Working Paper no. 277

When is mortgage indebtedness a financial burden to British households? A dynamic probit approach

Orla May and Merxe Tudela

October 2005

Bank of England

When is mortgage indebtedness a financial burden to British households? A dynamic probit approach

Orla May[*]

and

Merxe Tudela[**]

Working Paper no. 277

[*] Bank of England.
Email: orla.may@bankofengland.co.uk

[**] Bank of England.
Email: merxe.tudela@bankofengland.co.uk

This paper represents the views and analysis of the authors and should not be thought to represent those of the Bank of England. The authors are grateful to Wiji Arulampalam, Ian Bond, Charles Goodhart, Vassilis Hajivassiliou, Rob Hamilton, Stephen McKay, Jeffrey Wooldridge, Garry Young, seminar participants at the Bank of England, HM Treasury, Inland Revenue and 2005 British Household Panel Survey conference and two anonymous referees for fruitful discussions and helpful suggestions. This paper was finalised on 15 August 2005.

The Bank of England's working paper series is externally refereed.

Information on the Bank's working paper series can be found at www.bankofengland.co.uk/publications/workingpapers/index.html.

Publications Group, Bank of England, Threadneedle Street, London, EC2R 8AH; telephone +44 (0)20 7601 4030, fax +44 (0)20 7601 3298, email mapublications@bankofengland.co.uk.

©Bank of England 2005
ISSN 1368-5562

Contents

Abstract

Since the mid-1990s the volume of secured lending to households has expanded rapidly, both in absolute terms and in relation to household incomes. This paper examines the determinants of households' ability to service this stock of secured debt. It estimates a random effects probit model for the probability of households having mortgage payment problems. It is found that past experience of payment problems increases the probability that the household has difficulties servicing its secured debt today. At the household level, becoming unemployed, interest income gearing of 20% and above, high loan to value ratios and having a heavy burden of unsecured debt are all associated with a significantly higher probability of mortgage payment problems. Saving regularly and having unsecured debt which is not a problem are both associated with a significantly lower probability of mortgage payment problems. The only non-household-specific variable to have a significant effect is mortgage interest rates – the probability of payment problems increases with the level of mortgage interest rates. An aggregate measure of debt at risk is calculated. This has decreased between 1994 and 2002, as falls in the probability of mortgage payment problems have more than offset increases in the stock of mortgage debt outstanding. It is found that the fall in the probability of mortgage payment problems has been greatest among the most highly indebted households.

Key words: Mortgage debt; dynamic probit.

JEL classification: D14; C23, C35.

Summary

Since the mid-1990s the volume of secured lending to households has expanded rapidly, both in absolute terms and in relation to household incomes. In 2004, the stock of secured lending to households exceeded £850 billion (compared to around £400 billion in 1995) and represents the largest domestic on balance sheet exposure of UK-owned banks. The rates of arrears and write-offs on secured debt have fallen in recent years and, despite a slight pickup in the second half of 2004, are currently at historically low levels. But there is a risk these could rise further if households began to encounter problems servicing their mortgage debt.

This paper seeks to explain the determinants of mortgage payment problems using disaggregated data from the British Household Panel Survey (BHPS). By using disaggregated data, we can examine how both macroeconomic factors (such as interest rates and house prices) and household-level factors (such as employment status and saving behaviour) affect the probability of households meeting their mortgage commitments. Since the BHPS is a panel survey, it allows us to track the same individuals over time; so we can also examine the dynamics of mortgage payment problems. In particular, we can analyse whether changes in a person's circumstances (such as changes in income) and previous experience of payment problems affect their current ability to service mortgage debt.

The data confirm that the two most important household-level factors associated with mortgage payment problems are adverse changes in employment and the level of income gearing (the ratio of mortgage payments to household income). Becoming unemployed significantly increase the probability of mortgage payment problems. But the results show that if the household is persistently unemployed this is not associated with a higher probability of payment problems, presumably because the household can adjust consumption so that servicing the mortgage is no longer a problem. However, this result may be driven by the fact that there are only a small number of mortgagors in our sample who are unemployed for two or more years.

We find evidence of a positive relationship between income gearing and the probability of mortgage payment problems – a higher level of income gearing significantly increases the probability of payment problems. However, this relationship is only apparent when gearing passes 20% – below that level there is no significant effect on payment problems from income gearing.

The level of effective mortgage interest rates is also found to increase the probability of mortgage payment problems. This is the only non-household-specific variable that is found to have a significant effect. The aggregate level of unemployment has no independent effect beyond that identified at the household level.

The results also show that problems paying for secured debt are persistent. The experience of payment problems has a genuine behavioural effect upon the household in the sense that previous experience of problems increases the probability that the household will subsequently have difficulty servicing its mortgage. There are a number of possible explanations for this. Past experience of problems could affect access to credit if lenders use information about previous payment difficulties in their lending decisions. Alternatively, the experience of problems could lessen any stigma attached to payment difficulties and this could make the household less careful in avoiding these in the future. The evidence implies that policies addressing mortgage payment problems can have long-lasting effects.

We find no evidence for collateral effects: neither the amount of housing equity nor the presence of negative equity affects the probability of mortgage payment problems (although they will affect loss given default). This result is new and contrasts with previous work which has identified housing equity as a determinant of the aggregate level of mortgage arrears. This difference may be due to the sample period we use. The BHPS contains information on housing equity from 1993 onwards, so it does not allow us to directly measure the effects of falling house prices between 1990 and 1993 upon mortgage payment problems. It is possible that falling housing equity had already affected some mortgagors' ability to service their debts before 1993 and that these households would not appear in our sample.

We use the estimation results to construct a measure of mortgage debt at risk. Changes in the probability of payment problems and in the amount of secured debt held will both affect the amount of debt at risk. Over the sample period 1994 to 2002, we find that mean debt at risk has fallen. This implies that the probability of mortgage payment problems has fallen sufficiently to offset the effects of increasing mortgage debt over the same period. There is also evidence that mortgage debt is now concentrated in less risky households. This implies that the short-term financial stability risks associated with the stock of mortgage debt in 2002 are lower than in the mid-1990s.

1 Introduction

The biggest exposure of the UK banking system to the UK household sector comes through secured lending to households. The stock of secured lending to households has increased rapidly in recent years and in 2004 Q4 exceeded £850 billion. Secured debt represents the vast majority of institutions' exposures to the household sector – in 2004 around 80% of total lending to individuals was secured on houses. Changes in households' ability to service their secured debt may have implications for *financial stability* (if households with mortgage payment problems fall into arrears or default on their debts) and for *monetary policy* (if increases in the burden of mortgage debt cause households to cut back on consumption).

In this paper we analyse the determinants of financial risks from mortgage indebtedness and present estimates of the proportion of secured debt that is most at risk of default. Specifically, we use the British Household Panel Survey (BHPS) to study the determinants and dynamics of mortgage payment problems. We exploit the panel feature of the survey to identify causal relationships and evaluate the persistence of housing finance problems.

Much of the existing literature uses time-series or cross-sectional data to analyse the determinants of housing finance problems. To our knowledge only one published paper, Böheim and Taylor (2000), has used the BHPS to study the determinants of mortgage payment problems. They estimated a random effects probit model for the limited dependent variable that the household reports housing finance problems and found a significant positive association between previous experience of housing finance problems and current financial distress. However, a major shortcoming of Böheim and Taylor (2000)'s analysis is that they included a lagged dependent variable in their model without properly addressing the econometric problems this creates in relation to initial conditions and spurious state dependence. State dependence can be spurious if we do not control for unobserved heterogeneity and possible autocorrelation in the error term.

In the case of persistent mortgage payment difficulties over time, we need to understand the source of that persistence. There are two possible, distinct, explanations for this. First, if a household has experienced mortgage payment problems, those problems might entail subsequent constraints and conditions which alter the household's ability to meet its mortgage commitments. In this case, past experience would have a genuine behavioural effect in the sense that an otherwise identical household that had not experienced mortgage payment problems would behave differently from one that had experienced such problems. This is known as *true state dependence*.

Second, apparently identical households might differ in their propensity to incur mortgage payment problems. For econometric reasons, we have to distinguish between two components here. The first is related to the existence of unobserved household-specific attributes that are time-invariant, known as *unobserved heterogeneity*. The second is that household-specific attributes may be correlated over time. If this problem is not addressed properly, past episodes of mortgage payment problems might turn out to be significant solely because they are a proxy for autocorrelated unobservables.

We estimate a reduced-form model for mortgage payment problems that fully exploits the panel structure of the BHPS, allowing for state dependence and unobserved heterogeneity. We then use the estimated model to predict the probability that a household experiences problems meeting its mortgage payments. We identify the marginal effect of particular variables upon this probability and analyse how changing one of these variables would affect the probability of households experiencing difficulties. We also use the estimated model to construct a measure of mortgage debt at risk and examine its distribution.

The remainder of the paper is organised as follows. Section 2 summarises previous work on mortgage payment problems specifically and mortgage arrears more generally. Section 3 describes the data and gives a flavour of the persistence of mortgage payment problems. In Section 4 we briefly describe the econometric model. We present the main results in Section 5 and Section 6 concludes.

2 Related literature

The literature on mortgage payment problems and arrears is based strongly on empirical evidence, rather than underlying theoretical structure. The literature does however offer a broad theoretical framework for understanding mortgage payment problems and default.[(1)] It focuses on 'ability to pay' and 'equity' theories of default.

The 'ability to pay' theory suggests that households will have problems meeting their mortgage payments if their income flow is insufficient to meet these commitments without placing an undue burden on the household. Ability to pay considerations will matter only for liquidity constrained households: if liquidity constraints are not binding then the household could borrow further to smooth their income flow and alleviate any mortgage payment problems. The 'equity' theory of default suggests that households instead take a long-term view and base their default decision on a rational evaluation of the financial costs and benefits of continuing (or discontinuing) mortgage payments; the household will

(1) See Whitley, Windram and Cox (2004) for a fuller discussion.

default if this maximises its net financial return.

We focus here on reviewing the empirical literature on mortgage arrears and payment difficulties. We distinguish between those papers that use aggregate data to try to explain the general level of arrears and those that use household-level data and focus on individual characteristics to determine the likelihood of a household falling into mortgage payment difficulties.

2.1 Aggregate data

At the aggregate level, the literature has found that both 'ability to pay' and 'equity' variables have significant effects in explaining the level of mortgage arrears. Breedon and Joyce (1992) used a three-equation model of house prices, mortgage arrears and possessions to study the house price boom of the late 1980s and the subsequent sharp rise in mortgage arrears and possessions and falls in nominal house prices in the early 1990s. The authors found strong interactions between the three variables: arrears and possessions were related to house price movements through the latter's impact on the value of housing equity, and house prices were affected by the influence of possessions on housing demand.

Whitley *et al* (2004) also found that housing equity has a significant effect on mortgage arrears. But their empirical model of mortgage arrears implied that mortgage income gearing was the most significant explanatory variable.[(2)] Other significant variables included the unemployment rate and the loan to value (LTV) ratio on loans to first-time buyers. Arrears were negatively linked to first-time buyers' LTVs and the authors suggested that this could reflect supply-side behaviour by banks, given that they may be more willing to extend higher LTV loans to better credit risks.

2.2 Disaggregate data

The findings from aggregate data are useful in showing how macroeconomic factors can explain movements in the aggregate level of mortgage arrears. But household-level factors may be equally important in determining the level of mortgage arrears – idiosyncratic shocks can cause households to experience payment problems independent of macroeconomic factors. Disaggregate data allows us to study the determinants of mortgage arrears at the household level, thus capturing both macroeconomic and idiosyncratic factors.

(2) An earlier study by Brookes, Dicks and Pradhan (1994) also found that income gearing was the most important determinant of changes in aggregate mortgage arrears.

Coles (1992) used the results of a 1991 Council of Mortgage Lenders' (CML) survey to assess the factors contributing to mortgagors falling into arrears. The CML survey asked 20 UK lenders about the profile of mortgagors in arrears or whose property had been taken into possession. The survey took place in December 1991 (when mortgage arrears of twelve months or more peaked). Coles found that a high loan to value ratio was the most important single characteristic of loans going into arrears and properties being taken into possession. First-time buyers who had entered the market in 1988–89 were particularly exposed to this risk. He also found that income shocks (and income uncertainty) were important – unemployment and relationship breakdowns could each explain around 25% of arrears, and those in arrears were typically self-employed, working in an industry with exposure to the construction industry, or working in sales-orientated businesses (where commission made up a significant proportion of income).

Ford, Kempson and Wilson (1995) use the results of surveys of lenders and borrowers to study the characteristics of borrowers in arrears, compared to borrowers who were not in arrears. The authors also found that unemployment, income shocks and having bought property between 1988 and 1999 were important factors. Those in part-time work or who were self-employed were found to be particularly prone to income shocks and hence of falling into arrears. They found an association between arrears and relationship breakdowns, but note that the causality is unclear – there were very few cases in which arrears were directly caused by the relationship breakdown.

Burrows (1997) studied the determinants of mortgage arrears using a subsample of the 1994–95 Survey of English Housing. His sample comprised around 8,000 households with a mortgage of whom 1.9% were in arrears of three months or more, 4.1% were in arrears of any sort and 17.8% were either in arrears or having difficulties keeping up with their mortgage commitments. Burrows used a logistic regression to model the likelihood of households being in arrears of three months or more. The results suggested that households were more likely to be in arrears if they had a 100% mortgage, were employed part-time or unemployed or unable to work, worked in the private sector (relative to the public sector), or had bought their property between 1987–89. He also found some evidence of state dependence: those households containing members who had previously been subject to mortgage possession were more likely to be in arrears than other households.[(3)]

[(3)] Among the variables that were not significant in explaining the odds of being in arrears were: the age of the head of household, whether the head was a first-time buyer, the marital status of the head, the social class of the household head, his ethnicity, whether the property was bought under a right-to-buy scheme, the council tax band of the property and the region of residence.

Böheim and Taylor (2000) used the BHPS (1991–97) to identify the causes and consequences of falling into housing payment difficulties for both mortgagors and tenants. They estimated dynamic random effects probit models [4] for the probability of experiencing housing payment problems and for the probability of eviction, pooling mortgagors and renters together in each case. Unlike Burrows (1997), Böheim and Taylor found that age was important: households with older heads were less likely to experience housing finance problems. Higher household income and two-earner households were also less likely to experience housing finance problems. Past financial problems had a strong positive influence on the probability of experiencing current financial problems and the risk of eviction.

3 The data

In our estimates we use data from the BHPS. The BHPS is a panel of British individuals and households providing information on the social and economic characteristics of the British population. It was constructed to be representative of the British household population and consists of twelve waves as of 2004. [5] Waves are set at annual intervals, with wave one corresponding to 1991.

This initial wave consists of an equal-probability clustered sample of 8,167 addresses drawn from the Postcode Address File for Great Britain south of the Caledonian Canal (therefore excluding Northern Ireland and the North of Scotland). Non-residential or institutional addresses were excluded from the survey. The total number of interviews conducted at wave one was 10,264 – encompassing 5,505 households. In order to maintain the representativeness of the BHPS, all original members at wave one remain sample members at subsequent waves until they die (specific rules exist for following individuals who move addresses).

From wave seven onwards, a new sample was added to the BHPS when the BHPS began providing data for the United Kingdom European Community Household Panel (UKECHP). As a result it incorporates a subsample of the original UKECHP consisting of all sample households in Northern Ireland and all ‘low-income’ sample households in Great Britain. Furthermore, from wave nine onwards two additional subsamples were added to the original BHPS. These are the Scotland and Wales extension samples to permit

[4] Böheim and Taylor (2000) do not seem to control for the implications of autocorrelated errors in the estimation of the state dependence coefficient. Neither do they explain (or correct) for the possibility of the unobserved heterogeneity being correlated with the time-varying explanatory variables nor the initial conditions problem. For a discussion of these issues and their implications see Section 4.

[5] See Buck, Burton, Laurie and Lynn (2002) for a summary of the sample design and contents of the BHPS.

independent analysis of the two countries. At wave eleven a substantial new sample in Northern Ireland, the Northern Ireland Household Panel Survey (NIHPS), was added.[6] We exclude these booster samples from our analysis.

We restrict the sample to those households with a mortgage, as we are concerned with default on mortgage debt specifically (rather than housing payment problems in general).[7] Our unit of analysis is the household and all individual characteristics referred to correspond to those of the head of the household, except where otherwise indicated. We follow the head of household year on year as long as he or she remains a mortgagor and we have the relevant data. We therefore use an unbalanced panel, allowing the individual (the head of household) to both exit and enter the sample, but we only allow for one spell per individual.[8]

Table A: Incidence of mortgage payment problems

Year	Sample size	% with[a] problems	% 2^+ months in arrears	As % of those with problems		
				Cutbacks	Borrowing	2^+m in arrears
1991	2,265	16.5	4.1	80.6	20.7	25.3
1992	2,213	16.0	4.0	83.7	16.2	25.3
1993	2,091	13.1	3.9	83.2	13.5	30.2
1994	2,106	11.1	2.9	80.2	14.8	26.2
1995	2,027	9.6	1.7	79.1	16.1	18.3
1996	2,067	7.7	1.2	84.2	18.3	16.2
1997	2,077	7.5	0.9	81.3	18.3	12.0
1998	2,058	7.1	0.7	81.4	23.1	9.8
1999	2,037	5.8	0.8	71.2	16.2	14.0
2000	1,984	6.0	0.7	73.9	20.3	10.8
2001	1,969	5.3	0.9	70.4	16.7	16.6
2002	1,963	4.3	0.6	76.3	31.6	15.0

[a] Weights are used whenever descriptives are calculated. See Redwood and Tudela (2004) for further details of the weights used.

Sources: BHPS and Bank calculations.

At each interview every household's reference person[9] is asked about the household's housing payment costs and whether these represent a problem for the household. They are asked '*Many people these days are finding it difficult to keep up with their housing*

[6] We refer the reader to Taylor, Brice, Bruck and Prentice-Lane (2001) for a detailed description of the sample procedure.

[7] Households that move house remain in our sample so long as they remain mortgagors.

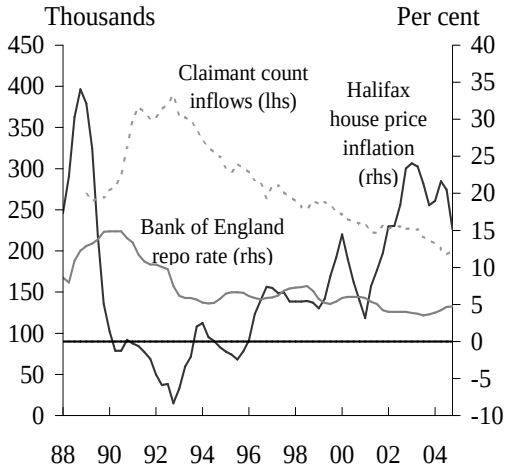
[8] Arulampalam, Booth and Taylor (2000) also use an unbalanced panel but they only allow individuals to exit the sample; as a result, all individuals have a common date of entry to the panel. The authors do so in order to facilitate estimation of initial conditions. We instead control for macroeconomic conditions around the entry date in the estimation of initial conditions and allow for entry at different dates (see Section 4 for further details). This may give rise to possibly non-random attrition, however it is difficult to correct for this in the sort of dynamic probit model we use.

[9] The principal survey respondent for the household.

payments. In the last twelve months would you say you have had any difficulties paying for your accommodation?' If the respondent answers yes, then he or she is asked whether the household has had to cut back on other household spending to meet their mortgage payments; whether it has had to borrow money to meet its housing payments; and whether the household has fallen two or more months in arrears on their housing payments in the past year. These questions are not mutually exclusive – the respondent can give positive answers to one or more questions.

As Table A shows, the proportion of mortgagors reporting difficulties in paying their mortgage has fallen steadily through the BHPS sample period, from 16.5% in 1991 to 4.3% in 2002. The proportion of all mortgagors who report being two or more months in arrears has also fallen – from 4% in 1991 to just 0.6% in 2002. This reflects the relatively benign macroeconomic conditions during our sample period, which was characterised by low nominal interest rates, stable or rising house prices and low inflows into unemployment (see Chart 1).

Chart 1: Macroeconomic conditions



Sources: Bank of England, Halifax and National Statistics.

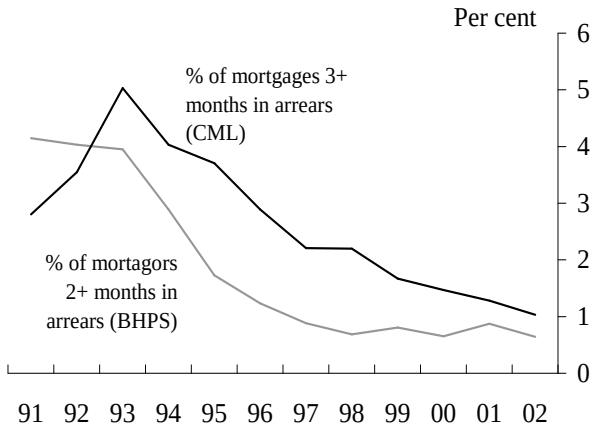
The vast majority of mortgagors alleviate mortgage payment problems by cutting back on consumption,[10] but in each year a significant minority of those with problems (on average 19%) needed to borrow further.[11] The proportion of mortgagors with payment

[10] Ford *et al* (1995) also found that cutting back or postponing consumption was the most common coping strategy adopted by mortgagors in arrears or whose homes had been possessed.

[11] The proportion of mortgagors who borrowed further to meet their mortgage commitments increased significantly between 2001 and 2002 and there was also a small increase in the underlying number of mortgagors who had borrowed further. In 2002 22 mortgagors out of 57 reporting payment problems borrowed further, compared to 16 out of 87 mortgagors in 2001. Although the results are not strictly comparable, a 2004 survey conducted by NMG Research on behalf of the Bank also found that a high proportion (29%) of mortgagors with payment problems had borrower further to ease their problems (see May, Tudela and Young (2004)).

problems who also report being two or more months in arrears on their mortgage has fallen throughout the BHPS sample period from 25.3% in 1991 to 15% in 2002 (see Table A). Very few mortgagors who report problems have not responded in some way – on average 13% of mortgagors who reported having problems did not cut back on consumption, borrow further or go into arrears, whereas 3.6% of those with problems reported all three.

Chart 2: Mortgage arrears: BHPS and CML



Sources: BHPS, CML and Bank calculations.

We can compare the BHPS measure of mortgage arrears with the CML aggregate measure of mortgage arrears to check the robustness of the BHPS results. The two measures are not strictly comparable as the BHPS measure shows the proportion of *mortgagors* who are *two* or more months in arrears, whereas the CML measure reflects the proportion of *mortgages* that are *three* or more months in arrears, but it is reassuring to see that they are positively correlated (see Chart 2). The CML measure of arrears peaks in 1993, slightly later than the BHPS series (possibly because it measures arrears of longer duration), but like the BHPS measure of arrears it has fallen steadily since 1993.

This is also reassuring from the point of view of econometric estimations (to be discussed in Section 4). Hausman (2001) points out that that if the left-hand side variable in a probit or logit model is misclassified and we estimate these models without allowing for misclassification, the result will be biased and inconsistent estimates. Our dependent variable is based on subjective responses (whether the household reports having difficulty paying for its accommodation) and is therefore subject to potential misclassification. However, Hausman, Abrevaya and Scott-Morton (1998) demonstrate that maximum likelihood estimation of this model provides consistent estimates if the combined probability of misclassification (ie classifying a household as having problems paying for its mortgage when it does not have such problems and classifying a household as not having problems paying for its mortgage when it does) is not so high that on average one

cannot tell which result actually occurred. This seems to be the case as shown in Chart 2: the BHPS seems to capture the general aggregate trend in mortgage arrears. Moreover, the downward trend in mortgage payment problems in the BHPS has also been observed in the Survey of English Housing, an annual survey comprising around 8,000 mortgagor households in England (see Robinson, Humphrey, Kafka, Oliver and Bose (2004)).

Table B: Persistence of mortgage payment problems

	Conditional proportions [a]			
Year	$P(y_t = 1 \mid y_{t-1} = 1)$	$P(y_t = 1 \mid y_{t-1} = 0)$	$P(y_t = 1 \mid y_{t-2} = 1)$	$P(y_t = 1 \mid y_{t-2} = 0)$
1992	0.58	0.08		
1993	0.57	0.05	0.42	0.08
1994	0.51	0.06	0.37	0.06
1995	0.44	0.05	0.37	0.06
1996	0.37	0.04	0.36	0.03
1997	0.43	0.05	0.28	0.05
1998	0.43	0.04	0.34	0.05
1999	0.48	0.03	0.38	0.04
2000	0.45	0.04	0.34	0.04
2001	0.43	0.03	0.36	0.03
2002	0.34	0.02	0.29	0.02

[a] Where $y_t = 1$ if the household had problems paying for its mortgage at time t, and zero otherwise. Therefore $P(y_t = 1 \mid y_{t-1} = 1)$ is the proportion having problems at t, conditional on having problems at $t - 1$.

Sources: BHPS and Bank calculations.

Table B gives a *prima facie* indication of the persistence of mortgage payment problems. The first column shows the proportion of mortgagors that report problems in a given year conditional on having reported problems the previous year. The results suggest that persistence has become less of an issue over the BHPS sample period, in the sense that the proportion of households that had problems in at least two consecutive years has declined (from nearly 60% in 1992 to around one third in 2002). By contrast, the proportion of mortgagors having problems in time t conditional on not having problems in time $t - 1$ is very low throughout the sample period (see second column of Table B). The third (fourth) column of Table B shows the proportion of mortgagors having problems in t conditional on (not) having problems in $t - 2$. These columns again indicate, based on the raw data, that there is persistence in having difficulties meeting mortgage commitments.

Table C shows the proportion of households who experience mortgage problems by selected characteristics. We can compare this to the proportion of all mortgagors that have payment problems (the bottom row of Table C) to get a sense of the characteristics of households in whom problems are concentrated.

Table C: Percent of mortgagors that have difficulties paying for their mortgage by selected characteristics

	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Female [(a)]	22.2	17.7	17.5	14.8	14.2	15.3	13.9	11.7	9.1	10.4	7.7	8.8
Minority ethnic	28.3	30.3	23.7	21.9	13.2	13.9	13.8	7.2	11.7	11.7	14.8	5.0
Low/no qualified	21.3	21.0	17.0	17.7	15.3	11.7	9.3	9.0	8.6	10.7	9.8	3.9
Self-employed	21.2	20.9	18.7	12.7	12.3	7.0	5.3	7.2	8.6	4.3	5.5	5.7
Unemployed	45.1	49.3	32.7	49.6	34.7	39.7	25.3	4.5	21.4	28.0	18.2	21.9
Full-time student	34.7	21.1	6.1	9.2	19.4	19.9	28.3	16.4	24.0	11.6	4.6	-
Disabled	33.8	30.0	32.1	30.3	28.9	29.6	20.6	21.1	18.8	19.3	13.3	4.3
Lost job	n/a	47.9	34.0	52.4	31.2	31.2	21.7	7.2	20.1	31.7	18.9	24.5
Unemployed in t and $t-1$	n/a	59.5	34.8	50.2	40.8	48.3	32.1	-	14.7	13.6	19.3	27.8
Partner is unemployed	25.6	20.4	27.0	21.8	9.7	8.2	26.8	18.7	6.9	-	15.2	-
Partner disabled	15.6	30.1	19.2	21.1	19.0	20.2	6.3	4.7	8.7	10.1	1.3	11.7
Relationship breakdown	n/a	26.7	32.8	30.8	33.3	28.2	13.8	25.3	8.0	17.1	8.3	18.4
Unsecured debt is somewhat of a burden	n/a	n/a	n/a	n/a	15.3	12.9	13.7	11.2	7.3	12.3	10.8	12.2
Unsecured debt is a heavy burden	n/a	n/a	n/a	n/a	39.3	29.7	43.2	32.8	32.8	29.4	27.5	26.1
Unsecured debt is somewhat of a burden, $t-1$	n/a	n/a	n/a	n/a	n/a	7.1	13.7	6.9	7.2	7.2	10.8	9.9
Unsecured debt is a heavy burden, $t-1$	n/a	n/a	n/a	n/a	n/a	27.6	38.1	37.1	37.5	25.9	14.8	18.8
All mortgagors	16.5	16.0	13.1	11.1	9.6	7.7	7.5	7.1	5.8	6.0	5.3	4.3

[(a)] Characteristics are those of the head of the household except where otherwise indicated.
Sources: BHPS and Bank calculations.

The results suggest that the proportion of households reporting problems is higher if the head of the household is female (comparing the first row in Table C to the bottom row). The proportion of households reporting problems is also higher if the head of the household is minority ethnic, has low/no qualifications, is a full-time student or long-term sick or disabled. Income shocks seem to be important: a higher proportion of households report problems if the head of household is currently unemployed, has become unemployed during the year, or has remained unemployed from the previous year. Some characteristics of the partner are also important, such as if he/she is disabled or unemployed. And the proportion of mortgagors reporting difficulties is also higher among those individuals who have experienced a relationship breakdown (such as a divorce or separation). There appears to be a correspondence between secured and unsecured debt problems: among households whose unsecured debt was a burden (heavy or somewhat of a burden) over the current or the previous year there is a higher proportion of households

reporting mortgage payment problems.

Both Burrows (1997) and Böheim and Taylor (2000) found that self-employed individuals had a higher probability of experiencing mortgage payment difficulties. The results in Table C suggest that this has not been the case in recent years – since 1998 the percentage of self-employed mortgagors reporting payment problems has been lower than that for all mortgagors.

4 Method

Our model of mortgage payment problems for individual i at time t is

$$y_{it}^{*} = x_{it}'\beta + \gamma y_{it-1} + \varepsilon_{it} \qquad i = 1, \ldots, I \quad t = 1, \ldots, T_i \qquad \textbf{(1)}$$

where y^{*} denotes the unobservable propensity to incur mortgage payment problems, x is a vector of observable covariates affecting y^{*}, β is the vector of coefficients associated with x and ε is the unobservable error term. An individual experiences mortgage payment problems ($y_{it} = 1$) if the latent propensity, y^{*}, exceeds a threshold, normalised to zero in this case. Given the high degree of persistence in mortgage payment problems (Table B), we also assume that the propensity of having mortgage payment problems depends on the experience of mortgage problems in the previous year (y_{it-1}). The inclusion of the lagged dependent variable on the right-hand side of **(1)** allows us to test for the presence of state dependence.

There are three possible distinct explanations for state dependence. First, if a household has experienced mortgage payment problems, then constraints and conditions relevant to the household meeting its mortgage commitments might be altered. In this case, past experience has a genuine behavioural effect in the sense that an otherwise identical household that did not experience mortgage payment problems would behave differently from one that did experience such problems. This is known as *true state dependence*. True state dependence could arise for a number of reasons. Supply-side factors could play a role if the household's experience of problems affects their future access to credit and the terms on which credit is available – for example if they fall behind with payments and this information is subsequently used to inform lending decisions. Demand factors could also cause true state dependence, for example if the stigma associated with defaulting upon payments is lessened by the borrower having experienced, and survived, previous episodes of payment difficulty. Alternatively, if households who report problems borrow further (as

shown in Table A), this could leave them more exposed to problems in future periods.[12]

Second, state dependence could arise because the household has experienced a single period of mortgage payment problems lasting for two (or more) years. This could yield a spurious estimate of state dependence if a large proportion of mortgagors with payment problems experience problems lasting more than one year. Previous work using the BHPS has been able to check whether sequential observations of unemployment (Arulampalam *et al* (2000)) or self-employment (Henley (2004)) formed discrete spells and so could circumvent this problem. Unfortunately, in the case of mortgage payment problems we have no way of identifying whether sequential observations represent different episodes of payment problems or whether they are observations at different points in the same spell. As a result, our estimates may over-state the extent of true state dependence.

Third, state dependence may arise because households differ in their propensity to incur mortgage payment problems. For econometric reasons, we have to distinguish between two components here. The first is related to the existence of unobserved household-specific attributes that are time-invariant – *unobserved heterogeneity*. The second component takes into account the fact that household-specific attributes may be correlated over time. If this problem is not addressed properly, then past episodes of mortgage payment problems might turn out to be significant solely because they are a proxy for these unobservables.

In order to identify *true* state dependence, we need to specify the error term correctly. We assume it has the following structure:

$$\varepsilon_{it} = \alpha_i + \eta_{it}, \qquad \eta_{it} = \rho\eta_{it-1} + \xi_{it} \qquad \textbf{(2)}$$

The individual-specific component (α_i) allows for *unobserved heterogeneity*, while the term η_{it} captures shocks correlated over time. In **(2)** α_i is treated as random, $\alpha_i \sim N(0, \sigma_\alpha^2)$, α_i and η_{it} are independent, the η_{it} are independent of x_{it} for all i and t, and $\xi_{it} \sim N(0, \sigma_\xi^2)$.

In a simple random effects model it is assumed that α_i is also independent of x_{it} for all i and t.[13] If this assumption is violated, maximum likelihood estimates will be inconsistent since the estimated β coefficients will pick up some of the unobservable α_i. For example (following Arulampalam *et al* (2000)), suppose that α_i represents individual responsibility

[12] This would however be captured by x if this included the total amount of household debt, however, as discussed in Section 5, we only observe the stock of unsecured debt in 1995 and 2000.
[13] This assumption is made by Böheim and Taylor (2000).

and being responsible makes the individual more likely to be employed and therefore less prone to incur mortgage payment problems. If the model does not allow for correlation between α_i and employment status, then it will suffer from omitted variable bias. We therefore do not impose the assumption that α_i is also independent of x_{it} for all i and t. Instead, following Chamberlain (1984), we model the dependence between α_i and x_{it} by assuming that the regression function of α_i is linear in the means of all time-varying covariates and therefore can be expressed as:

$$\alpha_i = a_0 + a_1' x_{i\cdot} + \upsilon_i, \qquad \textbf{(3)}$$

where a_0 is the intercept, $x_{i\cdot}$ is the mean value of x_{it} over time, and υ_i is a residual term that will act as the former individual-specific effect, α_i. We assume that $\upsilon_i \sim N(0, \sigma_\upsilon^2)$ and is independent of the x_{it} and the η_{it} for all i and t.

We can now write our model as:

$$y_{it}^* = x_{it}'\beta + \gamma y_{it-1} + a_1' x_{i\cdot} + \upsilon_i + \rho_i \eta_{it-1} + \xi_{it} \qquad i = 1, \ldots, I \quad t = 1, \ldots, T_i \qquad \textbf{(4)}$$

where we have absorbed the intercept a_0 into the β vector.

4.1 The initial conditions problem

A further problem arising from equation **(4)** is the 'initial conditions' problem.[14] This problem arises because the start of our observation period does not necessarily coincide with the start of the stochastic process that generates the sequence of observations of mortgage payment status. A large proportion of individuals in our sample had a mortgage prior to entering our sample and therefore were at risk of incurring mortgage payment problems before entering our sample. If an individual is already experiencing mortgage payment problems the first time we observe him, this may be due to his previous experience of problems (state dependence) or it may be due to observable and unobservable information prior to the date we first observe him. To account for this problem we follow Heckman (1981) and explicitly model the initial condition.[15]

---

[14] Ignored by Böheim and Taylor (2000).
[15] See also Arulampalam *et al* (2000) and Henley (2004).

We first specify a reduced-form equation for the initial observation (y_{i1}^*):

$$y_{i1}^* = \lambda' z_i + \omega_i \qquad \textbf{(5)}$$

where z is a vector of strictly exogenous instruments, which includes variables from the period in which we first observe the individual, pre-sample information and the vector of means x_i.(to allow for any correlation between the time-varying covariates and unobserved heterogeneity).[16] We assume that ω_i has variance σ_ω^2, and we allow for non-zero correlation, ϱ, between υ_i and ω_i as follows:

$$\omega_i = \theta \upsilon_i + \xi_{i1} \qquad \textbf{(6)}$$

By construction υ_i and ξ_{i1} are orthogonal to one another and ξ_{i1} is independent of x_{it} and $\theta = \varrho\sigma_\omega/\sigma_\upsilon$ and $var(\xi_{i1}) = \sigma_\omega^2(1-\varrho^2)$. The initial conditions equation is then:

$$y_{i1}^* = \lambda' z_i + \theta \upsilon_i + \xi_{i1} \qquad i = 1, ..., I \ \ and \ \ t = 1 \qquad \textbf{(7)}$$

Equations **(4)** and **(7)** could be estimated by maximum likelihood. However, this estimation procedure requires special software to be written. We follow Arulampalam *et al* (2000) and Henley (2004) and apply the two-step pseudo-ML estimator proposed by Orme (1997) in the spirit of Heckman's standard sample selection correction method, which is an approximation in the case of small values of ϱ.

In this way, equation **(6)** is transformed to:

$$\upsilon_i = \delta \omega_i + \mu_i \qquad \textbf{(8)}$$

where $\delta = \varrho\sigma_\upsilon/\sigma_\omega$ and $var(\mu_i) = \sigma_\upsilon^2(1-\varrho^2)$. We can now express equation **(4)** as:

$$y_{it}^* = x_{it}'\beta + \gamma y_{it-1} + a_1' x_{i\cdot} + \delta\omega_i + \mu_i + \rho_i \eta_{it-1} + \xi_{it} \qquad i = 1, \ldots, I \quad t = 1, \ldots, T_i \qquad \textbf{(9)}$$

[16] For identification purposes z_i should include some variables that are not in x_{it}.

As Orme (1997) notes, equation **(9)** now has two individual-specific random error components, ω_i and μ_i. Also, the assumption of bivariate normality of (ω_i, υ_i) implies that $E(\mu_i|y_{i1}) = 0$ and that $E(\omega_i|y_{i1}) = e_i$,[17] where $e_i = (2y_{i1} - 1)\varphi(\lambda' z_i)/\Phi(\{2y_{i1} - 1\}\lambda' z_i)$, the generalised probit residual from the probit estimation of equation **(5)**. So after estimating **(5)** we can generate the generalised probit error and this replaces ω_i in equation **(9)**. A formal test of the exogeneity of initial conditions is provided by a standard t-test of the significance of δ.

In the absence of autocorrelated errors, $\rho = 0$ in equation **(9)**, we can estimate **(9)** using standard random effects probit software. But the estimation of the full dynamic probit model as described in **(9)** where ρ is not necessarily zero requires the evaluation of T-dimensional integrals of Normal density functions. For values of T greater than three the computational burden makes the estimation of such models infeasible. We therefore resort to simulation methods and specifically we use the simulation estimation method of maximum smoothly simulated likelihood (MSSL) in conjunction with the Geweke-Hajivassiliou-Keane (GHK) simulator. The derived estimates are asymptotically efficient (see Geweke and Keane (2001) and Hajivassiliou (2002) for further details).

5 Results

5.1 The estimates

To derive the explanatory variables used in our baseline model we use BHPS data for the years 1992 to 2002. We lag all personal and economic time-varying individual characteristics by one period in order to ensure that we identify individual characteristics before the household experienced problems and so identify a true lead-lag relationship.[18] Similarly, we need to use two years of data to construct dummy variables for a relationship breakdown and for moving into unemployment: relationship breakdown in t is defined using the change in marital status from $t - 2$ to $t - 1$ (and similarly for change in employment status). As a result, the estimation sample is reduced by two years to 1994–2002.

Table D presents the results of our baseline model. The first column lists the variables

[17] The consistency of the estimates hinges on the assumption of bivariate normality of (ω_i, υ_i).

[18] The BHPS question about housing payment problems refers to problems '*in the last twelve months*' so we need to lag characteristics to ensure we are capturing the household's characteristics prior to its experience of problems. If we instead used contemporaneous variables for the individual characteristics, then, depending on the timing of the household's problems relative to the time it was interviewed, we could observe the household status subsequent to its experience of problems. In this situation it would be difficult to identify causal relationships.

included in the estimation.[19] The second and third columns present coefficients and t-statistics for the initial conditions probit as described in Section 4.1. The last two columns present the coefficients and t-statistics derived from the dynamic probit model that allows for unobserved heterogeneity and autocorrelated errors. It is these last two columns which we now focus on.

The results in Table D validate our estimation approach – there is evidence of unobserved heterogeneity. About 34% of the total variance is explained by the unobserved household-specific characteristics (α_i, which could include factors such as stigma or ability to manage household finances, or simply differences in attitude to what constitutes a problem and willingness to report any problems to the interviewer). In Böheim and Taylor (2000)'s study of housing payment problems, there was also evidence of unobserved heterogeneity – they found that 19% of the total variance was due to household-specific unobservables. The large effects identified in both studies demonstrate the importance of following households over time to study the incidence of mortgage payment problems and the adequate use of panel data methods.

In contrast to Böheim and Taylor (2000), we allow the error component to be autocorrelated over time to control for household-specific unobservables that might be correlated over time but a likelihood ratio test indicates that this term is not significantly different from zero at the 5% level.[20] The generalised probit error is significant at the 1% level, highlighting the relevance of modelling the initial conditions problem. In Section 4.1, we noted that the Orme two-step approach is an approximation in the case of small values of ϱ. Following Arulampalam *et al* (2000)'s approximation to calculate ϱ, we find the value of this parameter is 0.476 (and 0.390 for the results reported in Table E).[21]

After controlling for time-invariant unobserved heterogeneity and autocorrelated errors, we find evidence of persistence in mortgage payment problems: there is true state dependence. Böheim and Taylor (2000) also found that previous experience of housing finance problems (for renters and mortgagors) had a positive and significant impact on having problems today. The authors associated this result with evidence of poverty

[19] For a description of the variables see Table A.a in the appendix. Means and standard deviation of the same variables are presented in Table A.b.

[20] At the 10% level we can reject the hypothesis that this term is zero. Although the autocorrelated error is not generally statistically significant, it is interesting to note that it is negative. One explanation for the negative autocorrelated error term could be that if a household experiences mortgage payment problems at, say, time $t-1$, they then over-compensate in response (for example, by cutting back other consumption or being more careful in managing its finances) so that they are less likely to have problems paying for their mortgage in period t.

[21] In Arulampalam's work this parameter ranges from 0.182 to 0.555, depending on the variant of her model.

persistence: transitions from poverty are limited and associated with small moves within the income distribution. Burrows (1997) also found evidence of some state dependence: those households containing an adult who had previously been subject to mortgage possession were more likely to be in arrears than other households.[22]

Table D: Coefficient estimates — 1994–2002 sample, model 1

Variable	Initial conditions		Dynamic probit	
	coefficient	t-statistic	coefficient	t-statistic
Constant	-3.52^{***}	-4.07	-3.75^{***}	-24.66
			Lagged dependent variable	
Problems_{t-1}			0.76^{***}	4.92
			Loan to value ratios	
LTV 50%-69%	-0.03	-0.15	0.25^{*}	1.71
LTV 70%-89%	0.09	0.50	0.22	1.56
LTV 90+%	0.22	1.37	0.34^{**}	2.56
			Income gearing	
IG interest only$\geq 20\%$	0.44^{**}	2.42	0.20^{*}	1.76
IG principal$\geq 20\%$	1.09	1.63	0.37	1.08
			Other characteristics	
Saver	-0.34^{**}	-2.09	-0.19^{*}	-1.93
Lost job	0.18	0.51	0.57^{**}	2.53
Health problems	0.23	1.17	-0.16	-1.09
			Region of residence	
NE	-0.68^{**}	-2.31	0.11	0.50
Merseyside	-0.96^{**}	-2.28	-0.40	-1.01
York	-0.86^{***}	-3.09	0.04	0.21
EM	-0.13	-0.57	0.45^{**}	2.28
WM	-0.25	-1.08	0.39^{*}	1.95
E	0.04	0.16	0.40	1.58
London	-0.55^{**}	-2.11	0.18	0.82
SE	0.05	0.26	0.29^{*}	1.72
SW	-0.10	-0.47	0.44^{**}	2.24
Wales	-0.03	-0.14	0.27	1.17
Scotland	-0.54^{**}	-2.14	0.07	0.33
			Macroeconomic conditions	
House prices	3.25^{**}	2.03	-0.23	-0.27
Unemployment	0.15^{***}	3.52	0.04	1.59
Interest rates	1.71	1.42	1.71^{***}	2.92
			Other — initial conditions probit	
1987–89	0.38^{***}	2.58		
Post 1989	0.11	0.82		
Negative equity	0.16	0.64		
Relationship breakdown	-0.02	-0.07		
Dependents	0.28^{***}	2.82		
Low qualifications	0.29^{**}	2.10		

continued on next page

[22] In the sovereign context, Reinhart, Rogoff and Savastano (2003) also find evidence of state dependence – countries with worse track records in international capital markets suffer greater financial fragility due to increased borrowing costs at any given level of GDP.

Table D: ***continued***

Mid qualifications	0.01	0.12		
Male	−0.21	−1.63		
Non-white	0.08	0.26		
	Means of time-varying covariates			
IG interest only≥ 20%	0.53**	2.33	0.78***	4.30
IG principal≥ 20%	−0.95	−1.32	0.27	0.61
Saver	−0.40**	−2.02	−0.65***	−4.27
Lost job	0.65	1.07	0.36	0.67
Health problems	0.13	0.62	0.85***	4.62
Generalised residual			0.34***	4.23
Proportion of the total variance contributed by the panel-level variance component			0.34***	6.85
AR(1) error			−0.19	
Log-likelihood				−2763.54
Log-likelihood excluding the AR(1) error term				−2764.08
Number of observations	1,709			7,197
Number of households				1,709
Obs. per household: min.				1
Obs. per household: avg.				4.2
Obs. per household: max.				9

Notes: (i) The model allows for endogenous initial conditions which are estimated using a two-step procedure following Orme (1997). (ii) Correlation between the time-varying covariates and the unobservable heterogeneity is allowed for by including the time means of these variables. (iii) For details of the 'generalised residual' from the initial condition probit see text, page 20. (iv) ***, ** and * denote coefficient significant at the 1%, 5% and 10% significance levels respectively, for a two-sided test.

In an alternative specification (not shown) we interacted the lagged dependent variable with an age dummy variable in order to investigate whether the relationship between financial problems in previous and current years differs by age. Those results indicated that persistence in mortgage payment problems was greater among households in which the head was 35 years old or over than it was among households headed by younger individuals. That is, younger households are more capable of getting out of problems than those aged 35 or over. This might be linked to the fact that income growth is larger for younger households, which would tend to facilitate their exit from financial difficulties.[(23)] However a likelihood ratio test showed that the coefficients were not significantly different from each other, so we decided not to include an age dummy interaction in our baseline specification.

[(23)] Pooling all years together, the year-on-year percentage increase in income is about 12% for those aged 16–24 and 7% for 25–34 years old. For older households, the average year-on-year percentage increase in income is much smaller: for those aged 35–44 it is 5%, for 45–54 it is 4% and 3% for 55 years old and over. In absolute terms the increases in income year on year are also larger for younger households.

Loan to value ratios (LTV), defined as the ratio of the original mortgage (ie value of mortgage when first taken) to the original value of the house, are also significant, but only when we include them as a categorical variable. If we include loan to value ratio in levels, we do not find this variable to be significant. This points to a non-linear relationship between mortgage problems and LTV. Relative to the reference group of having an LTV less than 50%, having an LTV greater than 90% or of 50%–69% increases the probability of having problems (although in the latter case this is only marginally significant at the 10% level). Unlike Whitley *et al* (2004), we find no evidence that housing equity (defined as the current value of the mortgage to the current value of the house) explains mortgage payment problems. This may reflect the shorter sample period used in our estimations (Whitley *et al* (2004) estimate their model from 1985 to 2000) or it may simply reflect our use of disaggregated, as opposed to aggregate, data.[(24)]

A key variable for our analysis is the income gearing ratio (IG). The IG measures the cost of servicing mortgage debt relative to the income of the head of household (and that of his or her partner, if applicable). Mortgage servicing costs are reported by survey respondents and should comprise both interest payments and repayment of principal. We differentiate between these in our estimations to investigate whether mortgage interest payments and repayments of principal have different effects on the probability of reporting mortgage payment problems.[(25)] Since IG is time-varying, we also include the means of interest IG and principal IG as regressors to control for their possible correlation with the unobserved heterogeneity term.[(26)]

Interest IG had a positive significant coefficient when we included it as a continuous variable. But when we included interest IG as a categorical variable, we found that it was not significant at low values but that once IG reached higher levels it had a significant positive effect on mortgage problems. We tried using a number of different IG groupings in our estimations and found that 20% was the lowest level at which IG had a significant effect. Due to the small number of observations, we are unable to pin down the precise functional form of the relationship between interest IG and payment problems at high levels of gearing.[(27)] However, the data are consistent with interest IG having no effect upon problems up to a level of around 20% and then beyond this level there is a positive

(24) Among the other variables that do not help to explain mortgage payment problems (either in the initial conditions or dynamic probit) are gender, ethnicity, age cohort or receiving income support.
(25) We estimate interest and principal repayments by applying appropriate annuity formulae to the mortgage debt outstanding reported by the respondent.
(26) See page 17 in the method section for a technical explanation.
(27) In part because some households with high gearing are likely to drop out of our sample as they migrate into default.

relationship between the level of IG and payment problems.[28]

The coefficient of the IG variable (0.20) indicates the temporary effect of IG, whereas the sum of this and the coefficient of the mean across time of the IG variable (0.78) indicates the permanent or total effect of having interest IG of 20% or more. It is clear from the results in Table D that the total effect of interest IG is highly significant. IG related to repayment of principal being greater than 20% is not significant. This suggests that it is the payment of mortgage interest which affects the likelihood of mortgage payment problems, perhaps because interest payments are non-negotiable, whereas the household may be able to renegotiate the terms of principal repayments.

Böheim and Taylor (2000) did not include an IG measure in their regressions, but they did include household income in logs and found a negative association with housing finance problems (which the authors qualify as hardly unexpected) since the ability to meet housing payments is clearly related to the household's financial situation in general. Whitley *et al* (2004) found that IG is the most significant explanatory variable in the determination of the aggregate level of mortgage arrears.

Being a regular saver is significantly and negatively associated with mortgage payment problems, as expected.[29] Health problems also help to explain the likelihood of housing finance problems, but only the mean across time is significant: a temporary deterioration in health status fails to explain financial difficulties.

Having moved into unemployment in the previous year increases the likelihood of payment problems. Only the level of this dummy variable (ie the temporary effect of becoming unemployed) increases the probability of mortgage payment problems – the mean value of the variable is not significant. That is, it is the event of moving into unemployment and not the fact of being in unemployment which is a significant determinant of having difficulties paying for the mortgage.[30] Similar evidence is found in Böheim and Taylor (2000): the authors expected the severity of housing finance problems to increase with the duration of unemployment spells on the basis that individuals had to

[28] See Box on page 20 of June 2004 *Financial Stability Review* for descriptive analysis of the relationship between the level of mortgage income gearing and payment problems.
[29] In an alternative specification we instead defined savings in terms of the proportion of income saved. The results showed that saving a larger proportion of income reduces the probability of payment problems, but only the mean over time was significant. We chose to use the regular saver variable in our baseline model as this allowed us to use a larger sample.
[30] It is likely that this result is affected by the small number of mortgagors who remain unemployed in our sample (less than 0.1% of the sample in any given year). As a result, we may not not be able to estimate the relationship precisely. There may also be an issue with panel attrition – mortgagors who remain unemployed may drop out of the panel because they default.

rely on dissaving to maintain housing payments. But Böheim and Taylor instead found a negative relationship between unemployment duration and housing finance problems. They interpret this result as implying that households adjust expenditure and expectations as unemployment duration increases.

Regional effects are also important: households living in the East Midlands and the South West have a higher probability of reporting problems (significant at the 5% level) relative to those living in the North West. At the 10% level, the dummies for the West Midlands and the South East are also significant and positive.

We include current regional house price inflation, regional unemployment and effective mortgage rates among the explanatory variables in order to control for the effects of macroeconomic variables on the probability of experiencing mortgage payment problems. Of these three variables, only interest rates are significant and at the 1% level. Since IG is included with a one-year lag and effective mortgage rates are those of the current year, we could infer from these results that households did not consider the effect of rising rates on their ability to service mortgage debt.[(31)] We also interacted the effective mortgage rate variable with a year dummy to investigate whether greater transparency in monetary policy since the Bank of England was granted operational independence in 1997 has been associated with a smaller impact of monetary policy changes upon the probability of mortgage payment problems. Results from a likelihood ratio test indicated that the coefficients on the interest rate term were not significantly different, although the coefficient since 1997 was slightly larger.[(32)]

The initial conditions probit (results for which are shown in the first two columns of Table D) includes all the variables of the dynamic probit plus some additional variables for identification purposes.[(33)] Among the extra variables we include dummy variables for houses bought in 1987–89 and post-1989, with the base group being houses bought before 1987. Burrows (1997) argues that households who bought a house with a mortgage in the house price boom of 1987–89 might differ in behaviour from other households in the sample. He argues that these households may have had excessive expectations about the investment potential of housing and that lenders may have had more relaxed lending criteria during this period in a way not controlled for with the other explanatory variables.

(31) Böheim and Taylor (2000) also found that base rates have a positive and significant impact on housing finance problems.

(32) Evidence from financial markets on the impact of transparency upon predictability is somewhat mixed – Haldane and Read (2000) find a smaller market reaction to interest rate decisions post-independence, but over a longer time period Lasaosa (2005) finds that markets have reacted the same or more since 1997.

(33) These variables were at some stage included in the main regression but were dropped because they were not significant.

A simpler explanation is that these borrowers were particularly vulnerable to the sharp rise in interest rates from around 8%–9% in the middle of the 1987–89 period to around 15% at the end of 1989-beginning of 1990. In our initial conditions model, the dummy variable for houses bought in 1987–89 is positive and significant at the 1% level: the first time we observe that individual, a household taking their mortgage in 1987–89 is more likely to encounter financial difficulties relative to those buying their homes before 1987. Having dependent children and having no or low qualifications (relative to being highly educated) also increases the probability of having mortgage payment difficulties only in the first wave that we observe the individual.

We estimated the same model as in Table D including some additional dummy variables indicating whether the household finds its unsecured debt commitments are a heavy burden, somewhat of a burden or not a problem (relative to the reference group of not having unsecured debt). The BHPS only started reporting this information in 1995, so the estimation sample is reduced to the period 1996–2002 (since we include the unsecured debt related dummies with a one-year lag).[(34)] The results from estimating the model over this shorter period are reported in Table E.

In general terms, the estimates are very similar to those reported in Table D. The main difference is that the LTV dummies are no longer significant. This result seems to be due to the inclusion of the unsecured debt burden variables in the model and not to the use of a different sample. When we estimate the same model as in Table D for the sample of Table E (ie excluding the unsecured debt burden variables), we find that the dummy for an LTV of 90% or more remains significant.

Relative to not having unsecured debt, having unsecured debt commitments which are a heavy burden for the household in a consistent manner is positively and significantly (at the 1% level) associated with having problems paying for the mortgage. This is also true if unsecured debt is somewhat of a burden. Interestingly, if the household has unsecured debt and this is generally not a problem, then this significantly decreases the likelihood of the household having problems with its secured debt relative to not having unsecured debt at all. This result could be related to the pricing of unsecured debt – mortgagors with a good repayment history may have access to unsecured debt on better terms than mortgagors who have previously gone into arrears. Therefore mortgagors with a good repayment history may be less likely to find that debt a burden. Once we have controlled for the average

[(34)] At the suggestion of a referee, we also considered including the amounts of unsecured debt and financial assets reported as covariates. However, due to the large number of missing values for these variables – and for financial assets in particular – we decided not to include them in the regression (their inclusion reduces the sample by over 25% from 8,100 year-individual observations to around 5,900).

effect of the burden of unsecured debt commitments on the household, the temporary effect of unsecured debt burdens on mortgage payment problems is no longer significant.

Table E: Coefficient estimates — 1996–2002 sample, model 2

Variable	Initial conditions		Dynamic probit	
	coefficient	t-statistic	coefficient	t-statistic
Constant	-3.36^{***}	-3.71	-4.14^{***}	-20.36
		Lagged dependent variable		
Problems_{t-1}			0.51^{***}	2.43
		Loan to value ratios		
LTV 50-69%	0.23	0.97	0.22	1.27
LTV 70-89%	0.02	0.11	0.07	0.42
LTV 90+%	0.37^{*}	1.80	0.18	1.13
		Income gearing		
IG interest only$\geq 20\%$	0.34	1.47	0.28^{**}	2.02
IG principal$\geq 20\%$	0.59	0.88	0.07	0.18
		Other characteristics		
Saver	-0.36^{*}	-1.82	-0.08	-0.69
Lost job	0.30	0.50	0.37	1.08
Health problems	0.35	1.60	-0.12	-0.76
		Region of residence		
NE	-0.81^{**}	-2.38	0.32	1.22
Merseyside	-1.35^{**}	-2.33	-0.15	-0.30
York	-0.76^{**}	-2.51	0.25	1.08
EM	0.01	0.02	0.72^{***}	3.18
WM	-0.12	-0.51	0.55^{**}	2.35
E	0.28	0.90	0.68^{**}	2.35
London	-0.66^{**}	-2.03	0.29	1.12
SE	-0.01	-0.03	0.45^{**}	2.13
SW	-0.21	-0.79	0.55^{**}	2.26
Wales	-0.35	-1.21	0.15	0.57
Scotland	-0.46	-1.64	0.18	0.73
		Unsecured debt		
Heavy burden	0.26	0.74	0.08	0.40
Somewhat of a burden	0.16	0.65	0.04	0.28
Not a problem	-0.20	-0.88	-0.19	-1.40
		Macroeconomic conditions		
House prices	3.88^{**}	2.15	0.01	0.01
Unemployment	0.15^{***}	2.77	0.04	1.01
Interest rates	1.07	0.96	2.06^{***}	3.40
		Other — initial conditions probit		
1987–89	0.28	1.54		
Post 1989	0.07	0.46		
Negative equity	-0.26	-0.80		
Relationship breakdown	-0.09	-0.31		
Dependents	0.10	0.83		
Low qualifications	0.44^{***}	2.66		
Mid qualifications	0.14	1.00		
Male	-0.53^{***}	-3.95		

continued on next page

Table E: ***continued***

Non-white	−0.00	−0.01		
	Means of time-varying covariates			
IG interest only≥ 20%	0.34	1.16	0.58***	2.71
IG principal≥ 20%	−0.09	−0.12	0.48	0.95
Saver	−0.10	−0.41	−0.69***	−3.92
Lost job	0.88	1.13	0.74	1.11
Health problems	0.26	1.02	0.95***	4.46
Heavy burden	1.62***	3.57	1.63***	4.61
Somewhat of a burden	0.46	1.44	0.82***	3.59
Not a problem	0.18	0.65	−0.45**	−2.01
Generalised residual			0.30***	3.11
Proportion of the total variance contributed by the panel-level variance component			0.37***	6.34
AR(1) error			−0.10	
Log-likelihood				−2346.91
Log-likelihood excluding the AR(1) error term				−2346.92
Number of observations		1,661		6,375
Number of households				1,661
Obs. per household: min.				1
Obs. per household: avg.				3.8
Obs. per household: max.				7

Notes: (i) The model allows for endogenous initial conditions which are estimated using a two-step procedure following Orme (1997). (ii) Correlation between the time-varying covariates and the unobservable heterogeneity is allowed for by including the time means of these variables. (iii) For details of the 'generalised residual' from the initial condition probit see text, page 20. (iv) ***, ** and * denote coefficient significant at the 1%, 5% and 10% significance levels respectively, for a two-sided test.

The unobserved heterogeneity (α_i) still explains a large proportion of all variance, 37%, in this specification and the generalised residual is significant at the 1% level. As with our baseline model, the autocorrelated error term is not significant.

Since the addition of the unsecured debt variables in our estimations does not materially alter the results, we concentrate on the baseline model (model 1) estimated over the longer sample period 1994–2002 in the remainder of the paper.

5.2 Predicted probabilities

This section calculates the 'marginal effects' or changes in the probability of mortgage payment difficulties when particular characteristics are changed, with all other characteristics held constant. This makes the quantitative interpretation of the estimates presented in Tables D and E easier. Given the panel nature of our data, two issues arise in

calculating marginal effects. First, we need to distinguish between a permanent and a temporary change. Second, we need to take into account the unobservable individual-specific component when calculating the estimated predicted probabilities.

We follow Arulampalam and Booth (2000) and consider the temporary mean effect of changing x_t from $\dot{x}$ to $\ddot{x}$ on the probability that a randomly chosen individual will experience mortgage payment problems. This is given by:

$$\int [prob(y_t = 1|x_t = \ddot{x}_t, \alpha) - prob(y_t = 1|x_t = \dot{x}_t, \alpha)]\mu(d\alpha) \qquad \textbf{(10)}$$

The distribution of y_{it} conditional on $x_{i.}$ but marginal on α has a probit form given by:

$$prob(y_{it} = 1|x_{i.}) = \Phi\left[\frac{x'_{it}\beta + a'_1 x_{i.}}{\sqrt{\sigma_\upsilon^2 + \sigma_\xi^2}}\right] \qquad \textbf{(11)}$$

A consistent estimate of equation **(10)** is then given by:

$$\frac{1}{N}\sum_{i=1}^{N}\left\{\Phi\left[\frac{\ddot{x}'_{it}\beta + a'_1 x_{i.}}{\sqrt{\sigma_\upsilon^2 + \sigma_\xi^2}}\right] - \Phi\left[\frac{\dot{x}'_{it}\beta + a'_1 x_{i.}}{\sqrt{\sigma_\upsilon^2 + \sigma_\xi^2}}\right]\right\} \qquad \textbf{(12)}$$

where the parameters are replaced by their estimates.

5.2.1 State dependence

To see how much of the estimated probabilities of payment problems are attributable to true state dependence, we follow Arulampalam *et al* (2000) and compare the raw probabilities conditional on having mortgage payment problems in the previous period with the predicted probabilities derived from the baseline model (model 1). Table F summarises the results. The predicted probabilities are first calculated conditional on having mortgage payment problems in the previous period (column (d) in Table F), and second on not having experienced problems in the previous year (column (e)). These probabilities are calculated year on year and then averaged over each year. The difference between columns (d) and (e) then gives us the contribution of true state dependence.

Column (g) presents this same contribution but expressed as a percentage of the observed persistence in mortgage payment problems (given by column (c)).

Table F: Raw and predicted probabilities, model 1

Year	Raw data probabilities conditional [a] on			Predicted probabilities holding characteristics constant			
	$y_{t-1}=1$ (a)	$y_{t-1}=0$ (b)	(a)-(b) (c)	$y_{t-1}=1$ (d)	$y_{t-1}=0$ (e)	state dependence (f)=(d)-(e)	as % of (c) (g)=(f)/(c)
1994	0.51	0.06	0.45	0.28	0.09	0.19	41.18
1995	0.44	0.05	0.38	0.28	0.08	0.20	51.20
1996	0.37	0.04	0.33	0.22	0.06	0.16	47.82
1997	0.43	0.05	0.39	0.23	0.07	0.16	40.81
1998	0.43	0.04	0.38	0.21	0.06	0.14	37.74
1999	0.48	0.03	0.45	0.14	0.03	0.10	22.63
2000	0.45	0.04	0.41	0.15	0.04	0.11	25.97
2001	0.43	0.03	0.40	0.08	0.02	0.06	15.31
2002	0.34	0.02	0.31	0.13	0.03	0.10	32.37

[a] $y_{t-1}=1$ if the household had problems paying for its mortgage in the previous period. $y_{t-1}=0$ if the household did not have problems in the previous period.

The predicted probabilities show that over our sample period about 35% of the observed persistence in mortgage payment problems is due to state dependence. So it is clear that the persistence of mortgage payment problems is very high. This suggests that policy measures that reduce the number of households getting into mortgage payment problems in the first instance could have long-lasting effects. Indeed, the decrease in observed persistence observed in our results could be related to a change in lenders' policies in the early 1990s, when they began making earlier contact with borrowers in difficulty and making greater efforts to solve payment problems without recourse to possessing the home (see Ford *et al* (1995) for further details).

5.2.2 Personal and economic characteristics

The expected changes in the probability of having mortgage problems when a selected characteristic is changed (the marginal effects) are shown in Table G. For time-varying individual characteristics, the temporary and permanent effects of changing a particular characteristic will differ due to the inclusion of the means over time of these variables as regressors in the dynamic probit. We therefore present both temporary effects and permanent effects for the baseline model, although of course for time-invariant individual characteristics and for the macroeconomic variables the temporary and permanent effects

will be the same. The temporary effects are calculated keeping the means unchanged, as shown in equation **(12)**. To calculate the permanent effects we change the value of the variable in each year, so that the mean of the variable over time is equal to the new level of the variable.[35]

To put the marginal effects into perspective we should compare them with the mean probability of having mortgage payment problems. These probabilities are represented in the third column of Table H. Mean probabilities range from 4.4% in 1994 to 0.7% in 2002.

Having an LTV of 90% or more relative to having an LTV of less than 50% has a large positive effect on the probability of mortgage payment problems – it increases by about 6 percentage points (pp), holding other characteristics constant. A temporary increase of interest IG to 20% or more, relative to interest IG of less than 20%,[36] increases the probability of mortgage problems by 4pp. But the difference in probabilities of having mortgage payment problems for a household that has always had interest IG of 20% or more relative to a household that has always had interest IG of less than 20% is much larger – 21pp. This suggests that it is continually having a high IG rather than temporary increases in IG that matters in causing mortgage payment problems.

Going into unemployment between $t-2$ and $t-1$ increases the probability of mortgage finance difficulties by 13.5pp. The temporary effect of having health problems is to decrease the probability of mortgage payment problems slightly (the coefficient was not significant), but the total or permanent effect has the expected effect on the probability of having mortgage problems: it increases the probability by around 15pp.

By comparison, house price inflation has a relatively small effect on the probability of payment problems. A slowdown in house price inflation from 21% (UK average in 2004 Q2) to 10% increases the probability of mortgage payment problems by around 0.5pp. Changes in regional unemployment also have quite a small impact, but it is worth noting that if the unemployment rate were to return to 1992 levels (10.5%) in all regions relative to the 2004 Q2 UK average rate of 4.7%, this would increase the probability of mortgage problems by about 5pp.

---

[35] Alternatively, we could have changed only the most recent value of the variable and recalculated the means accordingly. In this case we would expect a smaller change in the probability of having housing finance difficulties.

[36] Principal IG was held at less than 20% in both cases.

Table G: Probabilities for selected characteristics, model 1

	Change in characteristic	Predicted probability	
		Temporary	Permanent
(a)	LTV$<50\%$	0.10	
(b)	LTV$\geq 90\%$	0.16	
	(b)-(a) in percentage points	5.99	
(c)	Interest IG$<20\%$	0.11	0.05
(d)	Interest IG$\geq 20\%$	0.15	0.27
	(d)-(c) in percentage points	4.04	21.46
(e)	No regular saver	0.15	0.18
(f)	Regular saver	0.11	0.05
	(f)-(e) in percentage points	-3.52	-13.05
(g)	Did not go into unemployment in previous period	0.14	0.13
(h)	Went into unemployment in previous period	0.27	0.38
	(h)-(g) in percentage points	13.53	24.53
(i)	Good health	0.16	0.10
(j)	Bad health	0.13	0.26
	(j)-(i)	-2.90	15.31
(k)	House price inflation 21% (UK average 2004 Q2)	0.14	
(l)	House price inflation 10%	0.14	
(m)	House price inflation 0%	0.15	
(n)	House price inflation -10%	0.15	
(o)	House price inflation -20%	0.16	
	(l)-(k)	0.47	
	(m)-(k)	0.90	
	(n)-(k)	1.35	
	(o)-(k)	1.80	
(p)	Unemployment rate 4.7% (UK average 2004 Q2)	0.14	
(q)	Unemployment rate 6%	0.15	
(r)	Unemployment rate 10.5% (1992 level)	0.19	
	(q)-(p) in percentage points	1.03	
	(r)-(p) in percentage points	4.98	
(s)	Monthly effective rate 0.46% (equivalent repo rate 4.75%)	0.10	
(t)	Monthly effective rate 0.48% (equivalent repo rate 5%)	0.11	
(u)	Monthly effective rate 0.53% (equivalent repo rate 5.5%)	0.12	
(v)	Monthly effective rate 0.58% (equivalent repo rate 6%)	0.14	
	(t)-(s) in percentage points	0.53	
	(u)-(s) in percentage points	1.93	
	(v)-(s) in percentage points	3.45	

Notes: (i) Temporary effects are calculated changing only the variable of interest; the mean of the variable over time is not changed. (ii) Permanent effects are calculated changing the value of the variable in levels and the value of the mean of the variable for every single point in time (alternatively we could have changed only the most recent value and recalculate the means). (iii) Permanent effects are only reported if different from temporary.

The marginal effect of changes in interest rates are more interesting because they allow us

to gauge the impact of monetary policy upon payment problems and of the three macroeconomic variables included in the baseline model they are the only variable with an independent statistically significant effect. If repo rates were to increase to 5% (relative to the base case of 4.75%), we would expect a change in the probability of having financial difficulties of 0.5pp (assuming that the increase in the repo rate is fully passed through to the interest rate paid by the household and that it does not switch the individual IG from less than 20% to 20% or more). The change in the probability is about 3.5pp if repo rates were to increase to 6%.

5.3 Debt at risk

To summarise the health of the stock of net secured lending to UK households, we construct an indicator that is determined by both the probability of households having problems servicing their mortgage debt and the size of such mortgages. We call this indicator 'debt at risk'. We calculate debt at risk by multiplying the estimated probability of having mortgage payment problems for each household by the outstanding value of its mortgage. We can then sum debt at risk across all households to get an aggregate value of mortgage debt at risk. An increase in debt at risk could arise either because of increases in the probability of having problems or because of increases in the value of mortgage debt outstanding, or a combination of these factors.

Table H: Debt at risk — mortgage payment problems, model 1

Year	Mean debt at risk (£)	Mean mortgage outstanding (£)	Mean probability	Debt at risk as % of mortgage outstanding	Index of concentration of debt at risk[(a)]
1994	1,835	34,604	4.40	5.30	1.20
1995	1,562	36,786	3.91	4.25	1.09
1996	806	37,925	2.25	2.12	0.94
1997	686	39,822	2.09	1.72	0.82
1998	1,028	43,156	2.26	2.38	1.05
1999	509	45,240	1.30	1.12	0.87
2000	393	48,334	1.00	0.81	0.81
2001	352	53,249	0.83	0.66	0.80
2002	244	56,495	0.69	0.43	0.63

[(a)] This is constructed as debt at risk as a percentage of mortgage outstanding scaled by the mean probability of having problems by year.

Table H presents mean debt at risk (across all households), mean outstanding mortgage debt, mean probability of mortgage payment problems and debt at risk as a percentage of mortgage debt outstanding[(37)] in each year for the baseline model. At the aggregate level mean debt at risk has declined between 1994 and 2002 (although it picked up slightly in

[(37)] Constructed as mean debt at risk divided by mean mortgage outstanding.

1998), and this in spite of the increase in the mean value of mortgage debt outstanding over the same period. The rise in mortgage outstanding has been more than compensated for by declining probabilities of having mortgage problems. As a result, debt at risk as a percentage of total mortgage debt outstanding has declined from 5.3% in 1994 to less than 1% since 2000. This suggests that the increase in mortgage debt has not been concentrated among riskier households and therefore that the increase in mortgage debt has not been associated with an increase in short-term risks to financial stability.

This last point is also seen in the index of concentration of debt at risk (see last column of Table H). The index of concentration of debt at risk measures how heavily mortgage debt is concentrated in households with the highest probability of having problems. It is defined as debt at risk as a percentage of mortgage outstanding scaled by the mean probability of having problems by year. Values of this index greater than one indicate that mortgage debt is concentrated in more risky households, while values of the index less than one imply that mortgage debt is concentrated in less risky households.[38] Over the sample period, the index has fallen by almost 50% from 1.20 in 1994 to 0.63 in 2002. This shows that mortgage debt is now concentrated in less risky households compared with the mid-1990s, possibly reflecting a change in lenders' policies.

Why has the mean probability of problems fallen from 4.4% in 1994 to 0.7% in 2002? We can identify which factors explain the decrease by looking at the marginal effects (given in Table G) and relating these to the mean value of each variable in each year. If the characteristics of households in our sample have shifted towards those characteristics which have a negative marginal effect upon the probability of payment problems, then this can account for the fall in the mean probability. Between 1994 and 2002 there was a reduction in the number of households in our sample who had become unemployed in the previous period (0.7% in 2002, compared to 2.2% in 1994) or who had IG above 20% (around 10% in 2002 compared to 19% in 1994). Both of these movements will have contributed to the fall in the mean probability.

Macroeconomic factors have also played a role in reducing the mean probability. The marginal effects in Table G identified that a higher rate of house price inflation, a lower

[38] For example, suppose that there are two households: Household A has a mortgage of £100,000 and household B has a mortgage of £10,000. If both households had a 50% probability of having payment problems, then total debt at risk is £55,000 (or 50% of total outstanding mortgage debt) and the mean probability of having a problem is 50%. In this case, the index of concentration is equal to one (since both households are equally risky). Suppose instead that household A has a 75% probability of having a problem, while household B has a 25% probability of having a problem. In this case the mean probability of having a problem is still 50%, but total debt at risk is higher – £77,500 or 77.5% of total mortgage debt outstanding. As a result, the index of concentration is now 1.55, since the household with the larger amount of debt is now more risky.

rate of unemployment and a lower level of effective mortgage rates would all be associated with a fall in the probability of payment problems. Between 1994 and 2002 all three of these conditions were satisfied – mean annual house price inflation for our sample has risen from around 0.5% in 1994 to 20% by 2002, the mean regional unemployment rate fell from 8% in 1994 to 3% in 2002 and effective mortgage rates also fell. Finally, the decline in the proportion of households who had experienced payment problems in the previous year (from 13% of the sample in 1994 to 5% in 2002) will itself have contributed to the fall in the mean probability.

Table I: Probability (%) of having problems paying for mortgage, model 1

	All mortgagors			Mortgagors with mortgage above the median			on the 70^{th} percentile			on the 90^{th} percentile		
	1994	1998	2002	1994	1998	2002	1994	1998	2002	1994	1998	2002
Mean	4.40	2.26	0.69	5.64	2.23	0.47	5.90	2.48	0.57	6.15	3.35	0.79
10^{th} pc[(a)]	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.01	0.01	0.00
25^{th} pc	0.03	0.02	0.00	0.03	0.02	0.00	0.03	0.02	0.00	0.04	0.03	0.00
Median	0.19	0.10	0.01	0.23	0.13	0.01	0.27	0.11	0.01	0.57	0.18	0.01
75^{th} pc	1.50	0.65	0.07	3.50	0.80	0.07	3.48	0.88	0.05	5.59	0.76	0.06
90^{th} pc	11.88	3.65	0.38	13.82	3.61	0.34	19.89	3.65	0.53	13.82	2.76	0.60

[(a)] pc stands for percentile.

It is important that we also consider how the probability of mortgage payment problems has changed across the distribution. As a way of illustration, Table I presents the distribution of the probability of having mortgage problems for the first, middle and latest years of our sample (ie for 1994, 1998 and 2002). Table I shows that the decline in the average probability of mortgage payment problems occurred at all points of the distribution. It also suggests that the decline in the estimated probability of having mortgage payment problems has been greatest among the most indebted households. For all mortgagors the mean probability has declined by 3.7pp between 1994 and 2002, whereas for the most indebted households (ie those with an outstanding mortgage at the 90^{th} percentile) this figure was 5.4pp. This is true for other points of the distribution as well: the 90^{th} probability percentile declined by 11.5pp for all mortgagors and by 13.2pp for the most indebted households.

6 Conclusions

This paper estimates a dynamic probit model for the probability of mortgage payment problems, allowing for state dependence and unobserved heterogeneity. We find evidence of true state dependence: past experiences of mortgage payment problems increases the probability that the household has difficulties servicing its secured debt today. We also find

that the extent of state dependence has fallen over the sample period.

The estimated model finds that inflows into unemployment are the most important explanatory variable in the sense that they have the largest marginal effect upon the probability of mortgage payment problems. A temporary move into unemployment in year $t-1$ increases the probability of mortgage payment problems in year t by 13.5 percentage points. Interest income gearing is also highly significant, but its relationship with mortgage payment problems is non-linear – households with interest income gearing of 20% or more have a significantly higher probability of incurring problems. Both temporary and permanent increases in interest income gearing to above the 20% threshold are found to be significant, but an increase in income gearing to be permanently above the 20% threshold has a much larger impact on the probability of payment problems. We find no significant effects on mortgage payment problems from housing equity.

The estimation results highlight the interdependence of secured and unsecured debt problems. Having unsecured debt and this being a heavy burden significantly increases the probability of having mortgage payment problems (relative to not holding unsecured debt), whereas holding unsecured debt and this not being a problem significantly reduces the probability of mortgage payment problems (relative to not holding unsecured debt). This suggests that unsecured debt may play an important role in smoothing secured debt payment problems for some households. Future research could investigate this further by jointly estimating the probability of secured and unsecured debt problems.

Appendix

Table A.a: Variable definitions

Variable	Definition
	Lagged dependent variable
Problems$_{t-1}$	Having problems paying for the mortgage at wave $t-1$
	Loan to value ratios
LTV<50%	Original loan to value ratio less than 50% *(reference group)*
LTV 50%-69%	Original loan to value ratio between 50% and 69%
LTV 70%-89%	Original loan to value ratio between 70% and 89%
LTV 90+%	Original loan to value ratio greater than 90%
	Income gearing
IG interest only$< 20\%$	Mortgage interest payments to income of household head and partner (if applicable) less than 20% at wave $t-1$ *(reference group)*
IG interest only$\geq 20\%$	Mortgage interest payments to income of household head and partner (if applicable) 20% or more at wave $t-1$
IG principal$< 20\%$	Mortgage payment of principal to income of household head and partner (if applicable) less than 20% at wave $t-1$ *(reference group)*
IG principal$\geq 20\%$	Mortgage payment of principal to income of household head and partner (if applicable) 20% or more at wave $t-1$
	Year house bought
Pre-1987	House bought before 1987 *(reference group)*
1987–89	House bought between 1987 and 1989
Post-1989	House bought after 1989
	Unsecured debt
Heavy burden	Unsecured debt is a heavy burden to the household at wave $t-1$
Somewhat burden	Unsecured debt is somewhat of a burden for the household at wave $t-1$
Not a problem	Unsecured debt is not a problem for the household at wave $t-1$
No unsecured	The household does not have unsecured debt *(reference group)*
	Qualifications
High qualifications	A-levels, nursing, teaching, first degree or higher degree *(reference group)*
Mid qualifications	GSE, commercial, O-levels or equivalent
Low qualifications	No qualifications, still at school, apprenticeship
	Region of residence
NW	Resides in North West at wave t *(reference group)*
NE	Resides in North East at wave t
Merseyside	Resides in Merseyside at wave t
York	Resides in Yorkshire & Humberside at wave t
EM	Resides in East Midlands at wave t
WM	Resides in West Midlands at wave t
E	Resides in Eastern England at wave t
London	Resides in Greater London at wave t
SE	Resides in South East England at wave t

continued on next page

Table A.a: *continued*

SW	Resides in South West England at wave t
Wales	Resides in Wales at wave t
Scotland	Resides in Scotland at wave t
	Other personal and economic characteristics
Negative equity	Mortgage outstanding exceeds value of house
Saver	Head of household is a regular saver (puts something away now and then in a bank, building society, or Post Office account other than to meet regular bills) at wave $t-1$
Lost job	Household head went into unemployment between wave $t-2$ and wave $t-1$
Health problems	The health of the head of the household is such that it limits the type of work or the amount of work he can do, at wave $t-1$
Relationship breakdown	Head of household went through a relationship breakdown (divorce, separation...) at wave $t-1$
Dependents	Head of household has dependent children (children under 16 or aged 16–18 and in school or non-advanced further education, not married and living with parent)
Male	Head of household is male
Non-white	Head of household belongs to minority ethnic group
	Macroeconomic conditions
House prices	Regional Halifax house price inflation at wave t
Unemployment	Regional rate of unemployment at wave t
Interest rates	Monthly effective mortgage rates at wave t

Table A.b: Variable means and standard deviations

Variable	Long sample		Short sample	
	Mean	Std. Dev.	Mean	Std. Dev.
	Lagged dependent variable			
Problems_{t-1}	0.07	0.25	0.06	0.23
	Loan to value ratios			
LTV 50%-69%	0.19	0.39	0.19	0.39
LTV 70%-89%	0.27	0.44	0.27	0.44
LTV 90+%	0.39	0.49	0.39	0.49
	Income gearing			
IG interest only$\geq 20\%$	0.13	0.33	0.12	0.32
IG principal$\geq 20\%$	0.01	0.12	0.02	0.12
	Region of residence			
NE	0.07	0.25	0.07	0.26
Merseyside	0.02	0.14	0.02	0.14
York	0.09	0.29	0.09	0.29
EM	0.08	0.28	0.09	0.28
WM	0.08	0.27	0.09	0.28
E	0.03	0.18	0.03	0.18

continued on next page

Table A.b: *continued*

London	0.09	0.29	0.08	0.28
SE	0.23	0.42	0.22	0.41
SW	0.08	0.27	0.08	0.26
Wales	0.05	0.21	0.05	0.21
Scotland	0.08	0.26	0.08	0.27
Unsecured debt				
Heavy burden			0.03	0.17
Somewhat burden			0.13	0.34
No burden			0.30	0.46
Macroeconomic conditions				
House prices	0.07	0.07	0.09	0.07
Unemployment	5.03	2.34	4.24	1.81
Interest rates	0.59	0.07	0.58	0.08
Other economic and personal conditions				
Saver	0.53	0.50	0.53	0.50
Lost job	0.01	0.11	0.01	0.09
Health problems	0.09	0.28	0.08	0.30
Initial conditions variables - first observation statistics				
1987–89	0.13	0.34	0.12	0.33
Post-1989	0.60	0.49	0.64	0.48
Negative equity	0.03	0.16	0.03	0.17
Relationship breakdown	0.03	0.17	0.03	0.17
Dependents	0.42	0.49	0.44	0.50
Male	0.81	0.39	0.82	0.39
Non-white	0.02	0.15	0.03	0.16
Mid qualifications	0.22	0.42	0.22	0.41
Low qualifications	0.13	0.33	0.12	0.32
Sample size	7,197		6,375	

Sources: BHPS and Bank calculations.

References

Arulampalam, W and Booth, A L (2000), 'Union status of young men in Britain: a decade of change', *Journal of Applied Econometrics*, Vol. 15, pages 289–310.

Arulampalam, W, Booth, A L and Taylor, M P (2000), 'Unemployment persistence', *Oxford Economic Papers*, Vol. 52, pages 24–50.

Böheim, R and Taylor, M P (2000), 'My home was my castle: evictions and repossessions in Britain', *Journal of Housing Economics*, Vol. 9, pages 287–320.

Breedon, F and Joyce, M (1992), 'House prices, arrears and possessions', *Bank of England Quarterly Bulletin*, pages 173–79.

Brookes, M, Dicks, M and Pradhan, M (1994), 'An empirical model of mortgage arrears and repossessions', *Economic Modelling*, Vol. 11, No. 2, pages 134–44.

Buck, N, Burton, J, Laurie, H and Lynn, P (2002), 'Quality profile. British Household Panel Survey. Waves 1 to 10: 1991–2000', ISER.

Burrows, R (1997), 'Who needs a safety-net? The social distribution of mortgage arrears in England', *Housing Finance*, Vol. 34, pages 17–24.

Chamberlain, G (1984), 'Panel data', in Griliches, S and Intriligator, M (eds), *Handbook of econometrics*, Amsterdam, North-Holland.

Coles, A (1992), 'Causes and characteristics of arrears and possessions', *Housing Finance*, Vol. 13, pages 10–12.

Ford, J, Kempson, E and Wilson, M (1995), 'Mortgage arrears and possessions: perspectives from borrowers, lenders and the courts', HMSO, London.

Geweke, J and Keane, M (2001), 'Computationally intensive methods for integration in econometrics', in Heckman, J J and Leamer, E (eds), *Handbook of econometrics*, Elsevier Science BV.

Hajivassiliou, V (2002), 'The method of Maximum Smoothly Simulated Likelihood (MSSL) for LDV models with simultaneity', London School of Economics, *mimeo*.

Haldane, A G and Read, V (2000), 'Monetary policy surprises and the yield curve', *Bank of England Working Paper no. 106*.

Hausman, J (2001), 'Mismeasured variables in econometric analysis: problems from the left and problems from the right', *Journal of Economic Perspectives*, Vol. 15, No. 4, Autumn, pages 57–67.

Hausman, J, Abrevaya, J and Scott-Morton, F M (1998), 'Misclassification of the dependent variable in a discrete response setting', *Journal of Econometrics*, Vol. 87, No. 2, December, pages 239–69.

Heckman, J J (1981), 'The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process', in Manski, C F and McFadden, D (eds), *Structural analysis of discrete data with econometrics applications*, The MIT Press, Cambridge, Massachusetts.

Henley, A (2004), 'Self-employment status: the role of state dependence and initial circumstances', *Small Business Economics*, Vol. 22, No. 1, February, pages 67–82.

Lasaosa, A (2005), 'Learning the rules of the new game? Comparing the reactions in financial markets to announcements before and after the Bank of England's operational independence', *Bank of England Working Paper no. 255*.

May, O, Tudela, M and Young, G (2004), 'British household indebtedness and financial stress: a household level picture', *Bank of England Quartely Bulletin*, Winter, pages 414–28.

Orme, C D (1997), 'The initial conditions problem and two-step estimation in discrete panel data models', University of Manchester.

Redwood, V and Tudela, M (2004), 'From tiny samples do mighty populations grow? Using the British Household Panel Survey to analyse the household sector balance sheet', *Bank of England Working Paper no. 239*.

Reinhart, C M, Rogoff, K S and Savastano, M A (2003), 'Debt intolerance', *Brookings Papers on Economic Activity*, pages 1–74.

Robinson, C, Humphrey, A, Kafka, E, Oliver, R and Bose, S (2004), 'Housing in England 2002/3: a report principally from the 2002/2003 Survey of English Housing', HMSO, London.

Taylor, M F, Brice, J, Bruck, N and Prentice-Lane, E (2001), *British Household Panel Survey user manual volume A: introduction, technical report and appendices*, Colchester: University of Essex.

Whitley, J, Windram, R and Cox, P (2004), 'An empirical model of household arrears', *Bank of England Working Paper no. 214*.