FINANCIAL HARDSHIP AND SAVING BEHAVIOUR: BAYESIAN ANALYSIS OF BRITISH PANEL DATA

Sarah Brown[1], Pulak Ghosh[2], Bhuvanesh Pareek[2] and Karl Taylor[1]

[1] Department of Economics, University of Sheffield,
9 Mappin Street, Sheffield, UK, S1 4DT; and IZA Bonn

[2] Department of Quantitative Methods and Information Systems,
Indian Institute of Management, Bangalore, India

Abstract: We explore whether a protective role for savings against future financial hardship exists using household level panel data. We jointly model the incidence and extent of financial problems, using a two part approach which allows different data-generating processes for experiencing financial hardship and the extent of financial hardship experienced. Such a two-part approach is important given the considerable inflation at zero when analysing financial problems. The model is estimated using a flexible Bayesian approach with correlated random effects and the findings suggest that: (i) saving on a regular basis mitigates against both the likelihood of experiencing, as well as the number of, future financial problems; (ii) state dependence in financial problems exists; (iii) interdependence exists between financial problems and housing costs, specifically higher housing costs are associated with an increased probability of experiencing financial hardship.

Key Words: Bayesian Modelling; Financial Hardship; Saving; Zero Inflation.

JEL Classification: C11; D12; D14; R20.

Acknowledgements: We are grateful to the Data Archive at the University of Essex for supplying the British Household Panel Survey waves 1-18, and Understanding Society waves 1-7. We would also like to thank Raslan Alzuabi for excellent research assistance and Daniel Gray and Alberto Montagnoli for valuable comments. We are also grateful to seminar participants at the University of Birmingham and participants at the 2018 Royal Economic Society annual conference, University of Sussex, for valuable comments and advice.

June 2018

1. Introduction

Since the 2008 global financial crisis, the low levels of savings held at the household level in many countries have led to considerable concern amongst policymakers regarding the potential financial vulnerability faced by households (Garon, 2012). Savings provide a financial buffer in the event of adverse events from illness and job loss (i.e. income shocks) through to washing machine and car break-downs (i.e. expenditure shocks). Furthermore, according to the Office for National Statistics (ONS), the UK savings ratio has fallen from over 15% in 1993 to approximately 5.4% in 2017 quarter 3. Low savings may lead to increased demand for high cost lending products, e.g. payday loans, which may exacerbate financial problems and lead to persistence in financial distress over time. The relationship between saving behaviour and financial distress is clearly complex and, although an extensive literature exploring saving behaviour exists,[1] limited attention has been paid in the economics literature to understanding the implications of a lack of savings for future financial wellbeing. We contribute to existing knowledge by evaluating the implications of saving on a regular basis for future financial wellbeing, focusing on the protective role of saving in the context of a large nationally representative UK data set.

Although the general consensus amongst policymakers appears to be that individuals are not saving enough for either the short-term or the long-term, e.g. see Crossley et al. (2012), only a limited number of studies in the economics literature have explored the implications of saving for future financial wellbeing. Given that life cycle theories on household consumption and saving behaviour predict that households will consume savings and assets when faced with financial hardship (see, for example, Browning and Crossley 2001, and Modigliani and Brumberg 1954), it seems interesting to explore from an empirical perspective whether and to what extent holding savings provides a buffer against future financial adversity.

[1] Given our focus on modelling financial hardship, it is beyond the scope of this paper to present a detailed review of the extensive literature on saving, see Browning and Lusardi (1996) for a comprehensive review of the literature on household saving.

A small yet growing literature exploring household financial hardship exists, which uses nationally representative household surveys (see, for example, Brown et al., 2014, and Giarda, 2013). However, with the exception of a few US studies (e.g. McKernan et al., 2009, Mills and Amick, 2010, and Gjertson, 2016), an explicit link has not been made in such studies to the potential protective role of saving in mitigating financial hardship. In contrast, these US studies highlight the potential protective role of saving amongst samples of low income households. For example, McKernan et al. (2009) use data from the 1996 and 2001 US Survey of Income and Programme Participation, which oversamples low income households, to explore whether assets reduce material hardship following an adverse event. Their findings suggest that, after controlling for income, asset poor families are 14 percentage points more likely to experience deprivation than non-asset poor families. Interestingly, they also find that approximately 40% of families experiencing negative events reduce their liquid assets. Mills and Amick (2010) use the same data source to explore whether holding modest amounts of liquid assets provides protection against financial hardship for low income households. For households in the lowest income quintile, their results suggest that holding liquid assets of up to $1,999 relative to holding zero assets reduces the incidence of material hardship by 5.1 percentage points.

In a similar vein, Collins and Gjertson (2013) analyse data from the Annie E. Casey Foundation's Making Connections project, which is a longitudinal study of families residing in disadvantaged neighbourhoods in 10 US cities. Their findings suggest that families that do save for an emergency are less likely to experience as many material hardships as those households which do not save. Although such studies are not able to discern the direction of causality, they do highlight some interesting associations between saving behaviour and subsequent financial hardship. More recently, Gjertson (2016), also using data from the Annie E. Casey Foundation's Making Connections project, presents evidence supporting a protective role for small amounts of saving against future financial hardship for this non-representative sample of low income US households. Thus, households holding even small amounts of saving may have a financial buffer against future shocks. Furthermore, the regression analysis of longitudinal data highlights the dynamic aspect of household finances with those households which saved for emergencies experiencing less financial hardship three years later.

Establishing a financial buffer for adverse effects has been found to be an important motivation for saving in large scale nationally representative data sets. For example, Le Blanc et al. (2016), who explore household saving behaviour in 15 euro-area countries, using the Household Finance and Consumption Survey 2010-11, find that ‘saving for unexpected events’ is reported to be the most important saving motive at the euro-area level by 53 percent of respondents. Furthermore, the importance of this saving motive is found to be prevalent across all countries regardless of institutional differences and differences in welfare systems. Similar findings supporting the importance of precautionary saving motives are reported by Kennickell and Lusardi (2005) using the US Survey of Consumer Finances.

We contribute to the existing literature by exploring whether a protective role for saving against future financial hardship exists in the UK. Specifically, we explore the effect of regular saving behaviour on future financial hardship using household level panel drawn from the British Household Panel Survey and Understanding Society. Households holding even small amounts of saving may have a financial buffer against future shocks, such as changes in working or overtime hours as well as poor health, which may affect ability to work. As stated by Despard et al. (2016), ‘households without sufficient savings are at greater risk for material hardship,’ p.4. In order to allow for the fact that housing costs, i.e. mortgage payments and rent, represent one of the main financial commitments of households, we model financial problems and housing costs jointly to allow for their potential interdependence. In addition, we make a methodological contribution by developing a flexible Bayesian framework which allows for the considerable inflation at zero when analysing financial problems in the context of a large scale nationally representative survey, i.e. a significant number of households do not experience financial hardship. Within our flexible Bayesian framework, we also allow for persistence in experiencing financial problems, which has been commented on in existing studies. Bayesian modelling techniques have only been applied to household finances in a small number of papers (see, for example, Brown et al., 2014, 2015, 2016). Given that the Bayesian approach allows flexible modelling in complex applications, such an approach seems to be ideally suited to modelling such financial behaviour.

2. Data

We investigate the existence, intensity and persistence of financial hardship in the UK, focusing on the protective role of saving, using longitudinal data over nearly a twenty year period, from the 1990s to 2016. This is explored at the household level using the British Household Panel Survey (BHPS) and its successor Understanding Society, the UK Household Longitudinal Survey (UKHLS). The BHPS took place from 1991 through to 2008 and was replaced by the UKHLS in 2009. Both surveys are nationally representative large scale panel data sets containing detailed information on economic and socio-demographic characteristics. The BHPS comprises approximately 10,000 annual individual interviews, with the same individuals interviewed in successive waves. In the first wave of the UKHLS, over 50,000 individuals were interviewed from 2009 through to 2011 and correspondingly in the latest wave available, wave 7, around 45,000 individuals were interviewed between 2015 and 2017 (hereafter referred to as 2016). A subset of individuals in the UKHLS can be linked to the BHPS thus forming a relatively long panel survey.

After matching the BHPS and UKHLS and incorporating lags, the estimation sample spans the period 1998 through to 2016. We focus on a sample of 13,700 individuals who are the head of household or are identified as the individual responsible for making financial decisions within the household (referred to as the head of household hence forth). These individuals are observed over time yielding an unbalanced panel comprising 69,472 observations, where they are present in the panel for 8 years, on average, and we focus on individuals aged between 18 and 65.

We consider how saving behaviour influences both the incidence and the extent of future financial problems. From 1996 onwards, information on the following types of financial hardship are available in the data: problems paying for accommodation; problems with loan repayments (specifically non-mortgage debt); problems keeping their home adequately warm; difficulty in being able to pay for a week's annual holiday; difficulty in being able to replace worn-out furniture; ability to buy new rather than second hand clothing; ability to eat meat, chicken, fish every second day; and ability to have friends or family for a drink or meal at least once a month. Figure 1 shows the distribution of the number of household financial problems, where around 60% of the sample report no problems and 40% report between 1 to 6 or more financial problems over the period. The number of financial problems, conditional on experiencing financial hardship, is regarded as a count outcome and, hence, we employ a Poisson estimator as detailed in Section 3 below. Information is also available in the data on the household's housing costs (i.e. mortgage repayments and rent), specifically the last monthly payment made. Figure 2 shows the distribution of the natural logarithm of monthly housing costs, which comprise mortgage debt repayments and rent, where around 30% of the sample did not incur any such costs. Out of this group who report zero housing costs, 70% own their home outright. Conditional on holding secured debt or renting, the distribution of monthly payments is approximately normally distributed and so the level of housing cost is modelled as a continuous variable. The proposed modelling approach is developed in Section 3 below.

Our focus lies in exploring the protective role of saving on a regular basis. A distinction is made in the existing literature between passive and active saving, where active saving relates to money set aside to be used in the future and passive saving refers to wealth accumulation due to asset appreciation. Active saving has been explored from an empirical perspective by a small number of studies, including for the UK: Guariglia (2001); Yoshida and Guariglia (2002); Guariglia and Rossi (2004); and Brown and Taylor (2016). Our measure of monthly saving, which is akin to active saving, is based on responses to the following question: "*Do you save any amount of your income, for example, by putting something away now and then in a bank, building society, or Post Office account other than to meet regular bills? About how much, on average, do you manage to save a month?*" We explore two measures of the head of household's saving behaviour: a binary indicator of saving on a monthly basis in the previous year and the average amount of monthly saving in the previous year.

In the empirical analysis, we include a comprehensive range of control variables in matrix $\boldsymbol{X}$ (defined below). These include head of household characteristics such as gender; white; age; household size (excluding the head); highest educational attainment – specifically degree, other high educational qualification (e.g. teaching or nursing), A levels, GCSE/O levels, or any other qualification, with no qualifications as the omitted category; and labour market status, i.e. employee, self-employed or unemployed, out of the labour market is the reference category. We also control for: the natural logarithm of monthly household equivalized income; the natural logarithm of annual household expenditure on water, gas and electricity; the natural logarithm of total monthly household expenditure on non-durable goods; government office regions (London is the omitted category); and year of interview (pre 2001 is the reference period). In modelling the incidence and the number of financial problems, we also condition upon whether the individual has had a change in their health between waves, on the assumption that an adverse health shock will influence financial problems but not monthly housing costs.[2] The change in health state is defined as a binary indicator for whether the individual has experienced a change (between t-1 and t) in one or more of the following conditions/problems: sight; hearing; heart (including blood pressure); mobility and arthritis; bronchitis; diabetes; depression; epilepsy; cancer; stroke; or any other condition.

Summary statistics are provided in Table 1 Panels A and B. Panel A provides summary statistics on the dependent variables, whilst Panel B reports descriptive statistics for the covariates. All monetary variables are measured in constant prices deflated to 1997 prices. Conditional on reporting financial problems, the average number reported is 1.82, whilst conditional on having non-zero monthly housing costs, the last monthly payment is 5.919 log units, which is

[2] For example, French (2018) reports a relationship between the financial strain of individuals, their mental and general health status in the UK and the Royal Society for Public Health (2018) provides recent evidence of an association between ill health and debt.

approximately £517.65, see Table 1 Panel A. Around 38% of the sample saved in the previous year and the average monthly amount saved was 1.86 log units, which equates to £95.40. Approximately 65% of heads of household are males, 18% have a degree as their highest educational qualification, 20% experienced an adverse change in health, and 64% are employees, see Table 1 Panel B.

3. Methodology

The Bayesian estimator developed here allows for inflation at zero for both household financial problems and monthly housing costs, as well as examining the number of problems (conditional on facing financial hardship) and the level of housing costs (conditional on having a mortgage and/or paying rent), whilst also allowing for state dependence and interdependence between the outcomes. Our primary interest lies in the role that saving behaviour plays in terms of mitigating both the likelihood and extent of future financial problems.

Our key dependent variable, the number of financial problems, takes integer values from 0 to 6. Given the considerable inflation at zero, we use a zero-inflated Poisson model for modelling financial problems. The measure of monthly housing costs, on the other hand, is a continuous variable with a point mass at zero representing no mortgage or rent payments. Hence, we also develop a semi-continuous model for monthly housing payments. Furthermore, given the well-documented life cycle patterns associated with household finances, age may not have a linear relationship with the dependent variables. Hence, we model the relationship with the head of household's age as nonlinear spline effects. Finally, given the number of explanatory variables, we develop a shrinkage prior to account for the high dimensionality of the regression model. The rest of this section presents our Bayesian approach designed to account for the modelling issues summarised above.

3.1 Model Specification: A Semi-parametric Joint Model

Our joint model consists of three components, specifically: a semi-parametric Poisson hurdle mixed model for the number of financial problems, our key outcome variable of interest; a semi-parametric semi-continuous model for monthly housing costs; and, finally, a Dirichlet process (DP) for the joint distribution of the latent random effects from the Poisson hurdle and the semi-continuous models.

Modelling the number of financial problems – a zero-inflated Poisson model

Let Y_{ht}^{f} be the number of financial problems reported by the h^{th} household in the t^{th} year, $h = 1,2,\dots,N, t = 1,2,\dots,T$, where N represents the number of households in the sample, and T denotes the number of years. In the context of reported financial problems, a large number of zeros are observed in Y_{ht}^{f}. Following Lambert (1992), Hall (2000), Dagne (2004) and Ghosh et al. (2006), we further assume that for each observed event count, Y_{ht}^{f}, there is an unobserved random variable for the state of financial distress, U_{ht}, where $P(U_{ht} = 0) = p_{ht}^{f}$ if Y_{ht}^{f} comes from the degenerate distribution, and $P(U_{ht} = 1) = 1 - p_{ht}^{f}$ if Y_{ht}^{f}~Poisson (λ_{ht}):

$$Y_{ht}^{f} = \begin{cases} 0 & \text{with probability } p_{ht} \\ \text{Poisson}(\lambda_{ht}) & \text{with probability } (1 - p_{ht}) \end{cases} \tag{1}$$

where $\text{Poisson}(\lambda_{ht})$ is defined by the density function $P\left(Y_{ht}^{f} = y_{ht}^{f}\right) = \exp(-\lambda_{ht})\lambda_{ht}^{y_{ht}^{f}}/y_{ht}^{f}!$. It should be noted that both the degenerate distribution and the Poisson process can produce zero observations. Such a formulation is often referred to as the zero-inflated Poisson (ZIP) distribution. It then follows that

$$\Pr\left(Y_{ht}^{f} = 0\right) = p_{ht}^{f} + \left(1 - p_{ht}^{f}\right)\exp(-\lambda_{ht}) \tag{2}$$

$$\Pr\left(Y_{ht}^{f} = y_{ht}^{f}\right) = \left(1 - p_{ht}^{f}\right)\left\{\exp(-\lambda_{ht})\,\lambda_{ht}^{y_{ht}^{f}}/y_{ht}^{f}!\right\}, \quad y_{ht} = 1,2,\dots \tag{3}$$

One could conceptualize the degenerate distribution as representing a “no financial problem” state with probability, p_{ht}^{f}, while the Poisson process represents an “active financial problem” state with λ_{ht} being the mean annual number of financial problems.

Since the annual event counts are simultaneously influenced by the state that the household is in during the year and the annual event rate given that it is in an “active” state, we consider simultaneous modelling of both λ_{ht} and p_{ht}^{f}. We assume the following logistic and log-linear regression models for p_{ht}^{f} and λ_{ht} to accommodate covariates and random effects as follows:

$$Y_{ht}^{f} \sim \left(1 - p_{ht}^{f}\right) 1_{\left(Y_{ht}^{f}=0\right)} + p_{ht}^{f} \text{Poisson}(\lambda_{ht}) 1_{\left(Y_{ht}^{f} \geq 0\right)} \tag{4}$$

$$\text{logit}\left(p_{ht}^{f}\right) = \gamma_1 y_{h,t-1}^{f} + \varsigma_1 y_{h,t-1}^{m} + \psi_1 S_{h,t-1}^{A} + \boldsymbol{X}'_{ht} \beta_1 + \eta_1 CH_{ht} + g^{p}(\text{age}_{ht}) + b_{h1} \tag{5}$$

$$\log(\lambda_{ht}) = \gamma_2 y_{h,t-1}^{f} + \varsigma_2 y_{h,t-1}^{m} + \psi_2 S_{h,t-1}^{A} + \boldsymbol{X}'_{ht} \beta_2 + \eta_2 CH_{ht} + g^{\lambda}(\text{age}_{ht}) + b_{h2} \tag{6}$$

where γ_1, γ_2 are the autoregressive coefficients for the lag effect of order 1 of y_{ht}^{f} and ς_1, ς_2 are the autoregressive coefficients for the lag effect of order 1 of the other dependent variable, housing costs, y_{ht}^{m}, capturing interdependence. The inclusion of such lags is particularly important given the persistence in financial problems over time reported in the existing literature. Saving behaviour is lagged by a year and is represented by $S_{h,t-1}^{A}$ with associated parameters ψ_1 and ψ_2. The lag is introduced to explore whether savings insulate against future financial hardship. In addition, from a modelling perspective, this approach serves to reduce the potential for reverse causality, since, as argued by Angrist and Pischke (2009), savings predate the outcome variables. As stated above, we compare the protective role of saving using the incidence of saving and the amount saved. The covariates in $\boldsymbol{X}$ are as defined above and have the associated regression coefficients β_1 and β_2 in the respective equations for the incidence of financial problems and the number of financial problems. Whether the head of household experienced a change in their health is defined by a binary variable CH_{ht} with associated parameters η_1 and η_2. The b_{h1} and b_{h2} are the random effects of p_{ht}^{f} and λ_{ht}, respectively. We discuss the distribution of the random effects terms below.

Given that the life cycle effects of household finances have been long established, the effects of some covariates, viz., age_{ht}, on p_{ht}^{f} and λ_{ht}, may not be linear. Thus, the effects of the head of household's age are modelled by unspecified non-parametric functions $g^{p}(\text{age}_{ht})$ and $g^{\lambda}(\text{age}_{ht})$. These unknown smoothing functions reflect the nonlinear effects of this covariate. We approximate the spline function $g(\text{age}_{ht})$, suppressing the superscripts, by a piecewise polynomial of degree τ. The knots $\widetilde{\omega} = (\widetilde{\omega}_1, \widetilde{\omega}_2, \dots, \widetilde{\omega}_m)$ are placed within the range of age_{ht}, such that $\min(\text{age}_{ht}) < \widetilde{\omega}_1 < \widetilde{\omega}_2 < \cdots < \widetilde{\omega}_m < \max(\text{age}_{ht})$. Then $g(\text{age}_{ht})$ is approximated by

$$g(\text{age}_{ht}) = \nu_1 \text{age}_{ht} + \nu_2 \text{age}_{ht}^{2} + \cdots + \nu_\tau \text{age}_{ht}^{\tau} + \sum_{c=1}^{C} u_c \gamma_c (\text{age}_{ht} - \widetilde{\omega}_c)_{+}^{\tau} \tag{7}$$

where $X_+ = x$ if $x > 0$, and 0 otherwise, $\nu = (\nu_1, \dots, \nu_\tau)$, $\widetilde{\omega}$ are vectors of regression coefficients in the polynomial regression spline. Note that there is no intercept in the polynomial regression to avoid lack of identification. We assume $u_c \sim^{idd} N(0, \sigma_u^2)$; $h = 1, \dots, C$.

In the above formulation, one of the important issues is the choice of the number of knot points and where to locate them. Following Ruppert (2002) and Crainiceanu et al. (2005), we consider a number of knots that is large enough (typically 5 to 20) to ensure desired flexibility, and $\widetilde{\omega}_k$ denotes the sample quantiles of age$_{ht}$ corresponding to probability $k/(m+1)$, but the results hold for other choices of knots. In our empirical application, the function of age is modelled with m=20 knots chosen so that the k^{th} knot is the sample quantile of age corresponding to probability $k/(m+1)$. However, if there are too few knots or they are poorly located, estimates may be biased, while too many knots will inflate the local variance. Thus, to avoid overfitting, following Smith and Kohn (1996), we incorporate selector indices, γ_c, that allow the spline coefficients to be included or excluded and that are defined for each knot. The γ_c are then drawn independently from a Bernoulli prior, viz., $\gamma_c \sim \text{Bernoulli}(0.5)$. By introducing this, we can select a subset of well supported knots from a larger space. For each knot point u_c, the γ_c will weight the importance of a particular knot point. In the entire set-up, $\nu_1, \dots, \nu_\tau$, are the fixed effect regression parameters, and the u_c's are the random coefficients. The spline smoother corresponds to the optimal predictor in a mixed model framework assuming $u_c \sim^{idd} N(0, \sigma_u^2)$; $h = 1, \dots, C$.

Modelling monthly housing costs – a semi-continuous model

As stated above, although our primary focus lies in analysing the relationship between regular saving behaviour and future financial problems, given that housing costs arguably represent one of the most important financial commitments held by households, our modelling structure allows for the interdependence between financial problems and mortgage/rent payments. Hence, in this section, we present a semi-continuous model for longitudinal data relating to the amount of monthly housing costs. Since in some years the household may not hold a mortgage or pay rent and hence will make no monthly payments, this dependent variable is also characterised by a mixture of zero and positive continuous observations. To formulate a model for the housing cost amount, let Y_{ht}^{m} be the monthly housing cost comprising the mortgage and/or rental payments of household h at year t.

Let R_{ht} be a random variable which denotes incurring monthly housing costs where,

$$R_{ht} = \begin{cases} 0, & \text{if } Y_{ht}^{m} = 0 \\ 1, & \text{if } Y_{ht}^{m} > 0 \end{cases} \tag{8}$$

with conditional probabilities

$$\Pr(R_{ht} = r_{ht}) = \begin{cases} 1 - p_{ht}^{m}, & \text{if } r_{ht} = 0 \\ p_{ht}^{m}, & \text{if } r_{ht} = 1. \end{cases} \tag{9}$$

For such semi-continuous data, we introduce an analogous semi-continuous model consisting of a degenerate distribution at zero and a positive continuous distribution, such as a lognormal (LN), for the nonzero values as follows:

$$Y_{ht}^{m} \sim (1 - p_{ht}^{m})^{1-r_{ht}} \{p_{ht}^{m} \times N(\log(Y_{ht}^{m}); \mu_{ht}^{m}, \sigma^2)\}^{r_{ht}} \tag{10}$$

$$\text{logit}(p_{ht}^{m}) = \gamma_3 y_{h,t-1}^{m} + \varsigma_3 y_{h,t-1}^{f} + \psi_3 S_{h,t-1}^{A} + \boldsymbol{X}'_{ht}\beta_3 + h^{p}(\text{age}_{ht}) + b_{h3} \tag{11}$$

$$\mu_{ht} = \gamma_4 y_{h,t-1}^{m} + \varsigma_4 y_{h,t-1}^{f} + \psi_4 S_{h,t-1}^{A} + \boldsymbol{X}'_{ht}\beta_4 + h^{\mu}(\text{age}_{ht}) + b_{h4} \tag{12}$$

where, r_{ht} is an indicator as defined above, μ_{ht}^{m} and σ^2 are the mean and variance of $\log(Y_{ht}^{m})$, respectively. The model given by equations (11, 12) is a semi-parametric counterpart of the correlated two-part model proposed for modelling financial problems. Saving behaviour, $S_{h,t-1}^{A}$, is included as a lag for the aforementioned reasons.

Correlation structure and heterogeneity – joining the models

The models detailed above contain information about household behaviour and are, therefore, inter-related. To obtain the complete picture and to account for heterogeneity across households, we combine these effects by correlating the multiple outcomes. However, since these outcomes are measured on a variety of different scales (viz., binary, Poisson, log-normal), it is not possible to directly model the joint predictors' effects due to the lack of any natural multivariate distribution for characterising such dependency. A flexible solution is to model the association between the different responses by correlating the random heterogeneous effects from each response. In our joint modelling approach, random effects are assumed for each response process and the different processes are associated by imposing a joint multivariate distribution on the random effects. Such a model not only provides a covariance structure to assess the strength of association between the responses, but also borrows information across the outcomes and offers an intuitive way of describing the dependency between the responses.

Let $\boldsymbol{b}_h = (b_{h1}, b_{h2}, b_{h3}, b_{h4})'$ be the vector representing the random effects associated with the h^{th} household. Typically, a parametric normal distribution is considered for $\boldsymbol{b}_h$. However, the choice of normality is often due to computational tractability, an assumption which may not always hold in reality. In addition, it provides limited flexibility because it is unimodal. This may result in misleading inferences relating to the magnitude of effects and the nature of heterogeneity. One common approach entails using a finite mixture of normal distributions as an alternative choice. However, rather than handling the very large number of parameters resulting from finite mixture models with a large number of mixands, it may be more straightforward to work with an infinite dimensional specification by assuming a random mixing distribution which is not restricted to a specific parametric family. Following Li and Ansari (2014), we propose here an enriched class of models that can capture heterogeneity in a flexible yet structured manner. In the context of the proposed class of models, an unknown distribution G of the random effects is assumed to be random and a DP is placed on the distribution of G. Then, the model for $\boldsymbol{b}_h$ can be written as

$$\boldsymbol{b}_h \sim G, \quad G \sim \mathrm{DP}(\alpha G_0) \tag{13}$$

where α is a positive scalar precision parameter and G_0 is a parametric baseline distribution. With such a non-parametric modelling of the random effects, the entire model turns out to be a semi-parametric model. We assume a multivariate normal distribution for G_0, i.e. $G_0 \sim \boldsymbol{N}(\boldsymbol{0}, \Sigma)$. Realisations of the DP are discrete with probability one, implying that the estimated $\boldsymbol{b}_h$ that will be drawn from G will be grouped into a cluster, thus allowing for possible multimodality in the distribution of $\boldsymbol{b}_h$. The discrete nature of the DP is apparent from the popular stick-breaking formulation pioneered by Sethuraman (1994). The stick-breaking formulation implies that $G \sim DP(\alpha G_0)$ is equivalent to

$$G = \sum_{q=1}^{\infty} \pi_q^D \delta_{\boldsymbol{b}_q}, \quad \boldsymbol{b}_q \sim G_0, \quad \text{and} \quad \sum_{q=1}^{\infty} \pi_q^D = 1 \tag{14}$$

where G is a mixture of countably but infinite atoms, and these atoms are drawn independently from the base distribution G_0, and $\delta_{\boldsymbol{b}}$ is a point mass at $\boldsymbol{b}$. An atom is like a cluster (i.e. a sub-group of random effects), $\boldsymbol{b}_q$ is the value of that cluster and all random effects in a cluster share the same $\boldsymbol{b}_q$. In equation 14, $\pi_q^D = V_h \prod_{l<q}(1 - V_l)$, which is formulated from a stick-breaking process, with $V_q \sim \text{Beta}(1, \alpha)$, is the probability assigned to the q^{th} cluster. For small values of α, $V_q \to 1$ and thus $\pi_q^D \to 1$, assigning all probability weight to a few clusters and thus the G is far from G_0. On the contrary, for large values of α, the number of clusters can be as many as the number of random effects implying that the sampled distribution of G is close to the base distribution of G_0. For practicality, researchers use a finite truncation to approximate G, i.e. $G \sim \sum_{q=1}^{Q} \pi_q^D \delta_{\boldsymbol{b}_q}$.

While the above formulation appears appropriate, there is an issue of identifiability within it in the sense that, although the prior expectation of the mean of G is 0, the posterior expectation can be non-zero and, thus, can bias inference (Yang et al., 2010; Li et al., 2011). In parametric hierarchical models, it is standard practice to place a mean constraint on the latent variable distribution for the sake of identifiability and interpretability. In a nonparametric DP, Yang et al. (2010) proposed using an entered DP to tackle the identifiability issue. Li et al. (2011) have shown the utility of an entered DP in modelling heterogeneity in choice models. Following Yang et al. (2010) and Li et al. (2011), we centre the DP to have zero mean. We estimate the mean and variance of the process, i.e., μ_G^j and Σ_G^j, at the j^{th} Bayesian Markov Chain Monte Carlo (MCMC) iteration as follows

$$\mu_G^j = \sum_{q=1}^{Q} V_q^j \prod_{l<q}\left(1 - V_l^j\right) \boldsymbol{b}_q^j \tag{15}$$

$$\Sigma_G^j = \sum_{q=1}^{Q} V_q^j \prod_{l<q}\left(1 - V_l^j\right)\left(\boldsymbol{b}_q^j - \mu_G^j\right)\left(\boldsymbol{b}_q^j - \mu_G^j\right)' \tag{16}$$

where V_q^j and $\boldsymbol{b}_q^j$ are the posterior samples from the uncentered process defined in equation 14 and $\left(\boldsymbol{b}_q^j - \mu_G^j\right)$ is the centered estimate for random effects at the j^{th} iteration. The above entered DP implies that $\text{E}(\boldsymbol{b}_h | G = 0)$ and $\text{Var}(\boldsymbol{b}_h | G = \Sigma_G)$.

3.2 Bayesian Methods

Under the joint model described by equations 4, 5, 6, 8, 10, 11 and 12, the likelihood of the observed data for the h^{th} household, denoted by $\boldsymbol{Y}_{h1},\dots,\boldsymbol{Y}_{hN}$, with $\boldsymbol{Y}_{ht}=\left(Y_{ht}^{f},Y_{ht}^{m}\right)'$ for $t=1,\dots,T$, based on the parameter set Ω and the random effects $\boldsymbol{b}_h$ is proportional to

$$L_i(\Omega,\boldsymbol{b}_h|\boldsymbol{Y}_{h1},\dots,\boldsymbol{Y}_{hT})=\prod_{t=1}^{T}\left[\left(1-p_{ht}^{f}\right)\right]^{I\left[y_{ht}^{f}=0\right]}\times\left[\frac{p_{ht}^{f}{\mu_{ht}^{f}}^{y_{ht}^{f}}e^{-\mu_{ht}^{f}}}{y_{ht}^{f}!\left(1-e^{-\mu_{ht}^{f}}\right)}\right]^{1-I\left[y_{ht}^{f}=0\right]}$$

$$\times\left(1-p_{ht}^{m}\right)^{1-r_{ht}}\{p_{ht}^{m}\times\mathrm{LN}(y_{ht}^{m};\mu_{ht}^{m};\sigma^2)\}^{r_{ht}}\times f(\boldsymbol{b}_h) \qquad (17)$$

To complete the Bayesian specification of the model, we assign priors to the unknown parameters in the above likelihood function. For the regression coefficients $\beta_1,\dots,\beta_4,\eta_1,\eta_2,\psi_1,\dots,\psi_4$, we assume shrinkage priors. We have a large number of covariates and, thus, a shrinkage prior will be beneficial. We use a LASSO prior on these sets of parameters. Suppressing the subscripts and assuming that each coefficient is a vector of order $k\times 1$, ϕ_k, and where the shrinkage parameters are denoted by the τ's, we use a LASSO prior as follows:

$$\phi_k\big|\sigma^2,\tau_1^2,\dots,\tau_p^2\sim N_p(0,\sigma^2\boldsymbol{D}_\tau) \qquad (18a)$$

$$\text{where } \boldsymbol{D}_\tau=\mathrm{diag}(\tau_1^2,\dots,\tau_{P'}^2) \qquad (18b)$$

$$\tau_1^2,\dots,\tau_{P'}^2\sim\prod_{p=1}^{P'}\frac{\lambda^2}{2}\exp\left(-\frac{1}{2}\lambda\tau_p^2\right) \qquad (19)$$

$$\lambda^2\sim\mathrm{Gamma}(a,b) \qquad (20)$$

$$\sigma^2\sim\pi(\sigma^2)=\frac{1}{\sigma^2} \qquad (21)$$

For the rest of the regression parameters, we assume a normal prior and the spline coefficients (ν) are also assigned a normal density prior. For each variance parameter, we assume an inverse-gamma (IG) prior and for the variance-covariance matrix in the baseline distribution of G, we assume an inverse Wishart prior. Finally, for the total mass α of the DP, we assume a uniform distribution.

4. Results

The protective role of saving

In this section, the results from estimating the model, in particular the estimated parameters in equations (5) and (6) and equations (11) and (12) are discussed. Our key focus is on: (i) whether saving acts as a buffer against future financial problems, i.e. focusing on the ψ's, a priori, we expect saving to have a protective role against future hardship, hence $\psi_1, \psi_2 < 0$; (ii) whether state dependence is apparent in observed financial problems, where the key parameters of interest are the γ's; (iii) finally, whether there is interdependence between financial problems and housing costs, where the parameters of interest are the ς's.

The results from estimating the model detailed in Section 3 are presented in Tables 2 and 3. Table 2 shows the correlations in the unobservable effects across the equations, i.e. the variance – covariance matrix. Where statistically significant, both the variance and covariance terms are positive. For example, positive correlations are found to exist in the unobservable effects between the incidence of financial problems and housing costs. The findings of interdependence across the different parts of the empirical model support the joint modelling framework, as ignoring such effects would result in less efficient estimates.

Table 3 provides Bayesian posterior mean estimates (BPMEs). The first three rows of Table 3 show the key parameter estimates of interest, i.e. those BPMEs associated with: the role of saving, the ψ's; dynamics, the γ's; and interdependence across equations for each of the outcomes, the ς's. Each panel of Table 3 is split into four columns: the first two columns relate to financial problems, our primary outcome of interest, namely the probability of being in financial hardship and the number of problems reported; and the final two columns show the estimates for housing costs, namely the probability of incurring housing costs and the amount of the monthly payments. In addition to identifying correlation in the unobservables, the flexibility of the two-part process is also evident when comparing the influence of the explanatory variables across the binary and the non-binary parts of the model, where in what follows it can be seen that some explanatory variables exert different influences across the two parts – in terms of statistical significance, magnitude and sign.

Focusing initially on the key parameters, ψ's, i.e. whether past saving behaviour plays a protective role against currently experiencing financial problems, it is apparent that the parameters on whether the head of household saved in the previous year are negative, i.e. $\hat{\psi}_1, \hat{\psi}_2 < 0$. For example, having saved in the previous year is associated with a 42 percentage point lower probability of currently having a financial problem, i.e. the 'Odds Ratio' OR= $\exp(\hat{\psi}_1) = \exp(-0.540) = 0.58$, and reduces the number of financial problems by approximately 28 percentage points, e.g. OR= $\exp(\hat{\psi}_2) = \exp(-0.325) = 0.72$. Hence, the act of saving regardless of the amount put aside serves to mitigate future financial hardship, hence acting as a financial buffer. These findings are consistent with the existing international literature which has revealed a protective role of savings against financial hardship, e.g. Collins and Gjertson (2013), Mills and Amick (2010), both for the US, Giardi (2013) for Italy, and the study by Le Blanc et al. (2016) which revealed that a key motive for saving in European countries was for unexpected events. In contrast to existing studies, our modelling framework separates each outcome into a two-part process, i.e. the probability of having a financial problem and the number of financial problems experienced, revealing that saving has a large influence on both the incidence and extent of future financial problems. In contrast, saving is unrelated to both the incidence and the extent of future housing costs.

With respect to financial problems, there is also evidence of positive state dependence, which is consistent with findings in the existing literature, e.g. Giardi (2013) and Brown at al (2014). The 'Odds Ratio' shows that households, which experienced financial hardship in the previous year, are nearly three times as likely to currently report a financial problem, i.e. OR= $\exp(\hat{\gamma}_1) = \exp(1.010) = 2.75$. Similarly, there is also evidence of positive state dependence in the number of financial problems experienced. Turning to housing costs, again there is evidence of state dependence, where a 1% increase in housing costs in the previous year is associated with around a 3 percentage point increase in housing costs (i.e. OR= $\exp(\hat{\gamma}_4) = \exp(0.030) = 1.03$), which is consistent with existing evidence, e.g. Burrows (1997). Households which experienced financial problems in the previous year have higher levels of monthly housing costs, i.e. $\hat{\varsigma}_4 > 0$. Having had housing costs in the previous year is unrelated to the extent of financial hardship. This finding might reflect a housing tenure effect in that those who own a home may face fewer financial problems due to the wealth effect associated with home ownership, e.g. Taylor (2011) and Burrows (2018).[3]

Figures 3 and 4 show the effects of the head of household's age, illustrated by spline function graphs of age on each outcome. The shaded grey area represents the 95 percent credible interval. Figure 3A shows the association between the head of household's age and the probability of reporting a financial problem, and Figure 3B reveals the relationship between age and the number of problems reported at the household level. Financial problems have been found to be more prevalent for those under 30 compared to other age groups in the existing literature, e.g. Atkinson et al. (2006), which is consistent with the results shown in Figure 3A, where the probability of experiencing a financial problem increases up until around age 25, it then decreases monotonically with the head of household's age. The head of household's age also has a significant effect on the number of financial problems reported at the household level, as can be seen from Figure 3B. Clear life cycle effects are evident where the association between the head of household's age and the number of problems experienced increases monotonically until age 45 and then decreases. Figure 4 reveals that life cycle effects also exist for housing costs. The probability of a household incurring housing costs increases with the head of household's age up to just before age 40 and then decreases monotonically, see Figure 4A, whilst the level of the monthly housing payments decreases monotonically with the head of household's age after age 40. The results endorse the importance of allowing for the non-linear effects of age on the outcomes, where the spline function reveals evidence of life cycle effects.

[3] This finding is confirmed when housing costs are modelling as comprising solely of mortgage repayments, i.e. excluding rent.

We briefly comment on the other control variables reported in Table 3. Larger households have a higher probability of experiencing financial problems, whilst households with male heads have a higher likelihood of having housing costs than their female counterparts but are less likely to experience financial problems. This latter finding is consistent with the existing literature, e.g. Brown et al. (2014) for the UK, Gjertson (2016) for the US and Giarda (2013) for Italy. Households with a white head have a lower probability of reporting financial problems. In terms of educational attainment, those heads of household who have obtained a degree as their highest qualification not only have a lower likelihood of experiencing financial problems but they also have fewer problems compared to those with no qualifications. For example, if the head of household has a degree, the probability of having financial problems is reduced by approximately 26 percentage points, compared to those with no qualifications, i.e. $\exp(\hat{\beta}_{1k}) = \exp(-0.306) = 0.74$. Such findings may reflect the possibility that highly educated heads of household are likely to be more financially literate and capable of managing their household finances, see Lusardi and Mitchell (2014).

With respect to labour market status, the relative probability of an unemployed head of household having financial problems is around 137 percentage points higher compared to a household with a head who is out of the labour market, given the OR= $\exp(\hat{\beta}_{1k}) = \exp(0.836) = 2.37$. Households with an employed or self-employed head experience higher monthly housing costs. A 1% increase in real equivalized monthly income is associated with a decrease in the number of financial problems by 8 percentage points, i.e. OR= $\exp(\hat{\beta}_{2k}) = \exp(-0.076) = 0.92$, and an increase in the level of housing costs by 7 percentage points, i.e. OR= $\exp(\hat{\beta}_{4k}) = \exp(0.067) = 1.07$.

We also condition the outcomes on household expenditure on utilities and non-durable goods. Both higher utility costs and expenditure on non-durable goods such as food are positively associated with the likelihood of experiencing financial problems, which is consistent with prior expectations. For example, a 1% increase in annual utility costs is associated with a 6 percentage point increase in the probability of experiencing financial hardship, given the OR= $\exp(\hat{\beta}_{1k}) =$

$\exp(0.056) = 1.06$. The results show that households with a head who has experienced a change in health have a higher probability of facing financial problems and experience more problems. For example, a deterioration in health increases the probability of having a financial problem by 15 percentage points, i.e. OR= $\exp(\hat{\eta}_{1k}) = \exp(0.138) = 1.15$.

In Table 3 Panel B, we present the results associated with regional and business cycle effects, where for the former London is the reference category and for the latter pre-2001 is the omitted period. There are generally no significant differences across regions for either the incidence or the extent of financial hardship, with the exception that households in Wales and the North East have a higher probability of experiencing financial problems than those living in London. Focusing on housing costs, there is heterogeneity across regions in terms of the incidence and the amount of monthly payments. For example, all regions not only have a lower likelihood of incurring housing costs but also generally the monthly costs are lower compared to London (the exceptions are for the East of England and the South East). The business cycle effects are interesting, in that for financial problems, after the recent financial crisis period both the incidence and extent of household financial hardship increased. For example, in 2012 a head of household was, OR= $\exp(\hat{\beta}_{1k}) = \exp(0.125) = 1.13$, around 13 percentage points more likely to experience a financial problem compared to pre-2001, ceteris paribus. Moreover, throughout the sample period, the time effects are statistically significant. Prior to the financial crisis compared to pre-2001, the number of problems fell each year, whilst for the most recent years, post-2010, the number of problems has increased (this is especially noticeable in 2012).

Table 4 presents the results of estimating the alternative specification (model 2) where the incidence of saving is replaced by the amount saved in the previous year. For brevity, we only report the key parameters of interest, i.e. those associated with savings behaviour (the ψ's), dynamics (the γ's) and interdependence (the ς's). The influence of the amount saved on the incidence and extent of financial problems is similar to that of model 1. A 1% increase in the amount saved is associated with a 12 percentage point lower probability of experiencing a financial problem, i.e. OR= $\exp(\hat{\psi}_1) = \exp(-0.132) = 0.88$, and reduces the number of financial problems by 8 percentage points, OR= $\exp(\hat{\psi}_2) = \exp(-0.084) = 0.92$. Hence, these findings further endorse the existence of a protective role of saving in mitigating future financial hardship.

The protective role of saving – an IV approach

In this subsection, we explore the robustness of our findings to using an alternative measure of saving. Specifically, to allow for the potential endogeneity of saving, we incorporate the fitted values of saving into the model, where savings are instrumented using information on the saving behaviour of the head of household as a child. Thus, the remaining analysis focuses on a sub-sample of 1,299 heads of household who are aged between 18 and 33. The approach follows Brown and Taylor (2016) and uses information recorded in the *Youth Survey*, which asks children aged 11-15 '*what do you usually do with your money?*' The possible responses were: *save to buy things*; *save and not spend*; and *spend immediately*. From the responses to this question, a binary indicator, S_h^C, is created, which shows whether the individual saved as a child. Saving as a child has been found to be a strong predictor of saving behaviour as an adult, e.g. Knowles and Postlewaite (2004) and Cronqvist and Siegel (2015). Hence, the sub-sample comprises relatively young adults as our estimation approach requires observing the head of household as a child and as an adult. This age group is particularly interesting given that the results shown in Figure 3A revealed that the incidence of financial problems increased with age for young heads aged below 30. Moreover, in the UK, financial problems are typically more prevalent amongst the young, see Kempson et al. (2004), Atkinson et al. (2006), Taylor (2011) and Brown et al. (2014). In addition, the UK House of Lords Select Committee on Financial Exclusion (2017) reports that young people are more susceptible to financial exclusion and that 51% of 18-24 year olds are worried about money on a regular basis and 1 in 5 individuals in this age group have experienced financial problems as a result of poor credit ratings.

To model saving behaviour, we use a two-stage least squares (2SLS) approach as follows: in the first stage we model saving behaviour during childhood, as this may be endogenous if included directly as a control for adult saving, $S_h^C = 1\left[\boldsymbol{Z'}_h\boldsymbol{\phi}_1 + \mathbf{EXP}_h^{P'}\boldsymbol{\pi}_1 + \nu_{1h} > 0\right]$; and in the second stage the saving behaviour of adults is modelled, $S_h^A = 1[\boldsymbol{Z'}_h\boldsymbol{\phi}_2 + \psi_1 S_h^C + \nu_{2h} > 0]$. Where S_h^A is either a binary indicator (i.e. whether they saved as an adult in the previous period) or the natural logarithm of the amount of monthly savings in the previous period. The vector of controls, $\boldsymbol{Z}_h$, includes permanent income (constructed following the approach of Kazarosian, 1997) and its volatility, and $\mathbf{EXP}_h^P$ is a vector of the financial expectations of the child's parent (who is the head of household) as the literature has found expectations to be related to saving behaviour, e.g. Souleles (2004), Brown and Taylor (2006), Puri and Robinson (2007) and Gerhard et al. (2018). From the 2SLS analysis, we obtain the fitted values of savings (either the incidence or amount), denoted by $\hat{S}_{h,t-1}^A$. The results from modelling savings behaviour, shown in Table A1 in the appendix, reveal that the probability and level of savings are positively associated with: whether the individual saved during childhood; educational attainment; permanent income and its volatility, which is consistent with existing evidence in the literature, e.g. Guariglia (2001). Moreover, the financial expectations of their parent, in particular financial pessimism, is a valid instrument of having saved as a child, and is positively related to whether the individual saved as a child.[4]

The results focusing on the sub-sample of young adults are shown in Table 5, where for brevity, we only report the key parameters of interest, i.e. those associated with savings behaviour (the ψ's), dynamics (the γ's) and interdependence (the ς's). Panels A through to D report the BPMEs for models 3 to 6 respectively. Models 3 and 4 shown in Panels A and B replicate the analysis of Tables 3 and 4 for the young adult sample, whilst in Panels C and D the results are based on instrumenting the incidence and the amount saved, respectively. Clearly, throughout each panel, the dynamic effects and interdependence between financial problems and housing costs are very similar in terms of magnitude of the BPMEs to that of models 1 and 2, shown in Tables 3 and 4.

[4] The instruments seem plausible from a theoretical point of view in that there is no obvious reason why the financial expectations of the parent, measured *ex ante*, should influence the current saving behaviour of their offspring now observed as young adults. Moreover, the instruments pass the Kleibergen-Paap (2006) test of under-identification, the weak instrument test of Stock et al. (2002) and Stock and Yogo (2005), and, in accordance with the exclusion restriction, the instruments are statistically insignificant in the adult saving equation.

The protective role of savings in mitigating the likelihood of future financial problems and the extent of such hardship is also evident for this sub-sample of young adults, see Table 5 Panels A to D, in that $\hat{\psi}_1, \hat{\psi}_2 < 0$. The results where saving behaviour is treated exogenously, show that the incidence of past saving, regardless of the amount, reduces both the probability of having a financial problem by 34 percentage points (OR= $\exp(\hat{\psi}_1) = \exp(-0.417) = 0.66$) and the number of financial problems by 18 percentage points (OR= $\exp(\hat{\psi}_2) = \exp(-0.197) = 0.82$), see Table 5 Panel A. Hence, we find further evidence that saving serves to mitigate future financial hardship. These effects remain when the likelihood of saving is instrumented, as can be seen from Table 5 Panel C, although the magnitudes fall to 32 and 7 percentage points, respectively. Consistent with the results of model 1 shown in Table 3, past saving behaviour has a larger effect on reducing the incidence of financial hardship than on the number of financial problems. Replacing the incidence of saving with the amount saved again reveals very similar results to the full sample, i.e. comparing Table 5, model 5, Panel C to Table 4 model 2, and this finding is robust to instrumenting the amount saved as can be seen from Table 5, model 6, Panel D.

Figures 5A and 5B show the effects of the head of household's age, illustrated by spline function graphs of age on the incidence and number of financial problems.[5] The shaded grey area represents the 95 percent credible interval. In contrast to the analysis of the full sample, the probability of experiencing financial problems increases monotonically with the age of the head of household, see Figure 5A. Conversely, whilst the head of household's age has a significant effect on the number of financial problems reported at the household level, as can be seen from Figure 5B, the effects are very similar for each age and are small in terms of magnitude (with a BPME of around 0.05) at less than 1 percentage point per year.[6]

[5] For brevity, we do not show the corresponding figures for housing costs.

[6] Interestingly, this group of young household heads appears to have been more adversely affected by the financial crisis, in that the probability of experiencing a financial problem is higher after 2008 and is larger in magnitude than for the full sample of adults (Table 3 Panel B). For example, in 2010 the probability of having financial problems for 18 to 33 year olds was nearly twice that of pre-2001, i.e. OR $= \exp(\hat{\beta}_{1k}) = \exp(0.581) = 1.79$.

Summary of Bayesian model performance and comparison to a classical ZIP estimator

In order to assess the performance of the estimated models, a number of summary statistics are analysed, specifically: (i) the deviance information criteria (DIC) proposed by Spiegelhalter et al. (2002), where a smaller DIC value denotes a better model;[7] and (ii) the log-pseudo marginal likelihood (LPML), with a larger LPML value denoting a superior specification.[8] Table 6 reports the DIC and LPML statistics for each of the estimated models. For the sub-sample of young adults, both the DIC and LPML statistics support those specifications where saving behaviour is instrumented (models 5 and 6). For both the full sample of adults and the sub-sample of young adults, the preferred specifications are where we condition on the propensity to save (models 1, 3 and 5), rather than the amount saved (models 2, 4 and 6).

Finally, to place our modelling contributions into context, we have contrasted the results from the joint Bayesian estimator with potential classical alternatives, although it should be acknowledged that there is no classical estimator that can accommodate multiple joint outcomes and two-part processes along with individual effects. Given that financial hardship is our primary outcome of interest, the classical estimator we compare our findings to is a Zero Inflated Poisson (ZIP) estimator.[9] This classical estimator allows covariates to have different effects on the incidence of financial problems and the number of financial problems indicating the extent of financial hardship. The results of the univariate ZIP analysis are given in Table A2 in the appendix where standard errors are clustered at the individual level as it is not possible to incorporate individual random effects, which is a disadvantage of the classical univariate ZIP approach. Two models are estimated, one including whether the individual saved on a monthly basis in the previous year and the second model conditioning on the amount saved on a monthly basis. Coefficients are reported

[7] Note that defining $\boldsymbol{D} = (Y^f, Y^m)$ to be the observed data and ϕ as the parameters of the model given the latent variable $\boldsymbol{b}$, $p(\boldsymbol{D}|\phi)$ is not a closed form. Hence, we follow the approach in Jiang et al. (2015) and Celeux et al. (2006), and calculate DIC($\boldsymbol{D}$), by first considering the DIC measure with "complete data" with $\boldsymbol{b}$ and then integrate out the observed $\boldsymbol{b}$.

[8] This is a summary measure of the conditional predictive ordinate (CPO), see Gelfand et al. (1992) and Jiang et al. (2015), which has been widely used for model diagnostics and assessment to determine the best model.

[9] Another potential alternative would be a multivariate dynamic panel Tobit model, see Wooldridge (2005). However, such an estimator is less flexible than the approach we develop here in that it is not possible to disentangle the incidence and the level of financial problems. Consequently, this approach would mask whether the effect of particular covariates is operating through the discrete or continuous part of the distribution, as well as being unsuitable for count data.

throughout, although it should be noted that these are not directly comparable to the Bayesian marginal effects. However, we can compare the sign and statistical significance of the covariates between the Bayesian and classical ZIP estimators.

It is reassuring to find that the results from the ZIP estimator reveal a protective role of saving, which is consistent with the results from the new Bayesian framework. However, there are also some counterintuitive results in that past financial problems and housing costs are both associated with a reduction in the incidence of experiencing financial problems, which is contrary to the findings from the Bayesian analysis. This suggests firstly that it is important to jointly estimate the outcomes, which is perhaps not surprising given that housing costs are an important financial commitment for many households. This is further supported by the positive correlations found in the variance-covariance matrix, see Table 2. Secondly, the difference between the sets of results highlights the importance of incorporating individual heterogeneity into the estimation framework by allowing for random effects.[10]

5. Conclusion

Our findings suggest that savings provide a financial buffer in the event of future hardship and are consistent with evidence from the US, which has generally been based on non-representative samples of low income families. In addition to contributing to the existing literature by exploring British panel data, we have made by a methodological contribution by developing a flexible Bayesian framework to examine the two-part process behind financial hardship, specifically the incidence and extent of financial problems, as well as allowing for the two-part process behind important financial commitments such as mortgage debt or monthly rent. Our modelling approach, which allows for correlated random effects, identifies interdependence between financial hardship and housing costs and between each of the associated two-part processes. The analysis also allows

[10] Similar results are found when the univariate ZIP model is applied to the sub-sample of young adults aged 18-33. These results are available on request. In particular, the protective role of saving remains when replaced with the fitted values derived from the IV analysis.

for persistence over time in financial problems revealing clear evidence of dynamic effects and the existence of interdependence between the outcomes.

Our findings relate to the widespread concern amongst policymakers in a number of countries regarding the relatively low levels of household saving. A protective role of saving is also found to exist for a sub-sample of young household heads and this is an important finding given the evidence from the UK House of Lords Select Committee on Financial Exclusion (2017) indicating that young adults are more likely to face financial exclusion. Our analysis also highlights the need to enhance financial literacy and promote the importance of 'putting money aside'. Indeed, influencing saving behaviour during childhood, i.e. in the formative years, may ultimately help to reduce the prevailing levels of financial vulnerability and stress experienced by households later in the life cycle.

References

Angrist, J. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricists Companion*. Princeton: Princeton University Press.

Atkinson, A., McKay, S., Kempson, E. and Collard, S. (2006). Levels of Financial Capability in the UK: Results of a Baseline Survey. FSA Consumer Research Paper 47. FSA: London. http://financialhealthexchange.org.uk/wp-content/uploads/2015/11/Levels-of-financial-capability-in-the-UK-Results-of-a-basline-survey.pdf

Brown, S. and Taylor, K. (2006). Financial Expectations, Consumption and Saving: A Microeconomic Analysis. *Fiscal Studies*, 27, 313-38.

Brown, S. and Taylor, K. (2016). Early Influences on Saving Behaviour: Analysis of British Panel Data. *Journal of Banking and Finance*, 62, 1-14.

Brown, S., Ghosh, P. and Taylor, K. (2014). The Existence and Persistence of Household Financial Hardship. *Journal of Banking and Finance*, 46, 285-98.

Brown, S., Ghosh, P. and Taylor, K. (2016). Household Finances and Social Interaction: Bayesian Analysis of Household Panel Data. *Review of Income and Wealth*, 62, 467-88.

Brown, S., Ghosh, P., Su, L. and Taylor, K. (2015). Modelling Household Finances: A Bayesian Approach to a Multivariate Two-Part Model. *Journal of Empirical Finance*, 33, 190-207.

Browning, M. and Crossley, T. (2001). The Life Cycle Model of Consumption and Saving. *Journal of Economic Perspectives*, 15, 3-22.

Browning, M. and Lusardi, A. (1996). Household Saving: Micro Theories and Micro Facts. *Journal of Economic Literature*, 24, 1797-855.

Burrows, R. (1997). Who Needs a Safety-Net? The Social Distribution of Mortgage Arrears in England. *Housing Finance*, 34, 17-24.

Burrows, V. (2018). The Impact of House Prices on Consumption in the UK: A New Perspective. *Economica*, 85, 92-123.

Celeux, G., Forbes, F., Robert, C. P., and Titterington, D. M. (2006). Deviance Information Criteria for Missing Data Models. *Bayesian Analysis*, 1, 651-73.

Collins, J. M., and Gjertson, L. (2013). Emergency Savings for Low Income Consumers. *Focus*, 30(1). https://www.irp.wisc.edu/publications/focus/pdfs/foc301c.pdf.

Crainiceanu, C. M., Ruppert, D., and Wand, M. P. (2005). Bayesian Analysis for Penalized Spline Regression using WinBUGS. *Journal of Statistical Software*, 14, 1-24.

Cronqvist, H. and Siegel, S. (2015). The Origins of Savings Behavior. *Journal of Political Economy*, 123, 123-69.

Crossley, T., Emmerson, C. and Leicester, A. (2012). *Raising Household Saving*. British Academy Policy Centre Publication. http://www.britac.ac.uk/policy/Raising-household-saving.cfm.

Dagne, G. A. (2004). Hierarchical Bayesian Analysis of Correlated Zero-Inflated Count Data. *Biometrical Journal*, 46, 653-63.

Despard, M. R., Grinstein-Weiss, M., Ren, C., Guo, S. and Raghavan, R. (2016). Effects of a Tax-Time Savings Intervention on Use of Alternative Financial Services among Lower-Income Households. *Journal of Consumer Affairs*, 1-25.

French, D. (2018). Financial strain in the United Kingdom. *Oxford Economic Papers*, 70, 163-82.

Garon, S. (2012). *Beyond Our Means: Why America Spends While the World Saves*. Princeton University Press.

Gelfand, A. E., Dey, D. K., and Chang, H. (1992). Model Determination using Predictive Distributions with Implementation via Sampling-based Methods (No. TR-462). *Stanford University Department of Statistics*.

Gerhard, P., Gladstone, J. J., and Hoffmann, A. (2018). Psychological Characteristics and Household Savings Behavior: The Importance of Accounting for Latent Heterogeneity. *Journal of Economic Behavior and Organization*, 148, 66-82.

Ghosh, S. K., Mukhopadhyay, P., and Lu, J. C. J. (2006). Bayesian Analysis of Zero-Inflated Regression Models. *Journal of Statistical planning and Inference*, 136, 1360-75.

Giarda, E. (2013). Persistency of Financial Distress amongst Italian Households: Evidence from Dynamic Models for Binary Panel Data. *Journal of Banking and Finance*, 37, 3425-34.

Gjertson, L. (2016). Emergency Saving and Financial Hardship. *Journal of Family and Economic Issues*, 37, 1-17.

Guariglia, A. (2001). Saving Behaviour and Earnings Uncertainty: Evidence from the British Household Panel Survey. *Journal of Population Economics*, 14, 619-34.

Guariglia, A. and Rossi, M. (2004). Private Medical Insurance and Saving: Evidence from the British Household Panel Survey. *Journal of Health Economics*. 23, 761-83.

Hall, D. B. (2000). Zero-Inflated Poisson and Binomial Regression with Random Effects: A Case Study. *Biometrics*, 56, 1030-39.

Jiang, B., Wang, N., Sammel, M. D., and Elliott, M. R. (2015). Modelling Short-and Long-Term Characteristics of Follicle Stimulating Hormone as Predictors of Severe Hot Flashes in the Penn Ovarian Aging Study. *Journal of the Royal Statistical Society*, *Series C*, 64, 731-53.

Kazarosian, M. (1997). Precautionary Savings: A Panel Study. *Review of Economics and Statistics*, 79, 241-47.

Kempson, E., McKay, S. and Willets, M. (2004). Characteristics of Families in Debt and the Nature of Indebtedness. Department of Work and Pensions Research Report 211.

Kennickell, A., and Lusardi, A. (2005). Disentangling the Importance of the Precautionary Saving Motive. *CFS Working Paper* No. 2006/15.

Kleibergen, F. and Paap, R. (2006). Generalized Reduced Rank Tests using the Singular Value Decomposition. *Journal of Econometrics*, 133, 97-126.

Knowles, J. and Postlewaite, A. (2004). Do Children Learn to Save from their Parents? University of Pennsylvania, Mimeo.

Lambert, D. (1992). Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics*, 34, 1-14.

Le Blanc, J., Porpiglia, A., Teppa, F., Zhu, J. and Ziegelmeyer, M. (2016). Household Saving Behaviour in the Euro Area. *International Journal of Central Banking*, 12, 15-69.

Li, Y., and Ansari, A. (2014). A Bayesian Semiparametric Approach for Endogeneity and Heterogeneity in Choice Models. *Management Science*, 60, 1161-79.

Li, Y., Müller, P., and Lin, X. (2011). Center-adjusted Inference for a Nonparametric Bayesian Random Effect Distribution. *Statistica Sinica*, 21, 1201-23.

Lusardi, A., and Mitchell, O. S. (2014). The Economic Importance of Financial Literacy: Theory and Evidence. *Journal of Economic Literature*, 52, 5-44.

McKernan, S. M., Ratcliffe, C., and Vinopal, K. (2009). Do Assets Help Families Cope with Adverse Events? The Urban Institute. Perspectives on Low-Income Working Families. Brief 10 November 2009. http://www.urban.org/sites/default/files/publication/33001/411994-Do-Assets-Help-Families-Cope-with-Adverse-Events-.pdf.

Mills, G., and Amick, J. (2010). Can Savings Help Overcome Income Instability? The Urban Institute. Perspectives on Low-Income Working Families. Brief 18 December 2010.

http://www.urban.org/sites/default/files/publication/32771/412290-Can-Savings-Help-Overcome-Income-Instability-.pdf.

Modigliani, F. and Brumberg, R. (1954). Utility Analysis and the Consumption Function: An Interpretation of Cross-section Data. In: K. Kurihara (Ed.) Post-Keynesian Economics, Chapter 15, pp. 388-436. New Brunswick, NJ: Rutgers University Press.

Puri, M. and Robinson, D. T. (2007). Optimism and Economic Choice. *Journal of Financial Economics*, 86, 71-99.

Royal Society for Public Health (2018). Life on Debt Row. www.rsph.org.uk.

Ruppert, D. (2002). Selecting the Number of Knots for Penalized Splines. *Journal of Computational and Graphical Statistics*, 11, 735-57.

Sethuraman, J. (1994). A Constructive Definition of Dirichlet Priors. *Statistica Sinica*, 4, 639-50.

Smith, M., and Kohn, R. (1996). Nonparametric Regression using Bayesian Variable Selection. *Journal of Econometrics*, 75, 317-43.

Souleles, N. S. (2004). Expectations, Heterogeneous Forecast Errors, and Consumption: Micro Evidence from the Michigan Consumer Sentiment Surveys. *Journal of Money Credit and Banking*, 36, 39-72.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van Der Linde, A. (2002). Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society*, *Series B*, 64, 583-639.

Stock, J. H., Wright, J. H., and Yogo, M. (2002). A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business and Economic Statistics*, 20, 518-29.

Stock, J. H. and Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In: D. W. K. Andrews and J. H. Stock (Eds.) Identification and Inference for Econometric Models, Chapter 5, pp.80-108. New York: Cambridge University Press.

Taylor, M. P. (2011). Measuring Financial Capability and its Determinants using Survey Data. *Social Indicators Research*, 102, 297-314.

Wooldridge, J. (2005). Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity. *Journal of Applied Econometrics*, 20, 39-54.

Yang, M., Dunson, D. B., and Baird, D. (2010). Semiparametric Bayes Hierarchical Models with Mean and Variance Constraints. *Computational Statistics and Data Analysis*, 54, 2172-86.

Yoshida, A. and Guariglia, A. (2002). Estimating Saving Functions in the Presence of Excessive-Zeros Problems. *The Econometrics Journal*, 5, 435-56.

FIGURE 1: Number of financial problems

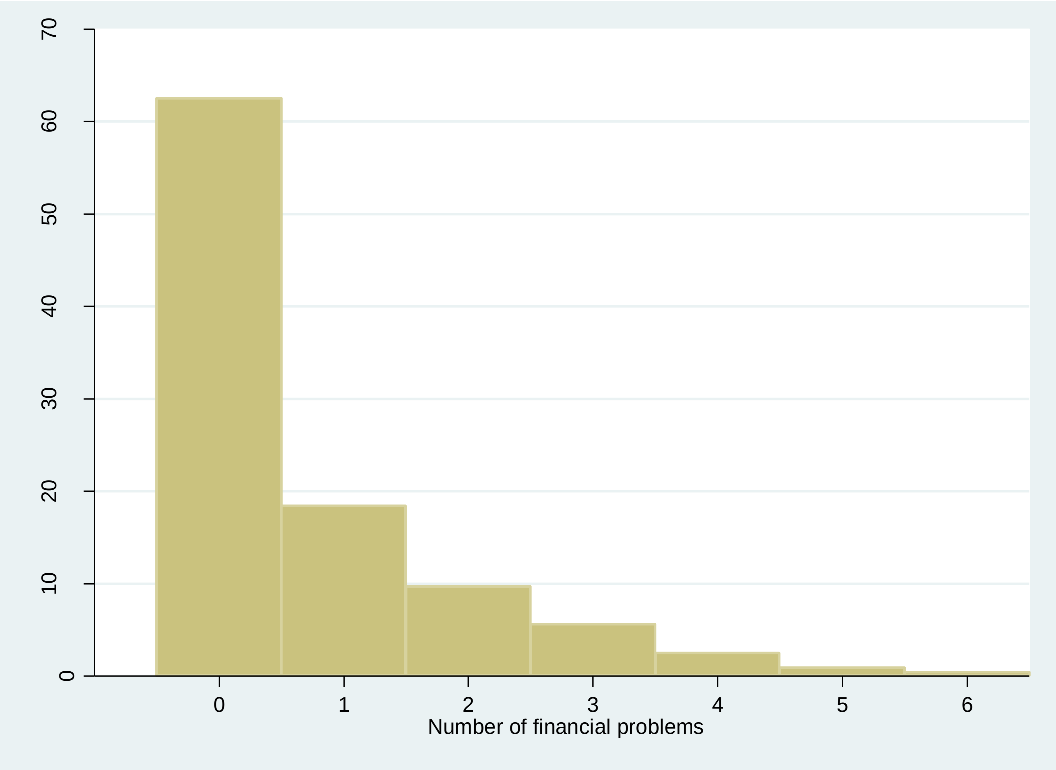


FIGURE 2: Natural logarithm of monthly housing costs (mortgage and rent)

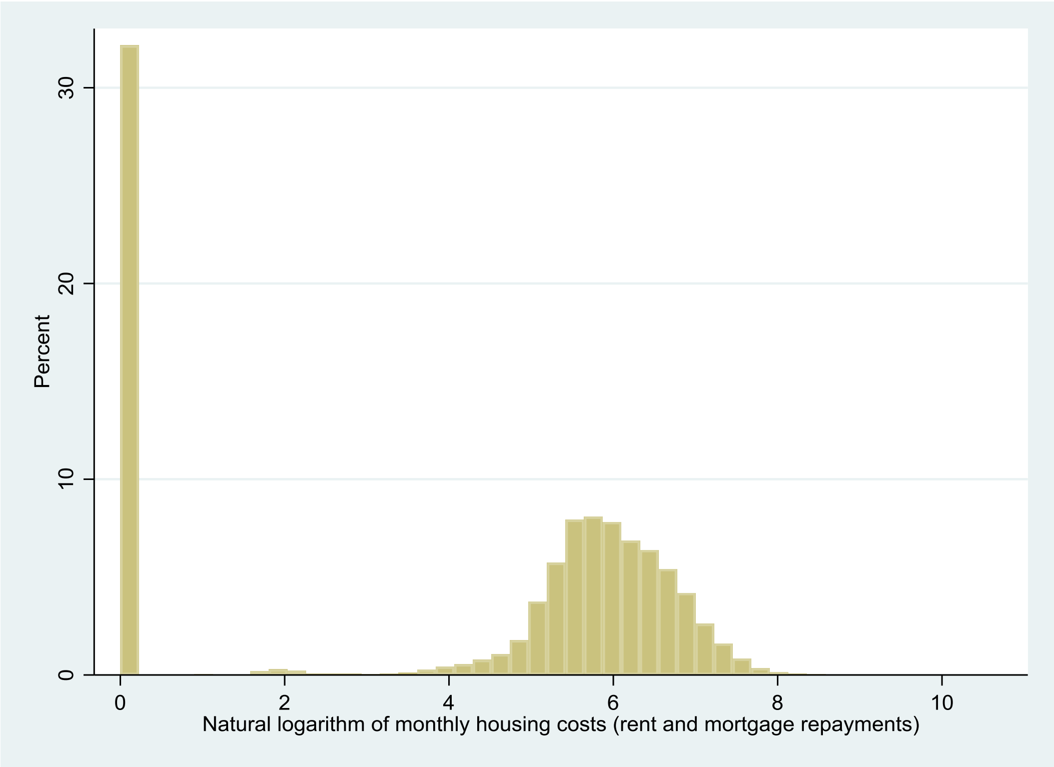
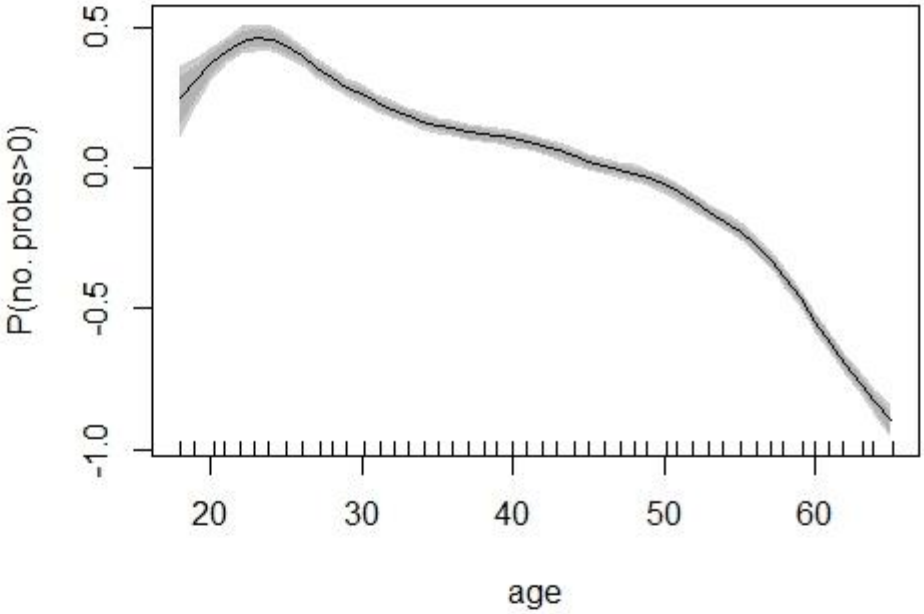
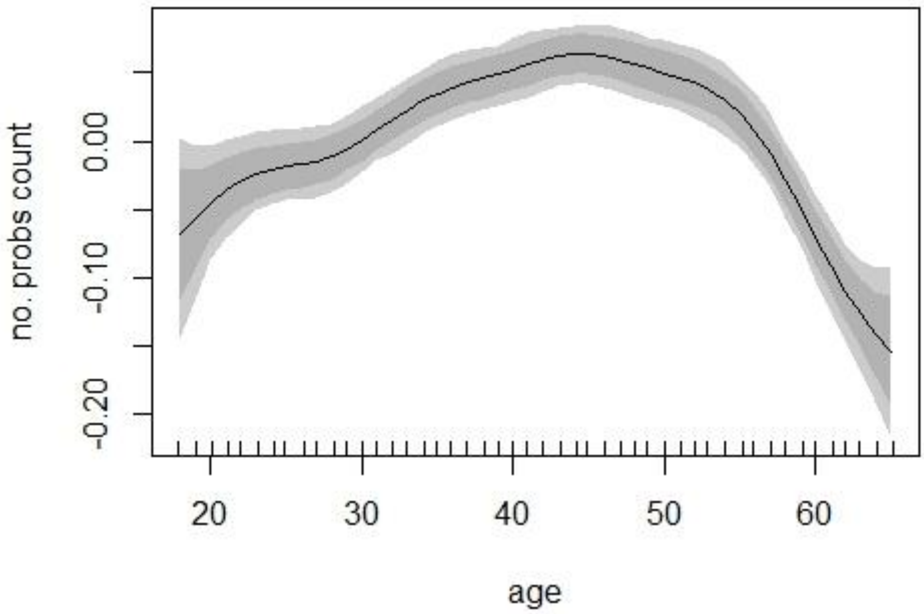


FIGURE 3A: Head of household age effects and the probability of having financial problems



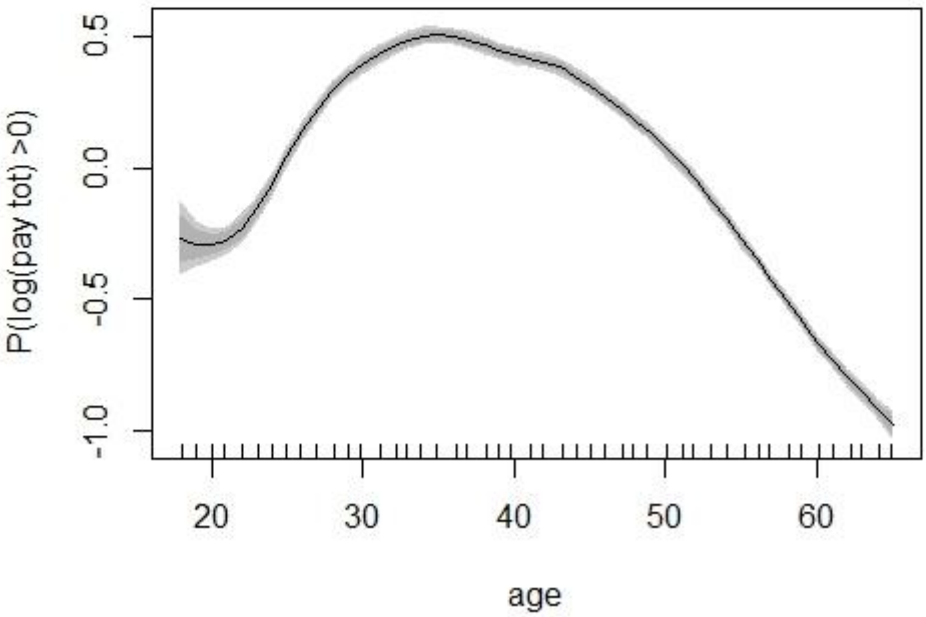
Note the vertical axis shows BPME for the probability of having financial problems.

FIGURE 3B: Head of household age effects and the number of financial problems



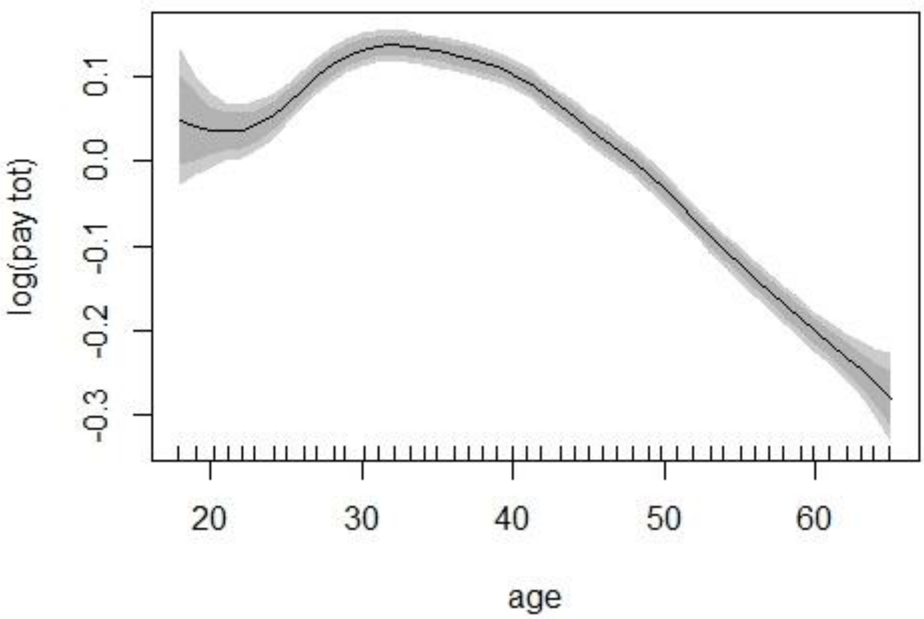
Note the vertical axis shows BPME for the number of financial problems.

FIGURE 4A: Head of household age effects and the probability of having housing costs



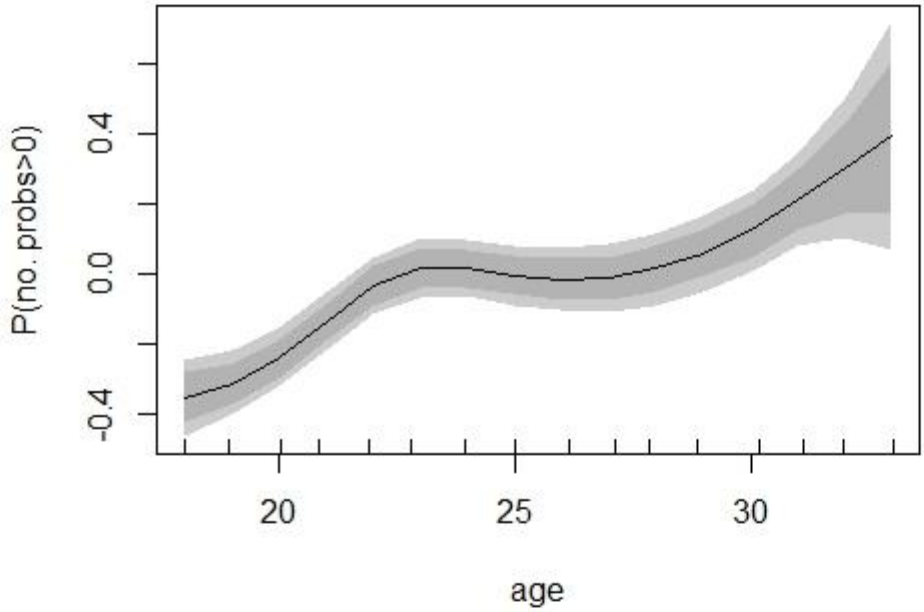
Note the vertical axis shows BPME for the probability of having housing costs (mortgage debt and/or rent).

FIGURE 4B: Head of household age effects and the natural logarithm of the amount of housing costs



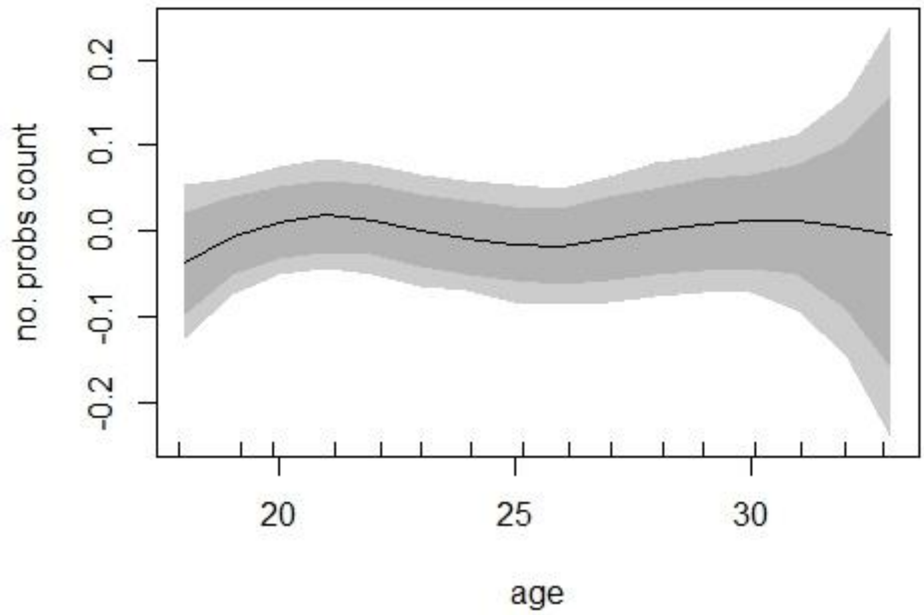
Note the vertical axis shows BPME for the log level of monthly housing costs (mortgage repayments and/or rent).

FIGURE 5A: Head of household age effects and the probability of having financial problems – sub-sample of young adults



Note the vertical axis shows BPME for the probability of having financial problems.

FIGURE 5B: Head of household age effects and the number of financial problems – sub-sample of young adults



Note the vertical axis shows BPME for the number of financial problems.

TABLE 1: Summary statistics

	MEAN	STD. DEV	MIN	MAX
PANEL A: Dependent variables				
Number of financial problems	0.556	1.034	0	6
Whether financial problems	0.306	–	0	1
Number of financial problems conditional upon non-zero	1.820	1.094	1	6
Natural logarithm monthly mortgage + rent	4.016	2.859	0	10.840
Whether monthly housing costs	0.678	–	0	1
Natural logarithm housing costs conditional upon non-zero	5.919	0.886	0.086	10.840
PANEL B: Control variables				
Whether saved last year, $S_{h,t-1}^{A}$	0.377	–	0	1
Natural logarithm of savings last year, $S_{h,t-1}^{A}$	1.864	2.484	0	9.561
Male	0.649	–	0	1
White	0.921	–	0	1
Age	44.111	11.881	18	65
Household size (excluding head of household)	1.795	1.244	0	4
Degree	0.178	–	0	1
Other higher qual., e.g. teaching or nursing	0.314	–	0	1
A levels	0.094	–	0	1
GCSE/O level	0.128	–	0	1
Any other qualification	0.055	–	0	1
Employee	0.640	–	0	1
Self-employed	0.112	–	0	1
Unemployed	0.034	–	0	1
Change in health	0.206	–	0	1
Natural logarithm monthly equivalized income	7.631	1.351	0	11.317
Natural logarithm annual utilities	6.365	1.902	0	10.032
Natural logarithm expenditure non-durable goods	5.708	1.077	0	9.337
Heads of Household (h)		13,700		
Observations (ht)		69,472		

TABLE 2: MODEL 1 – Variance-covariance matrix

VAR (binary financial problems) $\sum_{1,1}$	0.261	*
COV (binary financial problems and number of financial problems) $\sum_{1,2}$	-0.027	
COV (binary financial problems and binary housing costs) $\sum_{1,3}$	0.564	*
COV (binary financial problems and log housing costs) $\sum_{1,4}$	0.975	*
VAR (number of financial problems) $\sum_{2,2}$	0.030	*
COV (number of financial problems and binary housing costs) $\sum_{2,3}$	0.082	*
COV (number of financial problems and log housing costs) $\sum_{2,4}$	0.256	*
VAR (binary housing costs) $\sum_{3,3}$	1.332	*
COV (binary housing costs and log housing costs) $\sum_{3,4}$	2.452	*
VAR (log secured debt) $\sum_{4,4}$	5.414	*

* denotes statistical significance at the 5 per cent level.

TABLE 3: MODEL 1 – Estimated Bayesian marginal effects (posterior means) of the independent variables upon outcomes

	FINANCIAL PROBLEMS		HOUSING COSTS	
	Probability non-zero $\Pr(Y_{ht}^{f} \neq 0)$	Number (count >0) $\log(\lambda_{ht})$	Probability non-zero $\Pr(Y_{ht}^{m} \neq 0)$	Log amount >0 $\log(Y_{ht}^{m})$
Whether saved last year, $S_{h,t-1}^{A}$	-0.540 *	-0.325 *	-0.014	0.008
Financial problems last year, $y_{h,t-1}^{f}$	1.010 *	0.214 *	0.106 *	0.053 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.062 *	-0.029	-0.474 *	0.030 *
Male	-0.457 *	-0.122 *	0.218 *	0.013
White	-0.222 *	-0.106 *	0.093 *	-0.005
Household size	0.120 *	0.040 *	-0.193 *	0.057 *
Degree	-0.306 *	-0.126 *	0.175 *	0.395 *
Other higher qual., e.g. teaching or nursing	-0.018	-0.040 *	0.021	0.070 *
A levels	0.087 *	-0.075 *	-0.084	0.167 *
GCSE/O level	0.094 *	-0.026	-0.158 *	0.068 *
Any other qualification	0.068	-0.032	-0.083	-0.047 *
Employee	-0.286 *	-0.168 *	-0.400 *	0.259 *
Self-employed	-0.549 *	-0.215 *	0.042	0.329 *
Unemployed	0.836 *	0.161 *	-0.679 *	0.046 *
Natural logarithm monthly equivalized income	-0.032 *	-0.076 *	-0.242 *	0.067 *
Natural logarithm annual utilities	0.056 *	0.026 *	-0.125 *	-0.020 *
Natural logarithm expenditure non-durable goods	0.115 *	-0.209 *	-0.179 *	0.040 *
Change in health	0.138 *	0.089 *	–	–
Heads of household (h)		13,700		
Observations (ht)		69,472		

Notes: (i) * denotes statistical significance at the 5 per cent level.

TABLE 3 (Cont.): MODEL 1 – Estimated Bayesian marginal effects (posterior means) of the independent variables upon outcomes

PANEL B: Regional and Business Cycle Controls	FINANCIAL PROBLEMS: Probability non-zero $\Pr(Y_{ht}^{f} \neq 0)$	FINANCIAL PROBLEMS: Number (count >0) $\log(\lambda_{ht})$	HOUSING COSTS: Probability non-zero $\Pr(Y_{ht}^{m} \neq 0)$	HOUSING COSTS: Log amount >0 $\log(Y_{ht}^{m})$
Scotland	-0.070	0.072 *	-0.222 *	-0.146 *
Wales	0.119 *	0.118 *	-0.500 *	-0.164 *
North East	0.108 *	-0.003	-0.615 *	-0.120 *
North West	0.022	0.032	-0.534 *	-0.072 *
East Midlands	-0.048	0.038	-0.470 *	-0.083 *
West Midlands	0.057	0.025	-0.456 *	-0.045 *
East of England	-0.018	0.010	-0.294 *	0.064 *
South East	0.034	0.180 *	-0.176 *	0.198 *
South West	0.052	0.171 *	-0.215 *	0.037
2001	-0.011	-0.030	0.065	-0.040
2002	0.830	0.105 *	-1.027 *	-0.687 *
2003	0.543	-0.115 *	-1.840 *	-0.630 *
2004	0.177	-0.302 *	-1.586 *	-0.642 *
2005	0.193	-0.234 *	-1.558 *	-0.609 *
2006	0.023	-0.338 *	-0.972 *	-0.574 *
2007	-0.030	-0.256 *	-0.933 *	-0.453 *
2008	0.007	-0.243 *	-0.930 *	-0.344 *
2010	0.004	-0.213 *	-0.654 *	-0.287 *
2012	0.125 *	0.233 *	-0.539 *	-0.175 *
2014	0.123 *	0.087 *	-0.328 *	-0.108 *
2016	0.017 *	0.093 *	-1.212 *	-0.038 *
Heads of household (h)		13,700		
Observations (ht)		69,472		

* denotes statistical significance at the 5 per cent level.

TABLE 4: Estimated Bayesian marginal effects (posterior means) for key covariates – Alternative specification (Model 2)

	FINANCIAL PROBLEMS		HOUSING COSTS	
	Probability non-zero $\Pr(Y_{ht}^{f} \neq 0)$	Number (count >0) $\log(\lambda_{ht})$	Probability non-zero $\Pr(Y_{ht}^{m} \neq 0)$	Log amount >0 $\log(Y_{ht}^{m})$
PANEL A: MODEL 2 – Amount saved				
Natural logarithm savings last year, $S_{h,t-1}^{A}$	-0.132 *	-0.084 *	0.008	0.003
Financial problems last year, $y_{h,t-1}^{f}$	0.999 *	0.212 *	0.107 *	-0.054 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.062 *	-0.029 *	-0.475 *	0.029 *
Heads of household (h)		13,700		
Observations (ht)		69,472		

Notes: (i) * denotes statistical significance at the 5 per cent level; (ii) other controls as in Table 3; (iii) full results for model 2 are available from the authors on request.

TABLE 5: Estimated Bayesian marginal effects (posterior means) for key covariates – Sub-sample of young adults aged 18-33

	FINANCIAL PROBLEMS		HOUSING COSTS	
	Probability non-zero $\Pr(Y_{ht}^{f} \neq 0)$	Number (count >0) $\log(\lambda_{ht})$	Probability non-zero $\Pr(Y_{ht}^{m} \neq 0)$	Log amount >0 $\log(Y_{ht}^{m})$
PANEL A: MODEL 3 – Whether saved year				
Whether saved last year, $S_{h,t-1}^{A}$	-0.417 *	-0.197 *	-0.135	0.027
Financial problems last year, $y_{h,t-1}^{f}$	0.819 *	0.161 *	0.134 *	-0.074 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.054 *	-0.023 *	-0.219 *	0.018 *
PANEL B: MODEL 4 – Amount saved last year				
Natural logarithm savings last year, $S_{h,t-1}^{A}$	-0.103 *	-0.047 *	-0.032 *	0.011 *
Financial problems last year, $y_{h,t-1}^{f}$	0.821 *	0.161 *	0.129 *	-0.070 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.051 *	-0.022 *	-0.214 *	0.017 *
PANEL C: MODEL 5 – Whether saved last year, instrumented				
Instrumented whether saved last year, $\hat{S}_{h,t-1}^{A}$	-0.380 *	-0.076 *	-0.255 *	0.121 *
Financial problems last year, $y_{h,t-1}^{f}$	0.802 *	0.162 *	0.107 *	-0.058 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.045 *	-0.023 *	-0.209 *	0.014 *
PANEL D: MODEL 6 – Amount saved, instrumented				
Instrumented natural logarithm savings last year, $\hat{S}_{h,t-1}^{A}$	-0.138 *	-0.027 *	-0.090 *	0.043 *
Financial problems last year, $y_{h,t-1}^{f}$	0.812 *	0.161 *	0.105 *	-0.058 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.045 *	-0.022 *	-0.208 *	0.014 *
Heads of household (h)	1,299			
Observations (ht)	5,218			

Notes: (i) * denotes statistical significance at the 5 per cent level; (ii) other controls as in Table 3; (iii) full results for models 3 to 6 are available from the authors on request.

TABLE 6: Model selection

MODEL	DIC	LPML
FULL SAMPLE 18-65		
1: Whether saved last year	-21,122,174.1	-167,051.87
2: Amount saved last year	-21,729,541.7	-166,955.36
SUB-SAMPLE 18-33		
3: Whether saved last year	-439,177.4	-15,233.41
4: Amount saved last year	-471,423.4	-15,243.04
5: Whether saved last year, instrumented	-321,508.7	-15,202.71
6: Amount saved last year, instrumented	-367,621.6	-15,217.61

Note a smaller (larger) DIC (LPML) value denotes a superior model performance.

TABLE A1: Two-stage least squares analysis – obtaining fitted values for adult saving

FIRST STAGE SUMMARY – CHILD		
	WHETHER EVER SAVED AS A CHILD, S_h^C	
Instruments, $\mathbf{EXP}_h^P$		
Parent expects finances to get worse	0.333 *	
Parent expects finances to improve	-0.071 *	
SECOND STAGE – ADULT		
	WHETHER SAVES, S_h^A	AMOUNT SAVED, S_h^A
Ever saved during childhood, S_h^C	0.087 *	0.480 *
Male	-0.013	0.020
White	0.028	0.046
Degree	0.075 *	0.420 *
Other higher qual., e.g. teaching or nursing	0.040 *	0.183 *
A levels	0.050 *	0.213 *
GCSE/O level	-0.033	-0.242 *
Any other qualification	-0.113	-0.061
Employee	0.159 *	0.906 *
Self-employed	0.068	0.525 *
Unemployed	-0.101 *	-0.448 *
Natural logarithm of permanent income	0.039 *	0.226 *
Volatility of income	0.005 *	0.042 *
Test significance of $\mathbf{EXP}_h^P$, in S_h^A eq., F-statistic, *p-value*	0.70, *p=[0.495]*	0.27, *p=[0.766]*
Kleibergen-Paap χ^2-statistic, *p-value*	87.39, *p=[0.000]*	
Stock-Yogo F-statistic, *p-value*	54.57, *p=[0.000]*	

Notes: (i) * denotes statistical significance at the 5 per cent level; (ii) other controls include government office region and year dummies.

TABLE A2: Zero inflated poisson (ZIP) models of financial problems

	ZIP MODEL 1		ZIP MODEL 2	
	Probability non-zero $\Pr\left(Y^{f}_{ht} \neq 0\right)$	Number (count >0) $\log(\lambda_{ht})$	Probability non-zero $\Pr(Y^{m}_{ht} \neq 0)$	Log amount >0 $\log(Y^{m}_{ht})$
Whether saved last year, $S^{A}_{h,t-1}$	-0.294 *	-0.308 *	–	–
Amount saved last year, $S^{A}_{h,t-1}$	–	–	-0.069 *	-0.078 *
Financial problems last year, $y^{f}_{h,t-1}$	-1.389 *	0.681 *	-1.378 *	0.673 *
Natural logarithm of housing costs last year, $y^{m}_{h,t-1}$	-0.051 *	-0.007 *	-0.051 *	-0.007 *
Male	0.285 *	-0.122 *	0.278 *	-0.119 *
White	0.041	-0.098 *	0.031	-0.098 *
Household size	0.001	0.054 *	0.006	0.054 *
Degree	0.480 *	-0.133 *	0.451 *	-0.125 *
Other higher qual., e.g. teaching or nursing	0.136 *	-0.083 *	0.122	-0.081 *
A levels	0.192 *	-0.083 *	0.180 *	-0.082 *
GCSE/O level	0.156 *	-0.029	0.154 *	-0.029
Any other qualification	0.059	-0.007	0.060	-0.006
Employee	0.289 *	-0.198 *	0.277 *	-0.195 *
Self-employed	0.410 *	-0.250 *	0.399 *	-0.245 *
Unemployed	-0.661 *	0.106 *	-0.661 *	0.107 *
Natural logarithm monthly equivalized income	-0.050 *	-0.093 *	-0.057 *	-0.091 *
Natural logarithm annual utilities	-0.049 *	0.009 *	-0.048 *	0.009 *
Natural logarithm expenditure non-durable goods	-0.483 *	-0.276 *	-0.488 *	-0.275 *
Change in health	-0.190 *	0.127 *	-0.186 *	0.126 *
Heads of household (h)		13,700		
Observations (ht)		69,472		

Notes: (i) * denotes statistical significance at the 5 per cent level; (ii) other controls include government office region, year dummies and a quartic function in age.