Mortgage Borrower Income and Default Behavior

Hung Xuan Do[1], Harald Scheule[2 3]

Abstract

This paper analyses the impact of mortgagee income on default behavior in the context of residential mortgage loans. We find that level and sensitivity of default risk are greater for low borrower incomes due to thinner financial buffers such as discretionary income, funding ability and liquid assets. The finding persists for positive home equity indicating liquidation constraints. Further, we find that an increase in the default rate of borrowers is convex across levels of income reduction. This rate increases between 0.305 to 0.757 percentage points (0.126 to 0.225 percentage points) for every 1% increase in the proportion of low-income (high-income) borrowers facing a 15%+ to 35%+ income reduction. Our findings assist banks in anticipating future losses and providing lending products that are socially more equitable.

Keywords: Default behavior • Endogeneity • Financial Buffers • Home Equity • Household Liquidity • Income • Mortgage • Non-linearity • Propensity Score Matching • Social equity

JEL Classification: G21 • G28 • C19

[1] Corresponding author. School of Economics and Finance, Massey University, New Zealand, h.do@massey.ac.nz

[2] Finance Discipline Group, UTS Business School, University of Technology, Sydney, harald.scheule@uts.edu.au

[3] The authors would like to thank the participants of the seminars at Macquarie University, Massey University, University of Technology Sydney, and the 2019 Credit Risk Modelling Forum. Hung Do acknowledges research support provided by Massey University Research Funding, project number: 1000023037.

1. Motivation

Coibion et al. (2020) find that the current COVID-19 crisis results in income shocks to predominantly low-income households. These findings foreshadow the adverse impact of bank loan defaults through the income channel, where low incomes may imply greater chances of default as a greater relative income reduction may imply a greater inability to service minimum mortgage payments.

Banks are aware that borrower income affects the serviceability of loans, and that income is negatively related to default behavior of a borrower, an important parameter of credit risk. They include ratios such as debt-to-income as well as future stress scenarios in their underwriting standards. As a consequence, lower income borrowers receive lower credit volumes. However, default attachment levels should be comparable across income groups after loan approval. This paper finds that this prior does not hold and suggests socially responsible bank lending and prudential regulation should address this gap. Solutions may include changes to mortgage design or support of building financial buffers.

Further, our research enables banks to more accurately measure the probability of default, which is a key input into bank capital allocation, loan loss provisioning, and loan pricing. Ultimately, these changes will increase the resilience of the banking industry and hence, the financial system.

Consumer borrowers generally fail for two reasons: negative equity and illiquidity. In perfect financial markets both triggers need to happen to lead to a default event, which is known as the Double-Trigger Model (DTM). In economies that provide for recourse lending, borrowers continue to perform if only a single trigger hits. For example, borrowers with negative equity and liquidity continue to service loans from their ongoing income. Borrowers with positive equity and liquidity constraints continue to service loans by extracting home equity. We combine these observations and analyze the effect of borrower income on the sensitivity to negative income shocks based on unemployment, disability, and divorce. Figure 1 shows the implied incremental default rates for low- and high-income borrowers (above and below median income at a given time period) for different payment shock magnitudes.

Insert Figure 1 here

We observe greater default rate levels and a greater sensitivity to income shocks (slope) for low-income borrowers. We find that this is due to lower financial buffers to offset future serviceability shocks.

The paper makes four distinct contributions. First, we find that thinner financial buffers such as discretionary income, funding ability and liquid assets explain a greater sensitivity of default rates to income shocks. Second, the sensitivity of borrowers to income shocks is greater for low-income earners. Third, the effects persist for positive equity loans. This shows that borrowers with positive equity may default given adverse income shocks and that variations in the default rate persists for different income levels. Positive equity borrowers are constraint to liquidate their home equity when failing to service their mortgage loans. To some extent, this finding is consistent with the behavior of the first-time online borrowers reported in Bhanot (2017).[4] We observe from the Panel Study of Income Dynamics data over the waves 2009, 2011, and 2013 that approximately 60% of the 60-day delinquencies are associated with positive equity. Our findings indicate that a majority of the mortgagees are unable to pay rather than strategically default on their loans. Fourth, the relation between income shock and probability of default is convex. This implies that an increase in income shocks results in

[4] Bhanot (2017) studies the behavior of 4883 first time online borrowers and find that borrowers who default on their loans not due to dishonesty or behavioral biases, but primarily due to "they suffer from true financial hardship and are simply unable to pay".

probabilities of default that increase at an increasing rate. We provide various robustness checks for endogeneity, borrower and lender anticipation, as well as economic impact.

The paper is structured as follows. Section 2 reviews the literature in the areas of mortgage default risk modelling, serviceability, income and financial buffers and derives the research hypotheses. We analyze the impact of income shocks and income levels on financial buffers (H1), on default events (H2) and on default events for positive home equity (H3), and analyze the impact of the shock magnitude (H4). Section 3 provides an economic motivation for the default sensitivities by income and provides a decomposition of these sensitivities by financial buffers, as well as financial buffer sensitivities by income. Section 4 describes the data and the variable definitions. Propensity score matching is applied to ensure homogeneity between low- and high-income borrowers. Section 5 presents the empirical analysis as the main body of work and analyzes the four hypotheses with a number of robustness checks. Section 6 concludes and provides policy recommendations.

2. Research question

2.1. Double-Trigger Model: equity and serviceability as key drivers of mortgage default risk

Prior literature has analyzed negative equity and loan serviceability as key drivers of mortgage default, which is also known as the Double-Trigger Model (DTM). Liquidity is often not observed and included by proxies. Examples include Goldberg and Capone (2002) who use the debt coverage ratio as a proxy for liquidity, and Elul et al. (2010) who use the current utilization rate as a proxy for liquidity. More recent extensions to the DTM include Corradin (2014), Campbell and Cocco (2015), Laufer (2018) and Schelkle (2018), who have proposed life-cycle models consider (i) the uncertainty of a borrowers' incomes over their lifetime and

(ii) limits on the unsecured borrowing capacity. Negative equity is an essential condition for default, but the default likelihood increases with the severity of adverse life events. Most modern mortgage risk models include LTV (loan-to-value) as a proxy for negative equity (leverage) and DTI (debt-to-income) as a proxy for negative liquidity.

In perfect markets, a single equity or liquidity constraint does not trigger default. Borrowers with negative equity and no liquidity constraint continue to service loans from their ongoing income. Borrowers with positive equity and a liquidity constraint are able to delay payments by refinancing.

2.2. Liquidity constraints

Our research is mostly related to liquidity constraints, which are measured by exogenous serviceability shocks and may come from the income and expense channel. Income shocks have been analyzed by Gathergood (2009), Gerardi et al. (2018), and Cunningham et al. (2020).[5] Gathergood (2009) finds that unemployment, illness, and divorce shocks increase the probability for households to encounter payment difficulties, while an increase in the number of dependent children has no significant impact. Further, repayment difficulties are shown to be strongly correlated with ex ante repayment risk. This implies that the risk of experiencing an income shock may be observed at loan origination and lenders or borrowers may adjust the loan terms accordingly. Gerardi et al. (2018) analyze the impact of income shocks due to unemployment, divorce and disability and contrast the results to the effects of home equity. They find that many defaults are not linked to liquidity constraints, which may support the role of negative equity. Expense shocks have been analyzed in relation to the end

[5] Further, literature has analyzed credit improvements through positive income and negative expense shocks due to interest rate decreases post financial crises. See e.g., Fuster and Willen (2017) for hybrid mortgages at the rate reset, Tracy and Wright (2012) for prime adjustable-rate mortgages, Adelino et al. (2013), Haughwout et al. (2016) and Agarwal et al. (2011) for loan modifications. Cunningham et al. (2020) employ the fracking boom in Pennsylvania as a “positive” income shock that significantly increased the household income. This led to a substantial decrease in the mortgage default risk.

of draw period for home equity lines of credit (HELOCs), adjustable rate mortgages (ARMs) and liquidity conditions at rate reset. For example, Qi et al., (2020) find that the payment shock and refinance constraints at end of draw of HELOCs increases credit risk. Based on the US Survey of Consumer Finances, Johnson and Li (2013) find that ARM borrowers are more constrained than fixed-rate mortgage (FRM) borrowers. Zorn and Lea (1989) find that the probability of default among Canadian ARM borrowers rises with increases in mortgage rate and decreases with home equity. Chiang and Sa-Aadu (2014) find that theoretically, pay option ARMs are the optimal choice for constrained borrowers. Ambrose et al. (2005) analyze hybrid mortgages (i.e., mortgages with a period of fixed interest rates and a period of variable interest rates) and find relatively high rates of default at the conversion time from fixed to adjustable rate payments. Further, Mayer et al. (2009) highlight the positive causality of rate resets at the end of the teaser period and delinquency risk.

2.3. Financial buffers

Gerardi et al. (2018) find that 80% of these borrowers who cannot pay do not default. We speculate that this be explained by running down of financial buffers. Literature has focused on the role of home equity as a financial buffer, which can be extracted through second lien loans, i.e., home equity lines of credit. Chen et al. (2020) build a structural model for the economy to explain household liquidity management through refinance and include liquid assets next to house prices. Calem et al. (2011) find that line draws increase when unemployment increases and when borrowers are more liquidity constrained. Agarwal et al. (2006) find that borrowers originate HELOCs if they anticipate a decline in their future credit quality, and borrowers with lower credit scores at origination have lower initial line utilization than borrowers with greater credit scores.

2.4. Hypothesis development

The literature on financial buffers is slim and their role in impacting the sensitivity to income shocks and for different income levels has not been analyzed to date. Such an analysis requires income, consumption and wealth data, which is often not available for lenders. Consumer surveys may offer a valuable alternative sources but are limited in terms of observation counts and credit-relevant information as they focus on more holistic issues.

We therefore analyze the impact of income levels on financial buffers and default of mortgage borrowers, which is a gap in the exiting literature. Low-income borrowers have lower relative levels of financial buffers and hence, less room to (i) reduce (discretionary) consumption, (ii) access funding, and (iii) sell liquid assets. Thus, our first research hypothesis is:

H1: High-income borrowers reduce financial buffers more than low-income borrowers, given an adverse income shock.

Borrowers may default after these three buffers – consumption adjustment, borrowing capacity and liquid assets – are exhausted. The role of income on household debt and debt default has been analyzed and banks collect variables like the debt-to-income ratio at loan origination. The impact of income shocks on default risk has not been analysed by borrower income. Therefore, our second research hypothesis is:

H2: Low-income borrowers' default risk increases more than high-income borrowers' default risk given an adverse income shock.

We are interested to see whether H2 remains true if borrowers have positive equity which may be used to extract equity through loan refinance, second lien loans or house sales. This is a test to see whether our contributions are aligned with the DTM or a more substantial extension. We apply the key research question H2 to a sample of loans with positive home equity. Our third research hypothesis is:

H3: Low-income borrowers' default risk increases more than high-income borrowers' default risk given an adverse income shock and positive equity.

Finally, we are interested in whether the sensitivities to the payment shock of low- and high-income borrowers are sensitive to the magnitude of the payment shock, i.e., whether the relation is convex (exacerbating) or concave (saturating). The income shocks may have a greater increase, the greater the shock as existing financial buffers offset the shock to lowers extents. Hence, our fourth research question is:

H4: Impact shock sensitivity is greater for high-income shocks and low-income borrowers

Figure 2 summarizes the causality chain of mortgage default and our four research questions:

Insert Figure 2 here

3. Economic motivation

Our aim is to understand how income shocks, which is a key driver to credit default risk, affects the probability of default for the whole population and different income groups. We divide the population into a low- and a high-income group and develop a simplified model for our central hypotheses that income shocks result in a lower drawdown of financial buffers (H1), and a greater increase of default risk (H2 to H4) for low-income borrowers.

We denote the proportion of the low-income borrowers as τ and the proportion of the high-income borrowers as $1 - \tau$. We assume that the proportions of the low and high-income borrowers are stable over time, which means τ is time-constant. The default probability of the population can be represented as:

$$d_t = \tau d_t^{(l)} + (1-\tau) d_t^{(h)} \tag{1}$$

in which, $d_t^{(l)}$ and $d_t^{(h)}$ are the default probabilities of the low- and the high-income borrower groups, respectively.

In addition, we denote the proportions of the low- and high-income borrowers that face an income shock at time t as $\delta_t^{(l)}$ and $\delta_t^{(h)}$, respectively. Intuitively, we have $\delta_t^{(l)} > \delta_t^{(h)}$. From Equation (1) we can derive the expected default probability of the population at time $t+1$, given the income shocks at time t as follows:

$$\begin{aligned} E\left(d_{t+1} \middle| \delta_t^{(l)}, \delta_t^{(h)}\right) &= E\left(\tau d_{t+1}^{(l)} + (1-\tau) d_{t+1}^{(h)} \middle| \delta_t^{(l)}, \delta_t^{(h)}\right) \\ &= \tau E\left(d_{t+1}^{(l)} \middle| \delta_t^{(l)}\right) + (1-\tau) E\left(d_{t+1}^{(h)} \middle| \delta_t^{(h)}\right) \end{aligned} \tag{2}$$

where $E(d_{t+1}|\delta_t)$ is the expected default probability at time *t+1* given the adverse income shock at time *t*. *l* denotes the low-income group and *h* denotes the high-income group.

A certain proportion of each group may default after adverse income shocks. We denote the default rates that low- and high-income borrowers who face income shocks as $\rho^{(l)}$ and $\rho^{(h)}$. This can be respectively calculated as:

$$\rho^{(l)} = \frac{E\left(d_{t+1}^{(l)} \middle| \delta_t^{(l)}\right)}{\delta_t^{(l)}} \text{, or } E\left(d_{t+1}^{(l)} \middle| \delta_t^{(l)}\right) = \rho^{(l)} \delta_t^{(l)} \tag{3}$$

$$\rho^{(h)} = \frac{E\left(d_{t+1}^{(h)} \middle| \delta_t^{(h)}\right)}{\delta_t^{(h)}} \text{, or } E\left(d_{t+1}^{(h)} \middle| \delta_t^{(h)}\right) = \rho^{(h)} \delta_t^{(h)} \tag{4}$$

Given a 1% increase in the proportion of borrowers in a group facing adverse income shocks, Equations (3) and (4) show that the expected default rate in the next period increases by $\rho^{(l)}$% for the low-income group, and $\rho^{(h)}$% for the high-income group. Note that by definition, we have $0 \leq \rho^{(l)}, \rho^{(h)} \leq 1$ as the number of borrowers defaulting following an adverse income shock cannot be greater than the total number of borrowers facing the adverse income shock.

Let γ be the scale which measures the difference in the sensitivities to default of low- and the high-income groups. We can therefore link the sensitivities to default of the two groups via γ as, $\frac{\rho^{(l)}}{\rho^{(h)}} = \gamma$, or $\rho^{(l)} = \rho^{(h)}\gamma$. γ is the marginal relationship between the default sensitivities of low- and high-income groups, in which a 1% increase in the default probability of the high-income group is associated with γ% increase in the default probability of the low-income group, all else being equal.

We argue that low-income borrowers are more sensitive to an adverse income shock than high-income borrowers, as low-income borrowers have fewer financial buffers than high-income borrowers to overcome the adverse shocks, i.e., (i) less room to adjust their consumption expenses, (ii) less access to borrowing sources, and (iii) fewer liquid assets. Therefore, $\rho^{(l)} \geq \rho^{(h)}$, or $\gamma \geq 1$. We relate the marginal relationship between the default sensitivity of low- and high-income groups, γ, to the aforementioned financial buffers as follows:

$$1 \leq \gamma \leq \left(\frac{Discretionary\ expense^{(h)}}{Discretionary\ expense^{(l)}}\right)\left(\frac{Borrowing\ capacity^{(h)}}{Borrowing\ capacity^{(l)}}\right)\left(\frac{Liquid\ Asset^{(h)}}{Liquid\ Asset^{(l)}}\right) \tag{5}$$

The left equilibrium (i.e., $\gamma = 1$) holds when the difference in the financial buffers between the two income groups does not translate into the difference in their default sensitivities, or there is no difference in their financial buffers. Meanwhile, the right equilibrium holds when the differences in financial buffers between the two income groups fully translates into the differences in their default sensitivities.

4. Data

4.1. Data source, filters, and propensity score matching

We employ the dataset underlying Gerardi et al. (2018).[6] The data is based on the Panel Study of Income Dynamics (PSID) with Supplements on Housing, Mortgage Distress, and Wealth Data. The survey is the world's longest running household panel survey and is conducted every two years by the University of Michigan. The data provides information about household demographics, income and consumption, wealth, mortgage finance including their mortgage feature and performances. The survey includes the children of respondents as they form their own households and maintains its representative character over time.

After filtering, we obtain 7,404 observations, with 248 60-day-delinquency events in the three survey waves 2009, 2011 and 2013.[7] We chose the three waves as the Global Financial Crisis (GFC) resulted in a larger number of default events and limited the number of waves to ensure a homogeneous population and to limit time variation. We use the terms "household" and "borrower" interchangeably. Households generally consist of two adults with up to two incomes and the loss of income generally relates to these two household constituents.

As our focus is to investigate and compare an impact of adverse income shocks on low- and high-income groups, we firstly ensure the two groups are comparable in terms of borrower

[6] We are grateful and thank Kyle Herkenhoff for kindly sharing this data.
[7] The sample is filtered to include households with heads of family between 24 and 65 years old, LTV < 250% and who have not defaulted in prior surveys. For convenience, we refer the households included in the data to the residential mortgage borrowers in our subsequent discussions.

characteristics. We categorize household into two income groups based on whether income is below median (labeled "low-income") or above median (labeled "high-income") per wave to account for inflation:

$$inc\_b50_{i,t} = \begin{cases} 0, \text{ if household income is above median income} \\ 1, \text{ if household income is below median income} \end{cases} \quad (6)$$

We employ Propensity Score Matching (PSM) to perform a one-to-one matching between low-income (the treated group) and high-income households (the control group) using a rich set of borrower demographic variables including: household details, occupation, education level and race as shown in Table 1. The propensity score in our case represents the probability that a borrower is in the low-income group based on their personal characteristics. We employ radius matching with a caliper of 0.05 for the PSM process (see e.g., Smith, 2016). Out of 7,404 total households, we arrive at 2,445 observations for each income group over the three survey waves, leading to a total 4,890 observations in our final sample. We report the means and the difference in means of the matching variables between the two income groups in Table 1. After PSM, we find that the two groups show no significant difference in terms of borrower characteristics.

Insert Table 1 here

All regressions are based on weighting of the optimization algorithm (estimation least squares for OLS and 2SLS models and maximum likelihood for probit regressions) to reweight the data to the LTV distribution of the McDash/Equifax dataset, which is broadly considered to be a nationally representative loan-level mortgage servicing dataset.[8]

[8] Our main findings remain consistent with no reweighting considered.

4.2. Dependent and test variables

Our dependent variables in Hypothesis 1 are financial buffers and include discretionary expenses ratio, borrowing capacity ratio and liquid assets ratio.

To construct the discretionary expenses of households, we include the annual consumption for trips and vacations, and the annual consumption for other recreational activities. For example, for the survey conducted in 2013, the reported figures are the discretionary expenses consumed in 2012. We standardize these total discretionary expenses by the annual mortgage repayment (i.e., the annual interest and principal repayment) to represent the first financial buffer that the borrower may rely on following an adverse income shock. We argue that the financially constrained borrower would use this buffer to mitigate financial hardship.

The second financial buffer is the borrowing capacity. We measure the borrowing capacity of a borrower as one minus the ratio of the total outstanding of mortgage loans and all other debts over the total home value. Banks commonly lend up to 80% of home value through mortgage loans or HELOCs. The home value of mortgagees may increase over time due to mortgage amortizations (i.e., repayment) or house price increases (see e.g., Chen et al., 2020).

For the liquid assets that borrowers can readily exploit to cope with the adverse income shock, we include the reported cash and cash equivalent accounts (such as check and savings accounts), highly liquid short-term investments (e.g., money market funds and certificates of deposits), and highly liquid investments with virtually no credit risk (such as government saving bonds and the treasury bills). We standardize the liquid asset by annual mortgage payments (interest and principal payments) to obtain the liquid asset ratio as our third financial buffer.

In the remaining hypotheses, our dependent variable is the default indicator. Mortgage defaults of borrower *i* at time *t* are defined as follows:

$$Default_{i,t} = \begin{cases} 0, \text{ if no mortgage default reported} \\ 1, \text{ if mortgage default reported} \end{cases} \tag{7}$$

Note that the number of default events limits our study in terms of the experimental design, including the number of features that we can study.

Mortgage default is defined in the lending industry and prudential regulation based on a set of triggers which include delinquency, borrower insolvency and foreclosure (see Qi et al. (2021) for more details). We generally define default as payment delinquencies of more than 60 days and analyze a payment delinquencies of more than 90 days as a robustness check. We do not analyze foreclosures as the observation numbers are low.

Our main income shock is the unemployment event. We also consider other income shocks relating to divorce and disability by studying their combination as an additional variable that indicates if any of these three shocks occurs (*Shock3*). For robustness checks, we replace the unemployment shock with involuntary unemployment shock, which is discussed in more detail in Section 5. We dummy code the shock variables as follows:

$$Shock = \begin{cases} 0, \text{ if no income shock reported} \\ 1, \text{ if income shock reported} \end{cases} \tag{8}$$

4.3. Data description

Table 2 describes the dependent variables and test variables for default and non-default.

Insert Table 2 here

Defaulted households have lower financial buffers, lower incomes, greater combined LTV and greater income shock rates than non-default households.

We add borrower demographics, mortgage characteristics, state laws and economic conditions as control variables. Borrower demographics include the age of the household head, the number of children in a household, and dummies to indicate education levels, industries that the household head works in, and race. Mortgage characteristics include the combined LTV (i.e., all secured loans on a property over the value of a property), mortgage interest rate and indicator variables for a remaining maturity of the first lien loan greater than 15 years, the presence of a second lien loan, adjustable rate mortgage, whether a loan has been refinanced, and origination years.[9] State indicators include whether a state offers judicial foreclosure or recourse lending. The judicial process and statutory right of redemption may prolong the foreclosure and liquidation process. The recourse and judicial process are important default factors as they govern what happens to the borrower and collateral assets in the instance of a default (compare Qi and Yang, 2009). Finally, we include the economic conditions, as well as whether the state experienced a substantial house price appreciation and bust during the 2000s.[10]

Table 3 shows summary statistics for the main control variables for non-defaults and defaults. Defaulted households have lower education levels, greater interest rates, longer remaining maturities, more ARM loans and fewer refinanced loans.

Insert Table 3 here

Our research hypotheses focus on the role of high- and low-incomes. Table 4 describes the dependent variables and test variables for low (i.e., below median income) and high income

[9] Surveys dominate bank data as they are able to track borrowers over refinancing decision while banks can no longer observe loans after refinance.
[10] During the 2000s, there were four states in the US which experienced a significant housing boom including Arizona, California, Florida and Nevada.

(i.e., above median income). Low-income borrowers have a greater default rate, lower financial buffers, greater Combined LTV and a greater proportion of experienced shocks. For divorce, high incomes have a greater divorce rate than low incomes.

Insert Table 4 here

If we filter out the households with income shocks, we find that the difference in the default rate between the two income groups is insignificant.[11] Most default events are observed as a consequence of an income shock, as the default rates for borrowers experiencing an income shock are substantially higher and confirm the causality structure of our economic motivation.

We control for differences in levels of credit risk between the high- and low-income groups through our difference-in-difference setup (Effects 5 and 6 in Figures 3 and 4), the analysis of surprise income shocks as the difference between the shock the model-implied shock level in a robustness check, and the analysis of income shocks for income groups that are based on propensity score matching and clustering into high and low income levels in another robustness test.

Table 2 (test variables) and Table 3 (control variables) include tests for differences in means between defaulted and non-defaulted borrowers and Table 4 (test variables) for low- and high-income borrowers. Note that control variables are comparable for low- and high-income groups, as shown in Table 1, due to propensity score matching. Most test outcomes are significant. To summarize, this univariate analysis suggests that low-income borrowers have fewer financial buffers to rely on and are thus more exposed to income shocks.

[11] We do not report this figure in the table to conserve space. However, detail is available upon request.

5. Empirical analysis

5.1. Financial buffers

Prior to a default, borrowers try to reduce financial buffers to mitigate adverse income shocks. We now analyze whether borrowers with different income levels react differently to negative income shocks due to different levels of financial buffers.

We apply three difference-in-difference OLS regressions for the impact of income shocks on three financial buffers, discretionary expenses ratio, borrowing capacity ratio and liquid assets ratio.[12] All three financial buffers are metric and require linear models. The general model equation for financial buffers of borrower *i* at time *t* is specified as follows:

$$Financial\ Buffer_{i,t} = \alpha_0 + \alpha_s shock_{i,t} + \alpha_L inc\_b50_{i,t-1} + \alpha_{DID} inc\_b50_{i,t-1} \times shock_{i,t} + \alpha_C Controls \ + \varepsilon_{i,t} \tag{9}$$

As unemployment shocks account for the majority of the adverse income shock events (see Table 2 and 4), we investigate hypothesis H1 using unemployment shocks and combination shocks, (i.e., *Shock 3* as described in Section 4). These adverse income shocks are sufficient to be considered as exogenous to the borrowers' financial buffers and default process.[13] The main difference-in-difference estimator is α_{DID}, which measures the difference in behavior towards using financial buffers between low- and high-income borrowers (the first difference) after and before the adverse income shocks (the second difference).

[12] Difference-in-difference OLS regressions are common in the mortgage risk literature, see e.g., O'Malley, T. (2018).

[13] As argued in Gerardi et al. (2018), borrowers may be voluntarily unemployed, which does not rule out the chance that income shocks are endogenous to their defaults. We address this potential endogeneity by performing robustness checks using only involuntary unemployment shocks. We further consider divorce shocks may not be entirely exogenous to the borrowers. Therefore, we perform an additional robustness analysis using only involuntary unemployment and disability shocks as purely exogenous income shocks. All these robustness analyses provide consistent results with our main findings. Details are available upon request.

The parameters include the intercept α_0, standalone sensitivities α_S and α_L. The features include the standalone and difference-in-difference effects, the linear combination of control features weighted by the estimated weights ($Controls$) and the residuals ($\varepsilon_{i,t}$) which are assumed to be independent and identically normally distributed with mean zero.

Differences in means between the two income groups, and differences in means between after and before the income shock are computed from these sensitivities.

Table 5 to Table 7 shows the difference-in-difference model estimators for the three financial buffer (dependent) variables, including discretionary expenses, borrowing capacity and asset liquidity.

Insert Table 5 here

Insert Table 6 here

Insert Table 7 here

The left panel includes the unemployment shock, and the right panel shows the combination of all three shocks (i.e., variable *Shock 3*). We report the unemployment shocks separately as this is the most common and most interesting shock from an economic perspective. We find a consistent result that the difference-in-difference estimates are positive and significant in all cases. In addition, the impact of the unemployment shock is relatively comparable to that of the combined *shock3*, indicating its dominance among the three shocks. Note that the difference-in-difference estimates may be harder to interpret than standard OLS estimates. Figure 3 shows the underlying differences and the combined difference-in-difference effect. The relation of the numbers in Table 5 to Table 7 is consistent with this interpretation.

Insert Figure 3 here

Effect (5) shows the difference in a financial buffer between the low- and the high-income groups before the adverse income shock while Effect (6) shows the difference after the adverse income shock. The results for all three financial buffers consistently show that the low-income group has significantly fewer financial buffers than the high-income group before the adverse income shocks. For example, for the unemployment shock with controls, the gaps between the two groups before adverse income shocks are 0.06, 0.051, 0.55, for discretionary expenses, borrowing capacity and liquid assets, respectively. However, this gap reduces and becomes statistically insignificant after the borrowers experience adverse income shocks in the cases of discretionary expenses and borrowing capacity. This is due to a significant decrease in the financial buffers of the high-income group in coping with adverse income shocks, while there is no significant change in the financial buffers of the low-income group. For liquid assets, the gap remains statistically significant, which may indicate that borrowers try to cut their discretionary expenses and exploit their remaining borrowing capacity before selling their liquid assets following adverse income shocks.

Effect (7) shows the difference between a financial buffer after and before adverse income shock for the low-income group. Effect (8) shows the difference for the high-income group. The level of a financial buffer decreases with the adverse income shock, but to a greater degree for the high-income group. For example, for the discretionary expense ratio and the unemployment shock with controls, the effect reduces the financial buffer by 0.002 for the low-income group and by 0.06 for the high-income group.

The difference-in-difference estimator is equal to the difference between Effect (6) and Effect (5) or between Effect (7) and Effect (8). The effect is positive for all three financial buffers. For example, the effect is 0.078 for the discretionary expense ratio and the unemployment shock with controls.

The results of a “single” difference (i.e., individual effect (5), (6), (7), and (8)) and the difference-in-difference analyses confirms our Hypothesis 1, that the high-income borrower group can draw on significantly more financial buffers than the low-income borrower group. Therefore, the high-income borrower group is considerably better able to cope with the adverse income shocks, where cutting the discretionary expenses and utilizing the remaining borrowing capacity tend to be prioritized before selling liquid assets.

The finding that high-income borrowers can reduce financial buffers to a greater relative magnitude in anticipation of defaults is a key explanation as to why low-income borrowers are more exposed to income shocks.

As a robustness check, we have standardized the discretionary expenses and the liquid asset by the house value instead of the annual mortgage payments. We have also measured the proxy of borrowing capacity of the borrowers by dividing their home equity by the annual mortgage payment, which accommodates the case that the borrowers can exploit their home equity to cope with adverse income shocks. Our main findings remain consistent.[14]

5.2. Borrower defaults

Borrowers may default after financial buffers (e.g., consumption adjustment, borrowing capacity and liquid assets) are exhausted. Low-income borrowers have lower relative levels of financial buffers and hence less room to (i) reduce discretionary consumption, (ii) access funding, and (iii) sell liquid assets.

An important variable CLTV may be endogenous, as it is defined as the ratio of *self-reported* mortgage balance to the *self-reported* value of the house. Hence, inputs for the CLTV

[14] Results are available on request.

calculation depend on: (i) ability of household to make mortgage payments, and (ii) household's own estimate of house value. These decisions are driven by the household's unobservable characteristics, which may also drive default and therefore imply endogeneity. For example, if a household estimates their house value is well below the loan balance, they may decide to make a strategic default. We therefore apply difference-in-difference two-stage-least-squares (2SLS) regressions for the impact of income shocks on the likelihood of mortgage default to control for this endogeneity. In the first stage, we regress CLTV on house price appreciation (HPA) since origination, which is used as instrumental variable for CLTV:

$$CLTV_{i,t} = b_0 + b_{IV} HPA_t + b_X X_{i,t} + v_{i,t} \tag{10}$$

with residual $v_{i,t}$. $X_{i,t}$ collects all other explanatory variables shown in the second stage model, except for $CLTV$. HPA is calculated at the state level and, therefore, a macro-economic variable that is exogenous to the loan level CLTV.

In the second stage, we regress default $Default_{i,t}$ on the predicted CLTV, $\widehat{CLTV}_{i,t}$:

$$Default_{i,t} = \beta_{2SLS} + \beta_{2SLS,S} shock_{i,t} + \beta_{2SLS,L} inc\_b50_{i,t-1} + \beta_{2SLS,DID} inc\_b50_{i,t-1} \times shock_{i,t} + \beta_{2SLS,CLTV} \widehat{CLTV}_{i,t} + \beta_{2SLS,C} Controls + \epsilon_{i,t} \tag{11}$$

For the linear model, all effects can be interpreted in terms of the expected default rate or default probability. We compute differences in means between the two income groups, differences in means between after and before the income shock from the estimated sensitivities.

Table 8 shows the difference-in-difference results for a two-stage least square regression.

Insert Table 8 here

We compare two panels: 60- and 90-day delinquency and for a given panel in the first column the unemployment shock, and the second column the combination of all three income shocks with consistent results. As in the analysis for financial buffers, we find that the difference-in-difference estimates are positive and the unemployment shock dominates.[15]

[15] Note that our data reflects economic relief policies such as the Housing and Economic Recovery Act 2008 to reduce foreclosures by offering loans to households in hardship. The effects measured in this paper are after such effects, and without such interventions are likely to be substantially larger.

Figure **4** shows the separate constituent difference effects and the combined difference-in-difference effects.

***Insert

Figure **4** here***

Effect (5) shows the difference in expected default rate between the low- and the high-income group before the adverse income shock while Effect (6) shows the difference after the adverse income shock. After controlling for features, we observe that the difference between default sensitivity of the low-and high-income borrower groups is not significant before the adverse income shocks (between 0.1% and 1%). However, after adverse income shocks, the default rate of the low-income group is significantly greater than the high-income group and is between 2.9% and 5.1%.

Effect (7) shows the difference between the default rate after and before the adverse income shock for the low-income group. Effect (8) shows the difference for the high-income group. The level of default rate increases with the adverse income shock but to a greater degree for the low-income group. It increases by 4% to 6.3% for the low-income group, and by 0.4% to 1.7% for the high-income group, after controlling for other factors.

The difference-in-difference estimator is equal to the difference between Effect (6) and Effect (5) or between Effect (7) and Effect (8). The effect is 3.8% to 5.9% and significant at the 1% level. Overall, the results show that after experiencing adverse income shocks, an increase in the default probability for the low-income group is significantly greater than for the high-income group. The empirical evidence strongly supports our second hypothesis and the implication of our economic motivation that, $\gamma > 1$. In other words, the ratio between the default rate of the low- and the high-income group after the adverse income shock is greater than one.

Note that pre-payments prior to maturity are not a major concern for our data set as loan observations from household survey data are not right-censored. Prepayments generally occur in mortgage finance due to loan refinance and the households remain in the survey sample after a lender change.

The above model is a linear model with a binary dependent variable. Generally, non-linear models (probit and logistic regressions) are more appropriate as they ensure that estimated default probabilities are bounded between zero and one. Angrist and Pischke (2008) propose two-stage linear regressions to control for endogeneity. Non-linear regressions may result in an estimation bias.[16] Hence, we consider probit regressions only in robustness checks and confirm consistency for logistic regressions.[17] The model equation is:

$$P(Default_{i,t} = 1) = \Phi\big(\beta_0 + \beta_S shock_{i,t} + \beta_L inc\_b50_{i,t-1} + \beta_{DID} inc\_b50_{i,t-1} \times shock_{i,t} + \beta_{CLTV} CLTV_{i,t} + \beta_C Controls\big) \quad (12)$$

with the standard normal distribution function $\Phi(.)$, intercept β_0, the sensitivity of the standalone shock ($shock_{i,t}$) β_S, and low-income borrowers ($inc_{b50_{i,t-1}}$) β_L, the difference-in-difference sensitivity β_{DID}, the linear combination of control features weighted by the estimated parameters. Note that there is no error term, as probit models are formulated in expectations. They model the probabilities of the dependent variable, i.e., the probability of default $P(Default_{i,t} = 1)$. The area under the curve (AUC) is the area under the receiver operating characteristic curve (the relation between sensitivity and one minus specificity), which is a common measure of model fit with a value of 0.5 for a random model and a value of one for a perfect model.

Table 9 shows the difference-in-difference results for the probit regression.

[16] See also Wooldridge (2010) and Greene (2008).

[17] We have confirmed that logistic regressions, which are also non-linear regressions, result in identical signs and significances. Parameter estimates differ in absolute terms due to a different link function between the default probability and the linear predictor.

Insert Table 9 here

We compare two panels: 60- and 90-day delinquency. For each panel, we include the unemployment shock in the first column and the combination of all three shocks in the second column. We find highly consistent results for all settings. As in the main analysis for financial buffers, we find that the difference-in-difference estimates are positive and a dominant effect of the unemployment shock.

In a robustness check, we confirm that the difference and difference-in-difference estimators are consistent for involuntary unemployment shocks, i.e., situations where the head of a household loses a job due to exogenous shocks.[18]

5.3. The role of home equity

Under the double trigger theory and recourse lending, default occurs if borrowers have negative equity and are constrained by liquidity. However, in the empirical data after propensity score matching, we observe that about 57% of the 60-day delinquencies (84 out of 148) and about 53% of the 90-day delinquencies (46 out of 87) have positive equity (i.e., CLTV < 1).[19] Hence, we are interested to see whether borrowers may default despite positive equity, which is part of the borrowing capacity. We are effectively testing whether income has an impact after exhausting this buffer and whether our economic motivation as an extension of the DTM upholds. We sub-sample for positive equity observations and apply a two stage least squares model presented in Section 5.2, in which we use house price appreciation (HPA) since origination as an instrumental variable for CLTV.

[18] Results are available on request.
[19] For the before propensity score matching matched sample, we observe that 60% of the 60-day delinquencies (147 out of 248) and about 57% of the 90-day delinquencies (83 out of 146) have positive equity.

Table 10 shows the difference-in-difference results for a two-stage least square regression for mortgage defaults and positive equity (coded as one).

Insert Table 10 here

We compare two panels: 60- and 90-day delinquency and for a given panel in the first column the unemployment shock, and in the second column the combination of all three shocks with consistent results. As in the analyses of financial buffers and borrower defaults that account for both negative and positive equity, we find that the difference-in-difference estimates are positive and a dominant effect of the unemployment shock. It is worth noting that we find no significant impact of CLTV on the mortgage defaults with positive equity. We interpret this as an indication that borrowers may be unable to liquidate home equity due to owning illiquid assets or impediments in financial markets.

5.4. Non-linear income shock sensitivity

Next, we are interested in the sensitivities with the magnitude of the income shock, i.e., whether the relation is convex (exacerbating) or concave (saturatingp). The number of income shock observations drops if we filter for magnitudes, i.e., borrowers who experience an income shock that is greater than a threshold value. We define the income shocks considering the information of income drop magnitude as follows,

$$Income\ drop_{m,i,t} = \begin{cases} 0, \text{ if borrower's income drops less than } m \text{ \%} \\ 1, \text{ if borrower's income drops more than } m \text{ \%} \end{cases} \tag{13}$$

We perform analyses for thresholds of *m* with an increment of 5% (i.e., *m*=5,10,15, 20, 25, 30, 35). These income drops may be endogenous to the borrowers' default, for instance, there could be an unobservable borrower personality type driving both income drop and default. To address this potential endogeneity, we employ the income shock, *Shock 3*, used in previous sections as an instrument variable for the $Income\ drop_{m,i,t}$.[20] We also note from discussions in the previous section that HPA from origination at the state level is used as an instrument variable for the loan-level *CLTV*.

We apply a two-stage regression model, in which the main model (second stage) is specified as:

$$Default_{i,t} = \theta_{0,m} + \theta_{b50,m}\widehat{inc^{b50}_{m,i,t}} + \theta_{a50,m}\widehat{inc^{a50}_{m,i,t}} + \theta_{m,cltv}\widehat{CLTV_{i,t}} + \theta_{C,m}Controls + \eta_{i,t} \tag{14}$$

where $\eta_{i,t}$ is the residual, $inc^{b50}_{m,i,t} = inc\_b50_{i,t-1} \times Income\ drop_{m,i,t}$, and similarly, $inc^{a50}_{m,i,t} = (1 - inc\_b50_{i,t-1}) \times Income\ drop_{m,i,t}$. This modelling framework does not assist in straightforwardly explaining the difference-in-difference effect that we focus on in Section 5.2. However, its advantage is to conveniently link the empirical estimates with our economic motivation via the estimation of γ, which we highlight below.

As we discussed earlier, $Income\ drop_{m,i,t}$ and $CLTV_{i,t}$ can be endogenous to the $Default_{i,t}$. Therefore, we use the instrument variables to estimate $\widehat{inc^{b50}_{m,i,t}}$, $\widehat{inc^{a50}_{m,i,t}}$, and $\widehat{CLTV_{i,t}}$ in the first stage of the regressions as follows:

$$\begin{cases} inc^{b50}_{m,i,t} = \phi_1 + \phi_{1,z} z^{b50}_{m,i,t} + \phi_{1,C}controls + u_{1,it} \\ inc^{a50}_{m,i,t} = \phi_2 + \phi_{2,z} z^{a50}_{m,i,t} + \phi_{2,C}controls + u_{2,it} \\ CLTV_{i,t} = \phi_3 + \phi_{3,HPA} HPA_t + \phi_{3,C}controls + u_{3,it} \end{cases} \tag{15}$$

with residuals u_{it}. $z^{b50}_{m,i,t}$ is the instrument variable for $inc^{b50}_{m,i,t}$, such that $z^{b50}_{m,i,t} = inc\_b50_{i,t-1} \times shock_{i,t}$. Similarly, $z^{a50}_{m,i,t} = (1 - inc\_b50_{i,t-1}) \times shock_{i,t}$ is the instrument variable for $inc^{a50}_{m,i,t}$.

[20] For robustness checks, we also perform analyses using different exogenous income shocks as instrument variables for $Income\ drop^{m}_{i,t}$, including involuntary unemployment and a combination of involuntary unemployment and disability. Results remain consistent.

$\widehat{inc}_{m,i,t}^{b50}$ and $\widehat{inc}_{m,i,t}^{a50}$ are now continuous variables and bounded by 0 and 1. We can interpret them as the likelihood that the borrower in each group faces the income reduction ($\widehat{inc}_{m,i,t}^{b50}$ for the low-income group and $\widehat{inc}_{m,i,t}^{a50}$ for the high-income group). Their average values represent the proportion of borrowers in each income group who face income reduction.

Table 11 shows the estimation results for a two-stage least square regression.

Insert Table 11 here

Table 11 shows the parameter estimates ($inc_{m,i,t}^{b50} = inc\_b50_{i,t-1} \times Income\ drop_{m,i,t}$ and $inc_{m,i,t}^{a50} = \left(1 - inc\_b50_{i,t-1}\right) \times Income\ drop_{m,i,t}$) for the various levels of income reduction.

The estimates increase with the severity of income reduction at an increasing rate (i.e., a convex shape, see Figure 5).

The results are expressed relative to a zero-income shock and in terms of the proportional distribution of borrowers that experience income shocks. In particular, income shock sin excess of 15% result in higher default rates. More specifically, for a 1% increase in the proportion that the borrowers face a 15% plus income reduction, there is an increase of 0.305 percentage points (0.126 percentage points) in the default rate for the low- (high-) income group. When a 35% plus income reduction is considered, the default rate increases to 0.757 percentage points for the low-income group and 0.225 percentage points for the high-income group relative to a zero income shock.

The increase of default risk due to an increase in income reduction is convex and greater income reductions have greater impacts on default risk. We note that this effect is more pronounced for the low-income group.

To match these with our economic motivation presented in Section 3, we have $\delta_t^{(l)} = \frac{\sum_{i=1}^{L} \widehat{inc_{m,i,t}^{b50}}}{L}$, and $\delta_t^{(h)} = \frac{\sum_{i=1}^{H} \widehat{inc_{m,i,t}^{a50}}}{H}$, where L and H are the total number of borrowers in low- and high-income groups respectively. As a result, the $\theta_{b50,m} = \rho^{(l)}$ and $\theta_{a50,m} = \rho^{(h)}$, which results in $\gamma = \frac{\rho^{(l)}}{\rho^{(h)}} = \frac{\theta_{b50,m}}{\theta_{a50,m}}$. In other words, the implied gammas (γ) are the ratios between the default sensitivities of the low-income group and the high-income group (second and third row estimates of Table 11).

We present the trend of the implied gammas against the levels of income reduction in Figure 5. The implied gammas are all greater than one, which is consistent with the implication of our economic motivation and the empirical findings in the previous section. This marginal relationship between the default sensitivity of the two groups increases as the level of income reduction increases, until the income reduction reaches 30%. For income reductions which are greater than 30%, the marginal relationship starts to stabilize. This result provides evidence that the marginal changes in the default sensitivity of the two income groups increase in level due to income reduction and converge at 30% of income reduction.

Insert Figure 5 here

Gamma stabilizes somewhat for higher income shocks.

5.5. Surprise income shocks

We have tested involuntary unemployment with consistent results to unemployment. Further, borrowers may have different profiles and to some degree even anticipate involuntary unemployment as they work in more unemployment-prone industries (e.g., seasonal services). To test the income shock relative to the expected level of income shock, we compute a two-stage model that models the probability of an income shock in the first stage, and includes the gap between income shock and the probability of income shock as a test variable in the second stage.

In the first stage, we estimate a probit regression for the shock variable:

$$Pr\left(Shock_{i,t} = 1\right) = \Phi\left(\alpha_R + \beta_R X_{R,i,t-1}\right) \tag{16}$$

where $X_{R,i,t-1}$ is the income information and all other demographic information of the borrower *i* at time *t-1* that we discuss in the data description. We also predict *CLTV*, $\widehat{CLTV_{i,t}}$, following the model from Section 5.2.

In the second stage, we compute the surprise income shock $Shock_{i,t}^S$ as the difference of the shock dummy and the estimated shock probability $\Phi\left(\hat{\alpha}_R + \hat{\beta}_R X_{R,i,t}\right)$:

$$Shock_{i,t}^S = Shock_{i,t} - \Phi\left(\hat{\alpha}_R + \hat{\beta}_R X_{R,i,t}\right) \tag{17}$$

We then replicate our analysis in Section 5.2 using the surprise income shock in the regressions:

$$Default_{i,t} = \beta_R + \beta_{R,S} Shock_{i,t}^S + \beta_{R,L} inc\_b50_{i,t-1} + \beta_{R,I} inc\_b50_{i,t-1} \times Shock_{i,t}^S + \beta_{R,CLTV} \widehat{CLTV_{i,t}} + \beta_{R,C} Controls + v_{i,t} \tag{18}$$

For the linear model, all effects can be interpreted in terms of the expected default rate or default probability. We note that this framework is no longer a difference-in-difference analysis as the income shock is now a continuous variable. However, as seen from Table 12, the results are highly consistent with our main results using difference-in-difference analyses as well as the findings obtained in the analysis of income reduction levels. More specifically, the default sensitivity of the low-income group is significantly stronger than the high-income group given an increase in the surprise adverse income shock. The other models (H1, H3 and H4) tested for surprise shocks yield consistent results with our main analysis.

Insert Table 12 here

6. Conclusion and recommendations

This paper analyzes the impact of mortgagee income on the borrower behavior to default. The paper finds that financial buffers such as discretionary income, funding ability and liquid assets explain the greater sensitivity to income shocks for low-income borrowers. The effects persist for positive equity loans, and we hypothesize that borrowers may be unable to liquidate home equity due to impediments in financial markets. The income shocks are convex with regard to shock magnitude and default risk. This suggests a greater impact for low-income borrowers in instances where a larger fraction of household income is lost (e.g., unemployment in single income households) than in instances where a smaller fraction of income is lost (e.g., unemployment in multiple income households or demotions).

Mortgage lending is often based on a one size fits all lending process that may suggest that loans made to different income borrowers should have the same risk features. This paper finds that this does not hold and suggests socially responsible bank lending and prudential regulation should address this gap. Solutions may include a reduction in lending to lower incomes, changes to mortgage design or support of building financial buffers. For example, banks may include liquidity facilities in the design of financial contracts for low-income groups.

Alternatively, regulators may consider temporary relief programs by income levels with low incomes receiving greater support if financial hardship is temporary. The findings may support stratifying temporary relief programs by income levels with low incomes receiving a greater support if financial hardship is temporary. In response to COVID-19, many governments have established foreclosure moratoriums and rights to request forbearance (i.e., delay interest and principal payments) such as the Coronavirus Aid, Relief, and Economic Security (CARES) Act. Providing better financial solutions to low-income households may increase home ownership for this cohort and provide for better retirement outcomes. These solutions may include government-supported insurance against income shocks or liquidity facilities. Further, societies may offset the negative consequences for low-income borrowers by providing insurance against income shocks, liquidity facilities or outright affordable housing.

Lastly, the findings may assist banks in anticipating future losses and mitigating their impacts through strengthening their balance sheets in terms of reducing risk exposures and increasing capital buffers through earnings retentions.

References

Adelino, M., Gerardi, K., & Willen, P. S. (2013). Why don't lenders renegotiate more home mortgages? Redefaults, self-cures and securitization. *Journal of Monetary Economics*, 60(7), 835-853.

Agarwal, S., Ambrose, B. W., Chomsisengphet, S., & Liu, C. (2011). The role of soft information in a dynamic contract setting: Evidence from the home equity credit market. *Journal of Money, Credit and Banking*, 43(4), 633-655.

Agarwal, S., Ambrose, B. W., & Liu, C. (2006). Credit lines and credit utilization. *Journal of Money, Credit, and Banking*, 38(1), 1-22.

Ambrose, B. W., LaCour-Little, M., & Huszar, Z. R. (2005). A note on hybrid mortgages. *Real Estate Economics*, 33(4), 765-782.

Angrist, J. D., & Pischke, J. S. (2008). Mostly harmless econometrics: An empiricist's companion. *Princeton University Press*.

Bhanot, S. (2017). Cheap promises: Evidence from loan repayment pledges in an online experiment. *Journal of Economic Behavior and Organization*, 140, 246-266.

Campbell, J. Y., & Cocco, J. F. (2015). A model of mortgage default. *The Journal of Finance*, 70(4), 1495-1554.

Chen, H., Michaux, M., & Roussanov, N. (2020). Houses as ATMs: mortgage refinancing and macroeconomic uncertainty. *The Journal of Finance*, *75*(1), 323-375.

Chiang, Y. M., & Sa-Aadu, J. (2014). Optimal mortgage contract choice decision in the presence of pay option adjustable rate mortgage and the balloon mortgage. *The Journal of Real Estate Finance and Economics*, 48(4), 709-753.

Coibion, O., Gorodnichenko, Y., & Weber, M. (2020). Labor markets during the COVID-19 crisis: A preliminary view (No. w27017). *National Bureau of Economic Research*.

Corradin, S. (2014). Household leverage. *Journal of Money, Credit and Banking*, 46(4), 567-613.

Cunningham, C., Gerardi, K., and Shen, L. (2020). The double trigger for mortgage default: Evidence from the fracking boom. *Management Science*, 10.1287/mnsc.2020.3647.

Elul, R., Souleles, N. S., Chomsisengphet, S., Glennon, D., & Hunt, R. (2010). What" triggers" mortgage default? *American Economic Review*, 100(2), 490-94.

Epperson, J.F., Kau, J.B., Keenan, D.C. and Muller III, W.J., 1985. Pricing default risk in mortgages. *Real Estate Economics*, 13(3), pp.261-272.

Fuster, A., & Willen, P. S. (2017). Payment size, negative equity, and mortgage default. *American Economic Journal: Economic Policy*, 9(4), 167-91.

Gathergood, J. (2009). Income shocks, mortgage repayment risk and financial distress among UK households (No. 09/03), *Centre for Finance and Credit Markets.*

Gerardi, K., Herkenhoff, K. F., Ohanian, L. E., & Willen, P. S. (2018). Can't pay or won't pay? unemployment, negative equity, and strategic default. *The Review of Financial Studies*, *31*(3), 1098-1131.

Goldberg, L., & Capone, Jr, C. A. (2002). A Dynamic Double-Trigger Model of Multifamily Mortgage Default. *Real Estate Economics*, 30(1), 85-113.

Greene, W. H. (2008). The econometric approach to efficiency analysis. The measurement of productive efficiency and productivity growth, 1(1), in: The Measurement of Productive Efficiency and Productivity Change by Fried, H.O., Knox Lovell, C.A., and Schmidt, S.S., 92-250.

Haughwout, A., Okah, E., & Tracy, J. (2016). Second chances: Subprime mortgage modification and redefault. *Journal of Money, Credit and Banking*, 48(4), 771-793.

Johnson, K. W., & Li, G. (2010). The debt-payment-to-income ratio as an indicator of borrowing constraints: Evidence from two household surveys. *Journal of Money, Credit and Banking*, *42*(7), 1373-1390.

Johnson, K. W., & Li, G. (2014). Are Adjustable-Rate Mortgage Borrowers Borrowing Constrained?. *Real Estate Economics*, 42(2), 457-471.

Kau, J. B., Keenan, D. C., & Kim, T. (1994). Default probabilities for mortgages. *Journal of Urban Economics*, 35(3), 278-296.

Laufer, S. (2018). Equity extraction and mortgage default. *Review of Economic Dynamics*, 28, 1-33.

Mayer, C., Pence, K., & Sherlund, S. M. (2009). The rise in mortgage defaults. *The Journal of Economic Perspectives*, 23(1), 27-50.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, 29(2), 449-470.

O'Malley, T. (2018). The Impact of Repossession Risk on Mortgage Default. forthcoming *Journal of Finance*

Qi, M., & Yang, X. (2009). Loss given default of high loan-to-value residential mortgages. *Journal of Banking & Financ*e, 33(5), 788-799.

Qi, M., Scheule, H., & Zheng, Y. (2021). The Impact of Positive Payment Shocks, Borrower Liquidity and Refinance Constraints on Default Risk of Home Equity Lines of Credit at End of Draw, *Journal of Real Estate Finance and Economics*, 62(3), 423-454.

Schelkle, T. (2018). Mortgage default during the US mortgage crisis. *Journal of Money, Credit and Banking*, 50(6), 1101-1137.

Smith, J. D. (2016). US political corruption and firm financial policies. *Journal of Financial Economics*, 121, 350-367.

Tracy, J. S., & Wright, J. (2012). Payment changes and default risk: The impact of refinancing on expected credit losses. *Federal Reserve Bank of New York*, Staff Report No. 562.

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data. *MIT Press*.

Zorn, P. M., & Lea, M. J. (1989). Mortgage borrower repayment behavior: a microeconomic analysis with Canadian adjustable rate mortgage data. *AREUEA Journal*, 17(1), 118-136.

Figures and tables

Figure 1: Incremental implied default rate of low- and high-income borrower groups by levels of income reductions

Note: This figure shows the increase in default rate for income reductions in excess of a threshold value (reported on the x-axis). The estimated coefficients are extracted from Table 11.

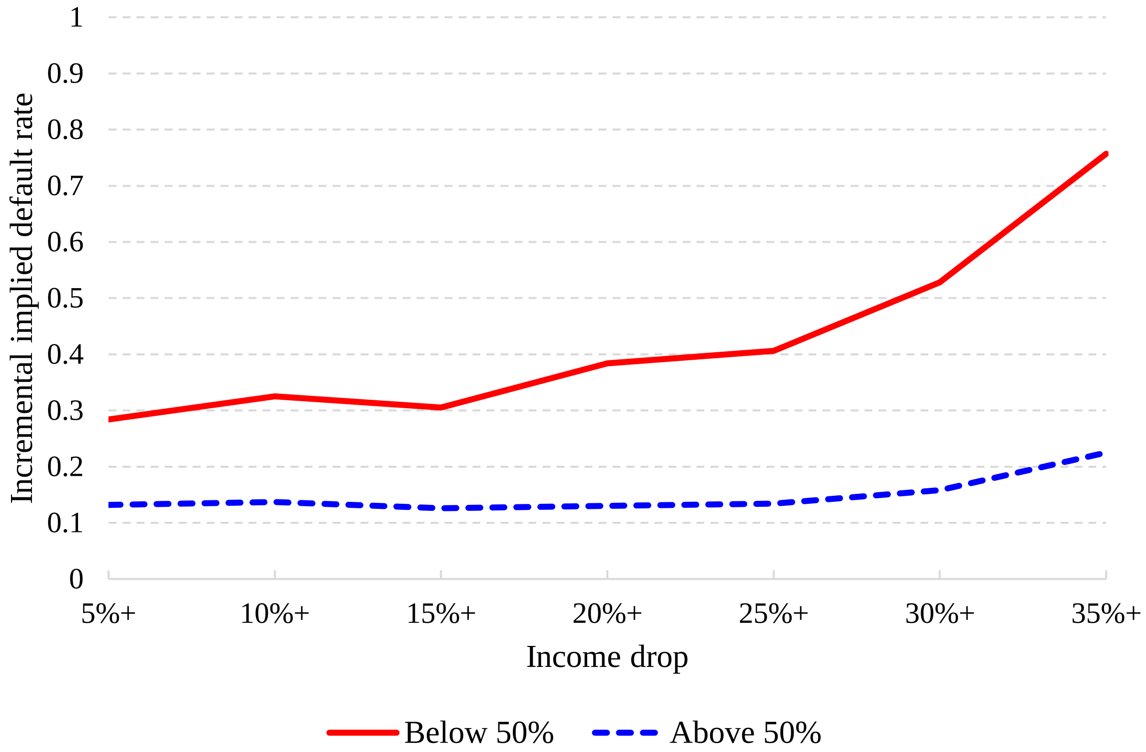

Figure 2: Hypotheses development

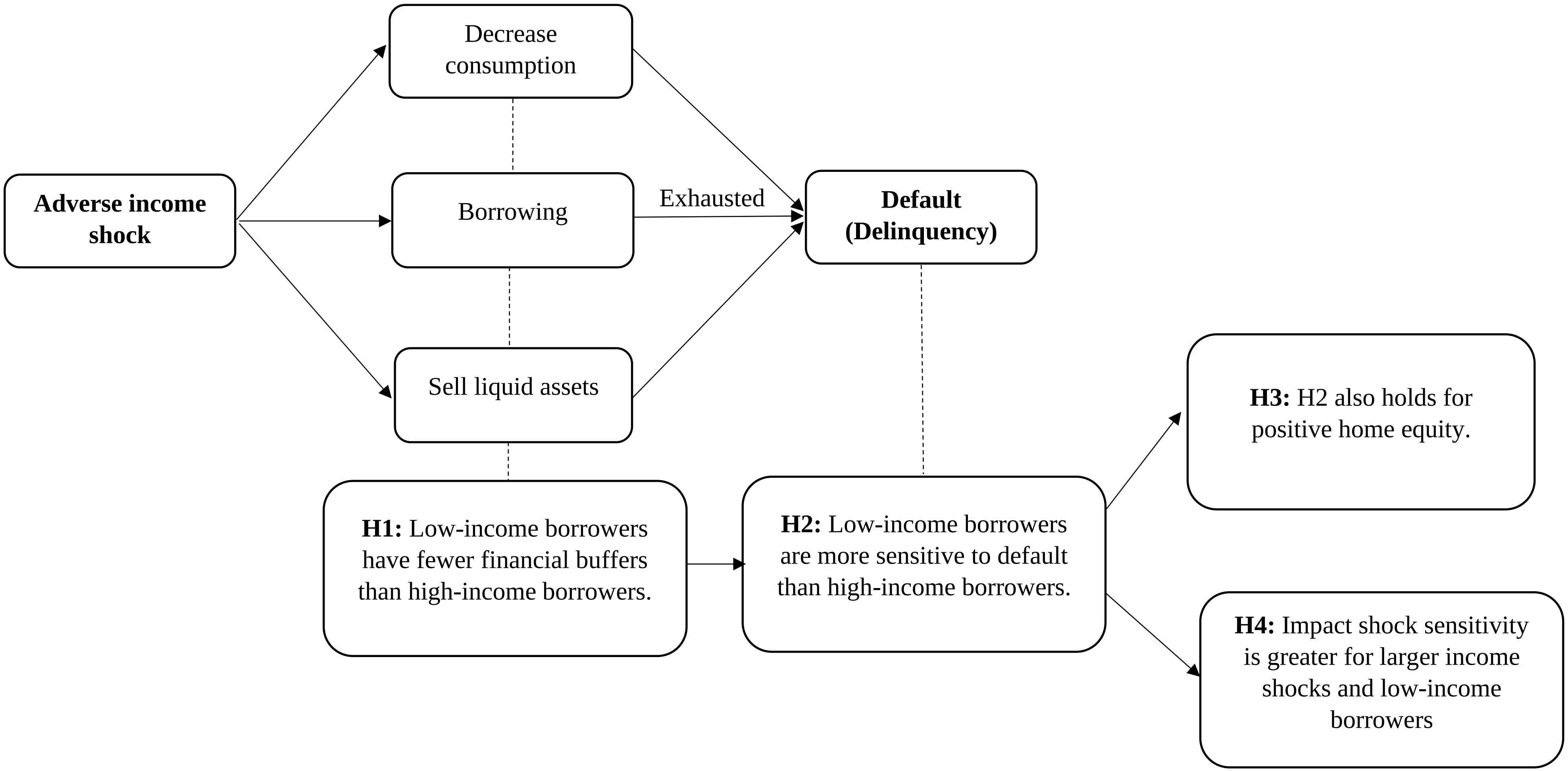

Figure 3: Development of difference estimators and difference-in-difference estimators for financial buffers

This figure illustrates the estimators of the DID analyses for the three financial buffers. It is compatible with Table 5 to Table 7. Effect (5) shows the difference in a financial buffer between the low- and the high-income groups before the adverse income shock, while Effect (6) shows the difference after the adverse income shock. The gap between the two groups closes with the adverse income shock. Effect (7) shows the difference between a financial buffer after and before the adverse income shock for the low-income group. Effect (8) shows the difference for the high-income group. The level of a financial buffer decreases with the adverse income shock but to a greater degree for the high-income group. The DID estimator is equal to the difference between Effect (5) and Effect (6) or between Effect (7) and Effect (8).

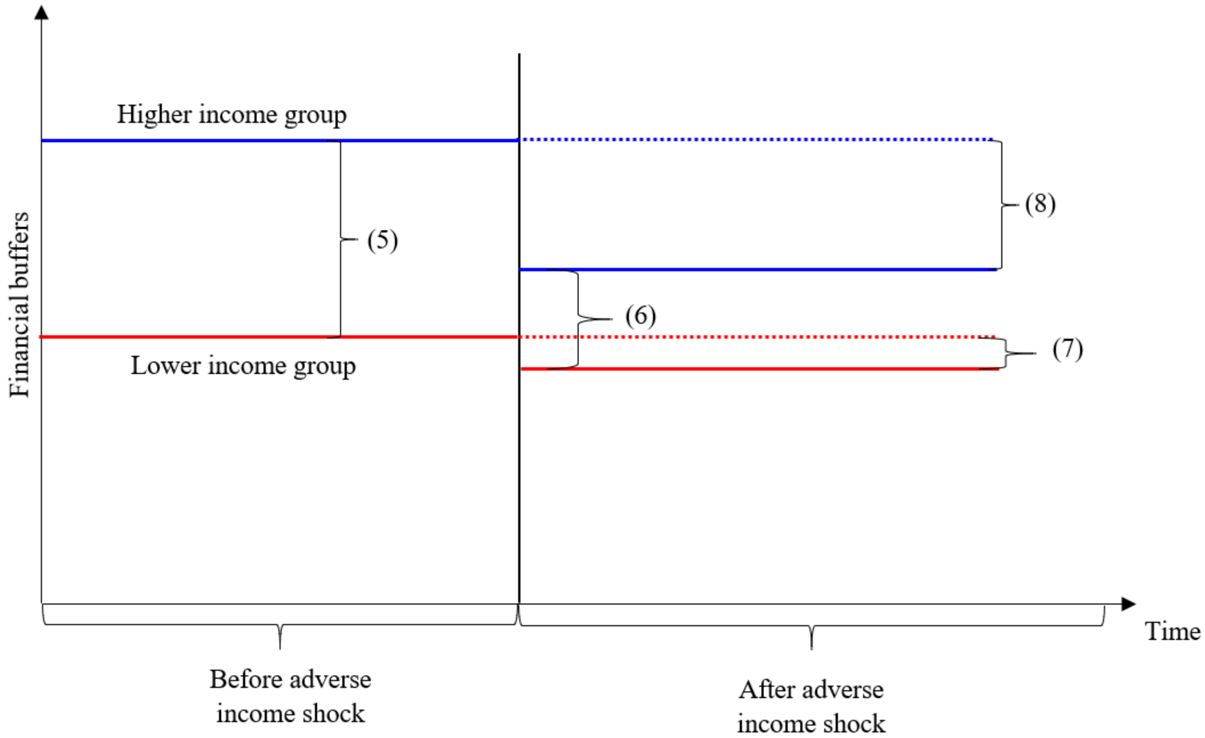

Figure 4: Development of difference estimators and difference-in-difference estimators for default

This figure illustrates the estimators of the DID analyses for default. It is compatible with Table 8-10. Effect (6) shows the difference in the default rate between the low- and the high-income group before the adverse income shock while Effect (5) shows the difference after the adverse income shock. The gap between the two groups closes with the adverse income shock. Effect (7) shows the difference between the default rate after and before the adverse income shock for the low-income group. Effect (8) shows the difference for the high-income group. The level of default rate increases with the adverse income shock but to a greater degree for the low-income group. The DID estimator is equal to the difference between Effect (5) and Effect (6) or between Effect (7) and Effect (8).

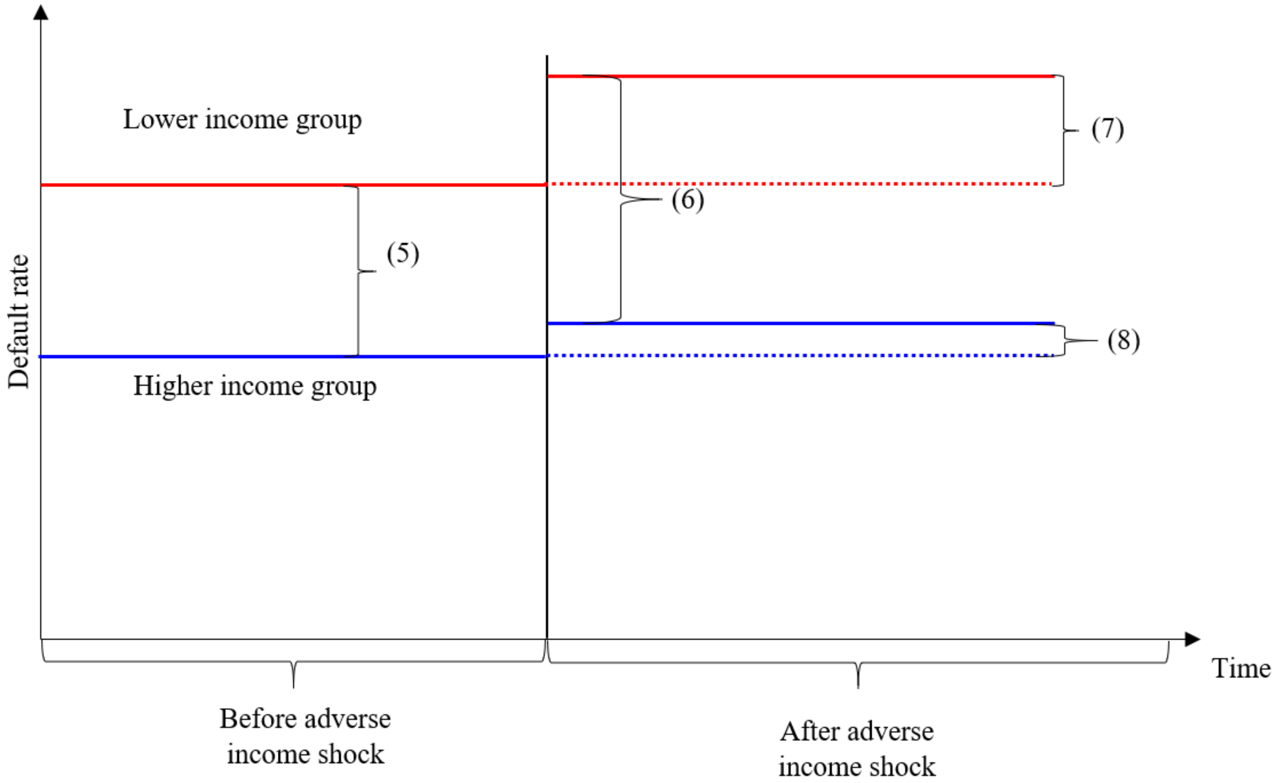

Figure 5: Implied gammas by level of income reduction

Note: This figure depicts the implied gammas (γ), which are calculated from the estimated coefficients of Table 11. γ reveals the marginal relationship between the default sensitivity of the low- and high-income groups for income reductions in excess of a threshold value (reported on the x-axis). The implied gamma is the ratio of default rate increase of the low- to the high-income group. γ is detailed in Section 3.

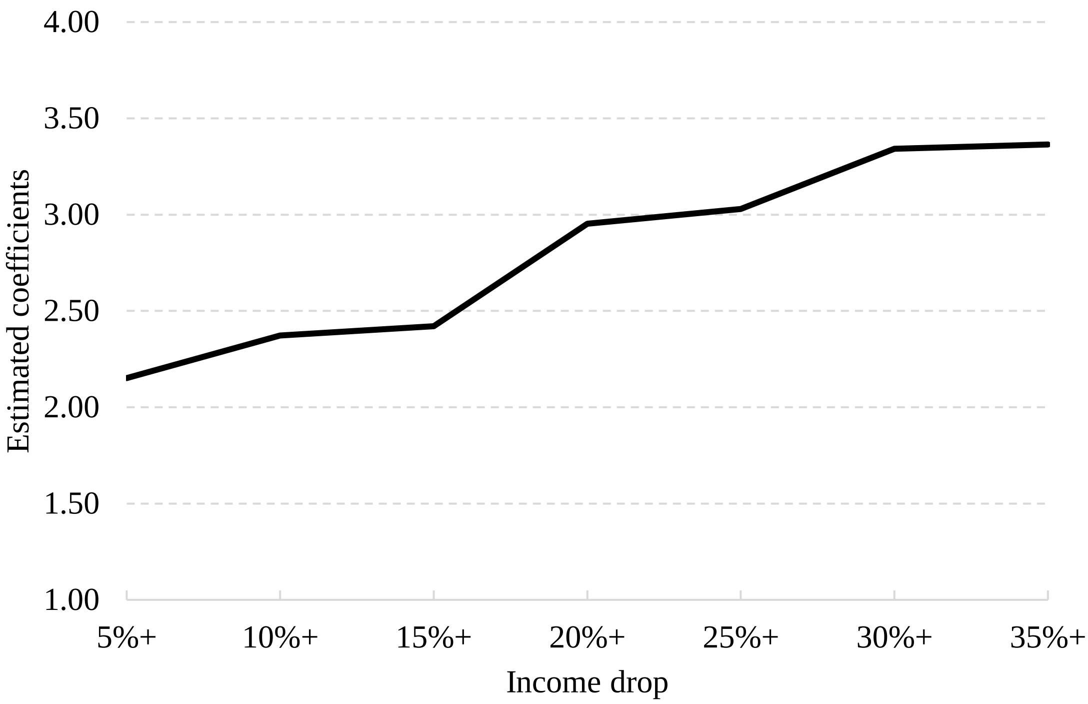

Table 1: Summary statistics of the matching variables by income groups

Note: This table shows the summary statistics of the borrower demographic variables by income groups after Propensity Score Matching (PSM). We employ these variables to perform a one-to-one PSM between low- and high-income groups with a threshold of 0.05. Out of 7,404 observations of the filtered sample, we obtain 2,445 observations for each income group. We present the means of these borrower demographic variables and their difference in mean by income groups. Standard errors are reported in parentheses. All mean values are statistically significant hence, we only report the statistical significance of the difference in means. ***, **, * indicate significance at the 1%, 5% and 10% level.

Borrower demographic	Low-income group	High-income group	Difference in mean
Household detail			
Age	45.43	45.796	-0.3654
	(0.222)	(0.201)	(0.3)
Number of children	0.833	0.85	-0.0167
	(0.023)	(0.022)	(0.032)
Occupation			
NAICS 1x code	0.02	0.019	0.0008
	(0.003)	(0.003)	(0.004)
NAICS 2x code	0.132	0.127	0.0053
	(0.007)	(0.007)	(0.01)
NAICS 3x code	0.158	0.178	-0.0202*
	(0.007)	(0.008)	(0.011)
NAICS 4x code	0.171	0.174	-0.0028
	(0.008)	(0.008)	(0.011)
NAICS 5x code	0.189	0.182	0.0074
	(0.008)	(0.008)	(0.011)
NAICS 6x code	0.145	0.143	0.0017
	(0.007)	(0.007)	(0.01)
NAICS 7x code	0.032	0.04	-0.0077
	(0.004)	(0.004)	(0.005)
NAICS 8x code	0.046	0.037	0.0092
	(0.004)	(0.004)	(0.006)
Education			
Less than high school certificate	0.055	0.044	0.0113*
	(0.005)	(0.004)	(0.006)
High school certificate	0.277	0.284	-0.0071
	(0.009)	(0.009)	(0.013)
College degree	0.288	0.3	-0.0123
	(0.009)	(0.009)	(0.013)
Higher than college degree	0.372	0.363	0.0088
	(0.010)	(0.010)	(0.014)
Race			
White	0.858	0.862	-0.0043
	(0.007)	(0.007)	(0.01)
Black	0.082	0.077	0.0057
	(0.006)	(0.005)	(0.008)

American Indian	0.006	0.003	0.0027
	(0.002)	(0.001)	(0.002)
Asian	0.015	0.02	-0.0051
	(0.002)	(0.003)	(0.004)
Other	0.035	0.031	0.004
	(0.004)	(0.004)	(0.005)
Number of observations	2,445	2,445	

Table 2: Summary statistics of main variables by default (60-day delinquency)

Note: the table reports means of the dependent and test variables and difference in mean by defaults and non-default. Standard errors are reported in parentheses. All mean values are statistically significant and hence, we only report the statistical significance of the difference in means. ***, **, * indicate significance at the 1%, 5% and 10% level.

Dependent variables	**Non-default household**	**Default household**	**Difference in mean**
Discretionary expense over mortgage	0.249	0.134	0.1144***
	(0.004)	(0.021)	(0.026)
Borrowing capacity	0.22	-0.13	0.3502***
	(0.006)	(0.032)	(0.034)
Liquid asset over mortgage	1.425	0.401	1.0238***
	(0.040)	(0.170)	(0.246)
Test variables			
Income ($)	103,473	72,296	31,176***
	(1,175)	(3,258)	(7,124)
Combined LTV	0.682	1.03	-0.3484***
	(0.005)	(0.030)	(0.029)
Divorce shock (Y=1, N=0)	0.019	0.044	-0.0252
	(0.002)	(0.017)	(0.012)
Disability shock (Y=1, N=0)	0.078	0.123	-0.0448
	(0.004)	(0.027)	(0.024)
Unemployment shock (Y=1, N=0)	0.104	0.318	-0.214***
	(0.004)	(0.038)	(0.028)
Combination of three shocks (shock3) (Y=1, N=0)	0.185	0.404	-0.2188***
	(0.006)	(0.040)	(0.035)
Observations	4,742	148	

Table 3: Summary statistics of control variables by default (60-day delinquency)

Note: the table reports means of the control variables and difference in mean by defaults and non-defaults. Standard errors are reported in parentheses. All mean values are statistically significant hence, we only report the statistical significance of the difference in means. ***, **, * indicate significance at the 1%, 5% and 10% level. Note that we do not report summary statistics of some borrower demographic and the origination years to conserve space.

Variable	Non-default household	Default household	Difference in mean
Borrower demographic			
Age	45.625	45.178	0.4472
	(0.153)	(0.775)	(0.931)
Number of children	0.84	0.917	-0.0771
	(0.016)	(0.091)	(0.098)
Less than high school certificate	0.047	0.127	-0.0792***
	(0.003)	(0.027)	(0.019)
College degree	0.372	0.227	0.1451***
	(0.007)	(0.035)	(0.043)
NAICS 1x code	0.02	0.006	0.0136**
	(0.002)	(0.006)	(0.012)
NAICS 6x code	0.145	0.081	0.0643***
	(0.005)	(0.023)	(0.031)
White	0.866	0.666	0.2001***
	(0.005)	(0.039)	(0.031)
Black	0.076	0.212	-0.1367***
	(0.004)	(0.034)	(0.024)
Mortgage characteristics			
Mortgage interest	4.784	5.392	-0.6072***
	(0.024)	(0.205)	(0.152)
Remaining 15 years and greater (Y=1, N=0)	0.656	0.854	-0.1983***
	(0.007)	(0.029)	(0.042)
Hold a second mortgage (Y=1, N=0)	0.173	0.217	-0.0436
	(0.005)	(0.034)	(0.034)
ARM (Y=1, N=0)	0.075	0.198	-0.1227***
	(0.004)	(0.033)	(0.024)
Refinance (Y=1, N=0)	0.497	0.514	-0.0168
	(0.007)	(0.041)	(0.044)
State laws			
Judicial (Y=1, N=0)	0.406	0.46	-0.0532
	(0.007)	(0.041)	(0.044)
Recourse (Y=1, N=0)	0.255	0.303	-0.0473
	(0.006)	(0.038)	(0.039)
Economic conditions			
House price appreciation since origination	-0.025	-0.062	0.0371***
	(0.001)	(0.007)	(0.007)
State housing bust in 2000s (Y=1, N=0)	0.163	0.256	-0.093**
	(0.005)	(0.036)	(0.033)
Observations	4,742	148	

Table 4: Summary statistics of main variables by income groups

Note: the table reports means of the dependent and test variables and difference in mean by income groups after one-to-one matching using PSM with matching variables as borrower demographic presented in Table 1. Standard errors are reported in parentheses. All mean values are statistically significant hence, we only report the statistical significance of the difference in means. ***, **, * indicate significance at the 1%, 5% and 10% level.

Dependent variables	**Low-income group**	**High-income group**	**Difference in mean**
Default	0.031	0.023	0.008*
	(0.003)	(0.003)	(0.005)
Discretionary expense over mortgage	0.208	0.283	-0.0748***
	(0.005)	(0.006)	(0.008)
Borrowing capacity	0.185	0.236	-0.0509***
	(0.008)	(0.007)	(0.011)
Liquid asset over mortgage	1.12	1.678	-0.5578***
	(0.048)	(0.062)	(0.079)
Test variables			
Income ($)	70,349	134,928	-64,578.961***
	(1,176)	(1,743)	(2,103)
Combined LTV	0.702	0.679	0.023**
	(0.007)	(0.006)	(0.009)
Divorce shock	0.008	0.032	-0.0238***
	(0.002)	(0.004)	(0.004)
Disability shock	0.095	0.064	0.0309***
	(0.006)	(0.005)	(0.008)
Unemployment shock	0.134	0.085	0.0491***
	(0.007)	(0.006)	(0.009)
Combination of three shocks (shock3)	0.214	0.167	0.0472***
	(0.008)	(0.008)	(0.011)
Observations	2,445	2,445	

Table 5: Effect of low income and income shocks on discretionary expenses, linear regression

Note: This table presents the difference-in-difference results, which explain the impact of low income and adverse income shocks on discretionary expenses using OLS regressions. Standard errors are reported in parentheses. ***, **, * indicate significance at the 1%, 5% and 10% level.

		Discretionary expenses over mortgage payments			
		Unemployment shock		Shock 3	
		No control	With controls	No control	With controls
Parameter estimates	Intercept (1)	0.285***	0.107	0.293***	0.108
		(0.008)	(0.114)	(0.009)	(0.114)
	inc_b50 (2)	-0.062***	-0.06***	-0.069***	-0.067***
		(0.012)	(0.012)	(0.012)	(0.012)
	shock (3)	-0.079***	-0.08***	-0.089***	-0.084***
		(0.029)	(0.029)	(0.021)	(0.021)
	inc_b50*shock (4)	0.07*	0.078**	0.08***	0.084***
		(0.037)	(0.037)	(0.029)	(0.029)
Difference in means between two income groups	Low - High (bef. shock)	-0.062***	-0.06***	-0.069***	-0.067***
	(5) = (2)	(0.012)	(0.012)	(0.012)	(0.012)
	Low - High (aft. shock)	0.008	0.018	0.01	0.016
	(6) = (4) + (2)	(0.035)	(0.035)	(0.027)	(0.026)
Difference in means between after and before income shock	Low: aft. shock - bef. shock	-0.009	-0.002	-0.01	0.000
	(7) = (3) + (4)	(0.024)	(0.024)	(0.02)	(0.02)
	High: aft. shock - bef. shock	-0.079***	-0.08***	-0.089***	-0.084***
	(8) = (3)	(0.029)	(0.029)	(0.021)	(0.021)
Difference in Difference	**Diff in Diff**	**0.07***	**0.078****	**0.08*****	**0.084*****
	(9) = (4) = (7) - (8)	(0.037)	(0.037)	(0.029)	(0.029)
Controls	Borrower demographics	No	Yes	No	Yes
	Fix year effect	No	Yes	No	Yes
Model fit	R-Square	0.013	0.046	0.017	0.049
	Number of Obs.	2,573	2,573	2,573	2,573

Table 6: Effect of low income and income shocks on borrowing capacity, linear regression

Note: This table presents the difference-in-difference results, which explain the impact of low income and adverse income shocks on borrowing capacity using OLS regressions. Standard errors are reported in parentheses. ***, **, * indicate significance at the 1%, 5% and 10% level.

		Borrowing Capacity			
		Unemployment shock		Shock 3	
		No control	With controls	No control	With controls
Parameter estimates	Intercept (1)	0.243***	-0.572***	0.249***	-0.575***
		(0.008)	(0.094)	(0.009)	(0.094)
	inc_b50 (2)	-0.056***	-0.044***	-0.066***	-0.052***
		(0.012)	(0.011)	(0.012)	(0.012)
	shock (3)	-0.081***	-0.083***	-0.078***	-0.09***
		(0.028)	(0.026)	(0.021)	(0.02)
	inc_b50*shock (4)	0.071*	0.058*	0.089***	0.074***
		(0.037)	(0.034)	(0.029)	(0.027)
Difference in means between two income groups	Low - High (bef. shock)	-0.11***	-0.051***	-0.117***	-0.057***
	(5) = (2)	(0.009)	(0.01)	(0.01)	(0.01)
	Low - High (aft. shock)	-0.071**	-0.023	-0.056***	-0.011
	(6) = (4) + (2)	(0.028)	(0.027)	(0.021)	(0.02)
Difference in means between after and before income shock	Low: aft. shock - bef. shock	-0.01	-0.025	0.01	-0.016
	(7) = (3) + (4)	(0.023)	(0.022)	(0.019)	(0.018)
	High: aft. shock - bef. shock	-0.081***	-0.083***	-0.078***	-0.09***
	(8) = (3)	(0.028)	(0.026)	(0.021)	(0.02)
Difference in Difference	**Diff in Diff**	**0.071***	**0.058***	**0.089*****	**0.074*****
	(9) = (4) = (7) - (8)	(0.037)	(0.034)	(0.029)	(0.027)
Controls	Borrower demographics	No	Yes	No	Yes
	Fix year effect	No	Yes	No	Yes
Model fit	R-Square	0.006	0.155	0.007	0.156
	Number of Obs.	4,784	4,784	4,784	4,784

Table 7: Effect of low income and income shocks on liquid assets, linear regression

Note: This table presents the difference-in-difference results, which explain the impact of low income and adverse income shocks on liquid asset over mortgage using OLS regressions. Standard errors are reported in parentheses. ***, **, * indicate significance at the 1%, 5% and 10% level.

		Liquid asset over mortgages			
		Unemployment shock		Shock 3	
		No control	With controls	No control	With controls
Parameter estimates	Intercept (1)	1.728***	-1.356**	1.774***	-1.344**
		(0.058)	(0.663)	(0.061)	(0.662)
	inc_b50 (2)	-0.612***	-0.596***	-0.657***	-0.648***
		(0.083)	(0.082)	(0.087)	(0.086)
	shock (3)	-0.593***	-0.633***	-0.588***	-0.661***
		(0.2)	(0.197)	(0.15)	(0.148)
	inc_b50*shock (4)	0.625**	0.675***	0.598***	0.656***
		(0.259)	(0.254)	(0.202)	(0.199)
Difference in means between two income groups	Low - High (bef. shock)	-0.895***	-0.55***	-0.895***	-0.546***
	(5) = (2)	(0.068)	(0.075)	(0.072)	(0.078)
	Low - High (aft. shock)	-0.837***	-0.419**	-0.848***	-0.471***
	(6) = (4) + (2)	(0.201)	(0.2)	(0.153)	(0.153)
Difference in means between after and before income shock	Low: aft. shock - bef. shock	0.031	0.042	0.01	-0.005
	(7) = (3) + (4)	(0.164)	(0.161)	(0.135)	(0.133)
	High: aft. shock - bef. shock	-0.593***	-0.633***	-0.588***	-0.661***
	(8) = (3)	(0.2)	(0.196)	(0.15)	(0.148)
Difference in Difference	**Diff in Diff**	**0.625****	**0.675*****	**0.598*****	**0.656*****
	(9) = (4) = (7) - (8)	**(0.259)**	**(0.253)**	**(0.202)**	**(0.198)**
Controls	Borrower demographics	No	Yes	No	Yes
	Fix year effect	No	Yes	No	Yes
Model fit	R-Square	0.012	0.07	0.012	0.07
	Number of Obs.	4,743	4,743	4,743	4,743

Table 8: Effect of low income and income shocks on default (60-day and 90-day delinquency), 2SLS

Note: This table presents the difference-in-difference results, which explain the impact of low income and adverse income shocks on default rate using Two Stages Least Square regressions. Standard errors are reported in parentheses. ***, **, * indicate significance at the 1%, 5% and 10% level.

		60-day delinquency				90-day delinquency			
		Unemployment shock		Shock 3		Unemployment shock		Shock 3	
		No control	With controls	No control	With controls	No control	With controls	No control	With controls
Parameter estimates	Intercept (1)	-0.019	-0.192***	-0.019	-0.183***	-0.022**	-0.182***	-0.021*	-0.174***
		(0.014)	(0.053)	(0.014)	(0.053)	(0.011)	(0.043)	(0.011)	(0.043)
	inc_b50 (2)	0.001	-0.005	-0.001	-0.007	-0.004	-0.009**	-0.006	-0.01**
		(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
	shock (3)	0.035***	0.017	0.019**	0.005	0.013	0.004	0.002	-0.007
		(0.012)	(0.012)	(0.009)	(0.009)	(0.009)	(0.01)	(0.007)	(0.007)
	inc_b50*shock (4)	0.032**	0.045***	0.03**	0.038***	0.052***	0.059***	0.042***	0.048***
		(0.015)	(0.015)	(0.012)	(0.012)	(0.012)	(0.012)	(0.01)	(0.01)
	CLTV	0.057***	0.189***	0.057***	0.188***	0.052***	0.147***	0.053***	0.148***
		(0.02)	(0.043)	(0.02)	(0.043)	(0.016)	(0.034)	(0.016)	(0.035)
Difference in means between two income groups	Low - High (bef. shock)	0.001	-0.005	-0.001	-0.007	-0.004	-0.009**	-0.006	-0.01**
	(5) = (2)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
	Low - High (aft. shock)	0.033**	0.04***	0.029***	0.032***	0.048***	0.051***	0.036***	0.037***
	(6) = (4) + (2)	(0.014)	(0.014)	(0.011)	(0.01)	(0.011)	(0.011)	(0.009)	(0.008)
Difference in means between after and before income shock	Low: aft. shock - bef. shock	0.067***	0.062***	0.049***	0.043***	0.066***	0.063***	0.044***	0.04***
	(7) = (3) + (4)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.006)	(0.006)
	High: aft. shock - bef. shock	0.035***	0.017	0.019**	0.005	0.013	0.004	0.002	-0.007
	(8) = (3)	(0.012)	(0.012)	(0.009)	(0.009)	(0.009)	(0.01)	(0.007)	(0.007)
Difference in Difference	**Diff in Diff**	**0.032****	**0.045*****	**0.03****	**0.038*****	**0.052*****	**0.059*****	**0.042*****	**0.048*****
	(9) = (4) = (7) - (8)	**(0.015)**	**(0.015)**	**(0.012)**	**(0.012)**	**(0.012)**	**(0.012)**	**(0.01)**	**(0.01)**
Controls	Borrower demographics	No	Yes	No	Yes	No	Yes	No	Yes
	Mortgage characteristics	No	Yes	No	Yes	No	Yes	No	Yes

	State laws	No	Yes	No	Yes	No	Yes	No	Yes
	Economic conditions	No	Yes	No	Yes	No	Yes	No	Yes
Instrument variable	For CLTV	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination
Model fit	R-Square	0.015	0.061	0.011	0.057	0.018	0.055	0.012	0.049
	AUC	0.632	0.793	0.659	0.795	0.669	0.821	0.677	0.815
	Number of Obs	4,890	4,890	4,890	4,890	4,890	4,890	4,890	4,890

Table 9: Effect of low income and income shocks on default (60-day and 90-day delinquency), probit regression

Note: This table presents the difference-in-difference results, which explain the impact of low income and adverse income shocks on default rate using probit regressions. Standard errors are reported in parentheses. ***, **, * indicate significance at the 1%, 5% and 10% level. We note that the reported r-square is the pseudo-square.

		60-day delinquency				90-day delinquency			
		Unemployment shock		Shock 3		Unemployment shock		Shock 3	
		No control	With controls	No control	With controls	No control	With controls	No control	With controls
Parameter estimates	Intercept (1)	-3.049***	-4.433***	-3.061***	-4.307***	-3.251***	-13.583	-3.22***	-12.559
		(0.026)	(0.212)	(0.026)	(0.209)	(0.031)	(73.647)	(0.03)	(52.736)
	inc_b50 (2)	-0.01	-0.002	-0.044**	-0.041	-0.165***	-0.16***	-0.208***	-0.207***
		(0.021)	(0.023)	(0.022)	(0.025)	(0.026)	(0.03)	(0.027)	(0.031)
	shock (3)	0.455***	0.432***	0.288***	0.25***	0.324***	0.427***	0.083**	0.092**
		(0.036)	(0.04)	(0.03)	(0.034)	(0.043)	(0.05)	(0.038)	(0.044)
	inc_b50*shock (4)	0.268***	0.317***	0.344***	0.373***	0.634***	0.672***	0.707***	0.753***
		(0.045)	(0.049)	(0.039)	(0.044)	(0.053)	(0.061)	(0.049)	(0.055)
	CLTV	1.219***	1.277***	1.23***	1.278***	1.256***	1.356***	1.249***	1.324***
		(0.025)	(0.031)	(0.025)	(0.031)	(0.028)	(0.039)	(0.028)	(0.038)
Difference in means between two income groups	Low - High (bef. shock)	-0.01	-0.002	-0.044**	-0.041	-0.165***	-0.16***	-0.208***	-0.207***
	(5) = (2)	(0.021)	(0.023)	(0.022)	(0.025)	(0.026)	(0.03)	(0.027)	(0.031)
	Low - High (aft. shock)	0.258***	0.315***	0.299***	0.333***	0.469***	0.512***	0.499***	0.546***
	(6) = (4) + (2)	(0.04)	(0.044)	(0.033)	(0.036)	(0.046)	(0.054)	(0.04)	(0.046)
Difference in means between after and before income shock	Low: aft. shock - bef. shock	0.723***	0.749***	0.631***	0.624***	0.958***	1.099***	0.79***	0.845***
	(7) = (3) + (4)	(0.027)	(0.03)	(0.025)	(0.028)	(0.031)	(0.036)	(0.03)	(0.034)
	High: aft. shock - bef. shock	0.455***	0.432***	0.288***	0.25***	0.324***	0.427***	0.083**	0.092**
	(8) = (3)	(0.036)	(0.04)	(0.03)	(0.034)	(0.043)	(0.05)	(0.038)	(0.044)
Difference in Difference	**Diff in Diff**	**0.268*** **	**0.317*** **	**0.344*** **	**0.373*** **	**0.634*** **	**0.672*** **	**0.707*** **	**0.753*** **
	(9) = (4) = (7) - (8)	**(0.045)**	**(0.049)**	**(0.039)**	**(0.044)**	**(0.053)**	**(0.061)**	**(0.049)**	**(0.055)**
Controls	Borrower demographics	No	Yes	No	Yes	No	Yes	No	Yes
	Mortgage characteristics	No	Yes	No	Yes	No	Yes	No	Yes

	State laws	No	Yes	No	Yes	No	Yes	No	Yes
	Economic conditions	No	Yes	No	Yes	No	Yes	No	Yes
Model fit	R-Square	0.5096	0.7436	0.4981	0.7361	0.4698	0.7082	0.4396	0.6848
	AUC	0.751	0.838	0.764	0.842	0.809	0.869	0.809	0.866
	Number of Obs.	4,890	4,890	4,890	4,890	4,890	4,890	4,890	4,890

Table 10: Effect of low income and income shocks on default (60-day delinquency), positive equity, 2SLS

Note: This table presents the difference-in-difference results, which explain the impact of low income and adverse income shocks on default with positive home equity using Two Stages Least Square regressions. Standard errors are reported in parentheses. ***, **, * indicate significance at the 1%, 5% and 10% level.

		60-day delinquency with positive equity				**90-day delinquency with positive equity**			
		Unemployment shock		Shock 3		Unemployment shock		Shock 3	
		No control	With controls	No control	With controls	No control	With controls	No control	With controls
Parameter estimates	Intercept (1)	0.004	-0.061	0.003	-0.055	-0.004	-0.064**	-0.005	-0.056*
		(0.01)	(0.039)	(0.01)	(0.039)	(0.008)	(0.031)	(0.008)	(0.031)
	inc_b50 (2)	-0.001	-0.004	-0.002	-0.005	-0.003	-0.005*	-0.003	-0.006*
		(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
	shock (3)	0.011	0.007	0.011*	0.007	0.015**	0.012*	0.007	0.004
		(0.008)	(0.009)	(0.006)	(0.006)	(0.007)	(0.007)	(0.005)	(0.005)
	inc_b50*shock (4)	0.037***	0.042***	0.027***	0.03***	0.036***	0.039***	0.026***	0.029***
		(0.011)	(0.011)	(0.008)	(0.009)	(0.009)	(0.009)	(0.007)	(0.007)
	CLTV	0.01	0.059*	0.01	0.056*	0.015	0.042*	0.015	0.04
		(0.014)	(0.031)	(0.014)	(0.031)	(0.011)	(0.025)	(0.011)	(0.025)
Difference in means between two income groups	Low - High (bef. shock)	-0.001	-0.004	-0.002	-0.005	-0.003	-0.005*	-0.003	-0.006*
	(5) = (2)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
	Low - High (aft. shock)	0.036***	0.038***	0.025***	0.025***	0.034***	0.034***	0.024***	0.023***
	(6) = (4) + (2)	(0.01)	(0.01)	(0.008)	(0.008)	(0.008)	(0.008)	(0.006)	(0.006)
Difference in means between after and before income shock	Low: aft. shock - bef. shock	0.049***	0.049***	0.038***	0.037***	0.051***	0.052***	0.033***	0.033***
	(7) = (3) + (4)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
	High: aft. shock - bef. shock	0.011	0.007	0.011*	0.007	0.015**	0.012*	0.007	0.004
	(8) = (3)	(0.008)	(0.009)	(0.006)	(0.006)	(0.007)	(0.007)	(0.005)	(0.005)
Difference in Difference	**Diff in Diff**	**0.037*** **	**0.042*** **	**0.027*** **	**0.03*** **	**0.036*** **	**0.039*** **	**0.026*** **	**0.029*** **
	(9) = (4) = (7) - (8)	**(0.011)**	**(0.011)**	**(0.008)**	**(0.009)**	**(0.009)**	**(0.009)**	**(0.007)**	**(0.007)**
Controls	Borrower demographics	No	Yes	No	Yes	No	Yes	No	Yes
	Mortgage characteristics	No	Yes	No	Yes	No	Yes	No	Yes

	State laws	No	Yes	No	Yes	No	Yes	No	Yes
	Economic conditions	No	Yes	No	Yes	No	Yes	No	Yes
Instrument variable	For CLTV	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination
Model fit	R-Square	0.011	0.037	0.01	0.036	0.019	0.043	0.012	0.035
	AUC	0.608	0.802	0.666	0.816	0.667	0.841	0.708	0.84
	Number of Obs.	4,890	4,890	4,890	4,890	4,890	4,890	4,890	4,890

Table 11: Effect of low income and magnitude of income shocks on default (60-day delinquency), 2SLS

Note: This table presents the estimation results about the marginal impact of an increase in the income reduction likelihood on the default rate using Two Stages Least Square regressions. Standard errors are reported in parentheses. ***, **, * indicate significance at the 1%, 5% and 10% level.

		60-day delinquency						
		5%+ income drop	**10%+ income drop**	**15%+ income drop**	**20%+ income drop**	**25%+ income drop**	**30%+ income drop**	**35%+ income drop**
Parameter estimates	Intercept	-0.201***	-0.216***	-0.213***	-0.219***	-0.206***	-0.214***	-0.23***
		(0.054)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.057)
	$\widehat{inc}_{m,i,t}^{b50}$	0.284***	0.325***	0.305***	0.384***	0.406***	0.528***	0.757***
		(0.065)	(0.073)	(0.069)	(0.085)	(0.089)	(0.116)	(0.167)
	$\widehat{inc}_{m,i,t}^{a50}$	0.132***	0.137***	0.126***	0.13***	0.134**	0.158***	0.225***
		(0.044)	(0.047)	(0.049)	(0.049)	(0.052)	(0.061)	(0.081)
	CLTV	0.184***	0.184***	0.184***	0.184***	0.184***	0.184***	0.184***
		(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
Test for effect difference	Inc drop b50 vs a50	0.152***	0.188***	0.179***	0.254***	0.272***	0.37***	0.532***
		(0.035)	(0.043)	(0.044)	(0.058)	(0.064)	(0.085)	(0.119)
Controls	Borrower demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Mortgage characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	State laws	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Economic conditions	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Instrument variables	CLTV	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination
	$inc_{m,i,t}^{b50}$	inc_b50*shock	inc_b50*shock	inc_b50*shock	inc_b50*shock	inc_b50*shock	inc_b50*shock	inc_b50*shock
	$inc_{m,i,t}^{a50}$	inc_a50*shock	inc_a50*shock	inc_a50*shock	inc_a50*shock	inc_a50*shock	inc_a50*shock	inc_a50*shock
Model fit	R-Square	0.055	0.055	0.055	0.055	0.055	0.055	0.055
	AUC	0.788	0.788	0.788	0.788	0.788	0.788	0.788
	No of Obs	4,890	4,890	4,890	4,890	4,890	4,890	4,890

Table 12: Effect of low income and surprise effect of income shocks on default (60-day and 90-day delinquency), 2SLS

Note: This table presents the impact of low income and surprise income shocks on default rate using Two Stages Least Square regressions. Standard errors are reported in parentheses. ***, **, * indicate significance at the 1%, 5% and 10% level.

		All delinquencies				**Delinquencies with positive equity**			
		60-day delinquency		**90-day delinquency**		**60-day delinquency**		**90-day delinquency**	
		Unemployment shock	Shock 3	Unemployment shock	Shock 3	Unemployment shock	Shock 3	Unemployment shock	Shock 3
Parameter estimates	Intercept	-0.183***	-0.181***	-0.176***	-0.175***	-0.056	-0.053	-0.058*	-0.055*
		(0.053)	(0.054)	(0.043)	(0.043)	(0.039)	(0.039)	(0.031)	(0.032)
	inc_b50	0.002	0.002	0.000	-0.001	0.002	0.002	0.000	0.000
		(0.005)	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
	shock	0.021*	0.006	0.004	-0.008	0.008	0.007	0.012*	0.004
		(0.012)	(0.009)	(0.01)	(0.007)	(0.009)	(0.007)	(0.007)	(0.005)
	inc_b50*shock	0.039**	0.035***	0.059***	0.048***	0.04***	0.03***	0.04***	0.029***
		(0.015)	(0.012)	(0.012)	(0.01)	(0.011)	(0.009)	(0.009)	(0.007)
	CLTV	0.189***	0.187***	0.148***	0.148***	0.059*	0.056*	0.042*	0.041
		(0.043)	(0.043)	(0.034)	(0.035)	(0.031)	(0.031)	(0.025)	(0.025)
Controls	Borrower demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Mortgage characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	State laws	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Economic conditions	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Instrument variable	For CLTV	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination	HPA since origination
Model fit	R-Square	0.06	0.057	0.055	0.049	0.037	0.036	0.043	0.035
	AUC	0.792	0.794	0.822	0.815	0.799	0.812	0.843	0.839
	Number of Obs.	4,890	4,890	4,890	4,890	4,890	4,890	4,890	4,890