XX/RT/15

Páipéar Taighde Teicniúil
Research Technical Paper

Some defaults are deeper than others

Understanding long-term mortgage arrears

Robert Kelly and Fergal McCann

Banc Ceannais na hÉireann
Central Bank of Ireland
Eurosystem

Some defaults are deeper than others: Understanding long-term mortgage arrears

Robert Kelly[a] and Fergal McCann[a*]
[a] Financial Stability Division,
Central Bank of Ireland

July 2015, Preliminary Draft, Do not cite or circulate

Abstract

The 2007-2008 financial crisis yielded a significant number of delinquent mortgage loans, which ordinarily would have faced foreclosure and repossession. However, given the negative externalities of repossession, policy response has shifted towards forbearance and mortgage modification, which has led to longer spells in default for delinquent mortgage holders. It is therefore imperative to move beyond binary models of default towards an understanding of the factors that drive the depth of default spells. Exploiting a highly detailed dataset on financially distressed households in Ireland in 2012 and 2013, we are able to identify the impact of a range of *current* household-level information, generally not available in loan-level studies of mortgage default, on the probability of entering early and deep states of mortgage default. Our results suggest that consumer credit growth, large mortgage repayments and shocks to employment and incomes should trigger serious concerns among policy makers regarding subsequent stability in the mortgage market, with these affordability measures all shown to have differentially large impacts on entry to deep, relative to early-stage arrears.

Keywords: Mortgages, default, days past due, affordability.

*E-mail: robert.kelly@centralbank.ie; fergal.mccann@centralbank.ie. We thank, in alphabetical order, Christian Danne, Kevin Denny, Paul Devereux, Edward Gaffney, Terry O'Malley for helpful discussions, along with seminar participants at the Central Bank of Ireland, University College Dublin, NUI Galway and University College Cork. The views presented in this paper are those of the authors alone and do not represent the official views of the Central Bank of Ireland or the European System of Central Banks. Any remaining errors are our own.

1 Introduction

The importance of the mortgage market to the banking system[1] and the economy at large cannot be understated given the central role played by misguided mortgage lending in precipitating the 2007-2008 financial crisis. The fallout from this crisis was a tranche of borrowers with unaffordable loans. Globally, governments have responded through intervention, for example the Home Affordable Modification Program (HAMP) introduced in the US, which aimed to minimise the negative externalities associated with foreclosure (Campbell et al. (2011) , Guiso et al. (2013) and Mian et al. (2011)), and the Central Bank of Ireland's MART program.[2] Remarkably, while there is a large stock of literature investigating the causes of default, there is scant empirical evidence on the extent to which the group of defaulted borrowers are heterogeneous in their responses to equity and affordability shocks. An understanding of these differences is of vital importance in evaluating the likely effectiveness of modification policies such as HAMP and MART, and in identifying patterns that should trigger concerns for potential repayment difficulties in the mortgage market.

In this paper we move beyond the typical binary treatment of mortgage default to consider deeper levels of mortgage default as distinct states.[3] Specifically, in our baseline model we take a sample of roughly twenty thousand financially distressed households in Ireland, and model the probability of default (greater than three missed payments, or ninety days past due) and deep default (greater than twelve missed payments, or three hundred and sixty days past due) relative to the probability of being in the early stages of mortgage arrears. We show that our results are not simply explained by the duration since the onset of a negative economic shock, but that our explanatory factors capture the ability and willingness of households to repay their mortgage.

The results of our baseline model suggest that households experiencing an unemployment shock or a divorce have a two percentage point higher probability of deep default.[4] A ten per cent increase in the *current* monthly debt service ratio (DSR) leads to an increase in the probability of deep default of 0.35 to 0.5 percentage points, suggesting an important role for mortgage affordability. Borrowers' non-mortgage leverage is also shown to play an extremely important role in driving long-term mortgage distress, with an one-standard-deviation increase in non-mortgage debts (either measured as a ratio relative to total debts or relative to income) leading to an increase of between 1 and 3 percentage points in the probability of deep default. Lower household incomes are also shown to have explanatory

[1] Jorda et al. (2014) have shown that the relative importance of mortgage lending in the activity of retail banks has increased unrelentingly since the 1950s, to the point where mortgages represent the majority of bank lending in most developed economies.

[2] Mortgage Arrears Resolution Targets.

[3] See Table A.1 for a classification of the ways in which default is defined in the economics literature.

[4] The baseline probabilities of default and deep default in the estimation sample are 18 per cent and 16 per cent, respectively.

power in the deep default equation. Further, longer mortgage terms and higher mortgage interest rates are also shown to be associated with higher probabilities of both default and deep default. For each explanatory variable, the impact on early-stage default is always smaller than the impact on deep default, and in many cases is not statistically significantly different from zero. These findings provide a crucial insight for policy-makers designing responses to a mortgage arrears crisis: shocks to borrowers' ability to repay are crucial drivers of mortgage arrears, and are more likely to lead borrowers to deeper states of default, where any recovery to full repayments is extremely unlikely.[5]

In our baseline model, we find no impact of housing equity on the depth of mortgage default. Recent studies from Gerardi et al. (2013), Guiso et al. (2013) and Bhutta et al. (2010) suggest that affordability shocks such as unemployment and income shocks are the economically more important factor in explaining mortgage default, with extremely large falls in housing equity required before "strategic default" becomes likely.[6] In a model extension, we estimate the effect of a limited set of covariates on the depth of default, where the reference category is the full population of truly performing loans (whereas in our baseline model we focus on loans in arrears due to data availability). This model estimates that, in order for housing equity to have the same impact on 0-90, 90-360 and 360+ arrears as a one-standard deviation change in interest rates, our adjusted measure of the household's Loan to Value (LTV) position would have to be between 4 and 6 standard deviations higher. The analogous findings for the impact of self-employed versus salaried borrowers (where the former act as a proxy for higher income uncertainty and precarity) also suggest affordability is unquestionably more important in driving deeper mortgage arrears than the housing equity position. Such findings are consistent with the aforementioned strand of the literature.

The post-2008 economic and policy climate in Ireland provides an ideal environment for a study that differentiates mortgage defaults according to their depth of arrears. Firstly, the sheer scale of the mortgage arrears crisis has few historical precedents, with the number of accounts in arrears rising from roughly 50,000 to 150,000 between 2009 and 2013, with the peak level representing 18 per cent of all primary residential mortgages (Figure 1a). Further, and more importantly from the point of view of this study, the composition of households in mortgage arrears has shifted through the crisis, with half of all accounts in arrears being in arrears of greater than one year (deep default) by end-2013 (Figure 1b).

This build-up in the number of mortgages in deep default has been caused in part by the significant

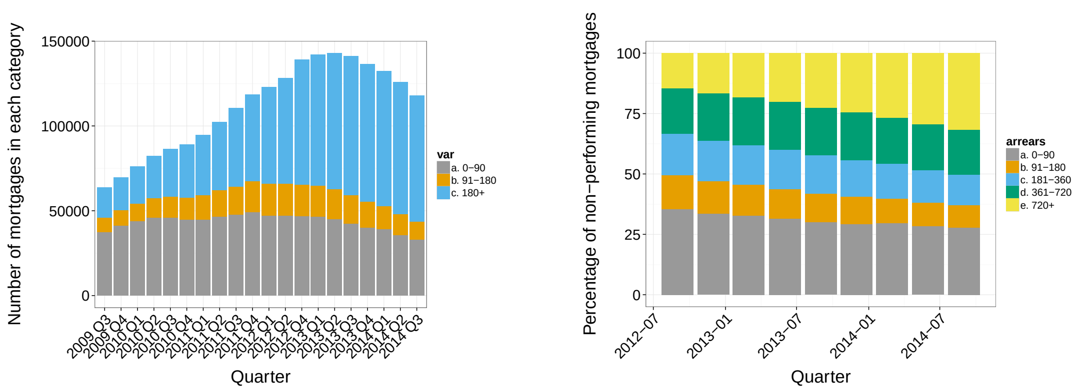
[5] Kelly and McCann (2015b) show that when borrowers have entered into arrears of greater than one year, the probability of any repayment is below 20 per cent, and falls even lower once borrowers enter arrears of more than two years.

[6] Strategic default is generally considered to be a default that is explained by a loan amount that is larger than the market value of the property (referred to as negative equity, where the loan to value ratio rises above 100 per cent).

policy uncertainty that existed in Ireland between 2009 and 2013. A legal judgment passed in 2009 rendered the repossession of homes in default impossible. Further, due to the scale of the crisis in Irish banks and public finances, and the state's role in recapitalizing the country's main mortgage lenders, the period was characterised by a high degree of uncertainty around the likely debt write-downs that might be received by distressed mortgage borrowers. These policy and political factors led to a situation where properties entered deeper states of mortgage arrears, with no move toward repossession on the part of lenders. It is highly likely that in jurisdictions with more clarity around the foreclosure process, a large number of these properties would have been repossessed, thus exiting the system and placing downward pressure on the aggregate number of accounts in arrears.[7]

Figure 1: The evolution of Irish mortgage arrears, 2009-2014

(a) Number of mortgages, 2009-2014 | (b) Arrears by DPD Category, 2012-2014



Market size: 760k primary residence mortgages. Source: Central Bank of Ireland; Residential Mortgage Arrears and Repossessions Statistics

The distinction between deep and early mortgage default has a number of crucial policy dimensions. Kelly and O'Malley (2014) and McCann (2014) have shown that the depth of mortgage arrears has an extremely strong negative association with the probability of loan cure (a return to full repayment). In the case of Ireland, Kelly and O'Malley (2014) show that the probability of loan cure for loans in default of three months is more than four times larger than the probability for loans in default of twelve months. These diminished cure probabilities have a number of important implications. From a prudential perspective, lower cure probabilities, especially if coupled with house price falls must be met with higher estimates of Loss Given Default (LGD), and subsequently higher loan provisions (Qi and Xiaolong (2009)). Lower cure probabilities also have social implications through their analogue, which

[7]In 2014, a large amount of the legal uncertainty around home repossessions was removed, leading to a heightened threat of repossession facing those in long-term mortgage arrears.

is a higher probability of entry to foreclosure for loans that are not successfully modified. Heightened foreclosures exert significant distress on the homeowners in question, have negative implications for house prices in the locality (Gerardi et al., 2012), affecting performance of other local area modifications (Been et al. (2013)) and place pressure on the public finances through the provision of social housing for those experiencing foreclosure.

Our paper builds on recent work that has exploited data on *current*, rather than at-origination measures of affordability such as household unemployment and income (McCarthy, 2014; Gerardi et al., 2013). Our study distinguishes itself from this previous work both in the focus on the depth of mortgage arrears, and in the nature of the dataset under study: both studies mentioned use survey data of between one and two thousand households, while our data set, on the other hand, contains information on twenty thousand households, with this information verified and audited by lenders before being used to assess the obligor's suitability for a modified mortgage.

The paper proceeds as follows: Section 2 explains our data sources; Section 3 describes our empirical approach and regression results; Section 5 outlines extensions to the baseline model, while Section 6 concludes.

2 Data

Two data sources are used to construct the file used in our baseline estimation. The first is the Central Bank of Ireland's Loan Level Data (LLD). These files contain information on all loans issued by Irish banks participating in the 2011 Financial Measures Programme (FMP). In the case of the Irish residential mortgage market, these lenders account for roughly two thirds of the total market, making this a particularly rich source of data. The data have been explained in detail by Kennedy and McIndoe-Calder (2012) and used subsequently in a number of mortgage default analyses (Kelly, 2011; Lydon and McCarthy, 2013; McCarthy, 2014; Kelly et al., 2014). The data are concerned mainly with the terms of the *mortgages*, with reliable information on *inter alia* current mortgage balance, bank, current interest rate, interest rate type, origination and maturity dates, current loan to value ratio (LTV), First Time Buyer status (FTB), and property values at origination and at December 2013. Certain characteristics of the *borrower* are also reported in the data, such as marital status, geographic location, employment group, income and joint versus single assessment. These variables are all collected at the mortgage origination date.

As is the case in the majority of studies on mortgage default, the LLD suffers from an important omitted variable problem, given that *current borrower characteristics* are relatively scarce in the data. This problem arises from the fact that, in managing their mortgage portfolios, lenders generally collect a large amount of information on borrowers at origination in order to inform the credit allocation decision, but do not follow up in detail on the borrowers' circumstances throughout the lifetime of the loan. This leads to an information gap, whereby most studies of mortgage default do not contain current information on factors as fundamental to the default decision as current employment status, income, indebtedness/leverage, household composition or marital status. Many studies of mortgage default proxy the "labour market" or "affordability" side of the mortgage default decision using regional economic conditions. Such an approach has been shown by Gyourko and Tracy (2014) to lead to a significant downward bias in the estimate of the effect of individual labour market outcomes on mortgage default.

In order to circumvent the information problems associated with the usage of data that focus mainly on loan and originating borrower characteristics, we exploit the Standard Financial Statement (SFS), a highly detailed data source on distressed borrowers. The completion of an SFS has been mandatory for any borrower engaging with their lender with a view to securing an alteration to their mortgage terms since 2012. In order to form the basis of an assessment of the borrower's debt sustainability, the SFS captures information on *inter alia* non-mortgage debt exposures, employment status, income, expenditure patterns, household composition and marital status. Using a unique loan identifier, SFS files can be linked to the associated mortgages in the LLD, meaning that an extremely rich data set on current loan and borrower information can be constructed for 20,645 mortgages.

The way in which the SFS data are collected presents two sources of bias. Firstly, given that by definition a borrower must be experiencing mortgage repayment difficulty before filling out an SFS with a bank, performing loans are hugely under-sampled in the SFS data. As a result, this dataset is not suited to the estimation of a standard default model where loans greater than 90 days past due are compared to those with no or early-stage arrears. However, where the purpose of the model is to understand the uniformity of default borrowers and hence predict borrowers' entry into *deeper states* of mortgage default, the SFS provides a wealth of important household balance sheet information, unavailable at such a scale to any previous study of which we are aware.

The second source of bias in the SFS data relates to the fact that, in order for SFS information to be available, the borrower must by definition have engaged with their lender after having experienced a negative shock. Given the policy context during our sample period discussed in Section 1, it is entirely plausible that non-engaging borrowers are a non-random sample of the population. Borrowers who suffered the worst shocks, or who experienced the biggest deterioration in their housing equity position, may be those that are least likely to engage with their bank.

Table A.2 provides some evidence on the extent of the bias. We compare loans with and without an SFS for loans in our three in-arrears categories, as well as across all loans in arrears. In making these comparisons, we are restricted to variables that are available for all loans in the LLD dataset. When observing the Current Loan to Value Ratio, there appears to be close to no difference between borrowers who have engaged by filling in an SFS and those who have not. The average CLTV among non-engaged borrowers is 95 across all loans in arrears, while the average for those with an SFS is 97. Borrowers who have filled out an SFS appear to have larger loans at December 2013, with this difference holding across all arrears buckets. Interest rates are lower among loans with an SFS, with this difference being driven by a higher share of tracker mortgages among those with an SFS (49 versus 34 per cent). Loans with an SFS appear to be slightly more likely to come from outside Dublin (25 versus 20 per cent). Finally, borrowers' age appears to have no influence on borrower engagement, with the average age among SFS and non-SFS loans being 45.9 and 46.6 years, respectively.

Whereas the LLD is a cross section of the full mortgage book of the four participating banks at December 2013, entries to the SFS data set vary in their timing. The SFS is filled out at the point of engagement between borrower and lender, with Table 1 reporting the distribution of SFS submission dates. 70 per cent of our observed SFS entries are in the calendar year 2012.

Table 1: Date of application, SFS data set

Date	Count	Share
Q1 2012	2,824	13.7
Q2 2012	3,527	17.1
Q3 2012	5,108	24.8
Q4 2012	2,870	13.9
Q1 2013	2,353	11.4
Q2 2013	1,882	9.1
Q3 2013	1,579	7.6
Q4 2013	502	2.4
Total	20,645	

2.1 Dependent Variable

The distribution of the depth of mortgage arrears among the 20,645 mortgages available in the SFS and LLD data is reported in Table 2. As one would expect given the nature of the SFS data-gathering process, loans without any arrears are severely under-represented in the SFS data set (84 versus 48 per cent). Given that the SFS data relates solely to mortgages in repayment difficulty, it is instructive to observe the share in each category *among those in arrears* across each data set. The columns $Share_{Arr}$ give the percentage of the non-zero DPD samples in each of our three arrears categories. Using this measurement, the SFS data appear to match much more closely the patterns observed in the LLD population data. The under-representation of deep default mortgages in the SFS sample (31.6 as opposed to 41.5 per cent) suggests that those who engage with their lender by filling out an SFS are less likely to be in deep default.

In our baseline empirical model, we amalgamate all those mortgages with zero to ninety days past due into an "early distress" category. The intuition for this grouping is that any borrowers filling out

Table 2: Dependent variable, LLD and SFS data sets

		LLD			SFS		
Category	DPD	Count	Share	$Share_{Arr}$	Count	Share	$Share_{Arr}$
Performing	0	224,500	83.71		9,902	47.96	
Early Arrears	1-90	12,797	4.77	29.3	3,472	16.82	32.32
Default	91-360	12,751	4.75	29.2	3,874	18.76	36.06
Deep Default	>360	18,141	6.76	41.5	3,397	16.45	31.62
Total		268,189			20,645		

the SFS with zero DPD are not "performing" in a similar way to the majority of zero-DPD borrowers in the full LLD population. Rather, these are borrowers who have engaged with their lender due to payment difficulty. A three-category multinomial logit model is specified where the probability of being in default and deep default is modelled relative to the reference category "early distress".

3 Empirical framework

At the core of our framework is a latent variable Y^*, which is decreasing in the likelihood that a household will repay its monthly mortgage payment M_t. All households begin their life as mortgage holders with a Y^*_0 that is consistent with a full monthly mortgage repayment. This condition is guaranteed to hold if we assume that banks' loan underwriting policies are such that all customers are given a mortgage that is consistent with repayment at origination, M_0. The empirically-observed dependent variable in our baseline model is the depth of arrears, DPD at the point of SFS engagement, T_{SFS}, which can take on three values (early distress, default and deep default). DPD rises by one month when a monthly repayment M_t is missed. However, DPD may also rise by some fraction F of one month when a household makes a partial payment $(1 - F)^* M_t$.

Between loan origination and T_{SFS}, a series of economic shocks will affect all households to varying degrees. Our dependent variable DPD is the realisation of Y^*, where Y^* can be influenced by:

1. The propensity of a household to be subject to a negative shock.
2. The nature of the shock.
3. The ability of the household to continue repayments, conditional on suffering a given shock.
4. The willingness of the household to continue with repayments.
5. The speed of engagement with the lender, once the household realises that its debts are unaffordable.
6. The time elapsed between the onset of the negative shock and December 2013.

We contend that the depth of arrears at T_{SFS} will be influenced by explanatory factors that are related to some or all of the above factors. In our baseline model, where a wide range of current household information is available to us, it is easy to imagine that household net income, unemployment status, the size of non-mortgage debts, the monthly debt service ratio (DSR), and household composition are all proxies for factors (1), (2) and (3) above. These variables, along with a measure of the household's equity position, may also influence factors (4) and (5), which relate to the willingness to repay, and the speed of engagement with the lender.

Factor (5), the speed with which a borrower engages with her lender, may also potentially drive differences between otherwise identical borrowers. Consider two households that suffer an identical shock, at an identical time, with an identical ability to pay. Household 1 engages with their bank after having missed twelve repayments, and fills out an SFS with a $DPD = 360$. Household 2, on the other hand, decides to approach his lender after having missed six repayments, and therefore is recorded in our SFS model as having a $DPD = 180$.

Figure 2: Schematic of DPD accumulation process for example households

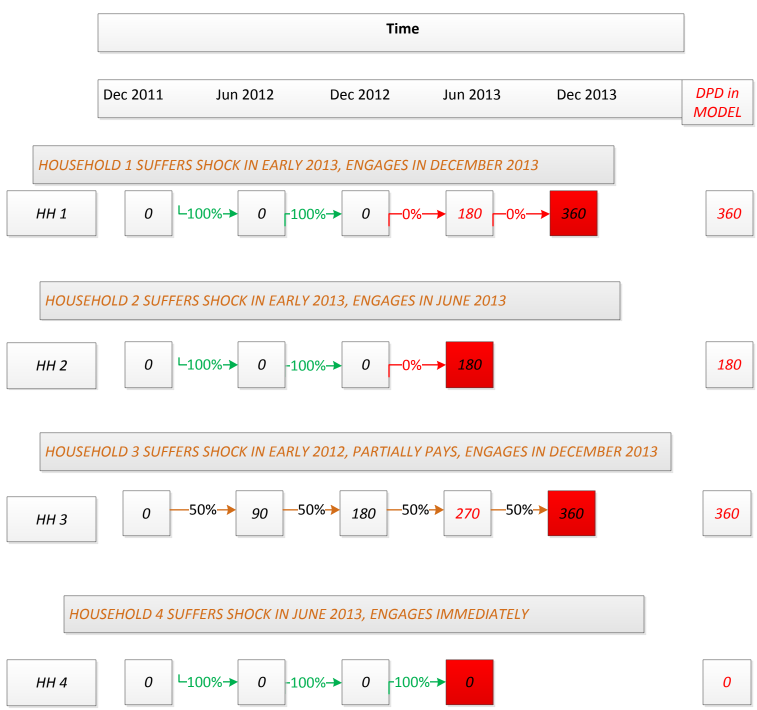


Figure 2 provides a visual representation of how households may end up with differing DPD values in our estimations. Household types 1 and 2 represent those discussed in the previous paragraph, where households differ only in the speed with which they decide to approach their bank. Household 3 has suffered a shock in early 2012, but has managed to pay half of the monthly repayment due in every month from then until T_{SFS}. This pattern suggests that this type of household varies crucially from households 1 and 2 in its ability to withstand the negative shock. The final type of household described in the schematic is one that, upon experiencing a shock, immediately approaches the lender to fill out an SFS. As shown in Table 2, this type of household accounts for two-fifths of all households filling out an SFS.

The final factor (6) underlying Y^* is the duration since the negative shock. It is important to acknowledge that the nature of our dependent variable is such that two households who have experienced an identical shock, and have an identical ability and willingness to repay, will have different DPD counts at T_{SFS}, depending on when the negative shock first affected the household. If the earlier onset of a shock is correlated with our household-level variables, for example because households in certain geographical areas or working in certain industrial sectors are more prone to negative shocks that hit specific sectors of the Irish economy at an earlier date, then the estimation of our multinomial model may be subject to omitted variable bias. For this reason, we include the time, in calendar months, since a household first entered arrears as a control variable in our baseline models. If Time in Arrears, $TinA$, is controlled for, we contend that the remaining effect of the explanatory variables on PD and PDD can be solely attributed to factors (1) to (4) above, given that $TinA$ captures both the time since initial shock (6), as well as acting as proxy for the willingness to engage (5). It should be noted that this estimation strategy is more onerous on the data than that typically employed in a binary default model, given that $TinA$, through its positive correlation with arrears balances, should be expected to reduce the explanatory power of the remaining specified variables.

Figure 3: Time in Arrears (months)

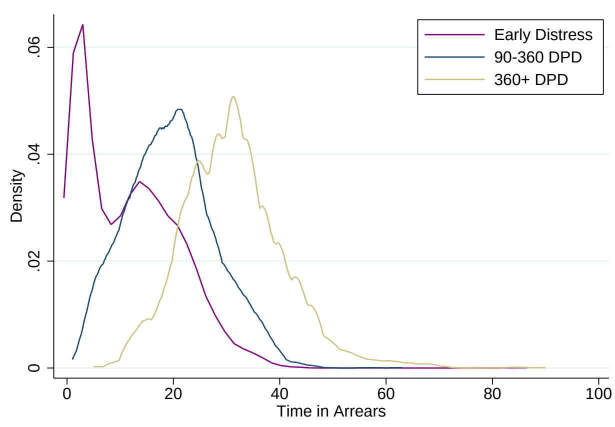


Figure 3 provides Kernel density plots of $TinA$ for each of the three groups comprising our dependent variable. As would be expected, $TinA$ is distributed further to the right for loans in deeper states of arrears. However, there is significant overlap in the $TinA$ distributions across the three groups. This overlap suggests that there are many households who, by virtue of duration alone, should have entered the deep default state, but have either partially paid, or only missed payments sporadically since the onset of the shock. Our estimation strategy rigorously isolates the impact of the explanatory variables on the (in)ability of the household to resist the movement into deeper arrears once the negative shock has been experienced.

The net result of the staggered engagement with the SFS process is a dataset which takes the form of a pooled cross section, with DPD_i being the realisation of the underlying propensity for delinquency level DPD_i^*, for loan i, which takes the values:

$$DPD_i = \begin{cases} 1 & 0 \leq DPD_i^* < 90; \\ 2 & 90 \leq DPD_i^* < 360; \\ 3 & DPD_i^* \geq 360 \end{cases}$$

In our baseline specification, the probability of the realised DPD indicator taking the value of 1 or 2 modeled as a function of the time in arrears and the underlying characteristics of the borrower, loan terms and dwelling controls:

$$Pr(DPD_i = 2|DPD_i = 3) = \mathbf{F}\left(TinA_i, \mathbf{X_i}, \mathbf{Z_i}\right) \tag{1}$$

where $\mathbf{X_i}$ is a vector of borrower-specific controls, $\mathbf{Z_i}$ a vector controls for loan characteristics. $\mathbf{Z_i}$ includes the loan vintage as a polynomial (months since the loan was originated), the term length, and the type and level of interest rate (binary indicators for standard variable rate, tracker, fixed-rate). Table 3 reports the mean and standard deviation for $TinA_i$, $\mathbf{X_i}$ and $\mathbf{Z_i}$ for our baseline model sample, while Table 7 acts as a reference point by reporting summary statistics among the entire population of LLD loans.

Previous studies have taken as standard the inclusion of the Current Loan to Value Ratio ($CLTV$) as a measure of housing equity. However, as elaborated on in Kelly and McCann (2015a), there is a mechanical reverse causality in the DPD-$CLTV$ relationship that is ignored by most researchers in this area. Due to this bias, we propose an alternate measure of housing equity which we term Adjusted Equity which we calculate as:

$$OLTV^* \frac{HP_{Dec13}}{HP_O} \tag{2}$$

Where $\frac{HP_{Dec13}}{HP_O}$ is the ratio of house prices in December 2013 to house prices at loan origination, and

Table 3: Summary statistics, SFS sample

Variable	Obs	Mean	Std. Dev.
FTB	20645	0.26	0.44
Term (Months)	20645	315.81	84.19
Adjusted Equity	20645	99.51	46.23
CLTV	20639	92.74	45.06
Loan Age	20645	92.31	30.59
Current Interest Rate	20645	2.94	1.55
Borrower Age	20645	38.31	8.71
Fixed Rate	20645	0.05	0.21
SVR	20645	0.45	0.50
Tracker	20645	0.50	0.50
Net Monthly Income (€000)	20645	2.88	1.56
Unemployment Shock	20645	0.30	0.46
Divorced Since Origination	20645	0.07	0.26
Monthly repayment income (DSR)	20645	0.33	0.28
Other Debt to Income	20645	2.60	6.42
Other Debt to Total Debt	20645	0.20	0.24
Single, No Children	20645	0.17	0.37
Single, 1/2 Children	20645	0.10	0.30
Single, 3+ Children	20645	0.02	0.15
Couple, No Children	20645	0.18	0.38
Couple, 1/2 Children	20645	0.34	0.47
Couple, 3+ Children	20645	0.20	0.40

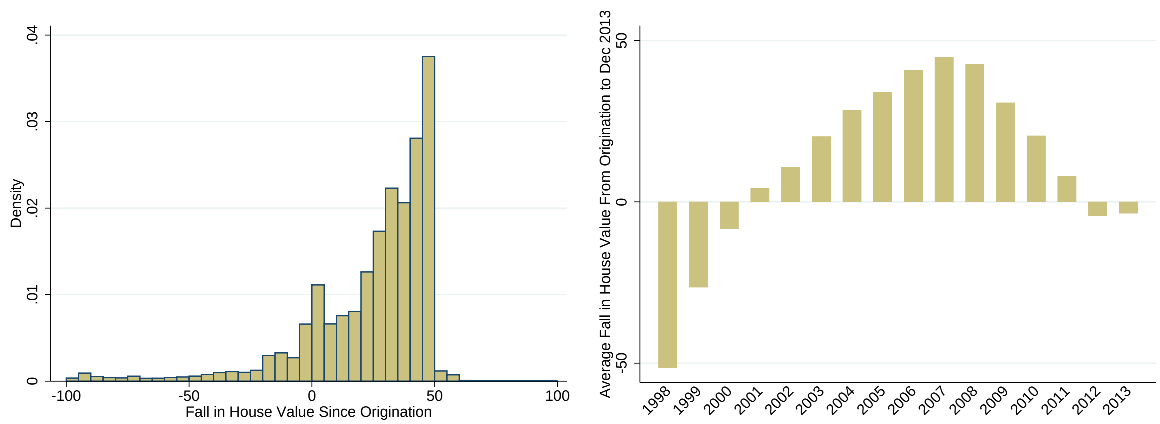
$OLTV$ is the originating loan to value ratio on the loan. This measure will not be biased by missed payments or the capitalisation of arrears, as would be the case with a traditionally-used $CLTV$. Rather, it simply measures the opening equity position, adjusted for any impact that house price movements have had on that position. In order for this Adjusted Equity measure to be interpreted meaningfully, we also control for loan age, loan term (both in months) and the interest rate on the loan, to fully capture the components of the amortization schedule which could lead to borrowers with an identical estimate of Adjusted Equity in fact differing in their equity position. Figure 4 gives more detail on house price changes in our dataset, with panel (a) showing that the majority of properties have experienced a house price fall since origination to December 2013. Panel (b) shows that the largest falls are for properties with mortgages originating in 2007, while loans originating in 1997 are associated with 60 per cent house price increases since origination.

The borrower-specific controls, $\mathbf{X_i}$, include borrower age modeled as a quadratic term and indicators for change in martial status, family composition and the current employment status of the borrower.[8] In the model sample, 7 per cent of households have experienced a divorce since origination, while 30 per cent of households are experiencing unemployment at the point of engagement T_{SFS}. In

[8]Unemployment shocks are measured as occurring where at least one individual in the household is not working. In cases where adults are not working but not unemployed in the statistical sense (e.g. they may be students, retired or ill), they are coded as being unemployed to reduce the number of categories in the data, while retaining the economic information as to whether or not income is being earned in the household.

Figure 4: The fall in house values since origination

(a) Histogram, December 2013 (b) Average fall by year of origination

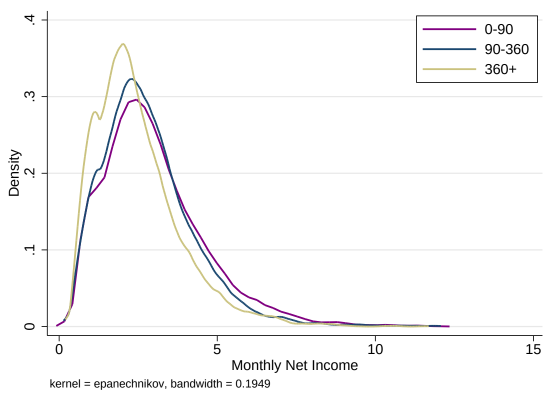


addition, current income is captured in the SFS data by observing all sources of household income, whether from salaries, self-employed income or welfare payments. The average after-tax monthly household income in our sample is €2,872. Figure 5a shows that there are important differences in income across the three arrears groups comprising our dependent variable: the income distribution for households in Deep Default is shifted clearly to the left of that for the two lower-arrears groups.

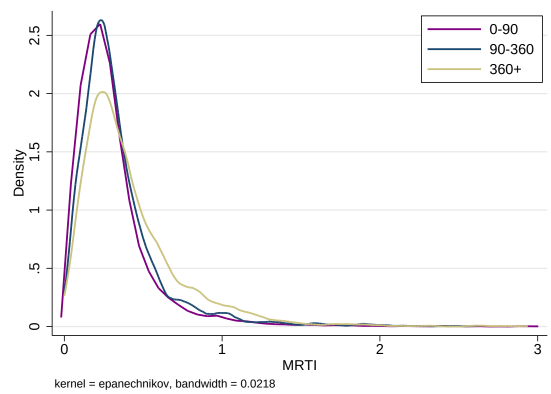
Mortgage payment burdens are captured using the ratio of mortgage repayments to income (DSR). This field, which uses the monthly repayment rather than a ratio of outstanding values to incomes, captures the recurring affordability of the mortgage and explicitly accounts for differing term structures and interest rates. Figure 5b shows that the mean value of 32 per cent masks a long right tail, with extremely unaffordable mortgages with DSRs of greater than 100 per cent being rare, but values up to 150 per cent existing in the data.

Non-mortgage debts constitute all reported Buy-to-Let mortgage, credit card, credit union, consumer loan and business debt. We measure these "other debts" in two ways in our regression models: firstly as a share of total debt (with a mean value of 19 per cent), and secondly as a ratio to annual household net income (with a mean value of 2.6 times). We calculcate both these measures using the total outstanding value of other debts, rather than a monthly repayment, given that many forms of consumer debt do not have a term structure or an associated monthly repayment. The distribution of other debts as a share of total debts is plotted in Figure 5c, with the plot showing that large shares of non-mortgage debt are relatively rare in the data set - most households' mortgages account for between 80 and 90 per cent of their total debt burden. However, for households in Deep Default, there is a significantly larger share with non-mortgage debts accounting for 30-40 per cent of their outstanding

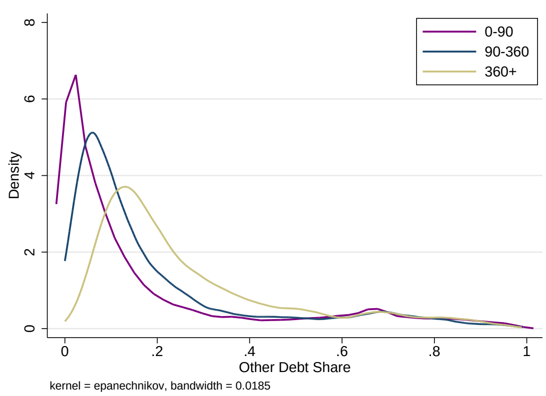
(a) Net monthly income, SFS sample.



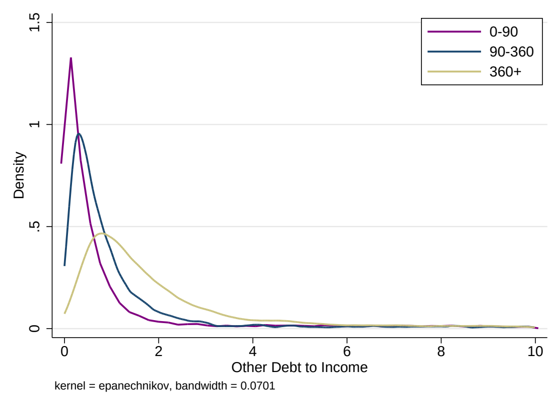
(b) DSR distribution, SFS sample.



(c) Non-mortgage debt to total debt, SFS sample.



(d) Non-mortgage debt to income, SFS sample.



debts. Similarly, in Figure 5d it is shown that most households have a non-mortgage debt value that is lower than one times their annual net income. Again, households in Deep Default are much more likely to have higher debt to income ratios, with ratios larger than two being relatively prevalent.

Table 4 reports tabulations and means for explanatory variables within each category of our dependent variable. Some important differences are clear in the raw data, with deeper-arrears mortgages being more prevalent among variable (SVR) and tracker mortgages than fixed-rate loans, among families experiencing a divorce since origination, and families with all or one adult not working. In terms of family structure, couples with one or two children have the lowest rates of deep default at 15.68 per cent, with the highest rates among single people with three or more children, at 25.5 per cent.

Analysis of the continuous explanatory variables reveals that monthly net income is shown to be over €500 lower among deep-default households than those in early distress. The mortgage repayment to income ratio is 40.5 per cent among deep-default mortgages, which differs importantly from early distress and early default mortgages (29.7 and 33.5 per cent, respectively). Households in deeper states

Table 4: Breakdown of key variables by arrears states, SFS sample

	Early Distress	**90-360**	**360+**
Total	64.8	18.8	16.5
Non-FTB	64.7	18.6	16.7
FTB	65	19.2	15.8
Fixed	79.9	12.1	8.1
SVR	62.9	19.2	17.9
Tracker	65	19	16
No Divorce	65.6	18.6	15.9
Divorce Since Origination	54.6	21.3	24.1
No Unemployment	68.6	18.1	13.3
Unemployment Shock	56	20.2	23.8
Couple, no children	61.1	20.4	18.5
Couple 1/2	68.2	17.9	13.9
Couple 3+	64.7	19.7	15.6
Single, no children	65.6	17.3	17.1
Single 1/2	61.1	19	19.8
Single 3+	53.3	21.1	25.6
Mean values for continuous variables			
Borrower Age	38.4	37.9	38.3
Vintage (Months)	89.9	93.9	100.1
Opening Term (Months)	313.4	325.4	314.5
Net Monthly Income	3,006	2,811	2,470
Equity Adjusted	98.1	104	100.1
CLTV	89.0	97.6	101.8
DSR	0.304	0.347	0.42
Interest Rate	2.9	2.99	3.06
Other Debts to Income	2.26	2.53	3.98
Other Debts to Total Debt	0.18	0.2	0.27

of mortgage arrears also appear to have accumulated higher non-mortgage debts: the ratio of non-mortgage debt to income is 3.62 among those in deep default, and below 2.5 for the other two groups, while the share of non-mortgage debts in total debts is 26.8 per cent for those in deep default, and below 20 per cent for the lower-arrears groups. Comparing our measure of Adjusted Equity with $CLTV$, it is clear that the relationship between the depth of default and housing equity is much less apparent when adjusting for the mechanical bias in the construction of $CLTV$. In order to fully ascertain whether this improved measure of equity has explanatory power, we must reserve judgement until mortgage term, vintage and interest rate have been controlled for in a formal regression specification.

4 Empirical results

In this section, we present results from a three-category multinomial model using the SFS data. From Section 3, we define a reference category, "early distress", which encompasses all households filling out an SFS with either zero or between 1 and 90 DPD. The probability of a loan being in default and deep default relative to being in early distress is estimated. The coefficients, presented in Table 5, cannot be directly compared to those of binary default models common to the literature, given that truly performing loans are not available in our data sample. Rather, we must interpret the results of the model of Table 5 as representing the effect of the $TinA_i$, $\mathbf{X_i}$ and $\mathbf{Z_i}$ on PD and PDD, conditional on having experienced some mortgage repayment difficulty. In the estimation sample, the PD is 18.76 per cent, with PDD being 16.45 per cent. All marginal effect estimates must be interpreted with these baseline probabilities in mind. The results of four models are presented in Table 5. The difference between the specifications is (i) in whether our Adjusted Equity measure is included, or whether $CLTV$ is included to increase comparability to previous literature (ii) in the way in which non-mortgage debts are captured in the data.

The first striking pattern in the model's results is that most of the variables included in the model do not explain entry into early-stage default. The vast majority of the statistically significant impacts observed in the model are found in the deep default equations. We have initial evidence from these patterns that where household affordability shocks and other factors drive borrowers into default, they have severe impacts that lead to the continued accumulation of large quantities of arrears.

Looking to the coefficients on $TinA$, one month in arrears is associated with a 0.3 percentage point increase in PD, and a 1.6 percentage point increase in PDD. These results intuitively suggest that our innovation in controlling explicitly for the duration since the onset of a negative shock has important explanatory power in all models. FTB mortgages are shown to be less likely to enter deep default, with the differential being between 1 and 2 percentage points in most models. Mortgages originated with a longer term are shown to be higher-risk in general, although the coefficient in the deep default equation does turn negative and statistically significant in model 3, where $CLTV$ is included as our measure of housing equity.

Standard Variable Rate and tracker mortgages are shown to have significantly higher probabilities of deeper states of arrears than fixed rate loans. The coefficients suggest that the impact of a tracker mortgage on PDD is to increase the probability by 14 percentage points relative to fixed rate loans, while the analogous effect of SVRs is smaller at 5 per cent. Beyond the impact of rate types, which may capture some underlying borrower heterogeneity in risk preferences, the interest rate on the loan has a positive association with credit risk, with a 100 bps rate increase associated with 1.9 percentage point increase in PD and a 3.5-3.9 percentage point increase in PDD.

Our use of an alternative measure of housing equity suggests that there is no relationship between equity and the depth of mortgage default. In both the PD and PDD equations, the coefficient on Adjusted Equity is positive, but always statistically insignificant across columns (1) to (4). By contrast, if we use $CLTV$, which contains a mechanical bias as discussed earlier, we find that there is an increase in PDD when $CLTV$ is higher.

Mortgage affordability, as measured by the ratio of monthly repayment to monthly household

Table 5: Multinomial Logit Results. Average Marginal Effects Reported

	(Model 1)		(Model 2)		(Model 3)		(Model 4)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	90-360	360+	90-360	360+	90-360	360+	90-360	360+
Time in Arrears	0.00327***	0.0161***	0.00340***	0.0157***	0.00343***	0.0156***	0.00356***	0.0152***
	(0.000160)	(0.000192)	(0.000162)	(0.000194)	(0.000166)	(0.000198)	(0.000168)	(0.000199)
FTB	0.0139**	-0.0117**	0.0148**	-0.00954*	0.0151**	-0.0197***	0.0159**	-0.0170***
	(0.00695)	(0.00491)	(0.00696)	(0.00489)	(0.00690)	(0.00474)	(0.00691)	(0.00473)
Term	0.000110**	0.0000512	0.0000995**	0.0000898***	0.000117**	-0.0000919***	0.000104**	-0.0000505
	(0.0000448)	(0.0000330)	(0.0000450)	(0.0000330)	(0.0000471)	(0.0000347)	(0.0000473)	(0.0000348)
SVR	0.0316**	0.0477***	0.0326**	0.0454***	0.0310**	0.0448***	0.0319**	0.0426***
	(0.0150)	(0.0121)	(0.0150)	(0.0119)	(0.0149)	(0.0120)	(0.0149)	(0.0119)
Tracker	0.0867***	0.140***	0.0885***	0.139***	0.0854***	0.127***	0.0872***	0.126***
	(0.0220)	(0.0153)	(0.0221)	(0.0154)	(0.0224)	(0.0161)	(0.0224)	(0.0161)
Curr Int Rate	0.0181***	0.0385***	0.0186***	0.0387***	0.0184***	0.0349***	0.0189***	0.0352***
	(0.00494)	(0.00378)	(0.00492)	(0.00375)	(0.00496)	(0.00388)	(0.00494)	(0.00384)
Equity Adjusted	0.0000560	0.0000566	0.0000586	0.0000932*				
	(0.0000735)	(0.0000542)	(0.0000734)	(0.0000536)				
CLTV					0.0000356	0.000610***	0.0000503	0.000626***
					(0.0000812)	(0.0000581)	(0.0000809)	(0.0000572)
Other Debt to Income	-0.000329	0.00229***			-0.000272	0.00213***		
	(0.000387)	(0.000261)			(0.000382)	(0.000253)		
Other Debt Share			-0.00340	0.112***			-0.00118	0.110***
			(0.0114)	(0.00800)			(0.0114)	(0.00797)
Monthly Net Income	-0.00136	0.000428	-0.00189	-0.00266	-0.000649	-0.00307*	-0.00127	-0.00614***
	(0.00216)	(0.00165)	(0.00224)	(0.00170)	(0.00217)	(0.00168)	(0.00226)	(0.00172)
Divorced	0.00980	0.0200***	0.00906	0.0232***	0.0105	0.0173**	0.00968	0.0206***
	(0.00988)	(0.00695)	(0.00985)	(0.00691)	(0.00982)	(0.00684)	(0.00980)	(0.00682)
Unemployment Shock	-0.00898	0.0295***	-0.00873	0.0301***	-0.00789	0.0304***	-0.00759	0.0309***
	(0.00604)	(0.00440)	(0.00604)	(0.00439)	(0.00604)	(0.00439)	(0.00605)	(0.00437)
DSR	0.0257***	0.0480***	0.0233**	0.0543***	0.0288***	0.0299***	0.0265***	0.0361***
	(0.00974)	(0.00704)	(0.00966)	(0.00692)	(0.00971)	(0.00685)	(0.00964)	(0.00674)
Single, 1/2	0.00872	0.0119	0.00902	0.0118	0.00846	0.0126*	0.00885	0.0124*
	(0.0100)	(0.00731)	(0.0100)	(0.00724)	(0.0100)	(0.00727)	(0.0100)	(0.00721)
Single, 3+	0.0255	0.0200*	0.0269	0.0200*	0.0264	0.0210*	0.0280	0.0209*
	(0.0191)	(0.0120)	(0.0192)	(0.0119)	(0.0192)	(0.0120)	(0.0193)	(0.0120)
Couple, Zero	0.0189**	0.00729	0.0200**	0.00688	0.0181**	0.00718	0.0194**	0.00657
	(0.00909)	(0.00654)	(0.00910)	(0.00647)	(0.00905)	(0.00649)	(0.00906)	(0.00640)
Couple, 1/2	0.00647	-0.00710	0.00689	-0.00592	0.00568	-0.00569	0.00634	-0.00472
	(0.00814)	(0.00587)	(0.00815)	(0.00582)	(0.00811)	(0.00586)	(0.00812)	(0.00579)
Couple, 3+	0.0133	-0.00267	0.0139	-0.00135	0.0126	0.00106	0.0135	0.00214
	(0.00925)	(0.00651)	(0.00927)	(0.00646)	(0.00923)	(0.00652)	(0.00924)	(0.00646)
Observations	20645	20645	20645	20645	20639	20639	20639	20639

Controls for county, loan age (quadratic), borrower age (quadratic) and bank included in all models
Robust standard errors in parentheses; * $p < .1$, ** $p < .05$, *** $p < .01$

income (Debt Service Ratio, DSR), is an important driver of both PD and PDD. A ten per cent increase in the DSR is associated with an increase of 0.25-0.3 and 0.4-0.6 percentage points, respectively. In addition, unemployment is shown to have an important effect on PDD. Across all models, a robust effect of close to 3 percentage points is found. Non-mortgage debts are associated with deeper states of arrears: a ten per cent increase in the ratio of non-mortgage debt to total debts leads to a 1.1 per cent increase in PDD. An increase of one in the ratio of non-mortgage debt to annual income leads to a .2 per cent increase in PDD.

The effect of an increase in monthly net household income of €1,000 is statistically insignificant in most models, which is perhaps unsurprising given that income is captured as the denominator in our measure of non-mortgage debts in Columns (1), (2), (5) and (6), and our Debt Service Ratio measure, as well as indirectly through our measure of unemployment. In models (3) and (4), where $CLTV$ is used as the measure of housing equity, there is evidence that, controlling for the above factors, a decrease in income of €1,000 per month is associated with a half a percentage point increase in PDD.

The bottom panel of Table 5 reports results for household composition. Relative to single borrowers without children, single people with three or more children, who represent just two per cent of the sample, have 2 percentage points higher PDD. In the majority of cases, family composition does not impact the depth of default. However, households experiencing a divorce since mortgage origination are significantly higher risk, with such a shock associated with PDD increases of 2 percentage points.

We can think about the relative economic magnitudes of our estimated effects by observing the impact of a one-standard-deviation increase in our continuous variables on PDD. The impact for the ratio of non-mortgage debts to income is 1.4 per cent, while for the ratio non-mortgage debts to total debt the impact is 2.8 per cent. The equivalent effect for the DSR is 1.3 to 1.5 per cent.

When combined with the impact of a switch in the dummy variables for divorce and unemployment, these findings suggest that affordability shocks, measured along a series of different measures, as well as non-mortgage indebtedness levels, all play a statistically as well as economically important role in predicting which households will enter deeper states of mortgage arrears. Crucially in the context of the debate around the "double trigger" effect of housing equity and affordability, our measure of housing equity appears to have no statistical effect on PDD in these models.[9] These findings can provide an important insight to policy-makers attempting to understand the process behind the accumulation of mortgage arrears in a financially distressed section of the population.

Table 6 reports whether or not the estimated coefficients of the models of Table 5 are statistically different across the PD and PDD equations. The table reports that our key household-level financial

[9]Given the recourse inherent in the Irish credit and bankruptcy system, it is perhaps unsurprising that housing equity is found not to have an effect on the depth of arrears, particularly in a model such as ours where a rich set of explanatory controls for affordability are available.

distress measures (unemployment, non-mortgage debts, divorce, DSR) all have significantly different effects on deep versus early default. Other variables with such a differential are $TinA$, the current interest rate, and interest rate type dummy variables. These findings provide strong evidence that where repayment shocks have occurred, they are more likely to lead borrowers towards deep levels of arrears, rather than to a state where arrears are accumulated at a slower pace due to partial or sporadic payments being made.

Table 6: Deep Default VS Default: Are coefficients significantly different?

	Model 1	Model 2	Model 3	Model 4
Time in Arrears	***	***	***	***
FTB	**	**	***	***
Term		*	***	*
SVR	***	***	***	**
Tracker	***	***	***	***
Curr Int Rate	***	***	***	***
Equity Adjusted			n/a	n/a
CLTV	n/a	n/a	***	***
Other Debt to Income	***	n/a	***	n/a
Other Debt Share	n/a	***	n/a	***
Monthly Net Income				***
Divorced	**	***	*	**
Unemployment Shock	***	***	***	***
MRTI	***	***	***	***
Single, 1/2				
Single, 3+				
Couple, Zero				
Couple, 1/2				
Couple, 3+				
* $p < .1$, ** $p < .05$, *** $p < .01$				

Given that the inclusion of an explicit measure of the duration since the onset of a shock is not common in the literature on mortgage defaults, we re-run all the specifications of Table 5 without our $TinA$ measure. The results of these specifications, reported in Table A.3 should therefore be more comparable to the extant cross-sectional default literature.

The results of Table A.3 suggest that there is important correlation between $TinA$ and our main explanatory variables, suggesting that a model that omits $TinA$ is likely to overestimate the importance of $\mathbf{X_i}$ and $\mathbf{Z_i}$. Average marginal effect estimates on income, non-mortgage debts, divorce, unemployment, DSR are all larger, sometimes by orders of magnitude, in this specification. Many effects, particularly in the PD equation, become statistically significant once $TinA$ is omitted. Particularly interesting is the fact that our measure of Adjusted Equity appears to be positively associated with deep default when $TinA$ is omitted. Further, many of the dummy variables for household composition appear to impact default in these models, suggesting that early onset of shocks was more prevalent in Ireland for more vulnerable family types.

5 Extension: A model for all loans

The baseline model of Section 4 contains a wealth of household-level explanatory factors that are often missing in studies of mortgage default that use loan-level data. The coefficient estimates on the depth of mortgage defaults from the model cannot be compared to those from the literature on binary mortgage default models because the SFS data set is only collected for households experiencing repayment difficulty.

As an extension to our analysis, we run a model utilising the population of owner-occupied mortgages at the subject banks. The model differs from the baseline model of Section 4 in a number of important ways. Firstly, the model described in equation 1 is extended to a four-category model, where all performing loans are considered the reference category, with the probabilities of early arrears (PEA), default (PD) and deep default (PDD) being estimated. The coefficients are no longer interpreted as the probability of entering deeper states of default conditional on experiencing some repayment difficulty, but rather are interpreted relative to loans making full repayments. Secondly, the set of explanatory factors available to us in running this model extension is restricted relative to the baseline model using the SFS-LLD data due to the fact that the majority of *current* household-level explanatory factors are not available in the LLD. Thirdly, the model discussed in this section is run for a cross-section of loans at December 2013, rather than the pooled approach that was required in the baseline model of Section 4.[10]

5.1 Expanded data set

The LLD data are limited in the range of X_i available, but provide a view of the population of mortgages at four large mortgage lenders, covering 70 per cent of the market at December 2013. Table 7 reports summary statistics for the X_i for the 259,643 mortgages that are included in the LLD models.

The share of First Time Buyers (FTB) in the data set is 42 per cent. The average loan age is 88 months, indicating that the average loan was originated in 2007, which corresponds with the peak of the pre-crisis Irish property and credit cycles. The average interest rate, which is a proxy for loan affordability, is 329 basis points, while the average borrower is 36 years of age. The split across interest rate types reveals that Standard Variable Rate (SVR) mortgages account for 52 per cent of the sample, with tracker mortgages accounting for a further 38 per cent. The share of the market held by fixed rate mortgage holders is small at 9.4 per cent.

[10]Loans that have received a modification before December 2013, and had their arrears capitalized, will be reported as having zero DPD at December 2013. These loans would distort an empirical model which treats zero-DPD loans as performing loans. For this reason, we omit the 22,000 loans that have been modified and have a DPD count of zero from the LLD model. Given that our SFS data are collected at the point of engagement between lender and borrower, before a modification has been offered, this problem does not arise in the SFS-LLD models reported earlier in the paper.

Information on the marital status of borrowers at the point of origination indicates that 49 per cent of mortgages were taken out by married couples, with 45.6 per cent taken out by single borrowers, which can include cohabiting couples. Individuals who were separated or divorced at origination account for 5 per cent of the sample.

Employment status at origination is also available in the LLD data set. Given the limited set of borrower-level fields available in the LLD, this variable acts as our best household proxy for affordability, given that self-employed borrowers have more uncertain income paths and were more likely to be affected by the recession in Ireland from 2008 onwards. The vast majority (77 per cent) of mortgages are originated to households where the primary borrower is in salaried employment, with 18 per cent of mortgages going to self-employed individuals.

Table 7: Summary statistics, LLD

Variable	Obs	Mean	Std. Dev.
FTB	259643	0.422	0.494
Adjusted Equity	259643	47.753	25.340
CLTV	259643	79.601	44.591
Loan Age (Months)	259643	88.188	42.479
Current Interest Rate	259643	3.286	1.545
SVR	259643	0.523	0.499
Tracker	259643	0.382	0.486
Term	259643	314.8	84.9
Additional BTL	259643	0.047	0.211
Borrower Age	259643	36.094	8.378
Married	259643	0.493	0.500
Separated/Divorced	259643	0.051	0.221
Single	259643	0.456	0.498
Employed	259643	0.770	0.421
Other	259643	0.051	0.220
Self-Employed	259643	0.179	0.384

5.2 Empirical results

Our baseline results for the four-category multinomial model are reported in Table 8. Average marginal effects, $\frac{\delta Y}{\delta X}$ are reported. The impact of each X variable reported here is a *percentage point* increase in the probability, where the probability of early arrears (PEA) in the sample data is 4.7 per cent, the probability of default (PD) is 4.75 per cent, and the probability of deep default (PDD) is 6.76 per cent. Model (1) differs from Model (2) in the way in which housing equity is accounted for: the former includes our measure of Adjusted Equity, while the latter includes the more commonly-used $CLTV$.

For a number of important explanatory factors, we find results that are consistent with the extant literature on mortgage default. Kelly et al. (2014) have previously found that First Time Buyers are less likely to default than borrowers purchasing their second or subsequent home. Our results confirm their finding across all states of default: for the probability of early arrears (PEA), probability of default (PD) the estimates range between -0.1 and -0.3 per centage points, while for the probability of deep default (PDD), FTBs have a probability of distress that is 1.9 or 3.3 lower than non-FTB purchasers, depending on the model chosen.

Previous studies have consistently found an impact of housing equity on the probability of mortgage default across multiple jurisdictions. As discussed earlier, we contend that the current loan to value ratio ($CLTV$) is an inappropriate and endogenous explanatory variable in a cross-sectional study of mortgage default. Using Adjusted Equity as an alternative proxy, we find that an equity position that is ten per cent weaker is associated with .01 per cent increase in PD and a .02 per cent increase in PDD. The impact of Adjusted Equity on early arrears is statistically insignificant. The coefficients on the unbiased measure of housing equity in the PD and PDD equations are roughly one-fifteenth the magnitude of the analogous coefficient on $CLTV$ in Columns (4), (5) and (6). Such a large magnitude difference suggests that it is of crucial importance to consider the endogeneity caused by the mechanical feedback between missed payments, arrears capitalizations and $CLTV$ when modelling the depth of default.

Affordability in our LLD model can be proxied by the current interest rate. A 100 basis-point increase in the interest rate at December 2013 is associated with a 1 per cent increase in both PA and PD, and a 2.5 per cent increase in PDD. Longer-term mortgages are also shown to be higher risk, while having additional Buy-to-Let borrowings is not shown to lead to higher credit risk. The type of interest rate product chosen by borrowers is also an important correlate of mortgage default. We find that, relative to fixed rate mortgages, variable rate loans are higher risk across all categories. The impact of SVRs and tracker mortgages on PDD is particularly pronounced, at 9.28 and 14.5 percentage points, respectively.

Relative to married people, borrowers who were divorced or separated at origination are between 0.3 and 2.3 per cent more likely to enter each state of arrears. Single borrowers have a lower credit risk than married couples, with the differential ranging beween 0.5 and 1 per cent for PEA and PD. These findings may mask the role of subsequent family size changes and post-origination divorces, both of which are explicitly controlled for in the SFS-LLD model of Section 4.

The type of employment of a borrower is also shown to matter, with self-employed borrowers having monotonically increasing probabilities of repayment difficulty, with PEA, PD and PDD premia of 1.1, 2.1 and 5.4 per cent relative to salaried employees. This pattern intuitively reflects the more uncertain income paths and precarious job tenure of self-employed people.

Loan age and borrower age enter the model with both their level and squared term included. For each probability, the predicted value peaks at around 90 to 95 months, which equates to a loan originating in late 2006 or early 2007. In terms of levels, the predicted PEA and PD follow a similar path, while for values of loan age beyond four years, the PDD is higher by three to four percentage points. These differentials begin to narrow once more for loans older than 15 years, which are relatively rare in the data set.

Further evidence of the relative impact of housing equity and affordability can be disentangled from this estimation. This exercise, despite a limited set of covariates, is extremely informative given that the reference category contains truly performing loans, and the data sample comprises the relevant population of mortgages. In order for our measure of Adjusted Equity to have the same impact on PD and PDD as a one-standard deviation change in interest rates, the value would have to be 140 and 116 per cent higher, respectively (where the standard deviation of Adjusted Equity is 44 per cent). Along the same lines, for housing equity to have the same impact as the differential between self-employed and salaried borrowers, Adjusted Equity would need to be 285 and 254 per cent higher, respectively. These magnitude estimates suggest that, relative to house price developments, this model's (albeit imperfect and incomplete) proxies for affordability are the economically more important driver of mortgage default. This is in line with recent research from Gerardi et al. (2013). Given that the set of proxies for affordability in this model is far less complete than that in the baseline SFS model, we can safely conclude that the above estimates of the relative importance of affordability over equity are if anything an under-estimate.

Table 8: Full Loan Level Model, Average Marginal Effects

	Model 1			Model 2		
	Early Arrears	Default	Deep Default	Early Arrears	Default	Deep Default
FTB	-0.0104***	-0.0137***	-0.0192***	-0.0122***	-0.0182***	-0.0326***
	(0.00109)	(0.00110)	(0.00127)	(0.00107)	(0.00107)	(0.00116)
Adjusted Equity	0.00000528	0.0000747**	0.000211***			
	(0.0000232)	(0.0000233)	(0.0000267)			
CLTV				0.000292***	0.000674***	0.00189***
				(0.0000132)	(0.0000134)	(0.0000162)
Curr Int Rate	0.0102***	0.0105***	0.0246***	0.00809***	0.00527***	0.00928***
	(0.00114)	(0.00110)	(0.00129)	(0.00115)	(0.00112)	(0.00127)
Term	0.000196***	0.000241***	0.000340***	0.000108***	0.0000357***	-0.000250***
	(0.00000735)	(0.00000744)	(0.00000868)	(0.00000835)	(0.00000838)	(0.00000953)
Additional BTL	-0.00418	0.00957***	0.0181***	-0.00588**	0.00382*	-0.00647**
	(0.00220)	(0.00188)	(0.00212)	(0.00220)	(0.00188)	(0.00205)
Reference Category: Fixed Interest Rate						
SVR	0.0231***	0.0314***	0.0928***	0.0213***	0.0244***	0.0658***
	(0.00208)	(0.00227)	(0.00346)	(0.00207)	(0.00224)	(0.00311)
Tracker	0.0420***	0.0495***	0.145***	0.0320***	0.0231***	0.0632***
	(0.00458)	(0.00457)	(0.00588)	(0.00461)	(0.00457)	(0.00551)
Reference Category: Married Borrower						
Single	-0.0100***	-0.00534***	0.00336**	-0.0104***	-0.00641***	-0.000261
	(0.00103)	(0.00103)	(0.00119)	(0.00103)	(0.00101)	(0.00111)
S/D	0.00377*	0.00878***	0.0229***	0.00392*	0.00882***	0.0206***
	(0.00182)	(0.00175)	(0.00198)	(0.00182)	(0.00174)	(0.00189)
Reference Category: Salaried Employed						
Self-Employed	0.0116***	0.0213***	0.0537***	0.0123***	0.0218***	0.0504***
	(0.00126)	(0.00118)	(0.00129)	(0.00125)	(0.00118)	(0.00122)
Other	0.00650**	0.0156***	0.0355***	0.00641**	0.0166***	0.0405***
	(0.00204)	(0.00186)	(0.00219)	(0.00204)	(0.00186)	(0.00206)
N	259643	259643	259643	259643	259643	259643

Loan age (quadratic), borrower age (quadratic), bank dummies and county dummies included in all specifications
Reference category: Performing Loans. Robust standard errors in parentheses.

6 Conclusion

The existing literature on mortgage defaults has identified a number of robust factors that explain households' missed mortgage payments. These studies have treated all defaulted mortgages as homogeneous by virtue of their use of binary models. The issue of homogeneity among defaulted borrowers is of new importance given the response of many developed economies to avoid the repossession model in favor of loan modification and restructuring. Using a unique dataset on Irish mortgage borrowers, we extend the current literature by treating mortgages in deep states of arrears (greater than one year past due) differentially to those in earlier stages of default. Such a distinction is crucial given that previous work has shown that mortgages in deeper default are less likely to ever begin repayment ("cure"). These lower cure probabilities lead to higher estimates of Loss Given Default and expected losses for mortgage lenders.

The dataset available allows us to estimate the effect of an extremely rich set of explanatory factors including interest rates, housing equity, unemployment, income, non-mortgage debt volumes, household composition and divorce. Our estimates suggest that these factors explain mortgage default in a direction consistent with previous literature. In all cases, the impact on the probability of deep default is found to be larger than that on entering earlier stages of default. These findings suggest that affordability shocks are extremely important, and when they occur, they have severe impacts which lead to rapid accumulation of large arrears balances. Further, we add to the recent literature that suggests that mortgage affordability has an economically more meaningful effect on default than that of housing equity, with estimates of the impact of the former orders of magnitude larger than those of the latter. As well as identifying patterns that can help in the early identification of impending growth in arrears, these findings are key to the design and efficiency of mortgage modification schemes which can involve a large amount of public spending.

References

Bajari, Patrick, Chenghuan Sean Chu, and Minjung Park, "An Empirical Model of Subprime Mortgage Default From 2000 to 2007," NBER Working Papers 14625, National Bureau of Economic Research, Inc December 2008.

Been, Vicki, Mary Weselcouch, Ioan Voicu, and Scott Murff, "Determinants of the incidence of U.S. Mortgage Loan Modifications," *Journal of Banking & Finance*, 2013, *37.*

Bhutta, Neil, Jane Dokko, and Hui Shan, "The Depth of Negative Equity and Mortgage Default Decisions," Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System (U.S.) 2010-35, Board of Governors of the Federal Reserve System (U.S.) May 2010.

Boheim, Rene and Mark P. Taylor, "My Home Was My Castle: Evictions and Repossessions in Britain," *Journal of Housing Economics, Elsevier*, December 2000, *9* (4), 287–319.

Campbell, John Y., Stefano Giglio, and Parag Pathak, "Forced Sales and House Prices," *American Economic Review*, August 2011, *101* (5), 2108–2131.

Elul, Ronel, Nicholas S. Souleles, Souphala Chomsisengphet, Dennis Glennon, and Robert Hunt, "What "Triggers" Mortgage Default?," *American Economic Review, American Economic Association*, May 2010, *100* (2), 490–94.

Foote, Christopher L., Kristopher Gerardi, and Paul S. Willen, "Negative equity and foreclosure: Theory and evidence," *Journal of Urban Economics*, September 2008, *64* (2), 234–245.

Gerardi, Kristopher, Eric Rosenblatt, Paul S. Willen, and Vincent Yao, "Foreclosure externalities: Some new evidence," NBER Working Papers 18353, National Bureau of Economic Research, Inc September 2012.

_ , Kyle F. Herkenhoff, Lee E. Ohanian, and Paul S. Willen, "Unemployment, Negative Equity, and Strategic Default," Working Paper, Federal Reserve Bank of Atlanta 2013-04, Federal Reserve Bank of Atlanta August 2013.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "The determinants of attitudes toward strategic default on mortgages," *Journal of Finance*, August 2013, *68* (4), 1473–1515.

Gyourko, Joseph and Joseph Tracy, "Reconciling theory and empirics on the role of unemployment in mortgage default," *Journal of Urban Economics*, 2014, *80* (C), 87–96.

Jorda, Oscar, Moritz Schularick, and Alan Taylor, "The great mortgaging: housing finance, crises and business cycles," NBER Working Papers 20501, National Bureau of Economic Research, Inc 2014.

Kau, James B., Donald C. Keenan, Constantine Lyubimov, and V. Carlos Slawson, "Subprime mortgage default," *Journal of Urban Economics*, September 2011, *70* (2-3), 75–87.

Kelly, Robert, "The Good, The Bad and The Impaired - A Credit Risk Model of the Irish Mortgage Market," Research Technical Papers 13/RT/11, Central Bank of Ireland November 2011.

_ and Fergal McCann, "Mortgage defaults and the loan to value ratio," mimeo forthcoming, Central Bank of Ireland 2015.

_ and _ , "Under the hood: a flow analysis of Irish mortgage arrears," Economic Letter 15/EL/XX, Central Bank of Ireland 2015.

_ and Terence O'Malley, "A Transitions-Based Model of Default for Irish Mortgages," Research Technical Papers 17/RT/14, Central Bank of Ireland November 2014.

_ , _ , and Conor O'Toole, "Do first time buyers default less? Implications for macro-prudential policy," Economic Letters 14/EL/14, Central Bank of Ireland December 2014.

Kennedy, Gerard and Tara McIndoe-Calder, "The Irish Mortgage Market: Stylised Facts, Negative Equity and Arrears," *Quarterly Bulletin Articles*, February 2012, pp. 85–108.

Lydon, Reamonn and Yvonne McCarthy, "What Lies Beneath? Understanding Recent Trends in Irish Mortgage Arrears," *The Economic and Social Review, Economic and Social Studies*, 2013, *44* (1), 117–150.

Mayer, Christopher, Karen Pence, and Shane M. Sherlund, "The Rise in Mortgage Defaults," *Journal of Economic Perspectives*, 2009, *23* (1), 27–50.

McCann, Fergal, "Modelling default transitions in the UK mortgage market," Research Technical Papers 18/RT/14, Central Bank of Ireland November 2014.

McCarthy, Yvonne, "Dis-entangling the mortgage arrears crisis: The role of the labour market, income volatility and housing equity," Research Technical Papers 02/RT/14, Central Bank of Ireland 2014.

Mian, Atif, Amir Sufi, and Francesco Trebbi, "Foreclosures, House Prices, and the Real Economy," IMES Discussion Paper Series 11-E-27, Institute for Monetary and Economic Studies, Bank of Japan September 2011.

Qi, Min and Yang Xiaolong, "Loss Given Default of High Loan to Value Residential Mortgages," *Journal of Banking & Finance*, 2009, *33*, 778–799.

A Appendix

Table A.1: Definitions of default in micro-level studies of mortgage default

Study	Country	Dataset	Definition
Gyourko and Tracy (2014)	USA	Lender Processing Services Inc. Applied Analytics	90 DPD
McCarthy (2014)	Ireland	Central Bank of Ireland Loan Level Data	90 DPD
Gerardi et al. (2013)	USA	PSID Supplement on Housing, Mortgage Distress and Wealth Data	60 DPD
Lydon and McCarthy (2013)	Ireland	Central Bank of Ireland Loan Level Data	90 DPD
Kau et al. (2011)	USA	Black Box Logic LLC	Foreclosure
Kelly (2011)	Ireland	Central Bank of Ireland Loan Level Data	Three categories: 0 ; 0-90; 90+ DPD
Elul et al. (2010)	USA	Loan Performance and Lender Processing Services and Equifax data	60 DPD
Bhutta et al. (2010)	USA	LoanPerformance, First American CoreLogic	90 DPD for two consecutive months
Mayer et al. (2009)	USA	First American LoanPerformance	"Seriously Delinquent", 90 DPD
Bajari et al. (2008)	USA	LoanPerformance	Foreclosure
Foote et al. (2008)	USA	Warren Group, Massacheusetts Registry of Deeds	Foreclosure
Boheim and Taylor (2000)	UK	British Household Panel Survey	Survey response on payment difficulty

Table A.2: Comparison of loans with and without an SFS by arrears bucket

Average LTV December 2013				
	0-90	90-360	360+	All Arrears
No SFS	83.85	92.99	105.75	95.85
SFS	88.47	95.3	104.24	97.12
Average Balance				
	0-90	90-360	360+	All Arrears
No SFS	128,122	143,793	154,484	143,823
SFS	173,343	182,566	190,622	183,454
Average Interest Rate				
	0-90	90-360	360+	All Arrears
No SFS	3.62	3.52	3.73	3.64
SFS	2.99	2.96	3.05	3
Dublin Share				
	0-90	90-360	360+	All Arrears
No SFS	0.28	0.27	0.23	0.25
SFS	0.21	0.21	0.19	0.2
Share of Trackers				
	0-90	90-360	360+	All Arrears
No SFS	0.33	0.35	0.35	0.34
SFS	0.48	0.51	0.49	0.49
Share Married at Origination				
	0-90	90-360	360+	All Arrears
No SFS	0.53	0.51	0.49	0.51
SFS	0.61	0.58	0.54	0.57
Average Age				
	0-90	90-360	360+	All Arrears
No SFS	45.29	45.72	46.44	45.91
SFS	46.91	46.39	46.74	46.66

Table A.3: Multinomial Logit Results. $TinA$ not included in model. Average Marginal Effects Reported

	(Model 1)		(Model 2)		(Model 3)		(Model 4)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	90-360	360+	90-360	360+	90-360	360+	90-360	360+
FTB	0.0107	-0.0101	0.0131*	-0.00217	0.00714	-0.0346***	0.0100	-0.0257***
	(0.00773)	(0.00673)	(0.00777)	(0.00683)	(0.00763)	(0.00612)	(0.00769)	(0.00624)
Term	0.000217***	0.0000337	0.000221***	0.000112**	0.000153***	-0.000463***	0.000153***	-0.000360***
	(0.0000498)	(0.0000443)	(0.0000499)	(0.0000440)	(0.0000531)	(0.0000493)	(0.0000532)	(0.0000486)
SVR	0.0812***	0.123***	0.0813***	0.118***	0.0797***	0.110***	0.0795***	0.105***
	(0.0159)	(0.0171)	(0.0159)	(0.0166)	(0.0159)	(0.0163)	(0.0159)	(0.0160)
Tracker	0.145***	0.204***	0.146***	0.204***	0.153***	0.158***	0.152***	0.160***
	(0.0232)	(0.0242)	(0.0230)	(0.0236)	(0.0238)	(0.0240)	(0.0237)	(0.0236)
Curr Int Rate	0.0393***	0.0385***	0.0397***	0.0406***	0.0370***	0.0286***	0.0375***	0.0311***
	(0.00549)	(0.00496)	(0.00548)	(0.00489)	(0.00550)	(0.00516)	(0.00549)	(0.00507)
Equity Adjusted	0.0000614	0.000194**	0.0000792	0.000290***				
	(0.0000831)	(0.0000779)	(0.0000832)	(0.0000768)				
CLTV					0.000321***	0.00209***	0.000352***	0.00207***
					(0.0000914)	(0.0000878)	(0.0000908)	(0.0000848)
Other Debt to Income	0.000398	0.00408***			0.000383	0.00362***		
	(0.000467)	(0.000344)			(0.000460)	(0.000337)		
Other Debt Share			0.0445***	0.246***			0.0476***	0.239***
			(0.0122)	(0.00949)			(0.0121)	(0.00946)
Monthly Net Income	0.0000620	-0.00873***	-0.00150	-0.0201***	-0.000972	-0.0184***	-0.00275	-0.0293***
	(0.00234)	(0.00227)	(0.00247)	(0.00245)	(0.00237)	(0.00236)	(0.00250)	(0.00251)
Divorced	0.0258**	0.0598***	0.0260**	0.0643***	0.0252**	0.0482***	0.0257**	0.0526***
	(0.0113)	(0.0108)	(0.0113)	(0.0108)	(0.0112)	(0.0101)	(0.0112)	(0.0101)
Unemployment Shock	0.00197	0.0428***	0.00193	0.0430***	0.00376	0.0478***	0.00388	0.0469***
	(0.00680)	(0.00648)	(0.00680)	(0.00642)	(0.00682)	(0.00640)	(0.00683)	(0.00633)
DSR	0.0448***	0.105***	0.0435***	0.109***	0.0342***	0.0538***	0.0330***	0.0578***
	(0.0114)	(0.00906)	(0.0114)	(0.00884)	(0.0113)	(0.00862)	(0.0112)	(0.00840)
Single, 1/2	0.0137	0.0341***	0.0140	0.0334***	0.0139	0.0365***	0.0144	0.0349***
	(0.0115)	(0.0107)	(0.0115)	(0.0106)	(0.0115)	(0.0105)	(0.0115)	(0.0103)
Single, 3+	0.0406*	0.104***	0.0409*	0.108***	0.0411*	0.106***	0.0417*	0.109***
	(0.0214)	(0.0212)	(0.0214)	(0.0210)	(0.0214)	(0.0206)	(0.0214)	(0.0204)
Couple, Zero	0.0371***	0.00997	0.0379***	0.0120	0.0356***	0.00991	0.0365***	0.0117
	(0.0106)	(0.00910)	(0.0107)	(0.00901)	(0.0106)	(0.00892)	(0.0106)	(0.00880)
Couple, 1/2	0.0113	-0.00630	0.0125	-0.00188	0.0105	-0.000934	0.0120	0.00289
	(0.00920)	(0.00814)	(0.00922)	(0.00807)	(0.00917)	(0.00803)	(0.00919)	(0.00794)
Couple, 3+	0.0311***	0.0151	0.0322***	0.0210**	0.0313***	0.0286***	0.0326***	0.0339***
	(0.0108)	(0.00970)	(0.0109)	(0.00970)	(0.0108)	(0.00975)	(0.0108)	(0.00970)
Observations	20645	20645	20645	20645	20639	20639	20639	20639

Controls for county, loan age (quadratic), borrower age (quadratic) and bank included in all models
Robust standard errors in parentheses; * $p < .1$, ** $p < .05$, *** $p < .01$