5/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

November 2015

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 high loan-to-value ratios, consumer credit growth, shocks to mortgage affordability and unemployment should all trigger serious concerns among policy makers regarding subsequent stability in the mortgage market, with these 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, Reamonn Lydon, Terry O'Malley and Gerard O'Reilly for helpful discussions, along with seminar participants at the Central Bank of Ireland, University College Dublin, NUI Galway, University College Cork and the C.R.E.D.I.T. conference, Venice 2015. 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.

Non-Technical Summary

The Irish financial, economic and ensuing mortgage arrears crises have been among the largest experienced in the developed world. The Irish experience has also been unique in that, due to legal and political uncertainty, home foreclosures have been extremely rare by international standards. For this reason, there has been an unprecedented build-up in mortgages in long-term mortgage arrears (LTMA), which we define in this paper as relating to loans with arrears greater than twelve monthly repayments.

In this paper we use December 2013 data on roughly twenty thousand Irish mortgages to extend the economics literature's understanding of mortgage default by modelling the LTMA group as having distinct characteristics which differentiate them from those in the earlier stages of mortgage default. An understanding of the way in which household circumstances differentially impact the entry to long term, as opposed to early-stage arrears, can provide policy makers attempting to alleviate the LTMA crisis with valuable new information. Further, the findings have important implications for the design of mortgage modification policies across jurisdictions, particularly where taxpayer funds have been committed to resolving banking crises.

The results of our baseline model suggest that households experiencing an unemployment shock or a divorce have a three and two percentage point higher probability of LTMA, respectively. We show that the affordability of a mortgage is a crucial determinant of LTMA, with the monthly debt service ratio (DSR, measured as the ratio of mortgage repayment to net income) being strongly associated with LTMA. However, we extend the literature's understanding of the role of affordability in mortgage default by showing that it is the *shock* to mortgage affordability which is the most important factor: when the change in DSR between origination and our sample period is included, it is this affordability shock which drives entry to LTMA, while the level of the DSR loses its statistical significance.

Borrowers' non-mortgage leverage is also shown to play an extremely important role in driving long-term mortgage distress, whether measured as a ratio of non-mortgage debts relative to total debts or relative to income. Lower household incomes are also shown to have explanatory power in the deep default equation. Further, longer mortgage terms and higher mortgage interest rates are also shown impact LTMA.

Finally, housing equity considerations are shown to play an important role, with high Loan to Value ratios being associated with higher probabilities of LTMA. In their totality, the results can be interpreted as assigning a role to housing market shocks, labour market shocks, mortgage affordability, borrowers' debt accumulation, and family circumstances in explaining the extremely high rates of LTMA experienced recently in Ireland.

1 Introduction

The importance of the mortgage market to the banking system[1] and the economy at large cannot be overstated 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 three and two percentage point higher probability of deep default, respectively.[4] We show that the affordability of a mortgage is a crucial determinant of deep mortgage defaults, with a one-standard-deviation increase in the monthly debt service ratio (DSR, measured as the ratio of mortgage repayment to net income) leading to a two percentage point increase in the probability of deep default. However, we extend the literature's understanding of the role of affordability in mortgage default by showing that it is the *shock* to mortgage affordability which is the most important factor: when the change in DSR between origination and our sample period is included, it is this affordability

[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.

shock which drives entry to deep default, while the level of the DSR loses its statistical significance.

Borrowers' non-mortgage leverage is also shown to play an extremely important role in driving long-term mortgage distress, with a 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 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 that housing equity has a similar impact on the depth of mortgage default to a household unemployment shock. 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] Our finding suggests that the "double trigger" hypothesis appears to hold when considering long-term mortgage arrears during the Irish crisis, with both equity and affordability problems playing a role.

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 policy uncertainty that existed in Ireland between 2009 and 2013. A legal judgment passed in 2009

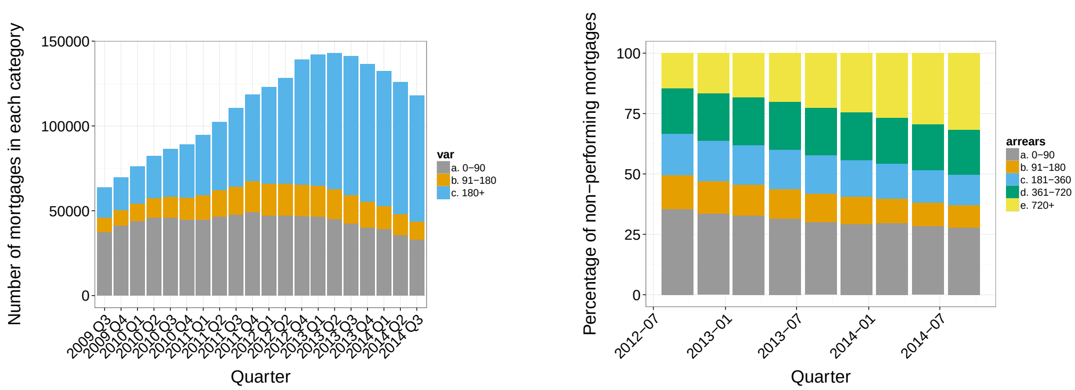
[5] Internal Central Bank of Ireland research shows that when borrowers have entered into arrears of greater than one year, the probability of any repayment is below 20 per cent, and falls even lower once borrowers enter arrears of more than two years.

[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).

rendered the repossession of homes in default extremely difficult, with the legal uncertainty only fully eradicated in 2013. 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, while Section 5 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 21,086 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	2,909	13.8
Q2 2012	3,599	17.07
Q3 2012	5,209	24.7
Q4 2012	2,914	13.82
Q1 2013	2,408	11.42
Q2 2013	1,912	9.07
Q3 2013	1,619	7.68
Q4 2013	516	2.45
Total	21,086	

2.1 Dependent Variable

The distribution of the depth of mortgage arrears among the 21,086 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 the SFS with zero DPD are not "performing" in a similar way to the majority of zero-DPD borrowers

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		10,120	47.99	
Early Arrears	1-90	12,797	4.77	29.3	3,541	16.79	32.3
Default	91-360	12,751	4.75	29.2	3,955	18.76	36.1
Deep Default	>360	18,141	6.76	41.5	3,470	16.46	31.6
Total		268,189			21,086		

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

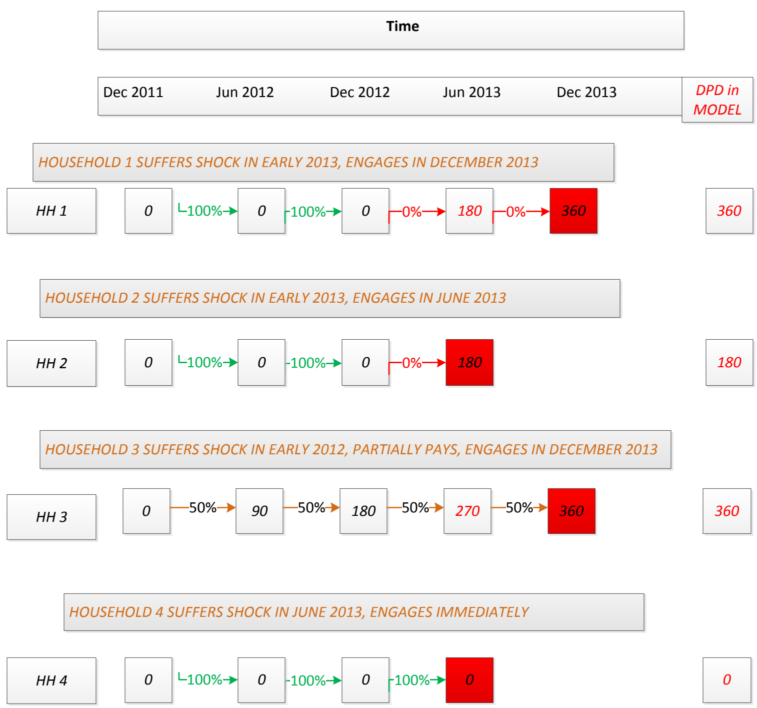


Figure 2 provides a visual representation of how households may end up with differing DPD values in our estimations. Household types 1 and 2 represent those discussed in the previous paragraph, where households differ only in the speed with which they decide to approach their bank. Household 3 has suffered 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.

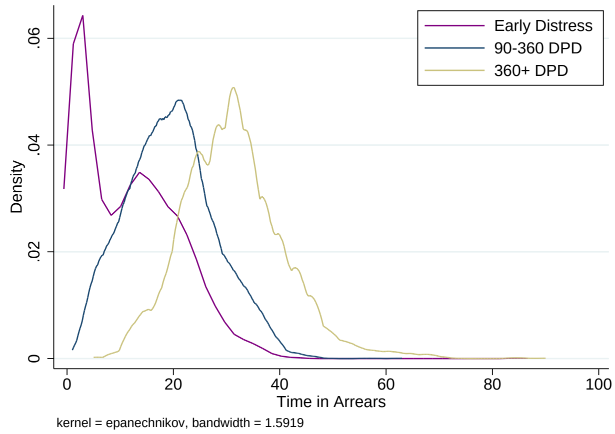
The final factor (6) underlying Y^* is the duration since the negative shock. It is important to acknowledge that the nature of our dependent variable is such that two households who have experienced an identical shock, and have an identical ability and willingness to repay, will have different DPD counts 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. Further, it may be that more financially vulnerable households have fewer resources available to withstand a shock and therefore enter arrears with greater frequency, for the same magnitude of shock, than those with greater savings or family resources to aid in continuing repayment.

For the reasons outlined above, we include the time, in calendar months, since a household first entered arrears as a control variable in our baseline models. This timing is directly observable due to the panel data nature of the LLD. 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 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

Figure 3: Time in Arrears (months)



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.

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 (2015), there is a mechanical reverse causality in the DPD-$CLTV$ relationship that is ignored by most researchers in this area. This bias is driven by the fact that, once a mortgage borrower stops payment, their outstanding balance no longer reduces along the monthly amortization schedule, while all performing loans continue to amortize as per contract terms. To exacerbate the effect, any arrears balance accumulated is often capitalised and added to the outstanding loan amount on the non-paying loans. This has the effect

Table 3: Summary statistics, SFS sample

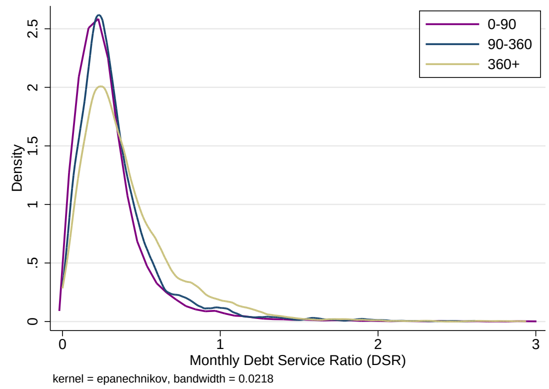
Variable	Obs	Mean	Std. Dev.
FTB	21,086	0.26	0.44
Term (Months)	21,086	315.81	84.19
Adjusted CLTV	21,086	86.24	44.1
CLTV	21,086	92.42	45.8
Loan Age	21,086	92.31	30.59
Current Interest Rate	21,086	2.94	1.55
Borrower Age	21,086	38.31	8.71
Fixed Rate	21,086	0.05	0.21
SVR	21,086	0.45	0.50
Tracker	21,086	0.50	0.50
Net Monthly Income (€000)	21,086	2.88	1.56
Unemployment Shock	21,086	0.30	0.46
Divorced Since Origination	21,086	0.07	0.26
Debt Service Ratio (DSR)	21,086	0.33	0.28
Δ DSR	19,941	0.049	0.248
Other Debt to Income	21,086	2.60	6.42
Other Debt to Total Debt	21,086	0.20	0.24
Single, No Children	21,086	0.17	0.37
Single, 1/2 Children	21,086	0.10	0.30
Single, 3+ Children	21,086	0.02	0.15
Couple, No Children	21,086	0.18	0.38
Couple, 1/2 Children	21,086	0.34	0.47
Couple, 3+ Children	21,086	0.20	0.40

that the numerator in the $CLTV$ will appear higher for non-paying loans in a cross-sectional regression setting, due mechanically to the fact that the loan is not paying. If not corrected for, this can lead to erroneous conclusions regarding the impact of housing equity on mortgage default. Due to this bias, we propose an alternate measure of housing equity which we term Adjusted $CLTV$, which adjusts downwards the observed CLTV values in the data set to account for the number of missed payments on defaulted loans.

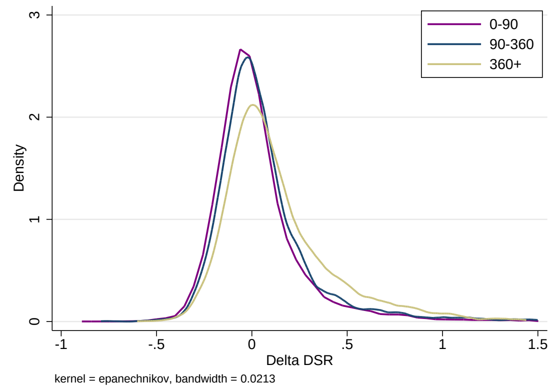
The borrower-specific controls, $\mathbf{X_i}$, include borrower age modeled as a quadratic term and indicators for change in martial status, family composition and the current employment status of the borrower.[8] In the model sample, 7 per cent of households have experienced a divorce since origination,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

[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.

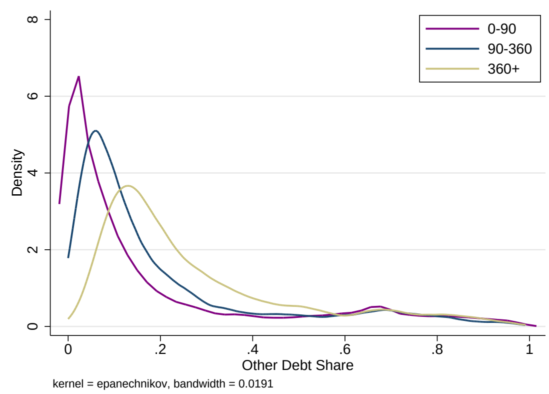
(a) DSR distribution, SFS sample.



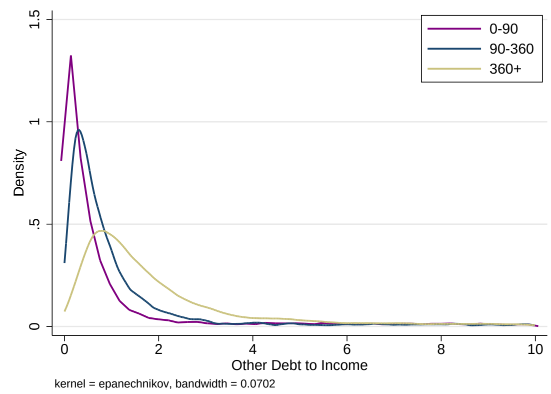
(b) Change in DSR since origination, SFS sample.



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



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



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 4a 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. The distribution is skewed further to the right among those in deep default, suggesting a weaker affordability position is associated with long-term arrears. Combining originating information and SFS information, we can construct a measure of the *shock* to affordability experienced by each household, which we refer to as Δ DSR. Such a combination of originating and current information provides an extremely useful indicator of mortgage distress which is rarely available to researchers in the area. Figure 4b shows that the distribution of these shocks for those in deep default sits clearly to the right of those in the other two groups.

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 calculate 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 4c, 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 debts. Similarly, in Figure 4d 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 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 $CLTV$ with the traditionally-used $CLTV$, it is clear that the unconditional relationship between the depth of default and housing equity is much less apparent when adjusting for the mechanical bias in the construction of $CLTV$.

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
Adjusted CLTV	86.2	87.7	84.7
CLTV	89.0	97.6	101.8
DSR	0.304	0.347	0.42
Δ DSR	0.025	0.059	0.135
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

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 $CLTV$ 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 models 1 and 3.

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 12-13 percentage points relative to fixed rate loans, while the analogous effect of SVRs is smaller at 4-4.3 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.

Both $CLTV$ and our adjusted measure are shown to have a positive impact on PDD and no impact on PD. A ten percentage point increase in $CLTV$ is estimated to lead to a 0.7 per cent increase in PDD, with a one-standard deviation change in both $CLTV$ and Adjusted $CLTV$ having a similar magnitude impact to an unemployment shock.

Mortgage affordability, as measured by the ratio of monthly repayment to monthly household 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 0.23-0.27 and 0.3-0.4 percentage points in PD and PDD, 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.

Monthly after-tax household income is found to impact the probability of deep default: a fall of €1,000 per month is associated with a 0.3-0.6 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

Table 5: Baseline regressions. 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	0.00343***	0.0159***	0.00356***	0.0156***	0.00343***	0.0156***	0.00355***	0.0152***
	(0.000157)	(0.000187)	(0.000159)	(0.000188)	(0.000163)	(0.000194)	(0.000165)	(0.000195)
FTB	0.0133*	-0.0179***	0.0140**	-0.0153***	0.0136**	-0.0185***	0.0143**	-0.0159***
	(0.00683)	(0.00471)	(0.00684)	(0.00470)	(0.00682)	(0.00470)	(0.00683)	(0.00469)
Term	0.000118**	-0.0000976***	0.000106**	-0.0000573	0.000115**	-0.0000849**	0.000103**	-0.0000438
	(0.0000478)	(0.0000355)	(0.0000479)	(0.0000356)	(0.0000467)	(0.0000342)	(0.0000469)	(0.0000343)
SVR	0.0311**	0.0433***	0.0321**	0.0409***	0.0308**	0.0434***	0.0318**	0.0411***
	(0.0148)	(0.0117)	(0.0148)	(0.0116)	(0.0148)	(0.0118)	(0.0148)	(0.0117)
Tracker	0.0841***	0.127***	0.0858***	0.126***	0.0838***	0.126***	0.0855***	0.125***
	(0.0222)	(0.0160)	(0.0223)	(0.0160)	(0.0222)	(0.0160)	(0.0223)	(0.0161)
Curr Int Rate	0.0184***	0.0359***	0.0188***	0.0361***	0.0186***	0.0350***	0.0190***	0.0352***
	(0.00491)	(0.00383)	(0.00489)	(0.00380)	(0.00491)	(0.00384)	(0.00490)	(0.00381)
Adjusted CLTV	0.0000194	0.000639***	0.0000323	0.000657***				
	(0.0000870)	(0.0000654)	(0.0000866)	(0.0000645)				
CLTV					0.0000286	0.000604***	0.0000426	0.000618***
					(0.0000788)	(0.0000563)	(0.0000785)	(0.0000556)
Other Debt to Income	-0.000363	0.00219***			-0.000351	0.00217***		
	(0.000382)	(0.000252)			(0.000381)	(0.000251)		
Other Debt Share			-0.00431	0.110***			-0.00398	0.109***
			(0.0112)	(0.00786)			(0.0112)	(0.00786)
Monthly Net Income	-0.00131	-0.00286*	-0.00185	-0.00582***	-0.00115	-0.00330**	-0.00170	-0.00626***
	(0.00214)	(0.00164)	(0.00223)	(0.00168)	(0.00215)	(0.00164)	(0.00223)	(0.00169)
Divorced	0.0110	0.0175**	0.0102	0.0208***	0.0112	0.0173**	0.0103	0.0206***
	(0.00975)	(0.00683)	(0.00973)	(0.00682)	(0.00975)	(0.00682)	(0.00973)	(0.00681)
Unemployment Shock	-0.00571	0.0282***	-0.00553	0.0288***	-0.00581	0.0284***	-0.00562	0.0290***
	(0.00598)	(0.00434)	(0.00599)	(0.00432)	(0.00598)	(0.00433)	(0.00598)	(0.00431)
Debt Service Ratio	0.0258***	0.0333***	0.0234**	0.0399***	0.0267***	0.0308***	0.0243**	0.0373***
	(0.00956)	(0.00675)	(0.00948)	(0.00665)	(0.00957)	(0.00675)	(0.00949)	(0.00665)
Single, 1/2=1	0.00928	0.0129*	0.00960	0.0127*	0.00920	0.0130*	0.00954	0.0128*
	(0.00993)	(0.00721)	(0.00993)	(0.00714)	(0.00992)	(0.00719)	(0.00992)	(0.00713)
Single, 3+=1	0.0209	0.0219*	0.0223	0.0219*	0.0211	0.0219*	0.0224	0.0218*
	(0.0189)	(0.0118)	(0.0190)	(0.0118)	(0.0189)	(0.0118)	(0.0190)	(0.0118)
Couple, Zero=1	0.0184**	0.00741	0.0195**	0.00687	0.0182**	0.00756	0.0193**	0.00700
	(0.00896)	(0.00643)	(0.00897)	(0.00635)	(0.00896)	(0.00643)	(0.00896)	(0.00635)
Couple, 1/2=1	0.00565	-0.00381	0.00620	-0.00291	0.00556	-0.00364	0.00613	-0.00276
	(0.00803)	(0.00581)	(0.00804)	(0.00576)	(0.00803)	(0.00581)	(0.00803)	(0.00576)
Couple, 3+=1	0.0105	0.00373	0.0115	0.00456	0.0104	0.00406	0.0114	0.00486
	(0.00912)	(0.00649)	(0.00913)	(0.00643)	(0.00911)	(0.00649)	(0.00913)	(0.00643)
Observations	21,086	21,086	21,086	21,086	21,086	21,086	21,086	21,086

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

Table 6: Income shock variable included. 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	0.00341***	0.0162***	0.00354***	0.0158***	0.00341***	0.0158***	0.00353***	0.0154***
	(0.000163)	(0.000194)	(0.000165)	(0.000195)	(0.000170)	(0.000202)	(0.000172)	(0.000202)
FTB	0.0112	-0.0179***	0.0121*	-0.0153***	0.0116*	-0.0187***	0.0125*	-0.0160***
	(0.00705)	(0.00490)	(0.00707)	(0.00490)	(0.00705)	(0.00489)	(0.00706)	(0.00489)
Term	0.000129***	-0.000107***	0.000119**	-0.0000647*	0.000126***	-0.0000938***	0.000116**	-0.0000518
	(0.0000497)	(0.0000370)	(0.0000499)	(0.0000371)	(0.0000485)	(0.0000356)	(0.0000487)	(0.0000357)
SVR	0.0336**	0.0432***	0.0348**	0.0404***	0.0334**	0.0433***	0.0345**	0.0406***
	(0.0154)	(0.0123)	(0.0154)	(0.0122)	(0.0154)	(0.0123)	(0.0154)	(0.0122)
Tracker	0.0918***	0.130***	0.0933***	0.128***	0.0915***	0.130***	0.0930***	0.128***
	(0.0228)	(0.0163)	(0.0229)	(0.0164)	(0.0228)	(0.0163)	(0.0229)	(0.0164)
Curr Int Rate	0.0197***	0.0368***	0.0201***	0.0367***	0.0199***	0.0360***	0.0203***	0.0359***
	(0.00505)	(0.00391)	(0.00503)	(0.00389)	(0.00505)	(0.00392)	(0.00503)	(0.00390)
Adjusted CLTV	0.0000371	0.000686***	0.0000454	0.000685***				
	(0.0000938)	(0.0000693)	(0.0000934)	(0.0000687)				
CLTV					0.0000446	0.000651***	0.0000536	0.000649***
					(0.0000851)	(0.0000599)	(0.0000849)	(0.0000594)
Other Debt to Income	-0.000286	0.00222***			-0.000275	0.00220***		
	(0.000405)	(0.000263)			(0.000404)	(0.000262)		
Other Debt Share			-0.00257	0.109***			-0.00219	0.108***
			(0.0118)	(0.00826)			(0.0118)	(0.00826)
Monthly Net Income	-0.00196	-0.00176	-0.00252	-0.00488***	-0.00180	-0.00216	-0.00238	-0.00526***
	(0.00224)	(0.00171)	(0.00233)	(0.00176)	(0.00224)	(0.00172)	(0.00234)	(0.00176)
Divorced	0.0121	0.0169**	0.0114	0.0199***	0.0122	0.0165**	0.0115	0.0195***
	(0.0102)	(0.00716)	(0.0102)	(0.00714)	(0.0102)	(0.00714)	(0.0102)	(0.00713)
Unemployment Shock	-0.00582	0.0285***	-0.00551	0.0292***	-0.00591	0.0287***	-0.00560	0.0294***
	(0.00621)	(0.00451)	(0.00622)	(0.00449)	(0.00621)	(0.00450)	(0.00621)	(0.00449)
Δ Debt Service Ratio	0.00339	0.0637***	0.000266	0.0387*	0.00155	0.0702***	-0.00154	0.0450**
	(0.0322)	(0.0229)	(0.0323)	(0.0228)	(0.0322)	(0.0230)	(0.0323)	(0.0228)
Debt Service Ratio	0.0181	-0.0112	0.0188	0.0190	0.0209	-0.0199	0.0215	0.0103
	(0.0322)	(0.0229)	(0.0323)	(0.0227)	(0.0322)	(0.0230)	(0.0324)	(0.0228)
Single, 1/2	0.00979	0.0122	0.0101	0.0122*	0.00977	0.0121	0.0101	0.0122*
	(0.0103)	(0.00744)	(0.0103)	(0.00738)	(0.0103)	(0.00742)	(0.0103)	(0.00736)
Single, 3+	0.0173	0.0237*	0.0185	0.0239**	0.0176	0.0235*	0.0188	0.0237*
	(0.0194)	(0.0122)	(0.0195)	(0.0121)	(0.0194)	(0.0122)	(0.0195)	(0.0121)
Couple, Zero	0.0164*	0.00852	0.0176*	0.00859	0.0163*	0.00840	0.0176*	0.00844
	(0.00933)	(0.00679)	(0.00934)	(0.00672)	(0.00932)	(0.00678)	(0.00934)	(0.00671)
Couple, 1/2	0.00456	-0.00374	0.00535	-0.00252	0.00453	-0.00374	0.00536	-0.00254
	(0.00834)	(0.00610)	(0.00835)	(0.00604)	(0.00834)	(0.00609)	(0.00835)	(0.00604)
Couple, 3+	0.00926	0.00392	0.0105	0.00489	0.00917	0.00408	0.0104	0.00502
	(0.00945)	(0.00677)	(0.00947)	(0.00670)	(0.00944)	(0.00676)	(0.00946)	(0.00670)
Observations	19941	19941	19941	19941	19941	19941	19941	19941

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

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.

In Table 6, we extend the analysis by replicating the model of Table 5 to incorporate the role of Δ DSR, our measure of the *shock* to mortgage affordability since origination. These estimates provide novel insights by showing that the level of mortgage affordability itself does not have a statistically significant impact on the depth of mortgage default once the affordability shock is controlled for. A ten percentage point increase in the DSR since origination leads to a 0.4 to 0.7 per cent increase in PDD, depending on the model specification. The estimated impact of FTB status, $TinA$, interest rates, interest rate type, unemployment, divorce and non-mortgage debts are all stable between Table 5 and Table 6, indicating that factors apart from our direct measure of mortgage affordability are not impacted by our inclusion of Δ DSR in the model.

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 of non-mortgage debts to total debt, the impact is 2.6 per cent. The equivalent effect for the DSR in Table 5 is 1.9 per cent, while the impact of Δ DSR in Table 6 is 1.6 per cent. Both $CLTV$ and our adjusted measure of housing equity have a one-SD impact of 2.8-2.9 percentage points on PDD, meaning that they are relatively large in magnitude, and as important as an unemployment shock in driving long-term mortgage arrears. 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.

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 binary default literature. Average marginal effect estimates on income, non-mortgage debts, divorce, unemployment, DSR are all larger, sometimes by orders of magnitude, in this specification than in Table 5. Many effects, particularly in the PD equation, become statistically significant once $TinA$ 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. On housing equity, we find that $CLTV$ is now estimated to significantly impact PD, and to have an MFX of .002 in the PDD equation, relative to an effect of .0007 in Table 5. This suggests intuitively that $TinA$ and housing equity are closely related, with the omission of $TinA$ from our specification leading to an erroneous tripling in the point estimate on $CLTV$. The results of Table A.3 suggest that there is important correlation between $TinA$ and our main explanatory variables, implying that a model that omits $TinA$ is likely to overestimate the importance of $\mathbf{X_i}$ and $\mathbf{Z_i}$.

5 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 present evidence that the "double trigger" impact is in operation when considering entry to deep mortgage arrears: housing equity is found to have a similar economic impact to an unemployment shock, and (in standard deviation terms) a larger impact than the level of non-mortgage debt outstanding. 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 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 of Table 5. Average Marginal Effects Reported

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
FTB	0.0153**	-0.0134**	0.0174**	-0.00501	0.00622	-0.0345***	0.00894	-0.0255***
	(0.00772)	(0.00659)	(0.00776)	(0.00670)	(0.00753)	(0.00606)	(0.00759)	(0.00618)
Term	0.000315***	-0.0000325	0.000317***	0.0000449	0.000153***	-0.000428***	0.000153***	-0.000327***
	(0.0000521)	(0.0000473)	(0.0000521)	(0.0000468)	(0.0000523)	(0.0000484)	(0.0000525)	(0.0000477)
SVR	0.0827***	0.121***	0.0830***	0.115***	0.0793***	0.110***	0.0792***	0.104***
	(0.0158)	(0.0168)	(0.0158)	(0.0163)	(0.0157)	(0.0161)	(0.0157)	(0.0158)
Tracker	0.154***	0.197***	0.154***	0.195***	0.150***	0.161***	0.151***	0.161***
	(0.0233)	(0.0241)	(0.0231)	(0.0235)	(0.0236)	(0.0239)	(0.0235)	(0.0235)
Curr Int Rate	0.0405***	0.0378***	0.0407***	0.0395***	0.0367***	0.0296***	0.0371***	0.0316***
	(0.00542)	(0.00494)	(0.00541)	(0.00487)	(0.00544)	(0.00510)	(0.00543)	(0.00502)
Adjusted CLTV	-0.000325***	0.000480***	-0.000305***	0.000559***				
	(0.0000970)	(0.0000922)	(0.0000968)	(0.0000900)				
CLTV					0.000334***	0.00202***	0.000363***	0.00200***
					(0.0000889)	(0.0000849)	(0.0000884)	(0.0000821)
Other Debt to Income	0.000329	0.00410***			0.000285	0.00372***		
	(0.000463)	(0.000339)			(0.000458)	(0.000333)		
Other Debt Share			0.0378***	0.243***			0.0430***	0.237***
			(0.0120)	(0.00932)			(0.0119)	(0.00930)
Monthly Net Income	0.00146	-0.0109***	0.000188	-0.0220***	-0.00149	-0.0190***	-0.00311	-0.0296***
	(0.00230)	(0.00226)	(0.00243)	(0.00243)	(0.00234)	(0.00232)	(0.00246)	(0.00247)
Divorced	0.0274**	0.0584***	0.0276**	0.0628***	0.0253**	0.0483***	0.0258**	0.0524***
	(0.0112)	(0.0107)	(0.0112)	(0.0106)	(0.0111)	(0.0101)	(0.0111)	(0.0101)
Unemployment Shock	0.00251	0.0414***	0.00253	0.0417***	0.00571	0.0455***	0.00583	0.0449***
	(0.00672)	(0.00640)	(0.00672)	(0.00635)	(0.00676)	(0.00633)	(0.00676)	(0.00626)
DSR	0.0540***	0.0992***	0.0530***	0.104***	0.0327***	0.0549***	0.0315***	0.0597***
	(0.0111)	(0.00888)	(0.0111)	(0.00868)	(0.0112)	(0.00852)	(0.0111)	(0.00831)
Single, 1/2	0.0136	0.0343***	0.0140	0.0336***	0.0147	0.0366***	0.0152	0.0353***
	(0.0113)	(0.0106)	(0.0113)	(0.0105)	(0.0113)	(0.0105)	(0.0114)	(0.0103)
Single, 3+	0.0351*	0.102***	0.0353*	0.106***	0.0366*	0.103***	0.0372*	0.106***
	(0.0209)	(0.0207)	(0.0209)	(0.0206)	(0.0210)	(0.0201)	(0.0210)	(0.0200)
Couple, Zero	0.0375***	0.00768	0.0382***	0.00971	0.0356***	0.00837	0.0365***	0.0103
	(0.0105)	(0.00896)	(0.0105)	(0.00888)	(0.0105)	(0.00881)	(0.0105)	(0.00871)
Couple, 1/2	0.0106	-0.00400	0.0117	-0.000127	0.0107	0.00126	0.0121	0.00484
	(0.00908)	(0.00807)	(0.00910)	(0.00799)	(0.00906)	(0.00797)	(0.00907)	(0.00789)
Couple, 3+	0.0264**	0.0181*	0.0274**	0.0235**	0.0295***	0.0301***	0.0306***	0.0352***
	(0.0106)	(0.00966)	(0.0106)	(0.00963)	(0.0106)	(0.00970)	(0.0107)	(0.00964)
Observations	21086	21086	21086	21086	21086	21086	21086	21086
Pseudo R^2								

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