The effect of borrowing constraints on consumer liabilities. Journal of Credit and Banking 25, 197-213.

Cox, P., Whitley, J., Brierley, P., 2002. Financial pressure in the UK household sector: evidence from the British Household Panel Survey. Bank of England Quarterly Bulletin 42, 410-419.

del Rio, A., Young, G., 2008. The impact of unsecured debt on financial pressure among British households. Applied Financial Economics 18, 1209-1220.

Faiella, I., 2008. Accounting for sampling design in the SHIW. Working Paper 662, Bank of Italy.

Georgarakos, D., Lojschove, A., Ward-Warmedinger, M., 2010. Mortgage indebtedness and household financial distress. Working Paper 1156, European Central Bank.

Guiso, L., Jappelli, T., 2002. Household portfolios in Italy. In: Guiso, L., Haliassos, M., Jappelli, T. (Eds.), Household Portfolios, Cambridge: MIT Press, 251-289.

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

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

Honoré, B. E., 1993. Orthogonality conditions for tobit models with fixed effects and lagged dependent variables. Journal of Econometrics 59, 35-61.

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

ISTAT, 2010. Rapporto sulla Coesione Sociale. ISTAT, Rome.

Jappelli, T., Pagano, M., di Maggio, M., 2008. Households' indebtedness and financial fragility. Working Paper 208, Centre for Studies in Economics and Finance.

May, O., Tudela, M., 2005. When is mortgage indebtedness a financial burden to British households? A dynamic probit approach. Working paper 277, Bank of England.

May, O., Tudela, M., Young, G., 2004. British household indebtedness and financial stress: a household level picture. Bank of England Quarterly Bulletin 44, 414-428.

Ministry of Labour and ABI, 2009 and 2010. Rapporto trimestrale - Indicatori di indebitamento, vulnerabilità e patologia finanziaria delle famiglie italiane.

http://www.lavoro.gov.it/Lavoro/md/AreaSociale/MonitoraggioValutazione/VulnerabilitaFamiglie/

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

Muzzicato, S., Sabbatini, R., Zollino, F., 2008. Prices of residential property in Italy: constructing a new indicator. Questioni di Economia e Finanza 17, Bank of Italy.

OECD, 2010. Economic Outlook. Vol. 58, OECD Publishing, Paris.

Orme, C.D., 2001. Two-step inference in dynamic non-linear panel data models. School of Economic Studies, University of Manchester, mimeo: http://personalpages.manchester.ac.uk/staff/chris.orme/

Pudney, S., 2008. The dynamics of perception: Modelling subjective wellbeing in a short panel. Journal of the Royal Statistical Society, Series A 171, 21-40.

Stewart, M.B., 2006. -redprob- A Stata program for the Heckman estimator of the random effects dynamic probit model. http://www2.warwick.ac.uk/fac/soc/economics/staff/academic/stewart/stata/redprobnote.pdf

Stewart, M.B., 2007. Inter-related dynamics of unemployment and low wage employment. Journal of Applied Econometrics 22, 511-531.

Whitley, J., Windram, R., Cox, P., 2004. An empirical model of household arrears. Working paper 214, Bank of England.

Wooldridge, J. M., 2005. Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. Journal of Applied Econometrics 20, 39–54.