Article:

Brown, S. orcid.org/0000-0002-4853-9115, Ghosh, P., Pareek, B. et al. (1 more author) (2021) The protective role of saving : Bayesian analysis of British panel data. Journal of Empirical Finance, 63. pp. 57-72. ISSN 1879-1727

https://doi.org/10.1016/j.jempfin.2021.05.002

THE PROTECTIVE ROLE OF SAVING: BAYESIAN ANALYSIS OF BRITISH PANEL DATA

Sarah Brown[1], Pulak Ghosh[2], Bhuvanesh Pareek[3] 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

[3] Indian Institute of Management, Indore, India

Abstract: We explore whether a protective role for savings against future financial hardship exists using household level panel data for a nationally representative sample of UK households. We jointly model the incidence and extent of financial problems, using a dynamic two-part approach allowing different data-generating processes for experiencing financial hardship and the extent of financial hardship experienced. Our results show that: (i) saving on a regular basis mitigates against 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, with higher housing costs 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 Editor and reviewer for valuable and constructive comments. We are also 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 Universities of Birmingham, Cardiff and Newcastle. Participants at the 2018 *Royal Economic Society Annual Conference*, University of Sussex, and the 2018 *Annual Scientific Meeting of the Italian Economic Association*, Bologna also provided valuable comments and advice.

[#] Corresponding author details – email: k.b.taylor@sheffield.ac.uk; tel.: +44 (0) 114 222 3420

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). To provide context, according to the Office for National Statistics (ONS), the UK savings ratio has fallen from approximately 15% in 1993 to 3.8% in 2018 quarter 3. 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, low or no savings may lead to increased demand for high cost lending products, such as 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 longitudinal 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, see, for example, Crossley et al. (2012), only a limited number of studies in the economics and finance 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

[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. Comprehensive reviews of the literature on household saving include: Browning and Lusardi (1996); Attanasio and Weber (2010) and Crossley et al. (2012). Clearly, there are a range of motivations behind saving behaviour, including alternatives to saving acting as a buffer. Browning and Lusardi (1996) discuss motivations for savings focusing on those listed by Keynes (1936) including: the precautionary motive whereby households hold a contingency fund in case of adverse future events; the life cycle motive to smooth income and consumption over the life cycle; and the inter-temporal substitution motive whereby households benefit from accumulating interest on savings.

seems important to explore from an empirical perspective whether and to what extent holding, as well as the amount of, savings provide a buffer against future financial adversity.

As stated above, whilst a large literature on saving behaviour exists, the basis of our contribution is not on precautionary savings *per se* but rather on how savings affect future financial hardship. In particular, we explore how savings may act as a financial buffer against future shocks, which could impinge on household finances. In what follows, we review the small, yet growing, literature on the relationship between financial hardship and the protective role of saving and, and then discuss our contribution in detail.

1.1 Overview of the literature on financial hardship and saving behaviour

Analysing the 2009 TNS Global Economic Crisis Survey, Lusardi et al. (2011) find that 'using savings' is the option most frequently used by US households experiencing shocks, with financial vulnerability exacerbated by the fact that many households hold little or no savings. This finding is repeated in every country examined in their international analysis, which includes the UK. They also argue that policymakers have tended to focus on incentives for building up assets for long-terms goals such as retirement and house purchase, with savings set aside for precautionary motives receiving no special treatment in terms of, for example, tax advantages. Furthermore, asset limits imposed on eligibility for social programmes and benefits in a number of countries serve as a disincentive to save. Their findings, which are based on analysis of households' subjective assessment of whether they would be able to cope with an unexpected need in the next month that required them to come up with $2,000, support a monotonic increase in the ability to deal with shocks and increasing levels of wealth.

Given such findings, it is surprising that only a small literature exists which explores household financial hardship using nationally representative household surveys (see, for example, Brown et al., 2014, and Giarda, 2013). Furthermore, with the exception of a small number of 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, rather than the population as a whole. 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 based on OLS regression analysis suggest that, after controlling for income, in a sample of households, which have experienced a negative event in the past such as job loss, onset of a health-related work limitation or parent leaving the family, asset poor families are 14 percentage points more likely to experience deprivation than non-asset poor families. 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, the results from their logit analysis suggest that holding liquid assets of up to $1,999 relative to holding zero assets is associated with a fall in 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. The project focuses on households with children. Their findings based on descriptive analysis suggest a negative relationship between whether the household saves for an emergency and the number of material hardships experienced by families. Although such studies are not able to discern the direction of causality, they do highlight some interesting associations between saving behaviour and financial hardship. More recently, Gjertson (2016), also using data from the Annie E. Casey Foundation's Making Connections project, presents evidence from a wide range of estimation approaches including random and fixed effects regression models supporting an inverse relationship between whether a household saved for an emergency and financial hardship for this non-representative sample of low income US households.

With respect to initiatives introduced to encourage saving, Grinstein-Weiss et al. (2014) focus on the US Refund-to-Savings (R2S) initiative, which was designed as an intervention to encourage saving at the time of filing a tax return for low and moderate income (LMI) households by making the saving of any lump sum refund automatic. In the US, the majority of taxpayers receive a refund from the Internal Revenue Service, where for LMI households this can be relatively large and is often perceived as an income windfall. The receipt of such a lump sum means that households need to make a saving decision. Whilst the empirical analysis is descriptive in nature, the findings reveal that those who save at the point of filing and those who have emergency savings experience fewer financial hardships.

More recently, Despard et al. (2018a) explore financial hardship amongst LMI households, focusing in particular on the role of liquid financial assets. They use US administrative tax records and surveys administered through the R2S initiative. The sample is restricted to tax filers who use online software to file their federal income tax returns in 2013. Adopting a structural equation modelling (SEM) approach, their results suggest that holding liquid financial assets mediates the relationship between the association of financial shocks and material hardship. However, there are a number of caveats with this study: firstly, the analysis only establishes associations rather than causation; secondly, the sample is potentially biased since it focuses on tax filers who may be more technologically skilled than average (given the online filing); and thirdly, the sample is more educated than the general US population and covers a higher proportion of whites. Despard et al. (2018b) extend the analysis to assess whether race moderates the mediating role of assets in the relationship between financial shocks and material hardship amongst LMI households. This study employs data from the two-wave 2013 Household Financial Survey. The results reveal that blacks have lower liquid financial assets compared to whites, and the mediating role of liquid financial assets is moderated for blacks and Latinos in comparison to whites.

1.2 Our contribution

The small number of existing studies in this area present some interesting statistical associations for the US. The findings are generally based on descriptive statistics and samples of households, which are non-representative samples of the population as a whole, specifically low income households. For example, Gjertson (2016) states that “due to the strategic site selection and data collection process the Making Connections data cannot be used to make inferences at the national level”, p.3. From a conceptual perspective, we make three important contributions to the existing literature. Firstly, we contribute to the existing literature by exploring whether a protective role for saving against future financial hardship exists within the context of the wider population. All households potentially face shocks such as job loss, ill health and unplanned expenditure. In addition, debt commitments are apparent in households beyond the poorest in society. Our analysis thus serves to establish the extent to which the protective role for saving exists beyond low income households, both in terms of whether an effect exists more generally in the population and the difference in the magnitude of any impact of saving on financial hardship between poor and non-poor UK households. Secondly, in contrast to much of the US literature which focuses on the effects of whether a household saves, we are able to explore the effects of the amount of regular monthly savings. Households holding even small amounts of savings 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. Thirdly, in contrast to much of the above literature, which focuses on contemporaneous relationships between saving behaviour and financial hardship, in order to shed light on causality, our use of panel data enables us to explore the impact of savings on future financial hardship. Finally, we are the first study to analyse the protective role of saving in the UK, which allows us to investigate whether the US findings hold more generally. Specifically, we explore the effect of regular short term monthly saving behaviour on future financial hardship using household level panel data drawn from the British Household Panel Survey and Understanding Society, over the period 1998 to 2016, which is considerably longer than the period explored in existing studies, as well as interestingly, spanning the 2008 financial crisis.

In addition to exploiting our rich panel dataset to make the four modelling contributions described above, we also 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.[2] Such zero inflation issues are less problematic when focusing on low income households in the context of the existing literature, where financial distress is more prevalent. Within our flexible Bayesian framework, we also allow for persistence in experiencing financial problems, which has been commented on in existing studies. To allow for the fact that housing costs represent one of the main financial commitments of households, we allow for the potential interdependence between experiencing financial problems and housing costs in our modelling approach. 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, and Feng et al., 2019). Given that the Bayesian approach allows flexible modelling in complex applications, such an approach seems to be ideally suited to modelling such financial behaviour. Specifically, as outlined above, the modelling structure of our application is quite complicated and the Bayesian approach offers some distinct advantages. Firstly, our Bayesian estimation approach, with the incorporation of the Markov chain Monte Carlo (MCMC) method (see Gelfand and Smith, 1990; Korteweg, 2013; Robert and Casella, 1999), is powerful and sufficiently flexible to deal with complex nonlinear problems such as ours, where the classical maximum likelihood approach can encounter severe computational difficulties (see Lopes and Carvalho, 2007). Secondly, the Bayesian approach allows the examination of the entire posterior distribution of parameters, thus quantifying uncertainty, and thereby avoiding the dependence on asymptotic properties to assess the sampling variability of the parameter estimates.

2. Modelling the Protective Role of Saving

Our focus lies on the role that saving behaviour plays in terms of mitigating both the likelihood and extent of future financial problems. Our dependent variable, the number of financial problems, takes

[2] Our approach is similar to Feng et al. (2019), who employ a two-part latent variable Bayesian model to explore financial literacy and household finances. However, in our analysis the model is a dynamic panel estimator applied to longitudinal data allowing for dependency in outcomes at both the intensive and extensive margins, i.e. temporal correlation, as well as a complex variance-covariance structure.

integer values from 0 to 6. Given the considerable inflation at zero, we use a zero-inflated Poisson model for modelling financial problems. The Bayesian estimator developed here allows for inflation at zero for household financial problems, as well as examining the number of financial problems experienced, conditional on facing financial hardship, whilst also allowing for state dependence and interdependence relating to housing costs. 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 non-linear 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.

Modelling the number of financial problems – a zero-inflated Poisson 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 to allow for the fact that housing costs represent one of the main financial commitments of households and are likely to be endogenous; and, finally, a Dirichlet process (DP) for the joint distribution of the latent random effects from the Poisson hurdle and the semi-continuous models. For brevity and given the focus of our paper, we discuss in detail the financial problems component of the model here and present the modelling details related to the housing cost component in the online technical appendix.

Let Y_{ht}^{f} be the number of financial problems reported by the h^{th} household in the t^{th} year, $h = 1,2,\ldots,N, t = 1,2,\ldots,T$, where N represents the number of households in the sample, and T denotes the number of years. In the context of 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 $Y_{ht}^{f}\sim$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

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

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

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 the covariates and random effects as follows:

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

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

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

where γ_1 and γ_2 are the autoregressive coefficients for the lag effect of order 1 of y_{ht}^{f} and ς_1 and ς_2 are the autoregressive coefficients for the lag effect of order 1 of housing costs, y_{ht}^{m}, capturing interdependence. The inclusion of such lags is particularly important given the persistence in financial problems over time discussed in the existing literature. Saving behaviour is lagged by a year and is represented by $S^A_{h,t-1}$ with associated parameters ψ_1 and ψ_2. The lag is introduced to explore whether savings insulate against future financial hardship. 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 defined in Section 3 below 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. The b_{h1} and b_{h2} are the random effects of p^f_{ht} and λ_{ht}, respectively. We discuss the distribution of the random effects terms below.[3]

Given that the life cycle effects of household finances have been long established, the effects of some covariates, viz., age_{ht}, on p^f_{ht} 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 non-linear 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 $\widetilde{\omega} = (\widetilde{\omega}_1, \widetilde{\omega}_2, \dots, \widetilde{\omega}_C)$ are placed within the range of age_{ht}, such that $\min(\text{age}_{ht}) < \widetilde{\omega}_1 < \widetilde{\omega}_2 < \cdots < \widetilde{\omega}_C < \max(\text{age}_{ht})$. Then $g(\text{age}_{ht})$ is approximated by

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

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 $u_c \sim^{idd} N(0, \sigma_u^2)$; $h = 1, \dots, C$. The online technical appendix provides full details on how the optimal number of knot points is found.[4]

---

[3] For brevity, the model used for housing costs, which follows the same structure but is specified for the less complex case of a semi-continuous dependent variable, is presented in the online technical appendix, where the continuous part of the distribution is log-normal.

[4] We have explored various values of τ, including $\tau = 1, 2, 3, 4$. Following Ruppert et al. (2003), we use the Akaike Information criterion (AIC) for model selection in order to select the optimal value of τ. The AIC reveals that τ =2 is the best model and, hence, we use τ =2. However, for completeness, we have also explored other values of τ and we find that the results did not change.

Correlation structure and heterogeneity – joining the models

The financial problems and housing costs models both 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 (financial problems), log-normal (housing costs), 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 $\boldsymbol{b}_h$ 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 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.[5] Full technical details are provided in the online technical appendix.

Bayesian Methods

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 $\boldsymbol{b}_h = (b_{h1}, b_{h2}, b_{h3}, b_{h4})'$ is proportional to

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

$$\times \left(1 - p_{ht}^{m}\right)^{1-r_{ht}} \{p_{ht}^{m} \times \mathrm{LN}(y_{ht}^{m}; \mu_{ht}^{m}; \sigma^2)\}^{r_{ht}} \times f(\boldsymbol{b}_h) \tag{8}$$

where r_{ht} is an indicator denoting whether monthly housing costs are incurred (see online technical appendix for full details). To complete the Bayesian specification of the model, we assign priors to the unknown parameters in the above likelihood function. For the regression coefficients, the $\beta's$ and $\psi's$, we assume shrinkage priors. We have a large number of covariates and, thus, a shrinkage prior will be beneficial where we adopt a LASSO prior on these sets of parameters (see the online technical appendix for full details).

3. 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

[5] It is important to acknowledge that the prior reflects the beliefs of the researcher. In typical random effects settings, one would not know or be able to test whether it follows a normal distribution as these are latent variables. This is particularly the case with heterogeneity of individuals or households, as in our application. Thus, the restriction of random effects to a normal distribution is regarded as limiting. Specifically, in applications in finance and economics, more flexibility is often preferred in the distribution of the random effects, such as a nonparametric structure (Gill and Casella, 2009). In addition, as noted by Burr and Doss (2005), random effects, unlike error terms, cannot be checked due to the absence of residuals. Thus, to avoid dependence on an unverifiable model assumption, a richer distribution structure on the random effects is assumed; the ideal being to provide a robust prior in such a way that, if the data supports unimodality, it is closer to the normal distribution or, if the data supports otherwise, it can assume multimodality. The DP is one such prior, which has been used for random effects for these reasons.

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 wave 7 (the latest available at the time of writing), around 45,000 individuals were interviewed between 2015 and 2016/2017 (hereafter referred to as 2016). A subset of individuals in the UKHLS can be linked to the BHPS thereby 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 explore 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 is 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 2 above.[6]

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, following the existing literature, we include a comprehensive range of control variables in matrix $\boldsymbol{X}$ (defined above). These include head of household characteristics such as gender; white; age; 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 whether the head of household has had a change in their health between waves (i.e. experienced an unexpected health shock).[7] 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;

[6] 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. Around 30% of the sample did not incur any such costs. Out of the group reporting zero housing costs, 70% own their home outright.

[7] 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.

hearing; heart (including blood pressure); mobility and arthritis; bronchitis; diabetes; depression; epilepsy; cancer; stroke; or any other condition. We also control for: household size (excluding the head of household); in order to further allow for household composition, the natural logarithm of monthly household equivalized income; the natural logarithm of annual household expenditure on water, gas and electricity; and the natural logarithm of total monthly household expenditure on non-durable goods. Finally, we also condition on government office regions (London is the omitted category) and year of interview (pre 2001 is the reference period).

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. 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. In terms of transitions in saving behaviour from one year to the next, amongst non-savers around 19.3% become savers in the subsequent year, whilst for savers approximately 30.4% do not save when next observed in the survey. Conditional on having non-zero monthly housing costs, the last monthly payment is 5.919 log units, which is approximately £517.65. 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.

4. Results

The protective role of saving

In this section, the results from estimating the model 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 $\psi_1, \psi_2 < 0$; and (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 2 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 two columns, showing the probability of being in financial hardship and the number of problems reported, respectively. 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, the 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 experiencing a financial problem, i.e. the 'Odds Ratio' OR= $\exp(\hat{\psi}_1) = \exp(-0.540) = 0.58$, and the number of financial problems is approximately 28 percentage points lower, 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 US literature, which has revealed a protective role of savings against financial hardship in the context of low income households. 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 the extent of future financial problems beyond low income households.

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. However, having incurred 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).[8]

Figure 2 shows 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 2A shows the association between the head of household's age and the probability of reporting a financial problem, and Figure 2B 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 aged under 30 compared to other age groups in the existing literature, e.g. Atkinson et al. (2010), which is consistent with the results shown in Figure 2A, 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 2B. Life cycle effects are evident

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

where the association between the head of household's age and the number of problems experienced increases monotonically until age 45 and then decreases. These 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.

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 are less likely to experience financial problems. This 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 report fewer problems compared to those with no qualifications. 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= $\exp(\hat{\beta}_{1k}) = \exp(0.836) = 2.37$. 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= $\exp(\hat{\beta}_{2k}) = \exp(-0.076) = 0.92$.

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 financial problems.

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. The business cycle effects are interesting, in that after the recent financial crisis both the incidence and extent of household financial hardship increased. For example, in 2012, a head of household was around 13 percentage points more likely to experience a financial problem compared to pre-2001, OR= $\exp(\hat{\beta}_{1k}) = \exp(0.125) = 1.13,$ 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 financial problems fell each year, whilst for the most recent years, post-2010, the number of financial 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 a reduction in the number of financial problems of 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 BHPS *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., for the US and Sweden, see Knowles and Postlewaite (2004) and Cronqvist and Siegel (2015), respectively. For the UK, Brown and Taylor (2016) find that having saved as a child increases the probability of saving during adulthood by approximately 12 percentage points, a sizeable effect.

The sub-sample we focus on here comprises relatively young adults as our estimation strategy requires observing the head of household as a child and as an adult. However, this age group (18-33) is particularly interesting given that the results shown in Figure 2A 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. (2010), 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.

To model saving behaviour, we use a two-stage least squares (2SLS) approach. In the first stage, we model saving behaviour during childhood (equation 9), as this may be endogenous if included directly as a control for adult saving and, in the second stage, the saving behaviour of adults is modelled (equation 10). The following depicts the empirical model:

$$S_h^C = 1\left[\boldsymbol{Z}'_h\boldsymbol{\phi}_1 + \mathbf{EXP}_h^{P'}\boldsymbol{\pi}_1 + \nu_{1h} > 0\right] \tag{9}$$

$$S_h^A = 1[\boldsymbol{Z}'_h\boldsymbol{\phi}_2 + \psi_1 S_h^C + \nu_{2h} > 0] \tag{10}$$

where S_h^A is either a binary indicator of 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. As the existing 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), the IV in the first stage is given by $\mathbf{EXP}_h^P$, a vector of the financial expectations of the child's parent (who is the head of household). From the 2SLS analysis, we obtain the fitted values of savings (either the incidence or amount), denoted by $\hat{S}_{h,t-1}^A$. In accordance with the existing literature, the results from modelling savings behaviour, shown in Table 5, reveal that both the probability and the amount of savings are positively associated with: whether the individual saved during childhood; educational attainment; permanent income and its volatility. 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. 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 instruments are strongly associated with the saving decision as a child and are arguably exogenous to their saving behaviour as an adult. Moreover, from a theoretical perspective they also seem plausible in that there is no obvious reason why the financial expectations of the parent, measured *ex ante*, would influence the current saving behaviour of their offspring when observed as young adults.[9]

The results focusing on the sub-sample of young adults are shown in Table 6, where for brevity, we only report the key parameters of interest, i.e. those associated with savings behaviour

[9] A potential caveat is where there is intergenerational correlation in financial attitudes. However, Brown and Taylor (2016) show that such correlation across generations is negligible.

(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 the magnitude of the BPMEs to that of models 1 and 2, shown in Tables 3 and 4.

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, is associated with a reduction in the probability of having a financial problem by 34 percentage points (OR= $\exp(\hat{\psi}_1) = \exp(-0.417) = 0.66$) and a reduction in the number of financial problems by 18 percentage points (OR= $\exp(\hat{\psi}_2) = \exp(-0.197) = 0.82$), see Table 6 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 6 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 6, model 4, Panel B to Table 4 model 2, and this finding is robust to instrumenting the amount saved, as can be seen from Table 6, model 6, Panel D.

Figures 4A and 4B 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. 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 3A. 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 3B, 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.[10]

Savings behaviour and future financial hardship – low versus high income households

As discussed in Section 1.1, much of the existing literature has focused on the protective role of savings in the context of low income households. To examine whether the above results are operating predominately for low income rather than high income households, we split the sample into two groups, households above or at the poverty line and households below the poverty line, where poverty is defined as having equivalized income less than 60% of the median household.[11] The results are shown in Tables 7 and 8. In Table 7, the full sample is split into two groups, those in poverty comprising 7,101 households (19,272 observations) and those not in poverty, i.e. at or above the poverty line, comprising 11,368 households (50,200 observations). In Table 8, we focus on the sub-sample of young adults, where we instrument saving behaviour, again there are two samples – those in poverty, where there are 712 households (1,515 observations) and those not in poverty, comprising 1,076 households (3,703 observations).

Focusing initially on the results in Table 7, two models are estimated: in Panel A, model 1 explores the effect of the incidence of saving on future financial problems; whilst in Panel B, model 2 explores the effect of the amount saved on future financial problems. We only report 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 7 is split into four columns: the first two focus on households in poverty, showing the probability of being in financial hardship (column 1) and the number of problems reported (column

[10] 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. $\mathrm{OR} = \exp\left(\hat{\beta}_{1k}\right) = \exp(0.581) = 1.79$.

[11] This definition is consistent with the measure of poverty used in official UK government statistics, for example, see www.gov.uk/government/statistics/households-below-average-income-199495-to-201718.

2), respectively; whilst the final two columns consider those households not in poverty, showing the probability of being in financial hardship (column 3) and the number of problems reported (column 4), respectively. The results in Panel A reveal that the incidence of past saving is associated with a reduction in the probability of having a financial problem, i.e. $\hat{\psi}_1 < 0$. For those in poverty, this is a reduction in the likelihood of a future financial problem of 43 percentage points, compared to 37 percentage points for those households not in poverty.[12] In terms of the extent of financial problems, i.e. the number reported, the comparable figures for those households below the poverty line and not in poverty are 31 and 21 percentage points, respectively.[13] Similarly, considering the extent of savings, i.e. the amount saved last year, Panel B, the reduction in the probability of financial problems (number of problems) for those households in poverty and those not in poverty is 25 and 23 (21 and 11) percentage points, respectively. Hence, our findings suggest that having a savings buffer is important for reducing the future incidence and extent of financial problems for all households, regardless of poverty status. As expected, the mitigating effects of saving are larger for households below the poverty line, although it is important to acknowledge that these effects remain sizeable for households above the poverty line.

Table 8 is constructed in the same way as Table 7, comprising four columns split across Panels A to D, and focuses on the sample of young adult heads of household. The first two panels replicate Table 7 for this sub-sample, whereas, in Panels C and D, the incidence of saving and the amount saved are instrumented, respectively. The results are consistent with those found for the full sample, in that the protective role of savings mitigates both the likelihood of future financial problems and the extent of such hardship for this sub-sample of young adults, see Table 8 Panels A to D, i.e. $\hat{\psi}_1, \hat{\psi}_2 < 0$, with these mitigating effects being apparent regardless of poverty status.

[12] Calculated as $\left[1 - \exp\left(\hat{\psi}_1\right)\right] \times 100$.
[13] Calculated as $\left[1 - \exp\left(\hat{\psi}_2\right)\right] \times 100$.

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 nationally representative UK panel data, we have made 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. 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 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 found to exist beyond low income households as well as for a sub-sample of young household heads. The latter 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.

Finally, we also examine whether the protective role of saving differs across low and high income households by splitting the sample into those below and those above the poverty line. This is important, in that the majority of the existing literature focuses on low income households, and furthermore our findings could be driven solely by a saving effect operating for low income households. However, our analysis reveals that, regardless of poverty status, having a savings buffer is important in reducing the future incidence and extent of financial problems for all households and not just those in poverty. Hence, our findings suggest that the protective role of saving behaviour exists beyond low income households.

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. (2010). Levels of Financial Capability in the UK. *Public Money and Management*, 27, 29-36.

Attanasio, O. and Weber, G. (2010). Consumption and Saving: Models of Intertemporal Allocation and Their Implications for Public Policy. *Journal of Economic Literature*, 48, 693-751.

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.

Burr, D. and Doss, H. (2005). A Bayesian semi-parametric model for random effects meta-analysis. *Journal of the American Statistical Association*, 100, 242-5.

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.

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.

Despard, M., Guo, S., Grinstein-Weiss, M., Russell, B., Oliphant, J. and de Ruyter, A. (2018a) The Mediating Role of Assets in Explaining Hardship Risk Among Households Experiencing Financial Shocks. *Social Work Research*, 42, 147-58.

Despard, M., Grinstein-Weiss, M., Guo, S., Taylor, S. and Russell, B. (2018b). Financial Shocks, Liquid Assets, and Material Hardship in Low and Moderate-Income Households: Differences by Race. *Journal of Economics, Race, and Policy*, 1, 205-16.

Feng, X., Lu, B., Song, X. and Ma, S. (2019). Financial Literacy and Household Finances: A Bayesian Two-Part Latent Variable Modeling Approach. *Journal of Empirical Finance*, 51, 119-37.

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, NJ: Princeton University Press.

Gelfand, A. E. and Smith A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of American Statistical Association*, 57, 398-409.

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.

Gill, J. and Casella, G. (2009). Nonparametric priors for ordinal Bayesian social science models: Specification and estimation. *Journal of the American Statistical Association*, 104, 453-64.

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

Grinstein-Weiss, M., Russell, B., Tucker, B. and Comer, K. (2014). Lack of Emergency Savings Puts American Households at Risk: Evidence from the Refund to Savings Initiative. CSD Policy Brief, 14-13. St. Louis, MO: Washington University, Center for Social Development.

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.

House of Lords (2017). Tackling Financial Exclusion: A Country That Works for Everyone? *Select Committee on Financial Exclusion HL Paper 132.*

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.

Korteweg, A. (2013). Markov Chain Monte Carlo Methods in Corporate Finance. In P. Damien, P. Dellaportas, N. Polson, and D. Stephens (Eds.), Bayesian Theory and Applications; Oxford University Press.

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.

Lopes, H. F. and C. M. Carvalho (2007) 'Factor Stochastic Volatility with Time-Varying Loadings and Markov Switching Regimes.' *Journal of Statistical Planning and Inference*, 137, 3082-91.

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.

Lusardi, A., Scheider, D. and Tufano, O. (2011). Financially Fragile Households: Evidence and Implications. *Brookings Papers on Economic Activity*, 42, 83-150.

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.

Robert, C. and Casella, G. (1999) *Monte Carlo Statistical Methods*. Springer, New York.

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.

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, Non-linear 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

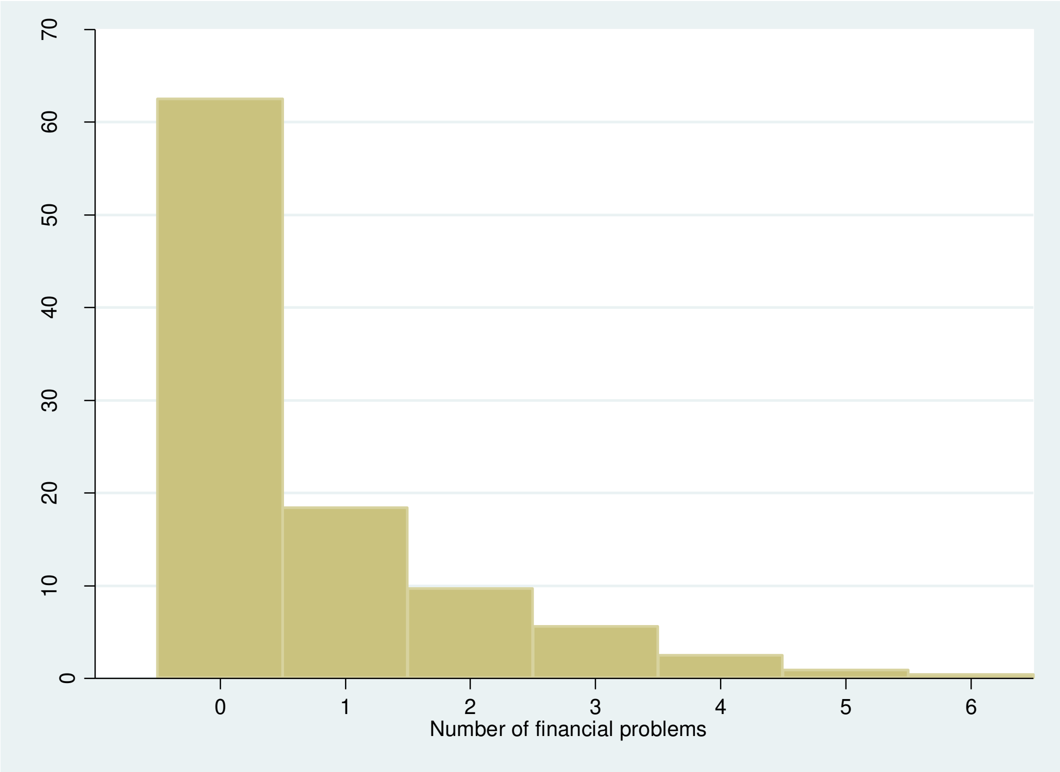
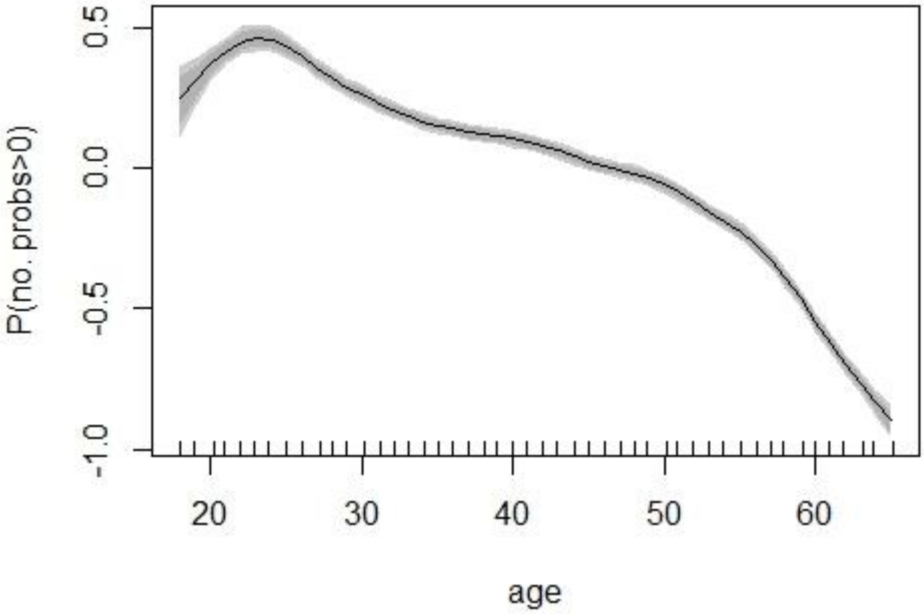
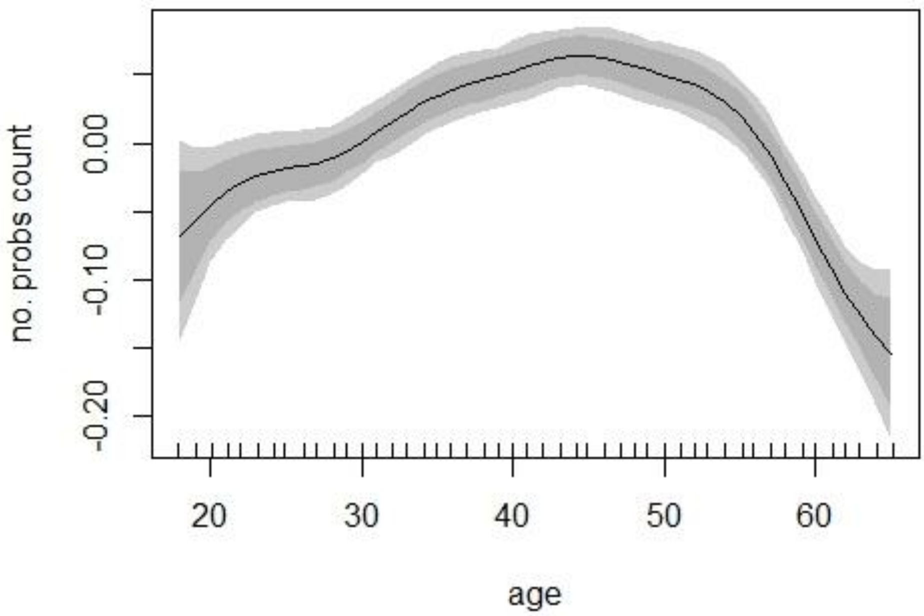


FIGURE 2A: Head of household age effects and the probability of having financial problems



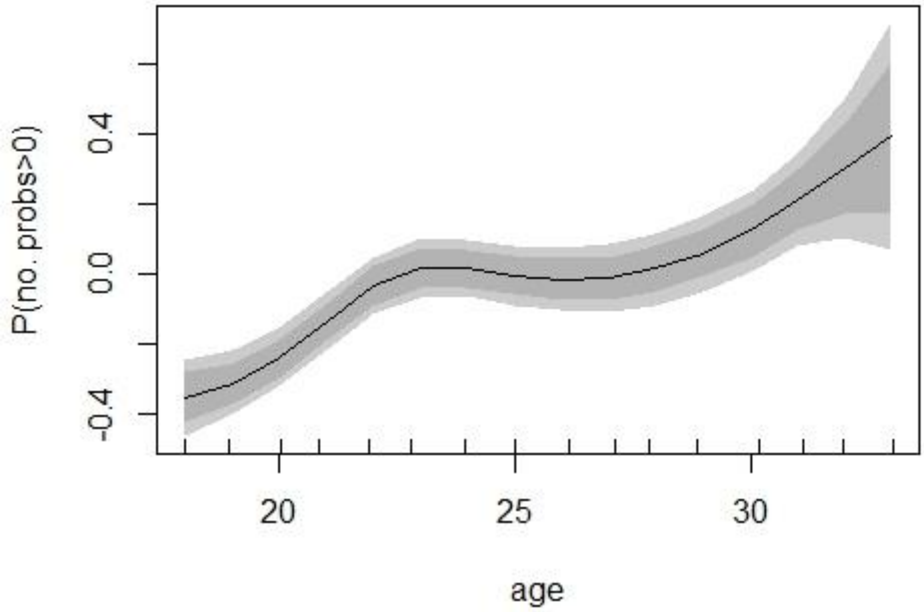
Note the vertical axis shows BPME for the probability of having financial problems.

FIGURE 2B: Head of household age effects and the number of financial problems



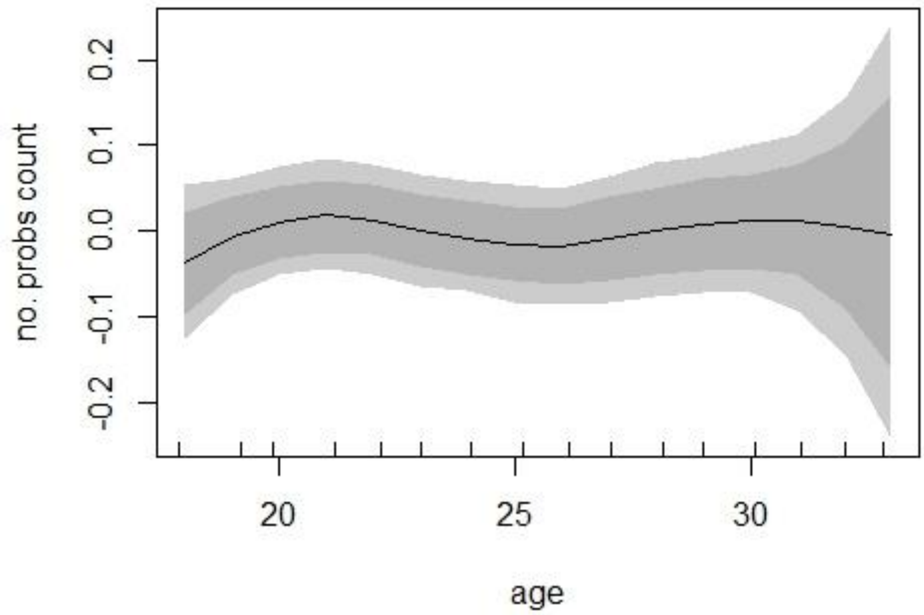
Note the vertical axis shows BPME for the number of financial problems.

FIGURE 3A: 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 3B: 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	0.386	–	0	1
Number of financial problems conditional upon non-zero	1.820	1.094	1	6
Natural logarithm housing costs	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)

	FINANCIAL PROBLEMS	
	Probability non-zero $\Pr(Y_{ht}^{f} \neq 0)$	Number (count >0) $\log(\lambda_{ht})$
Whether saved last year, $S_{h,t-1}^{A}$	-0.540 *	-0.325 *
Financial problems last year, $y_{h,t-1}^{f}$	1.010 *	0.214 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.062 *	-0.029
Male	-0.457 *	-0.122 *
White	-0.222 *	-0.106 *
Household size	0.120 *	0.040 *
Degree	-0.306 *	-0.126 *
Other higher qual., e.g. teaching or nursing	-0.018	-0.040 *
A levels	0.087 *	-0.075 *
GCSE/O level	0.094 *	-0.026
Any other qualification	0.068	-0.032
Employee	-0.286 *	-0.168 *
Self-employed	-0.549 *	-0.215 *
Unemployed	0.836 *	0.161 *
Natural logarithm monthly equivalized income	-0.032 *	-0.076 *
Natural logarithm annual utilities	0.056 *	0.026 *
Natural logarithm expenditure non-durable goods	0.115 *	-0.209 *
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)

	FINANCIAL PROBLEMS	
PANEL B: Regional and Business Cycle Controls	Probability non-zero $\Pr(Y_{ht}^{f} \neq 0)$	Number (count >0) $\log(\lambda_{ht})$
Scotland	-0.070	0.072 *
Wales	0.119 *	0.118 *
North East	0.108 *	-0.003
North West	0.022	0.032
East Midlands	-0.048	0.038
West Midlands	0.057	0.025
East of England	-0.018	0.010
South East	0.034	0.180 *
South West	0.052	0.171 *
2001	-0.011	-0.030
2002	0.830	0.105 *
2003	0.543	-0.115 *
2004	0.177	-0.302 *
2005	0.193	-0.234 *
2006	0.023	-0.338 *
2007	-0.030	-0.256 *
2008	0.007	-0.243 *
2010	0.004	-0.213 *
2012	0.125 *	0.233 *
2014	0.123 *	0.087 *
2016	0.017 *	0.093 *
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	
	Probability non-zero $\Pr\left(Y_{ht}^{f} \neq 0\right)$	Number (count >0) $\log(\lambda_{ht})$
PANEL A: MODEL 2 – Amount saved last year		
Natural logarithm savings last year, $S_{h,t-1}^{A}$	-0.132 *	-0.084 *
Financial problems last year, $y_{h,t-1}^{f}$	0.999 *	0.212 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.062 *	-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: 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 6: Estimated Bayesian marginal effects (posterior means) for key covariates – Sub-sample of young adults aged 18-33

	FINANCIAL PROBLEMS	
	Probability non-zero $\Pr(Y_{ht}^{f} \neq 0)$	Number (count >0) $\log(\lambda_{ht})$
PANEL A: MODEL 3 – Whether saved year		
Whether saved last year, $S_{h,t-1}^{A}$	-0.417 *	-0.197 *
Financial problems last year, $y_{h,t-1}^{f}$	0.819 *	0.161 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.054 *	-0.023 *
PANEL B: MODEL 4 – Amount saved last year		
Natural logarithm savings last year, $S_{h,t-1}^{A}$	-0.103 *	-0.047 *
Financial problems last year, $y_{h,t-1}^{f}$	0.821 *	0.161 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.051 *	-0.022 *
PANEL C: MODEL 5 – Whether saved last year, instrumented		
Instrumented whether saved last year, $\hat{S}_{h,t-1}^{A}$	-0.380 *	-0.076 *
Financial problems last year, $y_{h,t-1}^{f}$	0.802 *	0.162 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.045 *	-0.023 *
PANEL D: MODEL 6 – Amount saved last year, instrumented		
Instrumented natural logarithm savings last year, $\hat{S}_{h,t-1}^{A}$	-0.138 *	-0.027 *
Financial problems last year, $y_{h,t-1}^{f}$	0.812 *	0.161 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.045 *	-0.022 *
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 7: Estimated Bayesian marginal effects (posterior means) for key covariates – by poverty status

	FINANCIAL PROBLEMS			
	In poverty		Not in poverty	
	Probability non-zero $\Pr(Y_{ht}^{f} \neq 0)$	Number (count >0) $\log(\lambda_{ht})$	Probability non-zero $\Pr(Y_{ht}^{f} \neq 0)$	Number (count >0) $\log(\lambda_{ht})$
PANEL A: MODEL 1 – Whether saved last year				
Whether saved last year, $S_{h,t-1}^{A}$	-0.563 *	-0.373 *	-0.465 *	-0.231 *
Financial problems last year, $y_{h,t-1}^{f}$	0.854 *	0.172 *	0.962 *	0.217 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.093 *	-0.038 *	-0.091 *	-0.047 *
PANEL B: MODEL 2 – Amount saved last year				
Natural logarithm savings last year, $S_{h,t-1}^{A}$	-0.292 *	-0.235 *	-0.263 *	-0.116 *
Financial problems last year, $y_{h,t-1}^{f}$	0.850 *	0.172 *	0.957 *	0.217 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.094 *	-0.038 *	-0.080 *	-0.045 *
Heads of household (h)	7,101		11,368	
Observations (ht)	19,272		50,200	

Notes: (i) * denotes statistical significance at the 5 per cent level; (ii) other controls as in Table 3; (iii) full results for models 1 and 2 are available from the authors on request.

TABLE 8: Estimated Bayesian marginal effects (posterior means) for key covariates: Sub-sample of young adults aged 18-33 – by poverty status

	FINANCIAL PROBLEMS			
	In poverty		Not in poverty	
	Probability non-zero $\Pr(Y_{ht}^{f} \neq 0)$	Number (count >0) $\log(\lambda_{ht})$	Probability non-zero $\Pr(Y_{ht}^{f} \neq 0)$	Number (count >0) $\log(\lambda_{ht})$
PANEL A: MODEL 3 – Whether saved last year				
Whether saved last year, $S_{h,t-1}^{A}$	-0.468 *	-0.187 *	-0.213 *	-0.167 *
Financial problems last year, $y_{h,t-1}^{f}$	0.638 *	0.135 *	0.798 *	0.127 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.002	-0.022	-0.027	-0.032 *
PANEL B: MODEL 4 – Amount saved last year				
Natural logarithm savings last year, $S_{h,t-1}^{A}$	-0.244 *	-0.110 *	-0.152 *	-0.080 *
Financial problems last year, $y_{h,t-1}^{f}$	0.625 *	0.133 *	0.790 *	0.128 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.005	-0.022	-0.026	-0.032 *
PANEL C: MODEL 5 – Whether saved last year, instrumented				
Instrumented whether saved last year, $\hat{S}_{h,t-1}^{A}$	-0.334 *	-0.068 *	-0.242 *	-0.065 *
Financial problems last year, $y_{h,t-1}^{f}$	0.613 *	0.133 *	0.795 *	0.130 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.011	-0.021	-0.024	-0.031 *
PANEL D: MODEL 6 – Amount saved last year, instrumented				
Instrumented natural logarithm savings last year, $\hat{S}_{h,t-1}^{A}$	-0.249 *	-0.033 *	-0.192 *	-0.045 *
Financial problems last year, $y_{h,t-1}^{f}$	0.620 *	0.136 *	0.794 *	0.129 *
Natural logarithm of housing costs last year, $y_{h,t-1}^{m}$	-0.009 *	-0.022	-0.025	-0.031 *
Heads of household (h)	712		1,076	
Observations (ht)	1,515		3,703	

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.

ONLINE TECHNICAL APPENDIX

1. Modelling monthly housing costs – a semi-continuous model

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 housing costs. 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. 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. 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,

$$R_{ht} = \begin{cases} 0, & \text{if } Y_{ht}^{m} = 0 \\ 1, & \text{if } Y_{ht}^{m} > 0 \end{cases} \tag{A1}$$

with conditional probabilities

$$\Pr(R_{ht} = r_{ht}) = \begin{cases} 1 - p_{ht}^{m}, & \text{if } r_{ht} = 0 \\ p_{ht}^{m}, & \text{if } r_{ht} = 1. \end{cases} \tag{A2}$$

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:

$$Y_{ht}^{m} \sim (1 - p_{ht}^{m})^{1 - r_{ht}} \{p_{ht}^{m} \times N(\log(Y_{ht}^{m}); \mu_{ht}^{m}, \sigma^2)\}^{r_{ht}} \tag{A3}$$

$$\operatorname{logit}(p_{ht}^{m}) = \gamma_3 y_{h,t-1}^{m} + \varsigma_3 y_{h,t-1}^{f} + \psi_3 S_{h,t-1}^{A} + \boldsymbol{X}_{ht}^{'} \beta_3 + h^{p}(\text{age}_{ht}) + b_{h3} \tag{A4}$$

$$\mu_{ht} = \gamma_4 y_{h,t-1}^{m} + \varsigma_4 y_{h,t-1}^{f} + \psi_4 S_{h,t-1}^{A} + \boldsymbol{X}_{ht}^{'} \beta_4 + h^{\mu}(\text{age}_{ht}) + b_{h4} \tag{A5}$$

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 (A4, A5) is a semi-parametric counterpart of the correlated two-part model proposed for modelling financial problems (see equations 5 and 6). Saving behaviour, $S_{h,t-1}^{A}$, is included as a lag. Note that in equations A4 and A5 we do not include whether the head of household experienced a change in their health on the assumption that an adverse health shock will influence financial problems but not monthly housing costs.

Both the probability of and amount of housing costs are found to be associated with the incidence and extent of past housing costs (i.e. mortgage payments and/or rent). For example, 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). There is also positive interdependence between housing costs (both the incidence and extent) and having experienced financial problems during the previous year. However, contrary to the findings for financial problems, saving behaviour is unrelated to both the incidence and the extent of future housing costs.[14]

2. The optimal number of knot points

In equations (5) and (6), where we allow for non-linear age effects, 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 the desired flexibility, and $\widetilde{\omega}_k$ denotes the sample quantiles of age_{ht} corresponding to probability $k/(c+1)$, but the results hold for other choices of knots. In our empirical analysis, the function of age is modelled with c=20 knots chosen so that the k^{th} knot is the sample quantile of age corresponding to probability $k/(c+1)$. 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, which 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

[14] Full results are available upon request.

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 $u_c \sim^{idd} N(0, \sigma_u^2)$; $h = 1, \dots, C$.

3. Correlation structure and random effects

Let $\boldsymbol{b}_h = (b_{h1}, b_{h2}, b_{h3}, b_{h4})'$ denote a vector of random effects from each part of the empirical framework, i.e. equations 5, 6, A4 and A5. The model for $\boldsymbol{b}_h$ can be written as

$$\boldsymbol{b}_h \sim G, \quad G \sim \mathrm{DP}(\alpha G_0) \tag{A6}$$

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

$$G = \sum_{q=1}^{\infty} \pi_q^D \delta_{\boldsymbol{b}_q}, \quad \boldsymbol{b}_q \sim G_0, \quad \text{and} \quad \sum_{q=1}^{\infty} \pi_q^D = 1 \tag{A7}$$

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 A7, $\pi_q^D = V_h \prod_{l<q}(1 - V_l)$, which is formulated from a stick-breaking process, with $V_q \sim \mathrm{Beta}(1, \alpha)$, is the probability assigned to the q^{th} cluster. For small values of α, $V_q \rightarrow 1$ and thus $\pi_q^D \rightarrow 1$, 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

$$\mu_G^j = \sum_{q=1}^{Q} V_q^j \prod_{l<q} \left(1 - V_l^j\right) \boldsymbol{b}_q^j \tag{A8}$$

$$\Sigma_G^j = \sum_{q=1}^{Q} V_q^j \prod_{l<q} \left(1 - V_l^j\right) \left(\boldsymbol{b}_q^j - \mu_G^j\right) \left(\boldsymbol{b}_q^j - \mu_G^j\right)' \tag{A9}$$

where V_q^j and $\boldsymbol{b}_q^j$ are the posterior samples from the uncentered process defined in equation A7 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 $\mathrm{E}(\boldsymbol{b}_h | G = 0)$ and $\mathrm{Var}(\boldsymbol{b}_h | G = \Sigma_G)$.

4. LASSO priors

From equation (8) in the main text, 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:

$$\phi_k \left| \sigma^2, \tau_1^2, \ldots, \tau_p^2 \sim N_p(0, \sigma^2 \boldsymbol{D}_\tau) \right. \tag{A10a}$$

where $\boldsymbol{D}_\tau = \mathrm{diag}(\tau_1^2, \ldots, \tau_{P'}^2)$ (A10b)

$$\tau_1^2, \ldots, \tau_{P'}^2 \sim \prod_{p=1}^{P'} \frac{\lambda^2}{2} \exp\left(-\frac{1}{2} \lambda \tau_p^2\right) \tag{A11}$$

$$\lambda^2 \sim \mathrm{Gamma}(a, b) \tag{A12}$$

$$\sigma^2 \sim \pi(\sigma^2) = \frac{1}{\sigma^2} \tag{A13}$$

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.

5. Computational details

Although the posterior distributions are analytically intractable, the models detailed above can be fitted using Markov Chain Monte Carlo (MCMC) methods such as the Gibbs sampler (Gelman et al., 1992). However, since the full conditional distributions are not standard, a straightforward implementation of the Gibbs sampler using standard sampling techniques is not possible. However, sampling methods can be performed using, e.g., adaptive rejection sampling (ARS), metropolis hastings and/or blocked Gibbs sampling methods (Gilks and Wild, 1992).

In this paper, we have used a general program for Bayesian inference using Gibbs Sampling (WinBUGS) for computation (Spiegelhalter et al., 1996). This is a freely available Bayesian MCMC package. WinBUGS uses the Gibbs sampling algorithm to construct transition kernels for its Markov chain samplers. During compilation, WinBUGS chooses a method to draw samples from each of the full conditional distributions of the model parameters. Such sampling can be done univariately or in multivariate nodes. The sampling methods within WinBUGS include direct sampling using standard algorithms, derivative free adaptive rejection sampling (Gilks and Wild, 1992), slice sampling (Neal, 2000) and metropolis sampling (Gelfand and Smith, 1990) and blocked Gibbs Sampling. The first choice is always a standard density if it is available. This possibility arises when a full conditional is recognizable. For nonstandard but log-concave full conditionals, adaptive rejection sampling is used to sample from the full conditional (Gilks and Wild, 1992). WinBUGS checks if log-concavity is satisfied or not, and uses slice sampling (Neal, 2000), if it is not satisfied. The random walk Metropolis algorithm is also used by WinBUGS for nonconjugate continuous full conditionals. The samples from the posterior distribution obtained from the MCMC allow us to obtain summary measures of the parameter estimates and to obtain credible intervals (CIs) of the parameters of interest. The full WinBUGS code is available on request.

Gelfand, A. E. and Smith A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of American Statistical Association*, 57, 398-409.

Gelman A. and Rubin D. B. (1992). Inference from iterative simulation using multiple sequences (with Discussion). *Statistical Science*, 7, 457-511.

Gilks, W. R., and Wild, P. (1992). Adaptive Rejection Sampling for Gibbs Sampling. *Applied Statistics*, 41 92), 337-48.

Neal R. M. (2000). Markov chain sampling methods for Dirichlet process mixture models. *Journal of Computational and Graphical Statistics*, 9 (2), 249-65.

Spiegelhalter D., Thomas A., Best N. and Gilks W. (1996). BUGS: Bayesian inference using Gibbs sampling, version 0.50. MRC Biostatistics Unit: Cambridge.