

NBER WORKING PAPER SERIES

CAN'T PAY OR WON'T PAY? UNEMPLOYMENT, NEGATIVE EQUITY, AND
STRATEGIC DEFAULT

Kristopher Gerardi
Kyle F. Herkenhoff
Lee E. Ohanian
Paul S. Willen

Working Paper 21630
<http://www.nber.org/papers/w21630>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2015

We are grateful for comments by Gene Amromin, Jan Brueckner, Satyajit Chatterjee, Morris Davis, Andra Ghent, John Krainer, Edward Kung, Stuart Gabriel, Erwan Quintin, Joe Tracy, and Rob Valetta as well as for comments from seminar participants at the 2014 FRBSF-Ziman Center Housing Conference, 2014 HULM Conference at FRB Chicago, and 2015 AREUEA. Jaclene Begley and Lara Lowenstein provided excellent research assistance. Herkenhoff thanks the Ziman Center for Real Estate for support. The views expressed in this article are those of the authors, and not those of the Federal Reserve Banks of Atlanta, Boston, Minneapolis, the Federal Reserve System, or the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2015 by Kristopher Gerardi, Kyle F. Herkenhoff, Lee E. Ohanian, and Paul S. Willen. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Can't Pay or Won't Pay? Unemployment, Negative Equity, and Strategic Default
Kristopher Gerardi, Kyle F. Herkenhoff, Lee E. Ohanian, and Paul S. Willen
NBER Working Paper No. 21630
October 2015, Revised January 2017
JEL No. G21,G33,R3,R51

ABSTRACT

This paper exploits matched data from the PSID on borrower mortgages with income and demographic data to quantify the relative importance of negative equity, versus lack of ability to pay, as affecting default between 2009 and 2013. These data allow us to construct household budgets sets that provide better measures of ability to pay. We use instrumental variables to quantify the impact of ability to pay, including job loss and disability, versus negative equity. Changes in ability to pay have the largest estimated effects. Job loss has an equivalent effect on default likelihood as a 35 percent decline in equity.

Kristopher Gerardi
Federal Reserve Bank of Atlanta
1000 Peachtree St. NE
Atlanta, GA 30309
kristopher.gerardi@atl.frb.org

Kyle F. Herkenhoff
Department of Economics
University of Minnesota
kfh@umn.edu

Lee E. Ohanian
8283 Bunche Hall
UCLA, Department of Economics
Box 951477
Los Angeles, CA 90095
and NBER
ohanian@econ.ucla.edu

Paul S. Willen
Federal Reserve Bank of Boston
Research Department
600 Atlantic Avenue
Boston, MA 02210-2204
and NBER
paul.willen@bos.frb.org

1. Introduction

A large literature has studied the determinants of residential mortgage default, with a focus on the extent to which default occurs among borrowers who have the ability to pay their mortgage, but who choose to default for what are called strategic reasons related to negative equity, compared to default among borrowers who simply do not have the ability to pay their mortgage. Understanding the relative importance of these determinants of default is central for designing policies aimed at reducing the probability of a future wave of mortgage defaults and foreclosures, and for designing loss mitigation policies that reduce the negative economic impacts of future possible foreclosure crises on lenders and homeowners (see for example, Chatterjee and Eyigungor (2009), Foote et al. (2010), Adelino et al. (2013)).

Measuring a borrower's ability to pay fundamentally requires detailed, household-level data on borrowers' economic attributes, including their income, their employment status, and their balance sheet, as well as their mortgage characteristics and payment status. However, previous studies have lacked data on many of these variables, and have either omitted variables from the analysis, or have used regional-level data to proxy for household-level data.

This paper makes two contributions to the literature. First, it uses new data from the Panel Study of Income Dynamics (PSID) and the PSID supplemental housing survey, which provide detailed data on borrower incomes, employment status, balance sheets, and consumption, matched with household mortgage data. These data allow us to construct household budget sets and thus, provide the most comprehensive measures of ability to pay within the literature. This in turn enables us to analyze the relative importance of strategic motives in mortgage default decisions, versus ability to pay, in considerably more detail than the existing literature. In particular, the analysis provides the first estimates of how changes in borrower ability to pay affects the likelihood of default. Moreover, we are able to address the important question of how changes in ability to pay interact with changes in equity in driving default decisions.

The second contribution of the paper is to systematically study not only defaulters, but those who pay their mortgage. As we describe below, our findings for both those who default and those who pay are critical for understanding the mortgage default process and designing loss mitigation policies.

We begin by classifying defaulting borrowers in terms of their ability to pay in order to quantify the extent of strategic default in the PSID. Strategic default broadly refers to defaulters who have the ability to pay, but who default because their home value has fallen below their loan amount (Mian and Sufi (2009)). We develop a procedure to assess strategic default by first forming household budget sets, and then identifying the defaulting households with negative equity positions who could continue making their mortgage payments without having to reduce their consumption below a specific level. To assess the robustness of this procedure, we use three definitions of this reference consumption level, ranging from maintaining the same household consumption level as in the previous year, to the level of subsistence consumption as defined by the Veteran's Administration (VA). We compare these reference minimum consumption levels to borrower residual income, which is the difference between household resources and the mortgage payment.

These budget set comparisons suggest that both strategic motives and the lack of ability to pay are important in understanding household default decisions. We identify strategic motives in about 38 percent of the defaulting households, as this group has the ability to pay their mortgage without reducing their consumption from their pre-default level.

However, we also show that almost 30 percent of defaulting households have such low ability to pay that they would need to reduce consumption below subsistence levels to remain current on their mortgages, and that the remaining 33 percent of defaulting households would need to at least reduce consumption below their pre-default level to remain current.

While strategic motives are quantitatively important among defaulting borrowers, the budget set comparisons for all borrowers show that nearly 96 percent of low equity borrowers with the ability to pay remain current. Moreover, we show that the vast majority of borrowers

with very low ability to pay avoid default. Specifically, 80 percent of households that need to cut their consumption to subsistence levels to make their mortgage payments (“cant pay” borrowers) are current on their payments. This finding provides a simple explanation for why lenders rarely negotiate pre-emptive mortgage modifications with even very high risk borrowers, since most of these borrowers continue to pay (Foote et al. (2010), Adelino et al. (2013)).

Following this descriptive analysis, we quantify the relative importance of strategic motives versus ability to pay by analyzing how changes in home equity and in residual income affect the probability of default in a multivariate setting. We first fit linear probability and logit models of default on a rich set of covariates that allow us to control for a variety of economic and demographic factors.

To address some possible endogeneity issues, we next use the richness of the PSID to construct instruments for residual income and housing equity. To instrument for equity, we use the state-level house price appreciation since the purchase of the house. Instrumenting for residual income is more challenging. We therefore use three sets of instruments, and assess the robustness of the results across these specifications.

We exploit the long time series dimension of the PSID to construct household-level unemployment shocks to instrument for residual income. We focus on involuntary separations and control for previous unemployment spells to account for potential endogeneity concerns. The second and third instruments consist of two components that are motivated by previous research. The first component is a health disability shock to instrument for residual income, which follows Low and Pistaferri (2015). The third instrument is a Bartik-type state-level employment shock that is based on aggregate employment flows and industry shares at the state-level.

All of our instruments are strong predictors of residual income, and deliver similarly large estimates of the causal effect of residual income on mortgage default. Our IV estimates indicate that a 10 percent decline in residual income raises the probability of default by

between 1.1 and 2.5 percentage points.

To compare the magnitudes of residual income loss and changes in equity on default, we note that our reduced form estimates indicate that the effect of involuntary job loss on the default probability is equivalent to a 37 percentage point drop in equity. More broadly, we find that the estimated impact on the default probability of a one percent decline in residual income is about equal to the estimated impact of a one percentage point decline in equity. This means that a \$100 change in residual income for a household with \$1,000 available after paying the mortgage has an equivalent effect on the default probability as a \$500 change in the value of a home for a homeowner with \$50,000 of equity.

Regarding the importance of strategic motives, while approximately 38 percent of defaulters do have the ability to pay, we find that the estimated likelihood of default among low equity borrowers with the ability to pay is fairly low. Specifically, our IV estimates indicate that an increase in LTV from 75 percent to 125 percent raises the default probability of a high residual income borrower from about 3 percent to about 5 percent. However, we find that an increase in LTV from 75 percent to 125 percent raises the default probability for a low residual income borrower from 10 percent to 17 percent. This finding highlights a quantitatively important interaction between ability to pay and borrower equity in the pay/default decision.

Taken together, these findings have implications for the design of policies. In particular, they indicate that policies designed to reduce foreclosure by reducing monthly mortgage payments can be very effective, because these policies raise residual income. This applies to both low and high equity households, with the relative effect on the default probability being larger for high equity (low loan-to-value ratio) households, but the absolute effect being higher for the low equity (high loan-to-value ratio) households.

The paper is organized as follows. Section 2 discusses the approach in this paper within the context of some of key papers within the literature and describes in detail the PSID data used in the empirical analysis. Section 3 uses the PSID data to construct alternative measures

of residual income that we use to assess ability to pay, and provides cross-tabulations of ability to pay with defaulting and current borrowers. Section 4 presents regression estimates with a focus on quantifying the marginal contributions of residual income and homeowner equity. Section 5 discusses the implications of the results for economic policy and future research. Section 6 concludes.

2. Data

This section presents the data used in this analysis. The major data innovation in the analysis is the use of matched data on mortgage characteristics and status with borrower socio-economic and demographic variables. These matched data advance the literature in a number of ways.

One advance is on the measurement of household ability to pay. Measuring ability to pay in the literature has been very limited, and consequently little is known about the importance of this factor. On the one hand, anecdotal and limited survey results suggest that major life events such as job loss, illness and divorce are associated with mortgage default, (see Cutts and Merrill (2008) and Hurd and Rohwedder (2010)). However, previous quantitative studies of default have provided only weak evidence on the importance of these events due to the lack of household-level income, employment, and balance sheet data.¹ This lack of household-level data has led many researchers to use aggregate unemployment rate data and divorce rate data as proxies for household-level income shocks (e.g. Deng et al. (1996), Deng et al. (2000), Elul et al. (2010), and Bhutta et al. (2011)). These studies have found only weak correlations between these aggregate measures and default. More recently, Gyourko and Tracy (2014) analyze micro loan level data with county-level unemployment rates as a control. However, they adjust regional unemployment rate controls for attenuation bias, and this adjustment indicates a significant relationship between adjusted unemployment

¹To be clear, many administrative mortgage datasets do include some information on income and employment at the time a loan is originated, but to our knowledge, none of these datasets include information on these variables after origination.

rates and default. This evidence more broadly suggests a stronger relationship between income shocks and mortgage default than found in the earlier studies.

As described below, the PSID data on borrower mortgage information with borrower information on income, employment status, balance sheet data, and consumption enable us to measure ability to pay, and analyze its importance in default, in considerably more detail than in the previous literature.

Our enhanced measures of ability to pay also have important implications for defining and classifying strategic default. To see this, we note that the most prominent measures of strategic default in the existing literature are based on survey respondents who report whether or not they knew people who had the ability to pay their mortgage, but walked away from their homes during the crisis (Guiso et al. (2013)). In Online Appendix A we provide a comparison of our strategic default estimates to the literature, including comparisons of samples and methodologies.²

We use a very different approach to identifying strategic default by constructing household budget sets to measure ability to pay. Note that this approach to defining strategic default is considerably different from survey respondent subjective assessments of other's ability to pay. An important benefit of our approach is that it is scientifically reproducible across researchers, and thus can provide significant discipline in analysis. We therefore view this approach of classifying defaulters in terms of their ability to pay as an important advance relative to other studies.

2.1. Sample Construction

The primary data used in this study come from the 2009, 2011, and 2013 PSID Supplements on Housing, Mortgage Distress, and Wealth Data. We restrict the sample to mortgagor heads between the ages of 24 and 65 who report being in the labor force or being disabled. We also restrict the sample to households with LTV ratios below 250 percent that had not

²In Online Appendix A we focus on three studies in particular: Experian and Oliver Wyman (2009), Guiso et al. (2013), and Bradley et al. (2015).

defaulted as of a prior survey.³ These sample restrictions leave us with 7,404 households.⁴

2.2. Variable Definitions and Representativeness of the PSID

The unit of analysis in this study is the household. The household includes both the “head” and “spouse” as defined by the PSID, along with any children and other persons living in the primary residence. The primary measure of income is total household income, which is composed of the sum across household members of (1) wage and salary income; (2) transfer income (including social security, alimony and child support); (3) business income; and (4) interest and dividend income. This measure corresponds to the IRS definition of adjusted gross income less realized capital gains.⁵

Our measure of consumption includes expenditures on food, housing, clothing, health care, entertainment, and education. In Online Appendix B we show how the PSID consumption measures compare to the Consumer Expenditure Survey (CEX) measures as tabulated by the Bureau of Economic Analysis (BEA) from 2009-2013. We find that in general, consumption levels are quite similar across the two datasets, and for most expenditure categories, the trends in consumption are also quite similar.

The PSID provides information (i.e. interest rates and amounts) on all liens on the household’s principal residence (1st, 2nd, and 3rd mortgages). In addition, the survey includes the respondent’s estimate of the current market value of the principal residence. Table 1 compares mortgage statistics from our PSID sample with data from the 2009, 2011, and 2013 National American Housing Survey (AHS).⁶ In general, mortgage characteristics are

³The LTV requirement drops what appear to be misreported mortgage and home values (inclusion of these observations does not materially change the main results). Dropping households that reported being in default in a previous survey simply eliminates double counting.

⁴In Online Appendix A, we compare the sample selection criteria with previous studies of mortgage default. Relative to the existing literature, the sample is quite broad and, as we will show in the following section, appears to be representative of the population of mortgagors. It includes both fixed-rate and adjustable-rate mortgages, as well as older origination cohorts that have accumulated significant amounts of equity in their homes.

⁵In Online Appendix B we compare our PSID measure of average family income to what is reported by the Census, and show that they are very similar.

⁶The AHS is conducted biennially by the U.S. Census Bureau. It has a sample size of about 50,000

quite similar across the two datasets. The median outstanding principal balance is identical in both datasets in 2011 and within \$15,000 in 2009 and 2013. The median monthly mortgage payment is within \$200 in 2009 and \$100 in 2011 and 2013. The median mortgage interest rates, remaining terms, and LTV ratios (calculated for first liens only) are also extremely close in both datasets. Finally, the fraction of households with second mortgages and adjustable-rate mortgages (ARMs) is also similar across the two datasets, with slightly more households in the PSID reporting that they have ARMs and second mortgages compared to the AHS.

Figure 2 displays the distribution of housing equity in our PSID sample compared with the distribution in CoreLogic.⁷ According to CoreLogic, slightly more than 10 percent of properties in 2009 had greater than 25 percent negative equity, while slightly less than 4 percent did so in the PSID. While there could be many reasons for the divergence in equity estimates between the two databases, households tend to over-report house values as compared to actual selling prices by 5 percent to 10 percent (see Benítez-Silva et al. (2008)). While the PSID understates the amount of negative equity in the economy relative to CoreLogic estimates, we do not view this as a significant drawback of our analysis. In quantifying the role that negative equity plays in causing mortgage default, we believe that self-reported equity is the most appropriate measure. In choosing whether or not to default, households take into account their own perceived valuation of their home, which may or may not be derived in part from a third-party estimate (such as CoreLogic or Zillow). To put it another way, the value of using self-reported equity values is that only those households that believe that they are in positions of negative equity are flagged as having negative equity, and this is the group of households that we expect to be most sensitive to negative equity in

housing units and was designed to provide representative data on the U.S. housing and mortgage markets.

⁷The bottom panel of Figure 2 comes from the August 13, 2009 report entitled “Summary of Second Quarter 2009 Negative Equity Data from First American CoreLogic” http://www.loanperformance.com/infocenter/library/FACL%20Negative%20Equity_final_081309.pdf CoreLogic uses a national database of property transactions that covers 43 states to calculate their equity estimates, and thus their data should be quite representative of the U.S. population. CoreLogic uses administrative data on outstanding mortgage balances and estimates of housing values to compute equity, while we use reported mortgage balances and housing values in the PSID.

terms of default behavior.⁸ To further assess the performance of self-reported home values, Online Appendix J demonstrates that our LTV point estimates are consistent with existing results from proprietary loan level datasets.

Information on mortgage performance in the PSID is available beginning in the 2009 survey.⁹ Households were asked how many months they were behind on their mortgage payments at the time of the PSID interview. In the empirical analysis below we adopt a default definition that corresponds to two or more payments behind (at least 60 days delinquent), which is standard in the literature. Approximately 3.3 percent (248/7404) of our sample reported being at least 60 days late on their mortgage payment (using weights, this statistic falls to 2.7 percent).¹⁰ According to the National Delinquency Survey conducted by the Mortgage Bankers Association (MBA), the 60+ day delinquency rate calculated across all U.S. households with a mortgage in 2009 was 5.8 percent. When we include households that report being at least two payments behind in multiple surveys, the corresponding number in our PSID data is approximately 3.7 percent. One reason for this lower rate in the PSID is that it does not take into account mortgage delinquency associated with properties that are not primary residences (i.e. investment and vacation properties), whereas the MBA rate includes all mortgaged properties.¹¹

Finally, we use information on unemployment spells in our empirical analysis below. The PSID provides the employment status for both the head and the spouse over the previous calendar year as well as at the time of the interview. We discuss our exact unemployment variable definitions in detail in Section 4.1 where we present the results from our empirical

⁸In addition, it is likely the case that many households have information about the condition of their home and the state of their local housing market that is not captured in data-based estimates such as the CoreLogic numbers, which use zip code-level or county-level house price indices to estimate property values.

⁹There is some information on mortgage characteristics in PSID surveys prior to 2009, but there is no information on mortgage performance.

¹⁰Information on missed payments is only provided at the time of the interview making it impossible to measure the exact timing of the first missed payment. This means that we cannot identify, for example, borrowers who missed two or more payments at some point in the previous calendar year but cured by the time of the interview.

¹¹Unfortunately, the PSID does not include information on mortgage delinquency for properties that are not primary residences, so it is not possible to perform an apples-to-apples comparison.

models. Using the measure of employment status at the time of the survey yields an unemployment rate of 5 percent in our sample of mortgagors. For the years in question (2009, 2011 and 2013), the average of the headline unemployment rate reported by the Bureau of Labor Statistics was 8.5 percent.¹²

2.3. Summary Statistics

In Table 2, we provide summary statistics for our overall PSID sample as well as for the subset of mortgagor households who have defaulted on their loans.

Panel A of Table 2 reveals several key facts about the distributions of income and consumption for defaulters versus the population as a whole. First, defaulters have much lower levels of income than the population as a whole. The median income of defaulters (\$60,000) is 37 percent below the median of the full sample (\$94,000). While the entire distribution of income is lower for defaulters, the table shows that some defaulters do have considerable income; 10 percent of defaulters have pre-tax income of at least \$130,000. Households in default are much more likely to report a decline in income, as the median defaulter reports a seven percent fall in income over the two previous years compared to a six percent increase in income for the median household in the full sample. In addition, 42 percent of defaulters have experienced a drop in income exceeding 15 percent compared to only 19 percent for the whole sample. Differences in consumption are much smaller than differences in income (\$50,000 per year for defaulters compared to \$56,000 for the full sample, on average).

In Panel B we see that households in default are less educated and less likely to be married than the typical mortgagor. College graduates account for 45 percent of the sample but only 23 percent of defaults, while approximately 30 percent of household heads are not married, but account for 45 percent of defaults. Panel B also shows that the age distribution for defaulters and non-defaulters is quite similar.

Panel C shows that the distribution of LTV ratios is significantly higher for defaulters,

¹²It is well-known in the literature that homeowners are less likely to experience an unemployment spell compared to renters, which likely explains a significant portion of this gap.

a fact that has been well-documented in the literature. In Panel D we can clearly see that defaulters also have significantly less wealth. The median defaulter has only \$518 in liquid assets compared to the median household in the sample that has more than \$6,000.¹³ The gap in liquid wealth is especially large at the top of the distribution (90th percentiles of \$5,429 and \$57,000 respectively).

Finally, Panel E of Table 2 displays information on unemployment spells and disability shocks. It is clear from the panel that households in default are much more likely to have experienced a spell of unemployment. As of the survey date, 5 percent of the full sample of households reports being unemployed compared to 20 percent of the sample of defaulters. A similar pattern emerges for households that have experienced a disability. Only 1.5 percent of households in the full sample report having suffered a severe disability since the previous interview, compared to more than 6 percent of defaulters.¹⁴

3. Mortgage Affordability and Strategic Default

Since the mortgage foreclosure crisis that occurred in 2007 and the subsequent financial crisis and recession, the concept of strategic default has become a popular topic in the economics and finance literature. A major limitation of this literature however, is the lack of an economic framework to help distinguish between borrowers who strategically default and those who do not. As a result, there is significant disagreement about how to define strategic default, which has predictably led to very different estimates of its importance in the mortgage market.

In this section we develop a definition of strategic default that is linked to the economic concept of affordability. In the first part of the section, we establish a simple method for classifying mortgage payments into those that are “affordable” and those that are “unaffordable”

¹³Liquid assets include checking and savings accounts, money market funds, certificates of deposit, government savings bonds, and Treasury bills. Illiquid assets include equity and bond holdings, the value of automobiles, retirement accounts, and business income. Housing equity is not included in the measure of illiquid assets.

¹⁴A detailed description of how we construct disability shocks is provided in section 4.2 below.

and show that this classification yields significant differences in default rates across borrowers. In the final part of the section we relate this classification to the notion of strategic default, and use our PSID data to quantify its importance.

3.1. Identifying “Can Pay” and “Can’t Pay” Borrowers

We begin by proposing a classification system for mortgage defaults using a standard household budget constraint. Specifically, we define cutoffs for “unaffordable” and “affordable” mortgage payments based on the amount of disposable income available for a household to consume. Let c denote household spending on non-housing consumption in the year of default and h denote housing expenditures, which are financed with a mortgage with required payment, m . Assuming, for now, that a household has no wealth and cannot borrow in unsecured credit markets, the household budget constraint is given by:

$$c + h \leq y. \tag{1}$$

The household is faced with a choice of either making the mortgage payment m or defaulting, experiencing foreclosure, and subsequently paying rent r for a new home.¹⁵ Given a choice of m or r , the household’s residual income, $y - m$ or $y - r$, respectively defines its consumption, meaning that the household is choosing between the combination of paying the mortgage and consuming $y - m$ versus defaulting and consuming $y - r$. We assume that $m > r$ so a borrower can always increase non-housing consumption by defaulting. Even with perfect information about y , m and r , we cannot answer the question of whether a borrower should default without information about preferences, for example, over renting versus owning, or beliefs about the evolution of future house prices. But, even without such information, one can ask about the effect on residual income of the decision to make the mortgage payment and that is our focus in this section.

¹⁵A foreclosure severely impacts an individual’s credit score for seven years in the U.S., making it extremely difficult to obtain another mortgage to purchase a home during that period.

First, we define a mortgage as being *unaffordable* if the payment m leads to residual income that is below a subsistence level of consumption. We call this level c_{VA} because we use the Veteran’s Administration (VA) rules to measure subsistence. Formally:

$$\text{Unaffordability} \Leftrightarrow y - m < c_{VA}.$$

Intuitively, regardless of preferences, a mortgage payment is unaffordable if the household is unable to meet its basic necessities with its residual income.

Second, we define a mortgage as *affordable* if the household can maintain its level of consumption from the previous year c_{-1} , where we are assuming that the household chose not to default in the previous year.¹⁶ Formally:

$$\text{Affordability} \Leftrightarrow y - m > c_{-1}.$$

The idea here is that the fact that the household can maintain exactly its consumption level while paying the mortgage captures the popular notion of a borrower who “can afford his mortgage.” It is important to note that our definitions are not exhaustive as a mortgage could be neither unaffordable nor affordable if $c_{VA} < y - m < c_{-1}$, meaning that the residual income allows for consumption above subsistence levels but would require a reduction in consumption versus the previous year.¹⁷

While this is a simple framework to classify borrowers it is difficult to operationalize. To do so requires detailed data on both household consumption and income, in addition to mortgage debt, which previous studies on the topic have lacked. Fortunately, all of these variables are available in the PSID data. Our measure of household income, y , is the monthly average of after-tax income of the family unit, measured over the previous calendar year.¹⁸ Our measure of m is the sum of all mortgage payments, property taxes,

¹⁶Recall from our discussion above, that only first-time defaults are retained in the sample.

¹⁷It is possible for consumption to be *both* affordable and unaffordable if $c_{VA} > c_{-1}$ but we show that this is extremely rare in our data.

¹⁸Ideally, we would like a measure of residual income at the time of the survey to be consistent with the

and insurance associated with the family unit’s primary residence. $c(\text{VA})$ is a subsistence level of consumption defined by the VA that depends on the size and geographical location of the household.¹⁹ Our measure of consumption, c_{-1} , is the average monthly expenditures of the household, excluding mortgage related expenses, which is, as with income, measured over the previous calendar year.

Table 3 displays a set of simple cross tabulations using these definitions. In Panel A, column (1), we see that about 70 percent of all households in our sample have mortgage payments that are affordable based on our above classification. We refer to these households as “can pay.” In contrast, approximately 7 percent have unaffordable mortgage payments (i.e. their residual income is less than VA subsistence levels), and we refer to these as “can’t pay” households (column (3)). In column (2), we see that approximately 23 percent of households are in-between, meaning that they have enough income to pay their mortgages and consume more than subsistence levels, but not enough to maintain their previous levels of consumption. In Panels B and C of Table 3, we stratify the sample by LTV ratio. High LTV households ($LTV > 90$) are slightly less likely to be “can pay” (66.4 percent compared to 70.9 percent) and slightly more likely to be “can’t pay” (8.3 percent compared to 6.8 percent) than low LTV households ($LTV < 90$).

The default rates in Table 3 show that the can pay/can’t pay distinction has power. Focusing on Panel A, column (1) shows that out of more than 5,000 “can pay” households,

timing of our mortgage default variable. For this reason, we adjust income to account for the household’s employment status at the time of the survey. Specifically, if the head of household is not employed as of the survey date, we reduce y by the average monthly labor earnings of the head. If the spouse is not employed as of the survey date or the head was recently divorced, we reduce y by the average monthly labor earnings of the spouse. We do not make this adjustment in our regression analysis in section 4 since doing so would introduce a mechanical correlation between our measure of residual income and the unemployment instruments that we employ. We make no adjustment for households who are employed at the time of the survey, so for these households residual income is measured with a lag relative to their default decision. However, this timing discrepancy is unlikely to be a major issue as the vast majority of PSID interviews (about 80 percent) take place within the first six months of the calendar year.

¹⁹For more details see “Lenders Handbook - VA Pamphlet 26-7,” Ch.4 on underwriting loans which is available online, http://www.benefits.va.gov/warms/pam26_7.asp. This includes the VA definition of residual income as “Residual income is the amount of net income remaining (after deduction of debts and obligations and monthly shelter expenses) to cover family living expenses such as food, health care, clothing, and gasoline” (p. 55).

only 1.4 percent (74) default. In contrast, of the 531 “can’t pay” households, 10.7 percent (57) default, which implies that “can’t pay” borrowers are approximately 7 times more likely to default than “can pay” borrowers. Dividing the sample into high and low LTV samples yields even more dramatic differences. The least risky subsample, “can pay” households with low LTV ratios, account for more than half the sample (4,056 of 7,404 or 55 percent), and the table shows that only 0.7 percent of these borrowers default. In contrast, the most risky subsample, “can’t pay” households with high LTV ratios have a default rate approaching 20 percent. In other words, based solely on the ability-to-pay variables and LTV dichotomy we can identify groups with a 30-fold difference in default rates.

Comparing differences in default rates also yields important insights. First, the likelihood of default is very low for low LTV households – as column (4) shows, only 1.4 percent default – but for the “can’t pay” subsample of low LTV households the default rate soars to 7.5 percent, which is more than 10 times the default rate of “can pay” low LTV households. This is intuitive since the essence of “can’t pay” is that the borrower simply does not have the cash flow necessary to make the mortgage payments and maintain a minimal level of consumption. The fact that the household could enjoy a positive income shock or sell the house in the future does not matter, whereas for the “can pay” household, there is little point in defaulting as it has the cash flow necessary to make the payment without needing to sacrifice a significant amount of consumption. As one would expect, for the high LTV households, the effect of ability-to-pay on differences in default rates is smaller than for low LTV households (a five-fold difference instead of ten-fold).

While Table 3 is consistent with an important role for ability-to-pay, it also illustrates the limits of the framework. As discussed above, the fact that we find a 30-fold difference in default rates between “can’t pay” borrowers with high LTV ratios and “can pay” borrowers with low LTV ratios illustrates the importance of ability-to-pay. However, the flip-side of the fact that 20 percent of the “high-risk” households default is that 80 percent of them continue to make their payments. Indeed, the use of the phrase “can’t pay” to describe a subsample

of the population in which almost 90 percent (if we look at the whole sample including both high- and low-LTV households) *do* pay is, in a sense, a contradiction in terms. The issue is that while “ability-to-pay” is an easy concept to talk about it is not an easy concept to formalize. Equation (1) seems intuitive, but, it is based on the assumption that households must finance their current spending entirely out of current income. In reality, households can, potentially, finance spending either by borrowing or by drawing down accumulated savings. In other words, a more realistic version of equation (1) would look like:

$$c + m < y + a + b \tag{2}$$

where a is accumulated financial assets (i.e. wealth) and b is the maximum amount of (unsecured) credit that a household can access. Moving to such a formulation is not easy, especially in the context of strategic default. It seems reasonable to call a default strategic if a household has free cash flow that exceeds the cost of the mortgage. However, would it be equally appropriate to call a default strategic if the household could only “afford” the mortgage payment by drawing down its retirement savings or borrowing on credit cards? In other words, is default strategic unless the household has exhausted all of its savings and borrowed up to the maximum amount available on all available credit lines?²⁰

While we cannot tell for certain, it seems reasonable that the “can’t pay” households that *do* pay are using some combination of borrowing and drawing down savings, perhaps augmented by resources from their extended family. In principle, one could answer this question with data, but to assess the sources of funds for payments, one would need much higher frequency wealth information than the biennial data from the PSID.²¹

²⁰Adding more nuance here, by expanding the budget constraint, one could argue that a “can pay” household is diverting money from saving and, therefore, future consumption by making its monthly payment. If along some future path, such a lack of saving results in destitution, then some “can pay” households, as we have defined them, really cannot afford their mortgage payments.

²¹In Online Appendix C we incorporate information on assets and liabilities in the PSID to create a version of Table 3 that is based on equation (2). The results are broadly similar.

3.2. Quantifying Strategic Defaults

Our discussion above links the concept of affordability to mortgage default in a manner that makes it a natural definition for what has been termed “strategic” or “ruthless” default in the literature. The idea is that a household that chooses to default on its mortgage debt while having the ability to make its mortgage payment and maintain its level of consumption, has made a strategic decision. Such a definition is internally consistent with standard models of defaultable debt, as well as the popular notion of “ruthless” default, whereby a borrower defaults for purely investment considerations, as opposed to liquidity-related concerns. In this section we use our classification of “affordable” and “unaffordable” mortgage payments to quantify the extent of strategic default in the data.

Column (1) of Panel A of Table 3 shows that of the 196 defaulters in our sample, 74 defaulters or 38 percent, have affordable mortgages, in the sense that they could make their mortgage payment without reducing their consumption, implying that 38 percent of defaults are strategic. Column (3), however, shows that only 29 percent of mortgages are unaffordable in the sense that household consumption would drop below subsistence levels if the household makes the payment. Thus, the number of strategic defaults depends on exactly how one defines affordability. If one adopts the broad idea that a mortgage is only affordable if making the mortgage payment requires no reduction in household consumption, then the share of strategic defaults, 38 percent, is comparatively small. But using this definition implies that all other consumption takes priority over the mortgage. For example, if paying the mortgage requires that the household replace a luxury car with a more modest alternative, this definition would say that the household cannot afford the mortgage. At the other extreme, if one adopts the much stricter idea that a mortgage is unaffordable if making the mortgage payment will lead to a level of consumption that is below subsistence, then a comparatively large fraction of defaulters, 71 percent, are strategic.

Panels B and C show that strategic default is somewhat more common, using either definition, for high LTV borrowers and less common for low LTV borrowers. Whereas for

the whole population, our strategic default estimates ranged from 38 to 71 percent, for the high LTV sample, they range from 41 percent to 76 percent and for the low LTV sample from 33 percent to 64 percent. The response to LTV is consistent with the idea that high LTV ratios make households more likely to default even when they can afford their monthly payments.

In Online Appendix D, we consider an alternate definition of affordability based on the Qualified Residential Mortgage (QRM) guidelines, and we find consistent estimates of strategic default.

4. Default and Residual Income

The analysis in Section 3 above strongly suggests that many households default because they do not have the financial resources to continue making their periodic mortgage payments. For example, we found that approximately 29 percent of households in default do not have enough residual income to meet their basic consumption needs, and an additional 33 percent of defaulters do not have enough residual income to maintain their pre-default consumption levels. While illustrative, that analysis was entirely descriptive in nature. In this section, we attempt to measure the causal impact of residual income on mortgage default.

We begin by showing that the correlation between low residual income and mortgage default is very strong in the data, even after controlling for potentially confounding variables including housing equity, a detailed set of mortgage characteristics, geographic factors, and a detailed set of household demographic characteristics available in the PSID. We then use the richness of the PSID data to construct instruments for residual income and housing equity to directly address potential endogeneity bias. Our choice of instruments is based on the idea that negative shocks to household-level income should result in low residual income levels. We show that this is clearly the case in the data, as households in the bottom of the residual income distribution are much more likely to have experienced a recent negative income shock

compared to those in the top of the distribution. It then follows that any exogenous shock that affects household income is a good candidate for an instrument for residual income. We focus on two plausibly exogenous shocks: involuntary unemployment spells and disability.

4.1. OLS and Logit Results

We start our analysis by estimating OLS and logistic multivariate regressions of mortgage default on residual income, where we condition on a detailed set of household demographic characteristics, mortgage characteristics, as well as geographic controls (at the state-level). We calculate residual income using gross (before tax) household income that excludes capital gains, but includes all other sources of income including income from operating a business (see Section 2 for more details) less total mortgage expenses, including the first and second mortgage. We focus on the (natural) logarithm of residual income, in order to capture a potential non-linear relationship due to the existence of subsistence consumption levels.²² A minimum level of consumption that is required for survival, implies that the same increase in residual income for a household with very low levels of income should have a larger impact on its default decision compared to the decision of a household with very high levels of income.²³

24

Table 5 displays our baseline results. Columns (1) - (3) display OLS regression estimates (i.e. linear probability models) of an indicator of mortgage default (at least 60-days delinquent) on residual income, and columns (4) - (6) repeat the exercise using logistic regressions. In columns (1) and (4) we do not include any controls, while in the remaining columns we in-

²²There is a small issue in specifying the logarithm of residual income as a few households in our data have negative values of residual income. To deal with this issue, we winsorize the residual income variable choosing a threshold that corresponds to the first percentile of the residual income distribution \$3,810. We have experimented with alternative thresholds and find that the results reported below are not sensitive to this particular choice.

²³By taking the log of residual income, we are assuming that the impact of residual income on default is proportional. For example, the impact on default of an increase in residual income of \$50 for a household that starts with only \$100 in residual income will be the same as an increase in residual income of \$5,000 for a household that starts with \$10,000 in residual income.

²⁴In Online Appendix E we consider alternate specifications which use the ratio of m to y (i.e. the debt-to-income ratio), and we show that our baseline results are robust. Furthermore, in Online Appendix F, we consider income changes, and we show that our baseline results are robust.

clude numerous demographic, state-level, and loan-level controls.²⁵ In all columns we include the household’s (self-reported) LTV ratio at the time of the survey, as the prior literature has documented that home equity is a strong predictor of default.

The OLS coefficients associated with the logarithm of residual income should be interpreted as semi-elasticities. Thus, according to column (1), a 10% decrease in residual income is associated with a decrease in the likelihood of default of approximately 0.37 percentage points. In column (2) we see that the magnitude of the semi-elasticity drops somewhat when we include the control variables, but is still negative and statistically significant. In column (3) we include an interaction between the household’s LTV ratio and residual income. This specification is motivated by the double-trigger theory of mortgage default, which predicts that the combination of liquidity shocks and declines in home values generates large increases in mortgage defaults.²⁶ The coefficient associated with the interaction term is negative and statistically significant, which means that for higher LTV ratios (lower equity levels), decreases in residual income have more pronounced effects on default. For example, at an LTV ratio of 1 (no equity), a 10 percent drop in residual income will increase the default rate by 0.43 percentage points ($=(-.1)*.012-.055*1*(-.1)$), while at an LTV ratio of 1.5 (negative equity), a 10 percent drop in residual income increases the default rate by 0.7 percentage points ($=(-.1)*.012-.055*1.5*(-.1)$).

Columns (4) - (6) in Table 5 display estimates from logitistic regressions that correspond to the same specifications as columns (1) - (3). The logit coefficients are reported in the table without parentheses, while the standard errors are reported just below the coefficients (round

²⁵The demographic controls include 1-digit industry, year, race, education, marital status, and gender indicator variables as well as the age of the head of household and the number of children in the household. The mortgage controls include the mortgage interest rate as well as dummy variables for origination years, whether the mortgage is refinanced, the presence of a second mortgage, whether the term remaining is >15 years, whether or not the loan is refinanced. The state controls include indicator variables signifying if the state has a judicial foreclosure process, if the state allows lender recourse, and if the state is one of the “sand” states (Arizona, Florida, and Nevada) that experienced an especially dramatic housing boom and bust during the 2000s. In addition, changes in state-level house prices and unemployment are included. Online Appendix G includes a complete list and summary of the baseline set of controls.

²⁶Examples include Corbae and Quintin (2009), Garriga and Schlagenhauf (2009), Chatterjee and Eyigungor (2011), Campbell and Cocco (2011), Hedlund (2011), Schelke (2011), and Laufer (2012).

parentheses), and the average marginal effects (AME) are reported below the standard errors (square parentheses). The average marginal effects in columns (4) and (5) have the same sign and similar magnitudes to the OLS coefficients reported in columns (1) and (2). Column (6) includes the interaction term between LTV and residual income. Due to the non-linearity of the logit model, the AME associated with the interaction must be calculated at discrete points in the state space. We find that the AME associated with the interaction term, computed as the difference in the AME of LTV between the interquartile range of residual income, is about half of the magnitude of the corresponding estimate in the linear probability model. However, the magnitude of the interaction effect is sensitive to the points at which it is calculated, and in many instances the interaction effect is larger than the LPM model. As an additional check on the PSID data, in Online Appendix J, we compare our LTV point estimates to several recent papers on mortgage default, and we find that our point estimates are consistent with results from proprietary loan level datasets.

While the results in Table 5 support the findings in Section 3 and strongly suggest that residual income plays an important role in household-level mortgage default decisions, they are still descriptive in nature. Income is not randomly assigned to households, but is instead determined by decisions consciously made by the household as well as some outside, exogenous forces. Therefore, in order to make any causal inferences about the effect of residual income on default, we need valid instruments. As we mentioned above, our instrumental variables approach exploits the fact that many households with low residual income suffered major adverse shocks.

Table 4 provides details on the incidence of adverse shocks by quintile of residual income. The first row in the table shows that households in the bottom of the distribution are much more likely to have experienced a significant (30 percent or higher) drop in income since the previous interview compared to those in the top of the distribution. The next row in the table shows that unemployment is an important shock that leads to residual income differences as household heads in the bottom quintile of the distribution are more than three

times as likely to report a recent unemployment spell compared to those in the top quintile. There is also evidence in the table that disability shocks play a role in generating variation in residual income across households. We exploit these patterns below to generate plausibly exogenous instruments for household-level residual income.

4.2. IV Estimates based on Unemployment Shocks

Since unemployment spells are typically associated with significant income losses,²⁷ even accounting for unemployment insurance benefits, the literature has focused on unemployment as a potentially important driver of mortgage default. However, due to data limitations, previous studies have been forced to use aggregate unemployment (at the state, MSA, or county aggregation level),²⁸ and have found only weak correlations between unemployment and mortgage default.²⁹ The PSID asks detailed questions about employment and unemployment spells, which allow us to construct household-level unemployment shocks.

We define an unemployment shock as corresponding to a household head who reports being unemployed at the time of the PSID survey or who reports a spell of unemployment over the 12 months prior to the survey date. We also construct a spousal unemployment shock using the same definition. However, there are potential concerns that these shocks are not entirely exogenous to the household and more importantly to the decision of the household to default on its mortgage. For example, some unemployment spells are voluntary and initiated by the employee. If the same underlying factors (unobservable to us) that drive households to leave their jobs also drive them to default on their loans, then we would mistakenly attribute income loss from unemployment shocks to be driving increased mortgage default.

We address this potential endogeneity bias in a few different ways. First, we isolate job

²⁷See Saporta-Eksten (2014) and papers cited therein.

²⁸For example, Deng et al. (2000) include state-level unemployment rates in their mortgage default and prepayment models, and Demyanyk and Van Hemert (2011) include county-level unemployment rates as controls in their analysis of subprime mortgage defaults.

²⁹See Gyourko and Tracy (2014) for a detailed discussion of how using aggregate unemployment rates as proxies for individual shocks could mask the true relationship between unemployment and mortgage default.

losses due to involuntary separations, which are defined in the PSID to be either plant closures, strikes/lockouts, or layoffs. Involuntary unemployment spells are less likely to suffer from endogeneity bias since they are events that occur outside of the purview of the household. However, there could still be a concern that there is some unobserved heterogeneity across households that drives both involuntary job separations and default decisions. For example, perhaps impatient households who heavily discount the future might be more likely to default on debt and may also be more likely to be fired due to poor work habits. If this unobserved factor does not vary over time, then the panel dimension of the PSID allows us to address the issue. To do so we construct indicator variables based on the number of prior unemployment spells over the seven PSID surveys spanning 1994-2005, and include these variables in our control set (“Job Loss FEs”). If there is a time-invariant factor that causes both an increased propensity to experience job loss and an increased propensity to default, this additional control should take care of the issue. In addition to these measures, we follow the labor economics literature and construct a completely different employment-based instrument based on national-level industry employment flows and state-level industry shares, which we discuss below in Section 4.3.

In addition to the need to instrument for residual income, we also need to address the potential endogeneity of the LTV variable. The LTV variable is the ratio of the self-reported remaining mortgage balance to the self-reported value of the house, and thus, corresponds to the household’s estimate of its current equity position in the property. The total size of the mortgage and downpayment fraction at the time of origination are decisions made by the household.³⁰ These decisions may be correlated with some unobserved factors that are also correlated with the default decision, resulting in endogeneity bias. The example of heterogeneity in impatience discussed above is also applicable here as more impatient households are more likely to choose lower downpayments (i.e. higher LTV ratios at origination) and are also more likely to default on their debts. To address this issue we construct an instrument

³⁰In addition, the decision on how much mortgage debt to pay down over time is also under the household’s control.

for the LTV ratio that corresponds to the cumulative growth in the state-level housing price index from the year of home purchase to the current interview year. The PSID provides the state in which the household resides, and using the panel aspect of the PSID, we can identify the exact year when the household purchased its current home. Cumulative house price appreciation at the state-level is plausibly exogenous to the household’s mortgage default decision, but should be correlated with the LTV ratio variable.³¹

Table 6 displays the results of our instrumental variables analysis. Column (1) in the table corresponds to the simple OLS estimates, which are replicated from Table 5 (column (2)) for ease of comparison. Column (2) in Table 6 displays the estimation results when we use the head and spouse unemployment shocks to instrument for income loss and cumulative house price appreciation to instrument for LTV ratios (all columns in the table use the same instrument for LTV ratios). There is a sizeable increase in the magnitude of the coefficient associated with residual income in the IV specification compared to the OLS regression. Households that suffer a 10 percent loss in residual income that is caused by an unemployment shock are approximately 1.5 percentage points more likely to default on their mortgages. This effect is an order of magnitude larger than the OLS estimate displayed in column (1). The effect of the contemporaneous level of home equity on default also increases significantly when house price appreciation is used as an instrument, as the coefficient approximately doubles (from 0.078 to 0.189).

The increase in the estimated impact of residual income on mortgage default in the IV specification is both plausible and consistent with economic theory. The permanent income hypothesis predicts that permanent (or persistent) shocks to income have a significantly larger effect on consumption decisions compared to more transitory income shocks.³² The IV specification isolates losses in residual income due to unemployment shocks, which can be significant life events and thus, are likely to have persistent effects. In other words, the IV

³¹The assumption that is required for cumulative house price growth at the state-level to be a valid instrument is that the timing of home purchase and the choice of state to reside in is exogenous with respect to the mortgage default decision.

³²See Saporta-Eksten (2014) and Jarosch (2014) for job loss studies.

specification is likely isolating more permanent income shocks, which theory predicts should lead to a much larger impact on the propensity to default. This is exactly what we find.

The first stage results are reported in Table 7. The table shows the first stage results for all of the IV specifications in Table 6 (the column numbers correspond to the column numbers in Table 6). There are two first stage regressions since we are instrumenting for two endogenous variables, LTV and residual income, respectively. Panel A displays the estimates for the LTV ratio regression and Panel B shows the estimates for the residual income regression. Column (2) of Panel A shows that cumulative (state-level) house price appreciation is a very strong instrument for home equity, while Column (2) of Panel B shows that unemployment is an extremely strong instrument for residual income. Job loss experienced by the head of the household results in a 32 log point decline in residual income, whereas spousal job loss is associated with a 23 log point decline.

Column (3) in Table 6 displays results from the reduced form specification in which the default indicator is regressed directly on unemployment. Unemployment spells experienced by both the head of the household and the spouse are positively correlated with the incidence of default. Unemployed heads are approximately 5 percentage points more likely to default than employed heads, while households in which both the head and spouse experience an unemployment spell are more than 8 percentage points more likely to default. Head of household job loss has an equivalent effect on the likelihood of default as a 35 $(.053/.15)$ percent decline in home equity, and if both the head and spouse lose their job this has the same impact on default as a 55 $((.053+.03)/.15)$ percent decline in home equity. These results contradict many previous studies that have used aggregate data and found only a weak correlation between default rates and unemployment rates, and confirms the claim made by Gyourko and Tracy (2014) that using aggregate unemployment rates as a proxy for individual unemployment shocks can result in a serious attenuation bias.³³

In column (4) of Table 6 we modify the instrument set by substituting the unemployment

³³In fact, in unreported results we substituted the individual-level unemployment variables for state-level unemployment rates and found zero correlation between state-level unemployment and default.

variables with indicators of involuntary unemployment spells for both the head and spouse. In addition, we include a set of indicator variables corresponding to the number of prior involuntary unemployment spells to control for unobserved time invariant characteristics that may link job loss propensity to default. The coefficient associated with residual income remains negative and actually increases somewhat in magnitude compared to the coefficient obtained using the overall unemployment shock IV reported in column (2). The estimate is statistically significant at the 5 percent level, as we lose some power due to the smaller number of involuntary employment spells compared to overall employment spells in the data. Column (5) displays the reduced form estimates, which are very similar to those associated with the overall unemployment variables reported in column (3). The fact that our IV estimates remain quite similar when we focus on only involuntary unemployment spells and control for previous such spells assuages many of our concerns about endogeneity bias.

4.3. IV Estimates based on Bartik and Disability Shocks

While using involuntary unemployment spells and including prior unemployment episodes in the set of controls assuages many of our endogeneity concerns, it does not completely alleviate them. In this section we construct an alternative set of instruments for residual income, which we argue is less susceptible to endogeneity bias. Our alternative instrument set has two primary components, which are both motivated by previous research.

The first component in our instrument set is a disability shock that we construct based on the work of Low and Pistaferri (2015) (hereafter referred to as LP). In the course of calibrating a life-cycle model with disability insurance, LP estimates the effect of having a health disability that limits the ability to work on the decision to obtain employment and on the level of offered wages (conditional on employment) using PSID data. LP finds that disabilities significantly decrease the probability of working and also significantly decrease offered wages for households that choose to work.³⁴ These findings imply that health disabilities

³⁴Specifically, LP finds that a moderate disability decreases the likelihood of working by 27 percentage

result in significant declines in income. In addition, LP argues that health disabilities are exogenous shifters of employment probabilities and wages (and thus income), citing several studies to support such a position.³⁵

We largely follow the methods of Low and Pistaferri (2015) in identifying a household in which the head or the spouse has suffered a disability. Specifically, we use information from the following three PSID survey questions posed to both household heads and spouses: (i) *Do you have any physical or nervous condition that limits the type of work or the amount of work you can do?* If the respondent answers “Yes” the interviewer asks: (ii) *Does this condition keep you from doing some types of work?* where the possible answers are: “Yes”, “No”, or “Can do nothing”. Respondents that answer either “Yes” or “No” are then asked: (iii) *For work you can do, how much does it limit the amount of work you can do?* where the possible answers are given by: “A lot”, “Somewhat”, “Just a little”, or “Not at all”. If the answer to question (i) is “No” or the answer to question (iii) is “Not at all” then we assume that the respondent does not have a disability that limits her ability to work. We assume that the respondent has a severe disability if her response to question (i) is “Yes” and her response to question (ii) is “Can do nothing” or her response to question (iii) is “A lot”. We assume that the remainder of respondents have a moderate disability (i.e. they answer “Yes” to question (i) and either “Somewhat” or “Just a little” to question (iii)).

Using these definitions, we then construct dummy variables to identify a *change* in either the head’s or spouse’s disability status. The first set of dummy variables that we focus on indicate whether the head or the spouse has suffered any disability (moderate or severe) since the previous survey. The second set indicates whether the head or the spouse has suffered a severe disability since the previous survey. Summary statistics for the disability variables are displayed in Table 2.

points while a severe disability decreases the probability by 74 percentage points. Conditional on working, LP finds that a moderate disability reduces wages by 6 percentage points while a severe disability reduces wages by 18 percentage points.

³⁵These include (among others) Smith (2004) who found that income does not affect health status as long as one controls for education (which we do in our analysis), and Adda et al. (2009) who found that income innovations do not impact health status.

In addition to the disability shocks, we construct an additional instrument that is meant to isolate purely *exogenous* changes in employment status that influence residual income. The instrument is based on aggregate sectoral employment flows at the national level and industry shares at the state-level. Specifically, we use data from the Bureau of Labor Statistics (BLS) to construct the following Bartik state-level employment shock:

$$Bartik_{it} = \sum_j share_{i,j,t-k}^{empl} * \Delta empl_{j,t-k,t} \quad (3)$$

where i indexes the state, j indexes the 1-digit NAICs industry code, t indexes the current survey year (2009, 2011, or 2013), and k indexes the number of years over which the growth rates are computed. The Bartik shock is constructed by interacting national-level industry growth in employment, $\Delta empl_{j,t-k,t}$, with the state-level initial composition of employment in industry j , $share_{i,j,t-k}^{empl}$. Calculation of the national-level industry growth rates is performed using data from all states excluding i . Bartik shocks are used frequently in the labor literature to instrument for local aggregate demand shocks. The idea behind the Bartik shock is that employment in all states in all industries is affected by national industry-level employment movements, but movements in a given industry have a higher impact in a state where the industry employs a greater share of the population. For example, the Bartik shock calculation for Florida would place a lower weight on national employment changes in the financial activities industries than the Bartik shock calculation for New York. In our context, the Bartik variable is a natural choice for an instrument as state-level, labor demand shocks are unlikely to be correlated with individual default decisions except through their impact on the likelihood of job loss and, in turn, income loss.

We construct the Bartik variable over a two-year horizon (i.e. $k = 2$)³⁶ to maintain consistency with the biennial frequency of the PSID. Our measures of employment by industry and state are taken from the BLS. In particular, we use State and Area Employment, Hours,

³⁶We also estimated specifications using Bartik shocks constructed over a four-year horizon and found similar results.

and Earnings from the CES.

Column (6) in Table 6 displays the results when we instrument for income loss using the disability shock and the Bartik employment shocks. The coefficient estimate is -0.109 (statistically significant at the 5 percent level), which is similar in magnitude to the estimates we obtained using unemployment spells in columns (2) and (4). The first stage results displayed in Table 7, (column (6) in Panel B) show that the disability indicators are strong predictors of income loss, which is consistent with the findings in Low and Pistaferri (2015). The reduced form specification results reported in column (7) of Table 6 show that spousal disabilities have a similar impact on the default rate as unemployment, and head disabilities have a positive, but insignificant impact, on default. In column (8) of the table we consider only severe disability shocks in the instrument set, and find very similar estimates. There are some notable differences in the first stage estimates, as the severe disability shock variables are more strongly correlated with residual income.

IV diagnostics are displayed in the bottom two rows of Table 6. The disability and Bartik variables all exhibit strong first-stage results and easily pass weak IV tests. Since we have multiple instruments, we can also conduct over identification tests (the premise of these tests is to use each instrument one-at-a-time, and then check to see if the residuals are correlated with the excluded instrument) and in every specification, the instruments pass.³⁷

In Online Appendix H we show the results from several robustness exercises. For example, we consider the inclusion of state fixed effects in the IV specifications, we consider some alternative ways of specifying the instrumental variables (including different timing conventions for the Bartik shocks and alternative disability definitions), and we also substitute the continuous LTV variable with a dummy variable for negative equity ($LTV > 100$). The results are consistent across these permutations. Furthermore, in Online Appendix F, we consider a similar set of IV specifications for negative income growth shocks rather than

³⁷There is of course no explicit way of testing the exclusion restriction. The over-identification tests come the closest and are often portrayed as tests of the exclusion restriction, but they are not direct tests because they assume that at least one of the instruments is valid.

residual income and show that the results are robust.

5. Implications for Research and Policy

The foregoing analysis leads to five significant findings that we believe are relevant for both researchers and policy makers:

Finding 1: *Households with low residual income default more than households with high residual income.*

The cross tabs in Table 3 show that overall “can’t pay” borrowers are much more likely to default than “can pay” borrowers – approximately 7 times for the sample as a whole. The regression results in Section 4 confirm that this central finding still holds when we control for LTV ratios, a host of demographic and mortgage characteristics, as well as potential endogeneity bias. The top left panel of Figure 3 shows the predicted default probabilities across different LTV ratios for two different households using the coefficients from Column (6) in Table 5. The households are identical except that one has residual income of \$500 per month and the other has \$5000. The top right panel shows that the ratio of predicted default rates for low versus high residual income households ranges from above 8 for households with low LTV ratios to about 3 for households with very high LTV ratios.

Finding 2: *Households with high residual income are more sensitive to changes in LTV ratios than households with low residual income. Households with low LTV ratios are more sensitive to changes in residual income than households with high LTV ratios.*

The top left panel of Figure 3 shows the effect of LTV on the default hazard for high versus low residual income borrowers. For high residual income borrowers, increases in LTV

ratios have a much larger effect than for low residual income borrowers: an increase in the LTV ratio from 50 percent to 150 percent only increases the default hazard for low residual income borrowers by a factor of 3, whereas for high residual income borrowers, the default hazard increases by a factor of more than 6. It is important to stress that this does not mean that high residual income households default more, but rather reflects the fact that high residual income borrowers with low LTV ratios almost never default. Intuitively, negative equity is the only reason why a high residual income household would default, whereas for a low residual income household, negative equity only compounds an existing problem. As the bottom right panel of the figure shows, similar logic applies to the effect of residual income conditional on LTV ratios.

Finding 3: *Most households that suffer a shock will not default.*

Figure 3 shows a striking fact mentioned earlier in the paper: most borrowers do not default. Even for LTV ratios as high as 150 percent, the empirical estimates predict that only 22 percent of households with low residual income default, meaning that 78 percent continue making their mortgage payments.

Finding 4: *Many households that default can afford their mortgage payments.*

Table 3 shows that almost 38 percent of defaulters could afford their mortgage payments without any reduction in consumption relative to pre-default levels. As discussed above, this fits with the popular idea that a significant fraction of defaults were “strategic” in the sense that the household could continue to make payments without imminent sacrifice. We find that another 33 percent of households would have to cut consumption but could still maintain a subsistence level. In other words, slightly less than 30 percent of borrowers appear to default to avoid destitution.

Finding 5: *Very few households that can afford their mortgage payments default.*

Table 3 shows that the overwhelming majority of households that could afford to make their mortgage payments did, in fact, make their payments. If we focus on households with low equity who had enough income to make their mortgage payments without reducing consumption below pre-default levels we find that almost 96 percent are not in default.

Findings 4 and 5 at first blush appear to contradict each other. According to Fact 4, strategic default appears to be relatively frequent, but according to Fact 5, it seems like it is not a significant problem at all. The difference between the two findings has to do with denominators. The numerator in both instances is identical: the number of “can pay” defaulters. However the denominator corresponding to Fact 4 is the total number of defaulters in the sample, which is a small number, whereas the denominator corresponding to Fact 5 is the number of “can pay” households in general (both “can pay” defaulters and non-defaulters), which is a much larger number (about an order of magnitude larger than the number of defaulters). This difference between the number of defaulters versus the number of “can pay” households reconciles the apparent contradiction between Facts 4 and 5.

5.1. Implications for Research

Dating all the way back to Herzog and Earley (1970), researchers have studied the determinants of default using micro data. Although there is strong anecdotal evidence showing that unemployment and income shocks are important (see Cutts and Merrill (2008), for example), the evidence from micro-level data does not bear this out. As Gyourko and Tracy (2014) write, “Empirical models of mortgage default typically find that the influence of unemployment is negligible.” Gyourko and Tracy (2014) used simulation methods to argue that the weak relationship between income and employment and default in the data resulted from attenuation bias related to the use of aggregate (i.e. county-level and MSA-level)

unemployment indicators. Finding 1 of this paper provides direct evidence of the role of household-level income shocks in the default decision, and the analysis in Section 4 shows that unemployment and disability shocks are important drivers of the income shocks that generate mortgage defaults. Thus, our results confirm those of Gyourko and Tracy (2014)³⁸ and also provide some support for the double trigger model of mortgage default described in Foote et al. (2008) and elsewhere.³⁹

Our results also provide an explanation for recent empirical findings on the financial crisis, which documented an increase, relative to the pre-crisis period, in the the share of defaults from high-income households. Ferreira and Gyourko (2015) argue that the foreclosure crisis started out in subprime areas but, “quickly morphed into a much bigger and broader event dominated by prime borrowers losing their homes.” Adelino et al. (2016) find that, “for the 2003 mortgage cohort, the top quintile of the income distribution constituted only 13% of mortgage dollars in delinquency three years later, whereas for the 2006 cohort, the top income quintile made up 23% of the delinquencies three years out.” The link to the results in this paper can be seen in Finding 2, which shows that the relative effect of residual income on default weakens as LTV ratios rise. In other words, as house prices fall, defaults increase more for households with high residual income than for households with low residual income. Researchers going back as far as Mincer (1991) have shown that the hazard of job loss is much higher for low education workers who would be concentrated in low income areas, meaning that more households would have low residual income in low income areas in general.⁴⁰ As house prices fall, Finding 2 would imply that the default rate should therefore grow relatively more in high income areas, consistent with Adelino et al. (2016) and Ferreira and Gyourko (2015).

³⁸Indeed, consistent with previous findings in the literature, when we use aggregated unemployment instead of household-level unemployment shocks, we find a very weak connection between unemployment and default, which confirms the role of attenuation bias.

³⁹While our findings support one of the double-trigger model’s main predictions that household-level cash flow shocks are an important driver of default, they are not consistent with the prediction that such shocks are a necessary condition for default. Our findings on “can pay” defaulters show that a significant fraction of defaults appear to be driven by strategic motivations.

⁴⁰See Elsby et al. (2010) for more recent evidence.

As mentioned above, the question of the importance of strategic default has occupied both researchers and policy makers. For example, Guiso et al. (2013) conduct a survey and ask respondents, “Of the people you know who have defaulted on their mortgage, how many do you think walked away even if they could afford to pay the monthly mortgage?” Based on the answers, they conclude that in 2010, approximately 35 percent of defaults were strategic in nature. Finding 4 of our paper is that almost 38 percent of households in default could make their mortgage payments without reducing consumption. In that sense, our results confirm the importance of strategic default. However, it is important to stress again that according to Finding 5, just because strategic motivations appear to play an important role among defaulters, does not mean that strategic default was extremely prevalent. Table 3 shows that even among borrowers with low equity, almost 96 percent of borrowers who could afford their mortgage payments continued making them. Based on these findings, an important question for researchers to ask in future surveys could be, “Of the people you know who could afford to make their monthly mortgage payment, how many do you think walked away?”

Finally, one of the great puzzles in the mortgage literature is the low frequency of defaults in the data. In standard models, default is rampant at high LTV ratios. For example, Kau et al. (1993) using a contingent claims approach to modeling mortgage default and prepayment, calculate in their baseline parametrization, that 100 percent of borrowers with LTV ratios above 115 percent will default. Yet in the typical micro-level dataset, the share of borrowers who default with LTV ratios of at least 115 is in the single digits (see, for example, Deng et al. (2000)). The double-trigger theory proposed that once household financial stress was taken into consideration, one could reconcile the data and theory. However, the evidence is mixed. As already noted, Finding 1 shows that households with low residual income are more likely to default. Table 3 shows that “can’t pay” households with low equity are 30 times more likely to default than “can pay” households with high equity. But, the actual default rate, even for this high risk group is still only 20 percent or one-fifth of the default

rate generated by the model of Kau et al. (1993).

The open question for research is why many borrowers who face significant hardship with few options choose to continue making their mortgage payments, while other borrowers choose not to. One possibility along these lines is that many households have a strong moral aversion to defaulting on debt, especially mortgage debt. For example, Guiso et al. (2013) find evidence in survey data that many households consider strategic mortgage default to be an immoral practice, and as a result, are much less likely to engage in such behavior. Approximately 82 percent of the households in their survey report having a moral aversion to mortgage default, so it is conceivable that a large majority of our “can’t pay” households hold similar reservations against default.

Another possibility is that many households have a strong attachment to their homes, and thus, go to extreme measures to avoid default and foreclosure by drawing down their illiquid assets such as vehicles and retirement accounts.⁴¹ Finally, optimistic expectations of future house prices may play an important role. The theoretical literature tells us that house price expectations are an important determinant in the decision to default. While it’s hard to believe that optimistic expectations alone could explain why 80 percent of households with virtually no income or assets continue to make payments, such an explanation is at least consistent with the evidence on the sensitivity of default decisions to equity displayed in Figure 3.

5.2. Implications for Policy

During the foreclosure wave that swept the country before, during, and after the Great Recession, there were fierce debates among academics and policy makers about the best policies to mitigate foreclosures and prevent a future crisis. In this section we discuss how our empirical findings are relevant to this debate.

First, the signature foreclosure relief policies of the Obama Administration, the Home

⁴¹See Laufer (2012) and Schelkle (2011) for theoretical models which allow for strong attachment to homes or aversion to renting.

Affordable Modification Program (HAMP) and the Home Affordable Refinance Program (HARP), focused on reducing monthly mortgage payments. Finding 1 above shows that such an emphasis was sensible. Our regression estimates from Section 4 show that residual income has economically large effects on repayment behavior and that policies that increase residual income should be effective in reducing defaults.⁴² Policy makers then extended the logic of HAMP and HARP to new mortgage originations. For example, the Qualified Residential Mortgage (QRM) and Qualified Mortgage (QM) definitions created by the Dodd-Frank Act of 2010 emphasized that a borrower should only be allowed to obtain a mortgage if he or she has the “ability-to-repay”, defined as a mortgage payment that cannot exceed 43 percent of monthly income. One might think that Finding 1 confirms that QM and QRM should reduce defaults and foreclosures going forward. However, in Online Appendix I, we show that residual income has strong predictive power contemporaneously but residual income at origination has little power in predicting *future* defaults. The lack of predictive power of income at origination is well known in the mortgage industry (see the discussion in Foote et al. (2010)) and reflects the high volatility of household labor income. As discussed above, borrowers with low residual income were very likely to have suffered a recent labor income shock.

Findings 3 and 4 both imply major challenges in implementing foreclosure prevention policies and illustrate, at least partly, why the policies implemented during the crisis did not achieve as much as hoped. The central goal of a HAMP modification was to reduce a household’s monthly payment to an affordable level. Finding 4, however, shows that many borrowers in delinquency already have affordable payments. In Figure 2 we show this more clearly by calculating for each household in default the required mortgage payment reduction (in percentage terms) that is necessary to make the payment “affordable.” We consider three definitions of affordability in the figure. The first two correspond to the definitions in Table 3: the mortgage payment is affordable if the household can maintain its consumption level

⁴²Other research, including Fuster and Willen (2012) and Hsu et al. (2014) has made a similar point.

from the previous year, and the payment is affordable if the household can consume the VA subsistence level, respectively. The third definition considered in Figure 2 is that the mortgage payment must satisfy the QRM “ability-to-repay” standard. The left-most column in the figure shows that the payment reduction required to achieve affordability is zero for a large fraction of households in default. In other words, many delinquent households have too much income to qualify for a loan modification. Indeed, for the two most realistic definitions of affordability, QRM and VA subsistence, more than half of the delinquent borrowers already have “affordable” mortgages. At the other extreme, the right-most column of Figure 2 shows that many delinquent borrowers (between 15 and 40 percent depending on the affordability definition) require a 100 percent reduction (or more) in their monthly payment to achieve affordability. Complete payment reductions obviously present a problem for loss mitigation policies since the logic behind them is that the lender will recover more than through a foreclosure. Obviously, this would not hold if the borrower pays nothing. Figure 2 shows that the inframarginal borrowers are skewed toward small payment reductions which, according to our regression estimates, would only have small effects on their likelihood of repayment.

One idea that has been popular during and since the foreclosure crisis is to preemptively offer loan modifications to borrowers who are at increased risk of default. The logic is that such modifications are “win-win” because in addition to the obvious benefit to the borrower, the lender gains as the expected cost of reduced payments is believed to be less than the expected losses from foreclosures. Finding 3 illustrates why preemptive payment reductions were never popular with lenders, even when applied to borrowers at elevated risk of default. If we look at Table 3, we can find a group of borrowers with highly elevated default probabilities: “Can’t pay” borrowers with high LTV ratios are approximately 30 times more likely to default than “can pay” borrowers with low LTVs ratios. Finding 3, however, is that 80 percent of these high risk borrowers do not default. Thus, a loss mitigation plan that targeted only these high risk borrowers, would see 80 cents of every dollar of payment reduction go to a household that likely would have made its mortgage payment anyway. In addition, according

to the table, if payment reduction takes the residual income of “can’t pay” borrowers above the VA subsistence level (i.e. into column (2)), the default rate is approximately 9.5 percent. In other words, if lenders give payment reductions to 100 borrowers, for 80 it is basically a windfall as they likely would not have defaulted in the absence of the reduction, and only 10 defaults would actually be prevented. This suggests that the cost of a foreclosure must be 8 times higher than the payment reduction to make it profitable to the lender.

At the peak of the crisis, some commentators felt that policy makers needed to respond to the problem of strategic default. Finding 4 shows that many defaulters could make their mortgage payments but Finding 5 shows that at no point did it reach the epidemic proportions that some feared: the overwhelming majority of borrowers who could make their mortgage payments did make them as did many of those who, as far as we can tell, could not afford their payments.

6. Conclusion

To design efficient foreclosure mitigation policies, it is necessary to understand the sources of mortgage default. While there is broad agreement that a number of factors may potentially contribute to default, including housing equity, employment, and liquidity, due to data limitations, it has been impossible to directly test the relative importance of these factors. Using new PSID data, which includes detailed information on mortgagors’ budget constraints such as income and employment status, consumption expenditures, and assets, we measure the importance of ability to pay for the mortgage default decision and assess the significance of strategic default.

In our empirical analysis, we find a significant role for both factors. Among households in default, strategic motives appear to play an important role, as approximately 38 percent of households in default appear to have the ability to pay their mortgage without reducing their consumption below their pre-default levels. Furthermore, using an instrumental variables

design, we find that housing equity is an important predictor of mortgage default, holding residual income constant, which is also consistent with strategic motives playing an important role. At the same time we find an important role played by household-level income shocks including unemployment and disability shocks. Our IV estimates indicate that a 10 percent decline in residual income raises the probability of default by between 1.1 and 2.5 percentage points, depending on the particular instrument set.

An additional contribution of this paper is in showing the reluctance of most mortgage borrowers to default on their loans, even those borrowers who are experiencing significant financial distress. We show that approximately 80 percent of households that need to cut their consumption to subsistence levels in order to make their mortgage payments (“cant pay” borrowers) continue to make their mortgage payments. Furthermore, we find that nearly 96 percent of low equity borrowers with the ability to pay remain current on their loans. These findings provide a simple explanation for why lenders rarely renegotiate preemptively with borrowers who have extremely limited resources to pay their mortgage, since the majority of these borrowers will continue to pay.

While the PSID data have allowed us to analyze strategic default and the importance of income shocks for the mortgage default decision in more detail than prior studies in the literature, a significant amount of work remains to be done. We leave it to future researchers to develop higher-frequency asset, mortgage default, and consumption datasets, which will ultimately allow the profession to better understand default behavior.

References

- Adda, J., H.-M. Gaudecker, and J. Banks (2009). The impact of income shocks on health: evidence from cohort data. *Journal of the European Economic Association* 7(6), 1361–1399.
- Adelino, M., K. Gerardi, and P. Willen (2013). Why don't lenders renegotiate more home mortgages? redefaults, self-cures and securitization. *Journal of Monetary Economics* 60(7), 835–853.
- Adelino, M., A. Schoar, and F. Severino (2016). Loan originations and defaults in the mortgage crisis: The role of the middle class. *Review of Financial Studies*, hhw018.
- Benítez-Silva, H., S. Eren, F. Heiland, and S. Jiménez-Martín (2008). How well do individuals predict the selling prices of their homes? Levy Economics Institute Working Paper.
- Bhutta, N., J. Dokko, and H. Shan (2011). Consumer ruthlessness and strategic default during the 2007-2009 housing bust. Federal Reserve Board of Governors Working Paper.
- Bradley, M. G., A. C. Cutts, and W. Liu (2015). Strategic mortgage default: The effect of neighborhood factors. *Real Estate Economics* 43(2), 271–299.
- Campbell, J. and J. Cocco (2011). A model of mortgage default. Technical report, National Bureau of Economic Research.
- Chatterjee, S. and B. Eyigungor (2009). *Foreclosures and House Price Dynamics: A Quantitative Analysis of the Mortgage Crisis and the Foreclosure Prevention Policy*. Federal Reserve Bank of Philadelphia.
- Chatterjee, S. and B. Eyigungor (2011). A quantitative analysis of the us housing and mortgage markets and the foreclosure crisis. Federal Reserve Bank of Philadelphia Working Paper.
- Corbae, D. and E. Quintin (2009). Mortgage innovation and the foreclosure boom. Unpublished paper, University of Texas at Austin.
- Cutts, A. C. and W. Merrill (2008). Interventions in mortgage default: Policies and practices to prevent home loss and lower costs. *Borrowing to live: Consumer and mortgage credit revisited*, 203–254.
- Demyanyk, Y. and O. Van Hemert (2011). Understanding the subprime mortgage crisis. *Review of financial Studies* 24(6), 1848–1880.

- Demyanyk, Y. S., R. S. Kojien, and O. Van Hemert (2011). Determinants and consequences of mortgage default. *Available at SSRN 1706844*.
- Deng, Y., J. M. Quigley, and R. Order (2000). Mortgage terminations, heterogeneity and the exercise of mortgage options. *Econometrica* 68(2), 275–307.
- Deng, Y., J. M. Quigley, R. Van Order, and F. Mac (1996). Mortgage default and low down-payment loans: the costs of public subsidy. *Regional Science and Urban Economics* 26(3), 263–285.
- Elsby, M. W., B. Hobijn, A. Sahin, L. F. Katz, and R. Shimer (2010). The labor market in the great recession/comments and discussion. *Brookings Papers on Economic Activity*, 1.
- Elul, R., N. Souleles, S. Chomsisengphet, D. Glennon, and R. Hunt (2010). What triggers mortgage default? FRB of Philadelphia Working Paper No. 10-13.
- Experian and Oliver Wyman (2009). Understanding strategic default in mortgages part i. *Experian-Oliver Wyman Market Intelligence Report 2009*. Topical Report Series.
- Ferreira, F. and J. Gyourko (2015, June). A new look at the u.s. foreclosure crisis: Panel data evidence of prime and subprime borrowers from 1997 to 2012. Working Paper 21261, National Bureau of Economic Research.
- Foote, C., K. Gerardi, L. Goette, and P. Willen (2010). Reducing foreclosures: No easy answers. In *NBER Macroeconomics Annual 2009, Volume 24*, pp. 89–138. University of Chicago Press.
- Foote, C. L., K. Gerardi, and P. S. Willen (2008). Negative equity and foreclosure: Theory and evidence. *Journal of Urban Economics* 64(2), 234–245.
- Fuster, A. and P. Willen (2012). Payment size, negative equity, and mortgage default. *FRB of New York Staff Report* (582).
- Garriga, C. and D. Schlagenhauf (2009). Home equity, foreclosures, and bailouts. Manuscript, Federal Reserve Bank of St. Louis.
- Guiso, L., P. Sapienza, and L. Zingales (2013). The determinants of attitudes toward strategic default on mortgages. *The Journal of Finance* 68(4), 1473–1515.
- Gyourko, J. and J. Tracy (2014). Reconciling theory and empirics on the role of unemployment in mortgage default. *Journal of Urban Economics* 80, 87–96.
- Hedlund, A. (2011). Illiquidity and foreclosures in a directed search model of the housing market. Available at SSRN 1932908.

- Herzog, J. P. and J. S. Earley (1970). *Home Mortgage Delinquency and Foreclosure*. Cambridge, MA: National Bureau of Economic Research.
- Hsu, J. W., D. A. Matsa, and B. T. Melzer (2014, July). Positive externalities of social insurance: Unemployment insurance and consumer credit. Working Paper 20353, National Bureau of Economic Research.
- Hurd, M. D. and S. Rohwedder (2010). Effects of the financial crisis and great recession on american households. Technical report, National Bureau of Economic Research.
- Jarosch, G. (2014). Searching for job security and the consequences of job loss. *V University of Chicago, Job Market Paper*.
- Kau, J. B., D. Keenan, and T. Kim (1993). Transaction costs, suboptimal termination and default probabilities. *Real Estate Economics* 21(3), 247–263.
- Laufer, S. (2012). Equity extraction and mortgage default. Technical report, Working Paper. New York University.
- Low, H. and L. Pistaferri (2015). Disability insurance and the dynamics of the incentive insurance trade-off. *American Economic Review* 105, 2986–3029.
- Mian, A. and A. Sufi (2009). The consequences of mortgage credit expansion: Evidence from the us mortgage default crisis. *The Quarterly Journal of Economics* 124(4), 1449–1496.
- Mincer, J. (1991, September). Education and unemployment. Working Paper 3838, National Bureau of Economic Research.
- Saporta-Eksten, I. (2014). Job loss, consumption and unemployment insurance. *Unpublished manuscript*.
- Schelkle, T. (2011). Mortgage default during the us mortgage crisis. Working paper, London School of Economics.
- Smith, J. P. (2004). Unraveling the ses: health connection. *Population and development review* 30, 108–132.

Table 1: Comparing PSID and American Housing Survey Mortgage Data: 2009-2013

	2009		2011		2013	
	PSID	AHS	PSID	AHS	PSID	AHS
<i>Median:</i>						
Principal Remaining	120,000	106,909	120,000	120,000	130,000	121,324
Monthly Mortgage Payment	1,100	878	1,100	1015	1,085	997
Mortgage Interest Rate	5	6	5	5.3	4	4.5
Years Remaining on First Mortgage	24	23	22	22	21	21
LTV Ratio (1st Mort.)	65	63	68	71	69	70
<i>Fraction with:</i>						
Second Mortgage	0.21	0.21	0.18	0.13	0.15	0.10
ARM	0.09	0.06	0.08	0.07	0.08	0.06
Observations	2,640	51,969	2,462	49,734	2,295	47,741

Notes: This table compares mortgage summary statistics in the PSID and AHS surveys. AHS 2009 National Statistics are taken from Tables 3.13 and 3.15. AHS 2011 and 2013 National Statistics are taken from Table C-14A-OO. The PSID sample consists of household heads who are mortgagors, aged 24–65, and labor force participants (including those who are disabled) with combined loan to value ratios less than 250 percent in the 2009, 2011, and 2013 surveys. Households that report they are in default in a given year are subsequently dropped to avoid double counting.

Table 2: Sample Summary Statistics for PSID Household Heads.

	All Mortgagor Households				Households in Default			
	Mean	10 th Perc.	Median	90 th Perc.	Mean	10 th Perc.	Median	90 th Perc.
(A) Income and Consumption:								
Real Income	120,000	40,000	94,000	210,000	71,000	21,000	60,000	130,000
Real Income After Tax	87,000	34,000	72,000	150,000	58,000	23,000	53,000	110,000
Real Total Consumption	65,000	29,000	56,000	110,000	58,000	27,000	50,000	100,000
% Change in Income (2yr Period)	8	-31	6	59	-4	-55	-7	58
15% Income Drop or more (d)	0.19	0	0	1	0.42	0	0	1
30% Income Drop or more (d)	0.102	0	0	1	0.241	0	0	1
Log Residual Income	11.19	10.252	11.247	12.08	10.51	8.909	10.64	11.564
Residual Income	100,000	28,000	77,000	180,000	50,000	7,400	42,000	110,000
(B) Demographics:								
Married (d)	0.70	0	1	1	0.55	0	1	1
College Grad+ Education (d)	0.45	0	0	1	0.23	0	0	1
Less than High School (d)	0.06	0	0	0	0.16	0	0	1
Age	46.02	31	46	60	44.97	31	46	59
(C) LTV and Delinquency:								
Loan to Value Ratio	0.68	0.25	0.69	1	1.01	0.58	0.95	1.68
Default (60+ Days Late) (d)	0.03	0	0	0	1.00	1	1	1
Default (90+ Days Late) (d)	0.02	0	0	0	0.63	0	1	1
(D) Wealth								
Liquid Assets	26,000	0	6,214	57,000	4,630	0	518	5,429
Illiquid Assets	150,000	3,107	27,000	350,000	36,000	0	7,601	57,000
(E) Employment and Disabilities:								
Unemployed Head Last Year (d)	0.07	0	0	0	0.22	0	0	1
Unemployed Spouse Last Year (d)	0.04	0	0	0	0.10	0	0	1
Unemployed Head or Spouse Last Year (d)	0.11	0	0	1	0.28	0	0	1
Involuntary Unemployment, Head (d)	0.027	0	0	0	0.104	0	0	1
Involuntary Unemployment, Spouse (d)	0.007	0	0	0	0.006	0	0	0
Transition into Disability, Head (d)	0.040	0	0	0	0.060	0	0	0
Transition into Disability, Sp. (d)	0.036	0	0	0	0.056	0	0	0
Observations	7,404				248			

Notes: This table displays summary statistics for the sample of PSID households surveyed in 2009, 2011, and 2013. Default is defined as 60+ days late as of survey date (at least two missed payments). The sample includes heads of household who are mortgagors, aged 24–65, and labor force participants (including those who are disabled) with combined LTV ratios less than 250 percent. Households that report they are in default in a given year are subsequently dropped to avoid double counting. All \$ values are deflated by the 2013 CPI. PSID weights are used to calculate all summary statistics unless otherwise noted. Dummy variables are signified by (d).

Table 3: Measures of Strategic Default: How Many Defaulters Can Afford Their Payments?

	(1) "Can Pay" $c < y - m$		(2) $c > y - m > c(VA)$		(3) "Can't Pay" $y - m < c(VA)$		(4) Total
	#	share	#	share	#	share	#
	(i)	(ii)=(i)/(vii)	(iii)	(iv)=(iii)/(vii)	(v)	(vi)=(v)/(vii)	(vii)
A. Total							
Default	74	0.377	65	0.333	57	0.291	196
Population	5,173	0.699	1,704	0.230	531	0.072	7404
Default Rate (# Def./# Pop.)	0.014		0.038		0.107		0.027
B. $LTV > 90$							
Default	47	0.409	41	0.352	28	0.239	115
Population	1,117	0.664	428	0.254	140	0.083	1684
Default Rate	0.042		0.095		0.197		0.069
C. $LTV < 90$							
Default	27	0.330	25	0.306	29	0.364	81
Population	4,056	0.709	1,277	0.223	391	0.068	5720
Default Rate	0.007		0.019		0.075		0.014

Notes: This table displays statistics on strategic default measures calculated from the PSID. Income, y , is defined as average monthly after-tax family income. If the head of household is unemployed as of the survey date, then the head's labor earnings for that month are set to zero (likewise for the spouse). If the head is recently divorced, then spousal labor earnings are set to zero. Consumption, c , is defined as average monthly expenditures, and m is the monthly mortgage payment across all mortgages, plus associated property taxes and insurance. The Back-End After-Tax DTI is defined as the ratio of combined mortgage payments to after-tax family income less alimony, child care, and interest on other debts (imputed at a 12.73% per annum rate). VA subsistence consumption is defined using the VA residual income based on the region, number of children, and income of the mortgage as of the survey date.

Table 4: Adverse shock incidence by residual income quintile.

	Quintile of Residual Income (Gross Income-Mortgage Expenses)				
	1	2	3	4	5
30% Income Drop or more (d)	25.6%	12.2%	5.9%	4.7%	3.9%
Unemployed Head Last Year (d)	13.9%	7.1%	4.9%	4.0%	4.2%
Unemployed Spouse Last Year (d)	4.3%	6.1%	5.7%	3.2%	3.1%
Transition into Disability, Head (d)	6.0%	4.3%	3.7%	3.4%	2.7%
Transition into Disability, Spouse (d)	3.4%	3.9%	3.4%	3.6%	3.8%
Any Shock (Unemployment, Disability, or 30% Income Drop) (d)	40.9%	28.5%	20.0%	16.8%	16.1%

Notes: This table displays the fraction unemployed, recently disabled, and with a 30% income drop or more, split by quintile of residual income (defined as gross family income less mortgage expenses) for the sample of PSID households surveyed in 2009, 2011, and 2013.

Table 5: Baseline Results: Linear Probability Model Cols (1) to (3), Logit Coefficients Cols (4) to (6) (with AME in square brackets, interaction at interquartile range of residual income), Dependent Variable is 60+ Days Late Indicator.

	(1)	(2)	(3)	(4)	(5)	(6)
Loan to Value Ratio	0.071*** (7.61)	0.078*** (6.74)	0.680*** (4.34)	1.798*** (10.34)	1.755*** (8.93)	-1.446 (-0.77)
Log Residual Income	-0.037*** (-9.73)	-0.025*** (-5.74)	0.012 (1.32)	-0.932*** (-14.01)	-0.727*** (-8.19)	-0.993*** (-5.90)
Log Residual Income x Loan to Value Ratio			-0.055*** (-3.97)			0.305* (1.74)
Constant	0.399*** (9.03)	0.190*** (3.70)	-0.218** (-2.06)	5.289*** (7.12)	1.105 (0.73)	3.852* (1.82)
Observations	7,404	7,404	7,404	7,404	7,404	7,404
R-squared	0.046	0.077	0.084	0.134	0.189	0.190
Demographic Controls?	N	Y	Y	N	Y	Y
Mortgage Controls?	N	Y	Y	N	Y	Y
State Controls?	N	Y	Y	N	Y	Y

Notes: This table displays OLS estimation results of regressions of default on LTV ratios and residual income in Cols. (1) to (3). Cols (4) to (6) report logit coefficients, and the square bracketed terms are the average marginal effects. To compute the interaction we compute the difference in the LTV AME between the interquartile range of residual income. Residual Income is defined as gross family income less mortgage expenses. Default is defined as 60+ days late as of survey date (at least two missed payments). The sample includes all household heads in the PSID who are mortgagors, aged 24-65, and labor force participants (including those who are disabled) with combined LTV ratios less than 250 percent. Robust t-statistics are reported in parentheses and dummy variables are signified by (d). Level of statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 6: IV Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Loan to Value Ratio	0.078*** (6.74)	0.189*** (3.63)	0.150*** (3.16)	0.224*** (3.45)	0.154*** (3.25)	0.188*** (3.65)	0.149*** (3.14)	0.192*** (3.73)	0.150*** (3.15)
Log Residual Income	-0.025*** (-5.74)	-0.149*** (-4.47)		-0.240** (-2.35)		-0.109** (-1.97)		-0.116** (-1.97)	
Unemployed Head Last Year (d)			0.053*** (4.14)						
Unemployed Spouse Last Year (d)			0.030** (2.33)						
Involuntary Unemployment, Head (d)					0.055** (2.54)				
Involuntary Unemployment, Spouse (d)					0.007 (0.28)				
Bartik Instrument (2 Yr. Ch.)							-0.553 (-0.86)		-0.536 (-0.83)
Transition into Disability, Head (d)							0.012 (0.89)		
Transition into Disability, Spouse (d)							0.033** (2.33)		
Transition into Major Disability (d)									0.058 (1.36)
Transition into Major Disability, Spouse (d)									0.041 (1.09)
IV for LTV:	.	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase
IV for Income:	.	Job loss Head & Spouse	.	Invol. Job Loss, Head & Spouse	.	Disability, Bartik Shock	.	Severe Disability, Bartik Shock	.
Observations	7,404	7,404	7,404	7,404	7,404	7,404	7,404	7,404	7,404
Demographic Controls?	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mortgage Controls?	Y	Y	Y	Y	Y	Y	Y	Y	Y
State Controls?	Y	Y	Y	Y	Y	Y	Y	Y	Y
Job Loss FEs?	N	N	N	Y	Y	N	N	N	N
Jtest Pval Null Valid	.	0.580	.	0.280	.	0.182	.	0.630	.
Weak ID Pval Null Weak	.	0	0	0.000166	0	4.51e-07	0	1.18e-07	0

Notes: This table displays a set of estimation from regressions of default on LTV ratios and residual income. Default is defined as 60+ days late as of survey date (at least two missed payments). Residual Income is defined as gross family income less mortgage expenses. Col (1) repeats Col (2) from Table 5. Col (2) uses unemployment over prior year for head and spouse as IV for residual income and cumulative state HP growth as IV for LTV. Col (3) is the reduced form of Col (2). Col (4) uses involuntary job loss as of the survey date as the income IV, and Col (5) is the reduced form of Col (4). Col (6) uses general disability shocks and a 2 year Bartik instrument as the income IVs, and Col (7) is the reduced form of Col (6). Col (8) uses severe disability shocks as the income IV, and Col (9) is the reduced form of Col (8). The demographic controls include 1-digit industry, year, race, education, marital status, and sex dummies as well as age and number of children. The mortgage controls include dummies for origination years, whether the mortgage is refinanced, the presence of a second mortgage, whether the term remaining is >15 years, whether or not the loan is refinanced as well as the interest rate. The state controls include whether the state is judicial, recourse, or one of the sand states, as well as changes in state level house prices and unemployment. Job loss fixed effects are a set of dummies for the number of prior unemployment spells over the prior ten years to our study, from 1994-2005. Level of statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

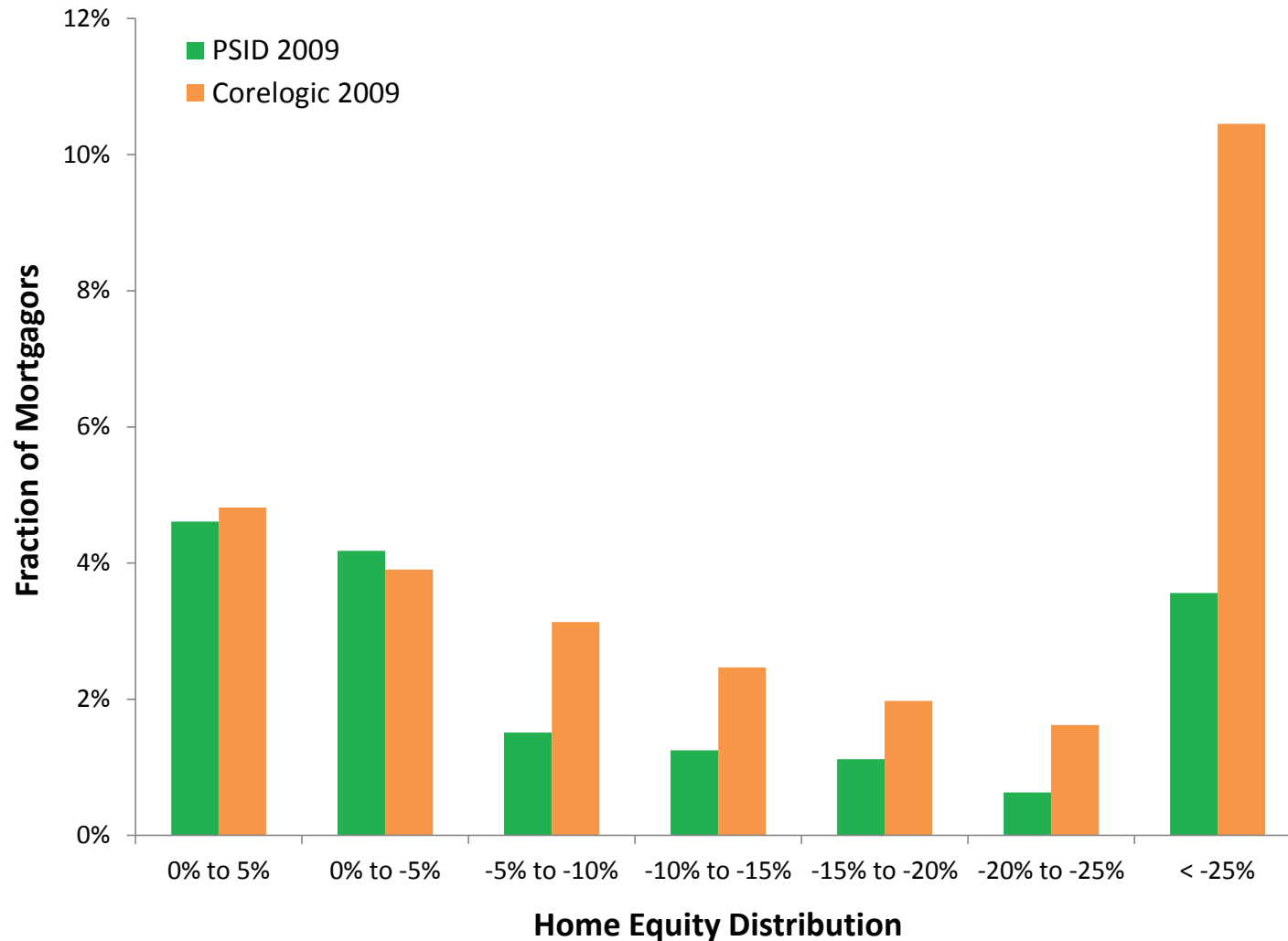
Table 7: IV Results, First Stages

A. LTV	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Cumulative State HP Growth from Purchase Date	.	-0.081***	-0.081***	-0.081***	-0.081***	-0.081***	-0.081***	-0.081***	-0.081***
	.	(-14.54)	(-14.54)	(-14.70)	(-14.70)	(-14.54)	(-14.54)	(-14.51)	(-14.51)
Unemployed Head Last Year (d)	.	0.020	0.020						
	.	(1.48)	(1.48)						
Unemployed Spouse Last Year (d)	.	0.026	0.026						
	.	(1.51)	(1.51)						
Involuntary Unemployment, Head (d)	.			0.025	0.025				
	.			(1.20)	(1.20)				
Involuntary Unemployment, Spouse (d)	.			0.013	0.013				
	.			(0.26)	(0.26)				
Bartik Instrument (2 Yr. Ch.)	.					0.438	0.438	0.489	0.489
	.					(0.51)	(0.51)	(0.57)	(0.57)
Transition into Disability, Head (d)	.					0.003	0.003		
	.					(0.15)	(0.15)		
Transition into Disability, Spouse (d)	.					0.015	0.015		
	.					(0.80)	(0.80)		
Transition into Major Disability (d)	.							0.048	0.048
	.							(1.43)	(1.43)
Transition into Major Disability, Spouse (d)	.							0.128**	0.128**
	.							(2.49)	(2.49)
Observations	.	7,404	7,404	7,404	7,404	7,404	7,404	7,404	7,404
R-squared	.	0.352	0.352	0.353	0.353	0.351	0.351	0.353	0.353
Demographic Controls?	.	Y	Y	Y	Y	Y	Y	Y	Y
Mortgage Controls?	.	Y	Y	Y	Y	Y	Y	Y	Y
State Controls?	.	Y	Y	Y	Y	Y	Y	Y	Y
Job Loss FEs?	.	N	N	Y	Y	N	N	N	N

B. Residual Income	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Cumulative State HP Growth from Purchase Date	.	-0.022*	.	-0.023*	.	-0.024*	.	-0.024*	.
	.	(-1.70)	.	(-1.69)	.	(-1.83)	.	(-1.80)	.
Unemployed Head Last Year (d)	.	-0.325***
	.	(-8.76)
Unemployed Spouse Last Year (d)	.	-0.239***
	.	(-6.44)
Involuntary Unemployment, Head (d)	.	.	.	-0.206***
	.	.	.	(-3.97)
Involuntary Unemployment, Spouse (d)	.	.	.	0.098
	.	.	.	(1.64)
Bartik Instrument (2 Yr. Ch.)	10.337***	.	10.118***	.
	(4.61)	.	(4.50)	.
Transition into Disability, Head (d)	-0.148***	.	.	.
	(-2.83)	.	.	.
Transition into Disability, Spouse (d)	-0.092**	.	.	.
	(-2.17)	.	.	.
Transition into Major Disability (d)	-0.335***	.
	(-2.99)	.
Transition into Major Disability, Spouse (d)	-0.223**	.
	(-2.45)	.
Observations	.	7,404	.	7,404	.	7,404	.	7,404	.
R-squared	.	0.335	.	0.326	.	0.321	.	0.322	.
Demographic Controls?	.	Y	.	Y	.	Y	.	Y	.
Mortgage Controls?	.	Y	.	Y	.	Y	.	Y	.
State Controls?	.	Y	.	Y	.	Y	.	Y	.
Job Loss FEs?	.	N	.	Y	.	N	.	N	.

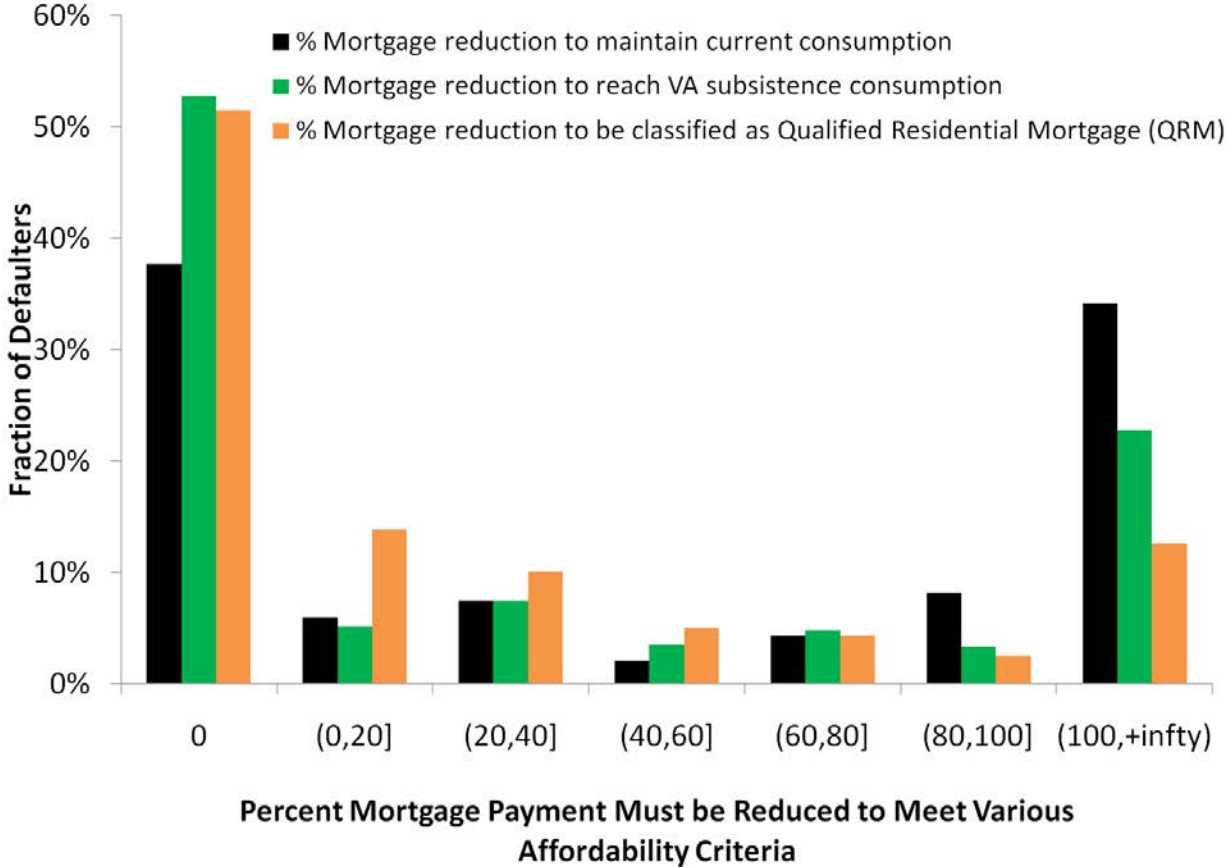
Notes: This table displays a set of estimation from regressions of default on LTV ratios and residual income. Default is defined as 60+ days late as of survey date (at least two missed payments). Residual Income is defined as gross family income less mortgage expenses. Col (1) repeats Col (2) from Table 5. Col (2) uses unemployment over prior year for head and spouse as IV for residual income and cumulative state HP growth as IV for LTV. Col (3) is the reduced form of Col (2). Col (4) uses involuntary job loss as of the survey date as the income IV, and Col (5) is the reduced form of Col (4). Col (6) uses general disability shocks and a 2 year Bartik instrument as the income IVs, and Col (7) is the reduced form of Col (6). Col (8) uses severe disability shocks as the income IV, and Col (9) is the reduced form of Col (8). Level of statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure 1. Comparing Equity Distribution in PSID and CoreLogic



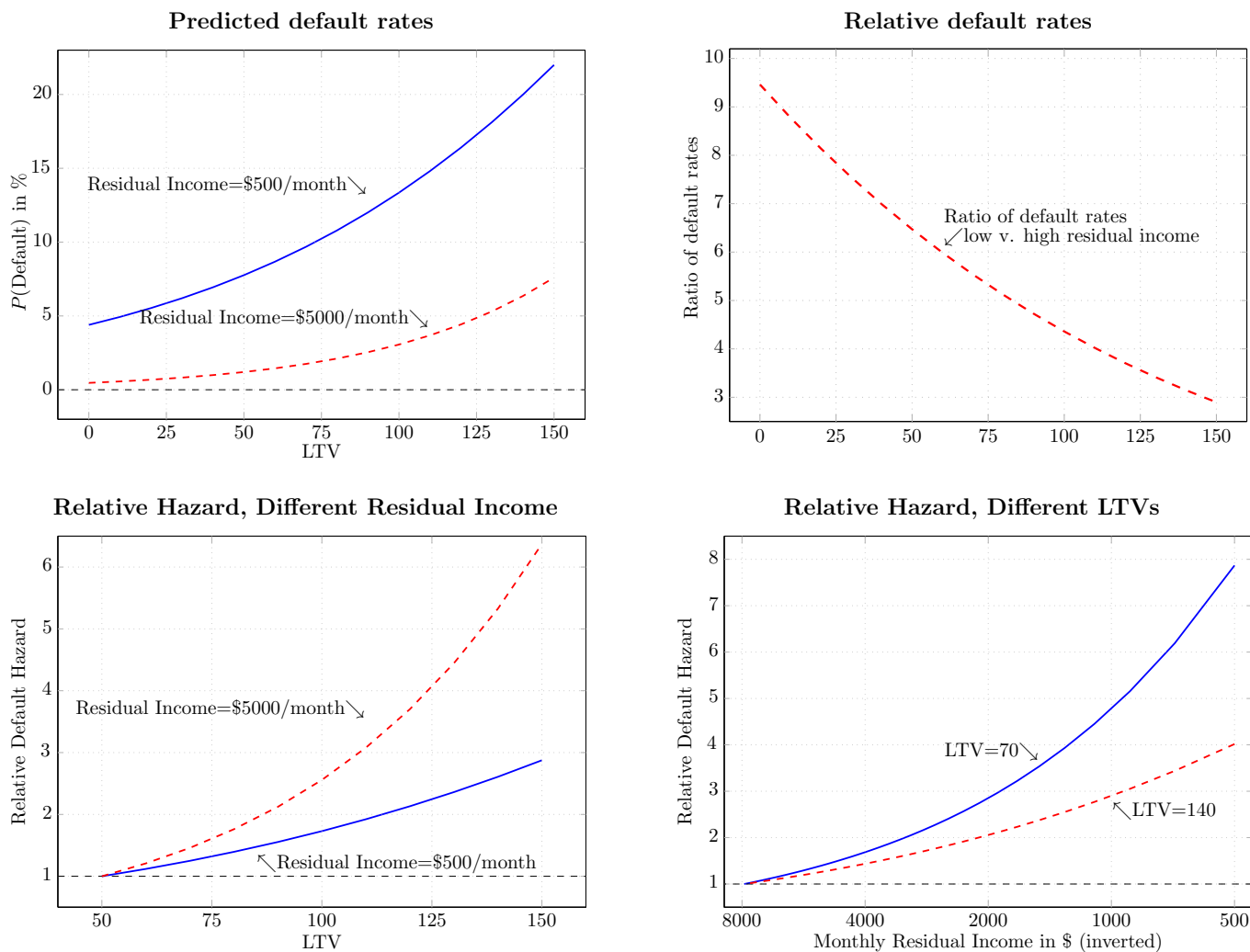
Notes: This figure compares the distribution of negative equity in the 2009 PSID to the distribution reported by CoreLogic in the third quarter of 2009. The PSID sample includes heads of household who are mortgagors, ages 24-65, labor force participants with combined loan to value ratios less than 250% in 2009. CoreLogic data taken as of 2009:Q3. "First American CoreLogic Releases Q3 Negative Equity Data," available from <http://www.recharts.com/reports/FACLNERQ32009/FACLNERQ32009.pdf>

Figure 2. Mortgage Affordability Measures Among Defaulters



Notes: Sample includes 60+ Days late defaulters from the PSID Main Sample, 2009-2013. Percent reduction calculated using 1st and 2nd mortgage. Payment reduction calculated as $\min((m_{max} - m)/m, 0)$, where m_{max} is the maximum payment a household can make and still have (i) the same consumption, (ii) VA residual consumption, and (iii) meet the definition of QRM. VA residual income defined in the text. QRM defined using 43% DTI after adjusting for insurance, taxes, other debt obligations, and alimony payments.

Figure 3. Predicted default rates based on Column (6) from Table 5



Notes: Sample includes 60+ Days late defaulters from the PSID Main Sample, 2009-2013. Percent reduction calculated using 1st and 2nd mortgage. Payment reduction calculated as $\min((m_{max} - m)/m, 0)$, where m_{max} is the maximum payment a household can make and still have (i) the same consumption, (ii) VA residual consumption, and (iii) meet the definition of QRM. VA residual income defined in the text. QRM defined using 43% DTI after adjusting for insurance, taxes, other debt obligations, and alimony payments.

Supplementary Appendix

“Can’t Pay or Won’t Pay?”

Unemployment, Negative Equity, and Strategic Default”

2016

Gerardi, Herkenhoff, Ohanian, and Willen

A. Existing Measures of Strategic Default

Table 8 compares our measures of strategic default to the existing literature’s definitions. We find a somewhat larger share of strategic defaults, 38%, relative to other studies whose estimates range from 19% to 35% depending on the year and method of measurement. What makes our measure unique is that it is model based, i.e. it is based on a mortgagor’s budget constraint, and we were able to measure each component of the budget constraint in the new PSID.

Table 8: Existing Measures of Strategic Default

Study	Experian/Oliver Wyman	Bradley, Cutts, Liu	Guiso, Sapienza, Zingales	Present Paper
Sample	Experian (Number of obs. undisclosed)	Equifax Merged w/ Payroll Data (EFX TWN) (N= 130k)	Chicago Booth Kellogg School Financial Trust Index, Q4-2008 to Q3-2010, (N= 1k)	PSID (N=7k)
Year	2004-Q4 to 2009-Q2	June 2008-June 2011	Q4-2008 to Q3 2010	2009 to 2013
Definition of Strategic Default	“[b]orrowers who rolled straight from 60 dpd to 180+ dpd, while staying less than 60 dpd on their auto loans and less than 90 dpd on their bank cards, retail cards, and other personal loans, for 6 months after they first went 60 dpd on their mortgage.”	Individuals with negative equity who transition from Current to 180+ Days Late with No Income Loss of 20% or More	Of the people you know who have defaulted on their mortgage, how many do you think walked away even if they could afford to pay the monthly mortgage?	Budget constraint definition: “What fraction of defaulters ‘can pay’, i.e. what fraction satisfy $c+m < y$ ”
Fraction strategic	19% in 2009, 18% in 2008	7% to 14.6%	25% to 35%	38%
Page Ref.	p. 4 and p.7	p.17	p.1 and p. 34 Fig. 5	Table 3

B. PSID Consumption Data and TAXSIM

In Table 9 we compare the entire PSID weighted sample of household heads (including renters and mortgagors) to the Consumer Expenditure Survey (CEX) as tabulated by the BEA.⁴³ The PSID data are treated as follows: each category is annualized, then aggregated to the line items below, and then the top 1% of positive values is winsorized, and only observations with annual food expenditure of at least \$500 are counted. The numbers below are reported at the family unit level in nominal terms. The main measures of food consumption align almost perfectly in both levels and trends. The expenditure on housing is quite different due to the fact that the PSID includes a category called ‘additions’ and this is a significant expenditure by many households. Most other line-items line up and follow similar trends, however healthcare was recoded in the 2013 PSID and falls significantly in 2013.

Table 9: PSID vs. CEX Expenditures Data (Source: PSID 2009-2013 Weighted)

Item	2009 CEX	2011 CEX	2013 CEX	CEX Notes	2009 PSID	2011 PSID	2013 PSID	PSID Notes
Avg. Annual Expenditures	49,067	49,705	51,100					
Avg. Annual Expenditures (Excluding Pension/Cash Contributions)	41,873	42,560	43,738		41,768	41,319	41,176	
Food + Alcoholic Beveridges	6,807	6,914	7,047		6,647	6,909	7,190	Food at home, away, delivered, and food stamps
Housing	16,895	16,803	17,148	Does not include additions to home	19,593	18,619	18,259	Mortgage Payments, Rent, Additions, Furnishings, Property Taxes and Insurance, Utilities
Apparel and services	1,725	1,740	1,604		1,307	1,153	1,144	Clothing Consumption
Transportation	7,658	8,293	9,004		7,032	7,503	7,555	Car repair, Gas, Parking, Trains, Cabs, Other Transp. Expenses, Car Insurance, Lease Outlays, Down payments, Loan payments, Outright Car Purchases
Health care	3,126	3,313	3,631		2,999	2,987	2,679	Health Insurance, Doctor, Hospital, Prescriptions
Entertainment	2,693	2,572	2,482		2,411	2,307	2,402	Trips and Recreation
Education	1,068	1,051	1,138		1,372	1,442	1,440	School Expenses, and Other School Exp.
Other Non-Aligned Consumption	1,902	1,875	1,685	Reading, Tobacco, Misc.	407	399	507	Child Care, Alimony

For TAXSIM computations, we base our code on the NBER TAXSIM code provided by Erick Zwick.⁴⁴ Table 10 summarizes PSID income per family compared to the comparable

⁴³Our measures come from the “Multi-year CEX Tables” entitled ‘Average annual expenditures and characteristics of all consumer units, Consumer Expenditure Survey, 2006-2012’ as well as the 2013-2014 version of the table. See <http://www.bls.gov/cex/tables.htm> for more details on the CEX tabulations.

⁴⁴The spouses pension variables were added later in the sample. For consistency we only focus on the

measure from the Census. Our Census measure is mean family income, Table H-6.⁴⁵ Table 10 shows that our measures of family income broadly align in levels with the Census measures, and our average tax burden per family is about 22% over this time period.

Table 10: PSID vs. Census Family Income and After-Tax Family Income (Source: PSID 2009-2013 Weighted)

	2009 Census	2011 Census	2013 Census	2009 PSID	2011 PSID	2013 PSID
Average Family Income	67976	69677	72641	72660	69000	73580
After TAXSIM Taxes	-	-	-	55700	54220	58020
N	117,538	121,084	122,952	9005	9235	9398

head's pension variables.

⁴⁵Table H-6. Regions--All Races by Median and Mean Income: 1975 to 2014' <https://www.census.gov/hhes/www/income/data/historical/household/>

C. Strategic Default with Assets

Table 11 computes measure of strategic default using asset information in the PSID. a is computed as the net financial assets of a household: the sum of checking, saving, money market accounts, government bonds, stocks, and other bonds, less an imputed 12.73% debt burden on all other unsecured debt obligations. 12.73% is the average credit card interest rate from 2009-2013 according to the Board of Governors. So in some cases, ability to pay of households may fall if they have negative net financial assets.

Table 11: Strategic Default with Assets

	Can Pay $c < y - m + a$		$c > y - m + a > c(VA)$		Can't Pay $y - m + a < c(VA)$		Total
	#	share	#	share	#	share	#
	(1)	(2)=(1)/(7)	(3)	(4)=(3)/(7)	(5)	(6)=(5)/(7)	(7)
A. All							
Default	95	0.485	40	0.205	61	0.309	196
Population	6184	0.835	570	0.077	655	0.088	7404
Default Rate	0.015		0.071		0.093		0.027
B. LTV>90							
Default	61	0.525	24	0.210	31	0.265	115
Population	1219	0.724	197	0.117	270	0.161	1684
Default Rate	0.050		0.123		0.113		0.069
C. LTV<90							
Default	35	0.429	16	0.199	30	0.372	81
Population	4965	0.868	373	0.065	384	0.067	5720
Default Rate	0.007		0.043		0.078		0.014

D. QRM Definitions of Strategic Default

Table 12 computes default rates among those who meet the QRM definition of affordability, and those who don't. We use the QRM guidelines to adjust income for taxes, insurance, alimony, and other debt obligations. If the ratio of combined mortgage payments to adjusted income is below 43%, the mortgage is deemed affordable. Applying this definition to our sample, Table 12 shows that there is a 5x difference in default propensity between those who meet the QRM definition of affordability (1.6%), and those who don't(9.2%). Among those with high LTVs (>90), the default rate among those who do not meet QRM affordability criteria is 17.9% relative to 4.0% for those who do. For those with positive equity, the level of default drops significantly for both groups.

Table 12: QRM Based Definitions of Strategic Default

	Can Pay		Can't Pay		Total
	Debt to Income<43%		Debt to Income>43%		
	#	share	#	share	#
	(1)	(2)=(1)/(5)	(3)	(4)=(3)/(5)	(5)
A. All					
Default	100	0.508	97	0.492	196
Population	6359	0.859	1045	0.141	7404
Default Rate		0.016		0.092	0.027
B. LTV>90					
Default	54	0.465	62	0.535	115
Population	1338	0.794	346	0.206	1684
Default Rate		0.040		0.179	0.069
C. LTV<90					
Default	46	0.571	35	0.429	81
Population	5021	0.878	699	0.122	5720
Default Rate		0.009		0.050	0.014

E. Baseline Regressions with DTI

Table 13 redoes Table 5 using the log of the debt to income ratio (i.e. $\log(m/y)$, the debt to income ratio, DTI) as the main independent regressor. Columns 1 through 3 report OLS coefficients, and Columns 4 through 6 report logit coefficients with average marginal effects in square parentheses below. As in Table 5, the interaction term is computed at the interquartile range for the logit specification. The coefficients can be interpreted as semi-elasticities. For example, column 1's point estimate implies that a 10% increase in DTI increases the default rate by .39 percentage points.

Table 13: Debt to Income Ratio Results: Linear Probability Model Cols (1) to (3), Logit Coefficients Cols (4) to (6) (with AME in square brackets, interaction at interquartile range of residual income), Dependent Variable is 60+ Days Late Indicator.

	(1)	(2)	(3)	(4)	(5)	(6)
Loan to Value Ratio	0.058*** (6.09)	0.071*** (6.06)	0.259*** (7.17)	1.568*** (8.51)	1.548*** (7.56)	2.341*** (4.57)
Log of DTI	0.039*** (8.47)	0.030*** (6.64)	-0.034*** (-3.48)	1.406*** (10.93)	1.110*** (7.61)	0.630** (2.02)
Log of DTI * LTV			0.103*** (6.39)	[0.047***] [0.043***]	[0.045***] [0.032***]	[0.043***] [0.033***]
Constant	0.066*** (5.30)	-0.019 (-0.69)	-0.134*** (-3.88)	-2.318*** (-8.45)	-4.024*** (-3.28)	-4.630*** (-3.61)
Observations	7,402	7,402	7,402	7,402	7,402	7,402
R-squared	0.036	0.077	0.093	-	-	-
Demographic Controls?	N	Y	Y	N	Y	Y
Mortgage Controls?	N	Y	Y	N	Y	Y
State Controls?	N	Y	Y	N	Y	Y

F. Income Changes and Non-Linearities

This section redoes our main analysis, but rather than use residual income, this table focuses on income shocks. In particular, we consider 2-year changes in gross family income between the PSID survey dates. Columns (1) through (4) of Table 14 show the non-linear impact of varying degrees of income loss on default. In columns (1) and (2) we include a series of indicator variables corresponding to various intervals in the income growth distribution: $(-\infty, -30\%]$, $(-30\%, -15\%]$, $(-15\%, -5\%]$, and $(-5\%, 0\%]$, with the omitted interval corresponding to any positive growth. The results reported in columns (1) and (2) show that income declines of more than 5% are significantly associated with increased mortgage default. Households that experienced negative income growth between 15% and 30% are 2 - 3 percentage points more likely to default compared to households that experienced flat or positive income growth, while households that suffered at least a 30% decline in income are more than 4 percentages more likely to default. Smaller declines in income (less than 5%) are not statistically significant predictors of mortgage default.

In column (3), and for the remainder of our analysis, we simplify the specification and include a single indicator variable for households that experienced a negative income shock of at least -15%.⁴⁶ Borrowers that saw their incomes decline by more than 15% were about 3 percentage points more likely to default compared to those that did not.

Table 15 illustrates the corresponding logit specifications which are comparable in sign, significance, and magnitude to our OLS estimates.

Table 16 displays the results of our IV analysis. Column (1) in the table corresponds to the simple OLS estimates, which are replicated from Table 14 (column (3)) for ease of comparison. Column (2) in the table displays the estimation results when we use the unemployment shock and recent divorce shock to instrument for income loss and cumulative house price appreciation to instrument for LTV ratios (all columns in the table use the same instrument for LTV ratios). There is a sizeable increase in the magnitude of the coefficient associated with income loss in the IV specification compared to the OLS regression. Households that experience a significant income loss that is caused by unemployment or divorce are approximately 26 percentage points more likely to default on their mortgages. The huge increase in the estimated impact of income loss on mortgage default in the IV specification is both plausible and consistent with economic theory. The permanent income hypothesis predicts that permanent (or persistent) shocks to income have a significantly larger effect on consumption decisions compared to more transitory income shocks. The IV specification isolates income losses due to unemployment and divorce shocks, which are

⁴⁶We chose this threshold based on the estimates reported in columns (1) and (2), where it appears that income growth becomes a significant predictor of mortgage default for declines between 5% and 15%. We do report results for alternative income growth thresholds of -5% and -30% in our analysis below.

both significant life events and thus, are likely to have persistent effects. In other words, the IV specification is isolating more permanent income shocks, which theory predicts should lead to a much larger impact on the propensity to default. Column (3) shows the reduced form of Column (2) where the default indicator is directly regressed on job loss and divorce indicators. In Column (4) of Table 16 we modify the instrument set by substituting for the unemployment variables with indicators of involuntary unemployment spells only (for both the head and spouse). In addition, we include a set of indicator variables corresponding to the number of prior unemployment spells as additional controls. The income loss coefficient decreases slightly (from 0.26 to 0.20), but is still very large in magnitude and statistically significant (at the 5 percent level). An income loss of at least 15% (between surveys) caused by an involuntary unemployment spell or divorce is estimated to increase the likelihood of default by 20 percentage points. Column (5) displays the reduced form regression results, where the default indicator is regressed directly on the involuntary unemployment shocks. The estimates are of comparable magnitudes with those in column (3).

Column (6) in Table 16 displays the results when we instrument for income loss using the disability shock and the Bartik employment shocks. We construct the Bartik variable over a two-year horizon (i.e. $k = 2$)⁴⁷ to maintain consistency with the biennial frequency of the PSID and our other results. We interact the Bartik variable with indicator variables corresponding to the industry in which the household head was employed at the beginning of the horizon.⁴⁸ In column (6) of Table 16, the coefficient estimate is 0.26 (statistically significant at the 5% level), which is very similar in magnitude to the estimates we obtained using unemployment spells and recent divorces as instruments (columns (2) and (4)). The first stage results displayed in Table 17, (column (6) in Panel B) show that the disability indicator is a strong predictor of severe income loss, which is consistent with the findings in Low and Pistaferri (2015). For space considerations we report the first stage estimates for the Bartik variables in the Appendix instead of Table 17.⁴⁹ The reduced form specification results reported in column (7) of Table 16 show that the disability variable has a slightly smaller direct impact on mortgage default compared to the the unemployment and divorce variables.

⁴⁷We also estimated specifications using Bartik shocks constructed over a four-year horizon and found similar results.

⁴⁸Interacting the Bartik variable with industry indicators allows the sensitivity of income loss to the exogenous, state-level, labor demand shocks to differ depending on the particular industry in which the individual is employed. We include a full set of industry fixed effects among the control variables (not in the instrument set).

⁴⁹Virtually all of the Bartik coefficients have the expected negative sign, so that positive state-level, labor demand shocks (i.e. increases in employment) are associated with a lower likelihood of significant income loss, however, they are not statistically significant, which suggests that they are not especially strong instruments for income loss at the household-level. However, it is clear from the weak instrument test p-values reported in Table 16 that the combination of the disability and Bartik variables constitute a strong set of instruments.

In column (8) of Table 16 we substitute the severe disability shock into the instrument set. Households that experience severe disability shocks are more likely to suffer more persistent income losses compared to households that suffer more moderate disability shocks, and thus we would expect the effect of income loss on default to increase as a result of this substitution. This is exactly what we find as the point estimate of the effect of income loss on mortgage default increases from 0.26 to 0.32.⁵⁰ In addition, the first stage results show that households that experience a severe disability shock are about twice as likely to experience an income loss of at least 15%, and the reduced form estimates (column (9)) show that they are also much more likely to default on their mortgage debt.

⁵⁰The difference between the two point estimates is not statistically significant however.

Table 14: Baseline Results: Linear Probability Model, Dependent Variable is 60+ Days Late Indicator.

	(1)	(2)	(3)	(4)
Loan to Value Ratio	0.082*** (8.60)	0.082*** (7.06)	0.083*** (7.10)	0.062*** (5.38)
Percent Income Change $\in (-\infty, -30]$ (d)	0.054*** (5.49)	0.041*** (4.39)		
Percent Income Change $\in (-30, -15]$ (d)	0.026*** (3.21)	0.020** (2.45)		
Percent Income Change $\in (-15, -5]$ (d)	0.022*** (2.82)	0.019** (2.46)		
Percent Income Change $\in (-5, 0]$ (d)	0.007 (1.03)	0.004 (0.70)		
Percent Income Change $< -15\%$ (d)			0.029*** (4.47)	-0.045** (-2.58)
LTV * Percent Income Change $< -15\%$				0.105*** (3.79)
Constant	-0.035*** (-5.42)	-0.076*** (-3.02)	-0.074*** (-2.92)	-0.057** (-2.28)
Observations	7,404	7,404	7,404	7,404
R ²	0.031	0.075	0.074	0.080
Demographic Controls	N	Y	Y	Y
Mortgage Controls	N	Y	Y	Y
State Controls	N	Y	Y	Y

Notes: This table displays OLS estimation results of regressions of default on LTV ratios and income growth. Income is defined as gross family income and growth in income is calculated between consecutive survey dates. Default is defined as 60+ days late as of survey date (at least two missed payments). The sample includes all household heads in the PSID who are mortgagors, aged 24–65, and labor force participants (including those who are disabled) with combined LTV ratios less than 250 percent. Robust t-statistics are reported in parentheses and dummy variables are signified by (d). Level of statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 15: Baseline Results: Logit, Dependent Variable is 60+ Days Late Indicator. Average Marginal Effects Reported.

	(1)	(2)	(3)	(4)
Percent Income Change $\in (-\infty, -30]$ (d)	0.063*** (5.15)	0.042*** (4.38)		
Percent Income Change $\in (-30, -15]$ (d)	0.032*** (3.05)	0.020** (2.37)		
Percent Income Change $\in (-15, -5]$ (d)	0.028*** (2.71)	0.025*** (2.61)		
Percent Income Change $\in (-5, 0]$ (d)	0.007 (0.79)	0.006 (0.66)		
Loan to Value Ratio	0.062*** (10.38)	0.049*** (8.33)	0.050*** (8.36)	0.050*** (8.34)
Percent Income Change $< -15\%$ (d)			0.025*** (4.40)	0.025*** (4.42)
LTV * % Income Ch. $< -15\%$ (d)				0.051*** (3.48)
Observations	7,404	7,404	7,404	7,404
Demographic Controls	N	Y	Y	Y
Mortgage Controls	N	Y	Y	Y
State Controls	N	Y	Y	Y

Notes: This table displays average marginal effects from logit regressions of default on LTV ratios and income growth. Income is defined as gross family income and growth in income is calculated between consecutive survey dates. Default is defined as 60+ days late as of survey date (at least two missed payments). The sample includes all household heads in the PSID who are mortgagors, aged 24–65, and labor force participants (including those who are disabled) with combined LTV ratios less than 250 percent. Robust t-statistics are reported in parentheses and dummy variables are signified by (d). Level of statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 16: IV Results: Dependent Variable is 60+ DL Indicator, 1st Endogenous Variable is 2-Year Income Change, 2nd Endogenous Variable is LTV. Col (1) is OLS, Cols (2) and (3) use unemployment and divorce as IVs for income. Cols (4) and (5) use invol. unemployment and divorce. Cols (6) and (7) use disability and Bartik shocks, and Cols (8) and (9) use severe disability and Bartik shocks. Cumulative HP growth is IV for LTV in all Columns.

66

Dependent Variable:	60+ Days Delinquent								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LTV Ratio	0.083*** (7.10)	0.167*** (3.15)	0.147*** (3.06)	0.175*** (3.46)	0.146*** (3.04)	0.172*** (3.35)	0.150*** (3.14)	0.178*** (3.35)	0.149*** (3.13)
Percent Income Change <-15% (d)	0.029*** (4.47)	0.264*** (4.26)		0.199** (2.45)		0.233** (2.27)		0.266** (2.13)	
Unemployed Head Last Year (d)			0.053*** (4.12)						
Unemployed Spouse Last Year (d)			0.031** (2.36)						
Recent Divorce (d)			0.034 (1.40)		0.034 (1.43)				
Involuntary Layoff (d)					0.035** (2.04)				
Involuntary Layoff, Spouse (d)					0.054* (1.88)				
Disability Shock (d)							0.018* (1.80)		
Severe Disability Shock (d)									0.051* (1.75)
IV for LTV Ratio:	.	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase
IV for Income:	.	Job Loss, Recent Divorce		Invol. Job Loss, Recent Divorce		Disability, Bartik Shock		Severe Disability, Bartik Shock	
Observations	7,404	7,404	7,404	7,404	7,404	7,404	7,404	7,404	7,404
R ²	0.074	.	0.069	.	0.067	.	0.061	.	0.062
Demographic Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mortgage Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
State Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
Control for Prior Unempl Spells	N	N	N	Y	Y	N	N	N	N
<i>IV Diagnostics</i>									
Over ID Pval, Null Valid	.	0.271	.	0.237	.	0.916	.	0.923	.
Weak ID Pval, Null Weak	.	0	0	3.49e-10	0	0.00252	0	0.00317	0

Notes: This table displays a set of estimation from regressions of default on LTV ratios and income loss. Default is defined as 60+ days late as of survey date (at least two missed payments). Income loss is defined as a drop in household income of at least 15% from the previous interview. The sample includes all household heads in the PSID who are mortgagors, aged 24–65, and labor force participants (including those who are disabled)

Table 17: First Stage IV Results: Col (1) is OLS, Cols (2) and (3) use unemployment and divorce as IVs for income. Cols (4) and (5) use invol. unemployment and divorce. Cols (6) and (7) use disability and Bartik shocks, and Cols (8) and (9) use severe disability and Bartik shocks. Cumulative HP growth is IV for LTV in all Columns.

Panel A: LTV Ratio								
Table 5 Column:	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Cumulative HPA (Since Purchase)	-0.080*** (-14.44)	-0.080*** (-14.44)	-0.080*** (-14.55)	-0.080*** (-14.55)	-0.081*** (-14.53)	-0.081*** (-14.53)	-0.081*** (-14.52)	-0.081*** (-14.52)
Unemployed Head Last Year (d)	0.019 (1.44)							
Unemployed Spouse Last Year (d)	0.026 (1.54)							
Recent Divorce (d)	0.053** (2.26)		0.053** (2.29)					
Involuntary Layoff (d)			0.007 (0.41)					
Involuntary Layoff, Spouse (d)			0.062 (1.41)					
Disability Shock (d)					0.012 (0.89)			
Severe Disability Shock (d)							0.091*** (2.94)	
Panel B: Income Loss								
	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Cumulative HPA (Since Purchase)	0.009 (1.09)	.	0.008 (0.94)	.	0.007 (0.82)	.	0.007 (0.80)	.
Unemployed Head Last Year (d)	0.139*** (6.48)
Unemployed Spouse Last Year (d)	0.122*** (4.98)
Recent Divorce (d)	0.235*** (5.69)	.	0.235*** (5.70)
Involuntary Layoff (d)		.	0.125*** (4.15)
Involuntary Layoff, Spouse (d)		.	0.042 (0.95)
Disability Shock (d)		.	.	.	0.065*** (3.33)	.	.	.
Severe Disability Shock (d)		0.149*** (3.37)	.
Observations	7,404	7,404	7,404	7,404	7,404	7,404	7,404	7,404
Demographic Controls	Y	Y	Y	Y	Y	Y	Y	Y
Mortgage Controls	Y	Y	Y	Y	Y	Y	Y	Y
State Controls	Y	Y	Y	Y	Y	Y	Y	Y

Notes: This table displays the first stage estimation results for IV specifications reported in columns (2) - (9) in Table 6. The sample includes all household heads in the PSID who are mortgagors, aged 24–65, and labor force participants (including those who are disabled) with combined LTV ratios less than 250 percent in 2009, 2011, and 2013. Robust t-statistics are reported in parentheses and dummy variables are signified by (d). Level of statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

G. Controls

Table 18 lists the baseline set of controls that are included in all of the main tables in the text.

Table 18: Controls

	Mean	Std.	Min	Max		Mean	Std.	Min	Max
NAICS Dummy 2	0.12	0.32	0	1	Second Mortgage Dummy	0.16	0.37	0	1
NAICS Dummy 3	0.17	0.37	0	1	Refi Dummy	0.47	0.50	0	1
NAICS Dummy 4	0.17	0.37	0	1	Refi Missing Dummy	0.00	0.04	0	1
NAICS Dummy 5	0.20	0.40	0	1	ARM Dummy	0.08	0.28	0	1
NAICS Dummy 6	0.15	0.36	0	1	ARM Missing Dummy	0.00	0.06	0	1
NAICS Dummy 7	0.04	0.19	0	1					
NAICS Dummy 8	0.04	0.20	0	1	Mortgage Interest Rate	4.81	1.98	0	23
NAICS Dummy 9	0.10	0.30	0	1	Mortgage Interest Rate Missing	0.05	0.21	0	1
Black	0.21	0.41	0	1	15+ Year Remaining on Mortgage Term Missing	0.02	0.13	0	1
American Indian	0.00	0.06	0	1	Origination Year 1992	0.00	0.06	0	1
Asian	0.01	0.12	0	1	Origination Year 1993	0.00	0.06	0	1
Pacific Islander	0.00	0.02	0	1	Origination Year 1994	0.01	0.08	0	1
Other	0.03	0.16	0	1	Origination Year 1995	0.01	0.08	0	1
Missing Race	0.01	0.08	0	1	Origination Year 1996	0.01	0.09	0	1
Age	44.00	10.50	24	65	Origination Year 1997	0.01	0.10	0	1
Male Dummy	0.85	0.36	0	1	Origination Year 1998	0.01	0.11	0	1
Married Dummy	0.74	0.44	0	1	Origination Year 1999	0.01	0.12	0	1
Less Than HS	0.25	0.43	0	1	Origination Year 2000	0.03	0.16	0	1
HS	0.27	0.45	0	1	Origination Year 2001	0.06	0.24	0	1
Some College	0.40	0.49	0	1	Origination Year 2002	0.08	0.27	0	1
College and More	0.01	0.10	0	1	Origination Year 2003	0.09	0.29	0	1
Number of Children	1.01	1.17	0	9	Origination Year 2004	0.11	0.31	0	1
2009 Dummy	0.36	0.48	0	1	Origination Year 2005	0.06	0.24	0	1
2011 Dummy	0.33	0.47	0	1	Origination Year 2006	0.03	0.17	0	1
2013 Dummy	0.31	0.46	0	1	Recourse Dummy	0.24	0.42	0	1
State House Price Growth Change	-0.02	0.08	-0.30523	0.22237	Judicial Dummy	0.40	0.49	0	1
State Unemployment Rate	0.08	0.16	-0.21622	0.636364	Sand States (CA, FL, AZ, NV)	0.14	0.34	0	1
Obs	7404								

H. Robustness

Table 19 displays robustness results for our main specifications in Table 6. Columns 1 and 2 of Table 19 include state fixed effects. These specifications yield consistent, although somewhat stronger, parameter estimates when compared to columns 4 and 6, respectively, of Table 6. Columns 3 and 4 of Table 19 use Bartik shocks that are constructed with 4 year and 1 year CES employment changes by state and industry, respectively. Our estimates are very close to Columns 6 and 8 in Table 6. Column 5 and Column 6 use a dummy for negative equity instead of a continuous variable, and for low LTVs, the dummy on negative equity implies a stronger effect of house price changes on default. An LTV of 1 in column 1 is associated with a 28% likelihood of default versus a 34% likelihood of default in Column (5). On the other hand, for higher LTVs, the relationship is reversed: an LTV of 1.2 in column 1 is associated with a 33% likelihood of default versus a 34% likelihood of default in column 5. Column 7 and Column 8 combine the head and spouse disability shocks to obtain more power, and again, we see similar results to the main table in the text. Additionally, in every case, the model passes over-identification tests at the 1%, 5%, and 10% levels.

Table 20 displays the first stages of the various regressions in Table 19, where each specification has two first stages corresponding to LTV and residual income. Panel A shows that cumulative house price growth is a strong instrument for LTV, and Panel B shows that the alternate instruments for income yield strong first stage results. In every case, the alternate sets of instruments pass weak identification tests.

Table 19: Robustness Results for Table 6.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Loan to Value Ratio	0.287*** (3.65)	0.255*** (3.67)	0.184*** (3.64)	0.190*** (3.62)			0.181*** (3.60)	0.194*** (3.77)
Log Residual Income	-0.242** (-2.30)	-0.178* (-1.94)	-0.099* (-1.91)	-0.124* (-1.95)	-0.289** (-2.53)	-0.098* (-1.76)	-0.094* (-1.85)	-0.116* (-1.96)
LTV>100 (d)					0.346*** (2.93)	0.297*** (3.30)		
IV for LTV:	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase	HPA Since Purchase
IV for Income:	Invol. Job Loss, Head & Spouse	Disability, Bartik Shock	Disability, Bartik Shock (4yr)	Disability, Bartik Shock (1yr)	Invol. Job Loss, Head & Spouse	Disability, Bartik Shock	Combined Disability, Bartik Shock	Combined Severe Disability, Bartik Shock
Observations	7,404	7,404	7,404	7,404	7,339	7,339	7,404	7,404
Demographic Controls?	Y	Y	Y	Y	Y	Y	Y	Y
Mortgage Controls?	Y	Y	Y	Y	Y	Y	Y	Y
State Controls?	Y	Y	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	N	N	N	N	N	N
Job Loss FEs?	Y	N	N	N	Y	N	N	N
Jtest Pval Null Valid	0.305	0.329	0.155	0.214	0.313	0.107	0.420	0.333
Weak ID Pval Null Weak	0.000225	0.00415	1.22e-07	1.83e-05	0.000425	6.81e-07	4.66e-08	2.90e-08

Notes: See Table 6 for additional notes. Col. 1 and Col. 2 include state FEs. Col. 3 and Col. 4 construct Bartik shocks using 4 year and 1 year employment changes by state and industry, respectively. Col. 5 and Col. 6 use a dummy for negative equity instead of a continuous variable. Col. 7 and Col. 8 combined the head and spouse disability shocks. Level of statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 20: First Stages of the Robustness Results for Table 6.

A. LTV First Stage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cumulative State HP Growth from Purchase Date	-0.076*** (-13.59)	-0.076*** (-13.42)	-0.081*** (-14.46)	-0.081*** (-14.53)	-0.081*** (-14.53)	-0.081*** (-14.53)	-0.081*** (-14.53)	-0.081*** (-14.52)
Bartik Instrument (2 Yr. Ch.)		0.933 (0.74)					0.424 (0.50)	0.470 (0.55)
Transition into Disability, Head (d)		0.004 (0.20)	0.003 (0.16)	0.003 (0.16)		0.003 (0.16)		
Transition into Disability, Spouse (d)		0.012 (0.65)	0.014 (0.80)	0.014 (0.80)		0.014 (0.80)		
Involuntary Unemployment, Head (d)	0.025 (1.21)							
Involuntary Unemployment, Spouse (d)	0.000 (0.00)							
Bartik Instrument (4 Yr. Ch.)			-0.070 (-0.14)					
Bartik Instrument (1 Yr. Ch.)				0.164 (0.11)				
Transition into Disability Head or Spouse (d)							0.012 (0.89)	
Transition into Severe Disability Head or Spouse (d)								0.091*** (2.96)
Observations	7,404	7,404	7,404	7,404	7,404	7,404	7,404	7,404
R-squared	0.372	0.370	0.351	0.351	0.351	0.351	0.351	0.352
Demographic Controls?	Y	Y	Y	Y	Y	Y	Y	Y
Mortgage Controls?	Y	Y	Y	Y	Y	Y	Y	Y
State Controls?	Y	Y	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	N	N	N	N	N	N
Job Loss FEs?	Y	N	N	N	Y	N	N	N
B. Income First Stage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cumulative State HP Growth from Purchase Date	-0.035*** (-2.65)	-0.034** (-2.53)	-0.026* (-1.92)	-0.023* (-1.71)	-0.019 (-1.46)	-0.022 (-1.61)	-0.025* (-1.83)	-0.024* (-1.81)
Bartik Instrument (2 Yr. Ch.)		5.605* (1.66)				10.463*** (4.64)	10.328*** (4.61)	10.092*** (4.49)
Transition into Disability, Head (d)		-0.134*** (-2.58)	-0.145*** (-2.78)	-0.146*** (-2.81)		-0.151*** (-2.88)		
Transition into Disability, Spouse (d)		-0.087** (-2.06)	-0.092** (-2.17)	-0.091** (-2.15)		-0.084** (-2.01)		
Involuntary Unemployment, Head (d)	-0.198*** (-3.85)				-0.217*** (-4.16)			
Involuntary Unemployment, Spouse (d)	0.094 (1.54)				0.081 (1.29)			
Bartik Instrument (4 Yr. Ch.)			6.240*** (4.85)					
Bartik Instrument (1 Yr. Ch.)				14.649*** (3.66)				
Transition into Disability Head or Spouse (d)							-0.126*** (-3.69)	
Transition into Severe Disability Head or Spouse (d)								-0.281*** (-3.83)
Observations	7,404	7,404	7,404	7,404	7,339	7,339	7,404	7,404
R-squared	0.347	0.339	0.321	0.320	0.325	0.319	0.321	0.321
Demographic Controls?	Y	Y	Y	Y	Y	Y	Y	Y
Mortgage Controls?	Y	Y	Y	Y	Y	Y	Y	Y
State Controls?	Y	Y	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	N	N	N	N	N	N
Job Loss FEs?	Y	N	N	N	Y	N	N	N

Notes: See Table 6 and Table 19 for additional notes.

I. Residual Income at Origination

Table 21 looks at the impact of residual income at origination on default propensities. Origination year is collected explicitly by the census. If the origination year is not a PSID survey year, we use the next nearest year ($t+1$) to compute residual income. For 300 households, we lack at least one piece of data to compute residual income at origination, and to maintain consistency, we winsorize the top and bottom 1% of residual income at origination. Comparing the coefficients in column (1) and column (2), residual income at origination is a much weaker predictor of default. The impact of residual income at origination on default is 4x less than contemporaneous residual income.

Table 21: Dependent variable is 60+ days late indicator. Col (1) repeats Col (2) from 5. Col (2) uses residual income at origination.

	(1)	(2)
Loan to Value Ratio	0.077*** (6.45)	0.081*** (6.74)
Log Residual Income	-0.026*** (-5.64)	
Log Residual Income at Origination		-0.006* (-1.67)
Observations	7,104	7,104
R-squared	0.076	0.069
Demographic Controls?	Y	Y
Mortgage Controls?	Y	Y
State Controls?	Y	Y

J. Comparison to Existing Loan Level Datasets

Table 22 compares our estimates of the impact of a 20% price drop, starting from an initial LTV of 100, to the prior literature. Using the logit estimates from Column (2), to be as comparable to the literature as possible, we find that a 20% price drop is associated with a .86% (.2*.046) increase in the default propensity. Our estimates are smaller than those obtained by focusing solely on subprime mortgages (e.g. Bhutta et al. (2011) and Demyanyk et al. (2011)), but somewhat higher than focusing only on fixed rate mortgages (e.g. Elul et al. (2010)). We find similar estimates to Gyourko and Tracy (2014) who use the most comparable sample. Gyourko and Tracy (2014) find an increase in default propensity of .89% for the same sized shock.

Table 22: Comparison of Samples and Comparison of Impact of LTV on Default Propensity

	Current Paper	Bhutta et al. (2011)	Gyourko & Tracy (2014)	Elul et al. (2010)	Demyanyk et al. (2011)
Sample	PSID 2009-2013, All Mortgages	Subprime 2006 Originations with CLTV>1 in AZ, CA, FL, NV	GNMA 2005-2012	Fixed Rate Mortgages Originated in 2005 or 2006	2004-2009 Subprime LP Mortgages
Default Measured During	2009-2013	2006-2009	2005-2012	2006-2009	2004-2009
Percentage Point Incr. in Default Rate from 20% Equity Drop, Starting from LTV of 100	1.08%	7.93%	0.89%	0.47%	4.40%
Type of Regression	Logit AME	Logit, Coeff. Reported Only	Re-LPM	Logit, MEM	Probit, MEM
Table	Table 5 Column (4)	Table 2, Column (1)	Table 1, Column (1)	Table 1, Column (3)	Table 4, Column (2)

Notes: Only a partial set of logit coefficients were reported in Bhutta et al (2011), so we can only compute the default propensity setting all other dummies to zero.