Mortgage Default with Positive Equity[*]

David Low[†]

NYU

Job Market Paper

This draft: January 16, 2015

Please see updated version on my website

Abstract

Over 80% of defaulting homeowners have positive equity in normal times. This paper develops a structural model of mortgage default in which selling a house takes time, which causes some homeowners to default even with positive equity. Calibrating the model with a standard income process, augmented with "disastrous" shocks, this paper shows that the considerable equity of many defaulters goes beyond what income shocks alone can explain. Including divorce and family size in the model allows it to match the aggregate default rate and the distribution of equity among defaulters. The estimated model suggests that recourse (which allows lenders to seize the assets of underwater defaulters) is ineffective and perhaps counterproductive, consistent with empirical research.

JEL Classification: D14, D91, E21, E60, G21, R21

*Keywords:* Mortgage default, home equity, liquidity constraints, divorce, families

[†]Email address: david.low@nyu.edu

1   Introduction

Why do homeowners default on mortgages? The literature offers two main answers: "strategic default" and "double trigger" default. "Strategic defaulters" walk away from a mortgage to maximize their wealth. A necessary (though perhaps not sufficient) condition for strategic default is that the home be worth less than the mortgage. In "double trigger" default, one trigger (e.g. unemployment) increases the value of liquidity while another trigger (e.g. a collapse in house prices) drives the homeowner underwater. Since the value of liquidity is high, the household chooses not to make its mortgage payments. Since it is underwater, the household prefers defaulting to selling the home. Hence, it defaults. Both paradigms of default predict that all defaulters are underwater, i.e. have Loan-to-Value ratios (LTVs) above 100.[1]

Figure 1: Distribution of Loan-to-Value Ratios of Defaulters and Mortgagors
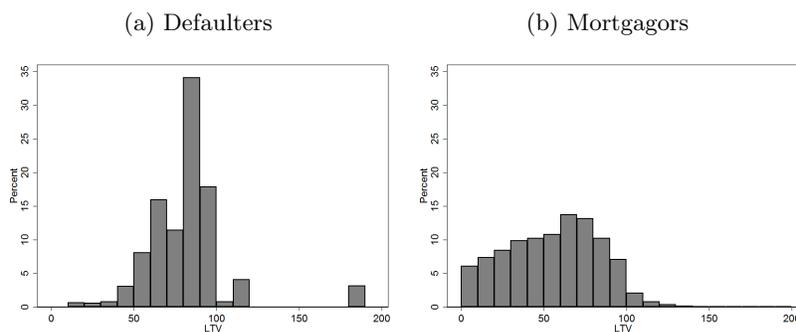
(a) Defaulters   (b) Mortgagors

Notes: The Loan-to-Value Ratio is defined as 100 times the ratio of total mortgage debt outstanding to house value. Data are from the 1998 & 2001 Survey of Consumer Finances (SCF). The literature standard is to count households as defaulters if they are at least two months behind on mortgage payments. However, the SCF does not ask how many months respondents are behind on mortgage payments. Therefore, I assume defaulters are households who: (1) were two months behind on debt payments within the last year, and (2) are currently behind schedule on their mortgage. Defaulters with LTVs below 10 or above 250 are dropped as outliers.

As shown in Figure 1a, many defaulters are in fact abovewater. In normal times over 80% of defaulters have positive home equity.[2] This is the result of two conflicting forces. The first, as noted by the "double trigger" literature, is that abovewater homeowners are less likely to default than underwater homeowners, in part because they may be able to sell their home to pay off their mortgage. The second force is that abovewater mortgagors outnumber underwater mortgagors, as evidenced by Figure 1b. In normal times, this second force is more powerful and abovewater defaulters outnumber underwater defaulters.

---

[1]Some papers in this literature note that homeowners with very little equity may also default, since they could lose money by selling their home after paying broker fees and other transaction costs. Plausible estimates of selling costs range from 6 to 10% of the value of the house, e.g. Bajari et al. (2013). Transaction costs are important to account for, but alone they cannot explain the abovewater default in the data.

[2]By "normal times" I mean times before the recent mortgage crisis, or the boom in house prices and credit that preceded it. Even during the recent mortgage crisis — a focus of much of the default literature, though not of this paper — between one third and one half of foreclosed homeowners had positive equity. More details are available in the appendix.

It is well-established that negative shocks (such as unemployment) drive many defaults. The more home equity defaulters have, the more financially costly their decision to default is, and the more severe these shocks must be for default to be optimal. A back-of-the envelope calculation suggests that the median abovewater defaulter stands to lose more than $22,000.[3] By neglecting the equity of defaulters, the literature downplays the magnitude of the shocks inducing default, while overestimating the effectiveness of policies that target negative equity, such as lender recourse.

This paper posits that some homeowners may default even if they have positive equity, since selling a house takes time. Still, matching the extent of equity among defaulters in the data requires powerful shocks. A version of my model that only includes income and equity as causes of default significantly underpredicts the home equity of defaulters, and can only explain 62% of the defaults in the data. The full model includes two additional causes of default. First, it includes divorce, which is perhaps the shock most often mentioned by the empirical literature after unemployment. Second, the model accounts for family size, which is a critical source of heterogeneity in default risk across households. Divorce generates default with equity because it reduces the income and liquid assets of the household, and because it reduces the incentive to avoid foreclosure as some home equity belongs to the ex-spouse. Family size generates default with equity because larger families have "more mouths to feed", and are less willing to reduce consumption to make mortgage payments. With these ingredients, the full model matches both the aggregate default rate and the distribution of home equity among defaulters.

This paper starts by showing that divorce and family size are both strong empirical predictors of mortgage default using data from the Panel Surve of Income Dynamics (PSID). Exploiting the panel dimension of the PSID, I control for a rich set of borrower characteristics at loan origination to accurately estimate the effect of unemployment, negative equity, divorce, and family size on default probability. To account for potential endogeneity, I instrument for divorce with the amount of time a couple has been together. Empirically, divorce has a stronger effect on default risk than unemployment. Adding three individuals to a household increases default risk by more than negative equity does. Critically, divorce and family size are also powerful aggregate drivers of default.[4] In the data, more foreclosures are associated with divorce than with negative equity.

Next the paper develops a lifecycle cash-in-advance model in which households are subject to income and marital status shocks, and family size evolves deterministically with marital status and age. Processes for income, marital status, and family size are calibrated outside the model to match

---

[3]This calculation assumes that the LTV distribution of defaulters is as given in Figure 1a. It also assumes that defaulters lose 25% of the value of their home (which Qi and Yang (2009) estimate the costs of foreclosure to be), or all their home equity, whichever is less.

[4]There is an important distinction between individual and aggregate drivers of default. Many studies, including this one, show that negative equity increases the probability that a household defaults. But, in normal times, negative equity does not cause many defaults because few homeowners have negative equity.

the data. Renters allocate expenditures between consumption, rent, and liquid assets. Home buyers choose levels of consumption and liquid assets, along with the size of their new home and mortgage. Mortgages are priced endogenously by competitive, risk-neutral banks. Homeowners choose levels of consumption and liquid assets, and also make a discrete choice. They can pay their mortgage and stay in their home. They can also sell their home to extract equity, or default to avoid making mortgage payments. Liquidity-constrained homeowners may default even if they have positive equity, because house sales are not processed immediately.[5]

I estimate this model using data from the Survey of Consumer Finances (SCF) and the PSID. The estimated model matches key targeted moments, including the aggregate default rate and the LTV distribution of defaulting homeowners. This distribution is very informative of the model parameters. As detailed later, it cleanly identifies both the elasticity of substitution between consumption and housing and the quantitative significance of family size. The model also matches key untargeted moments, including the proportion of defaulters undergoing divorce and the median LTV of mortgagors over the lifecycle. The model somewhat overpredicts the homeownership rate of single males.

The model generates valuable insights on lender recourse. This policy, which is intended to discourage default, allows lenders to seize the non-housing assets of defaulters to cover the difference between the outstanding mortgage balance and the value of the home. This difference is positive for underwater defaulters, so under recourse they may lose assets if they default.[6] Several structural models find that recourse is effective at discouraging default, e.g. Quintin (2012), Laufer (2013), Hatchondo et al. (2014), Corbae and Quintin (2014), and Li et al. (2014b). This is primarily because, in these models, almost all defaulters are underwater.

One result of my model is that recourse is ineffective at discouraging default for two distinct reasons. First, abovewater defaulters are not subject to recourse, because their home is worth more than the outstanding mortgage balance. Second, abovewater defaulters have very few other assets to seize. If they had other assets, they would spend them instead of defaulting. Therefore recourse has no effect on abovewater homeowners, who make up the vast majority of defaulters. Recourse does discourage underwater homeowners from defaulting in the model; there are just very few of

---

[5]This timing assumption is intended to capture, in a simple way, informational asymmetries in the housing market that prevent homeowners from lowering the price of their house to increase the probability of sale. Specifically, potential buyers may infer that an underpriced home has an undisclosed flaw. Guren (2014) provides evidence that decreasing the list price of a home below market value only slightly increases its probability of sale.

[6]Recourse typically allowers lenders to recover the difference between the outstanding mortgage balance and the fair market value of the home, *not* the sale price of the home at a foreclosure sale. As Ghent and Kudlyak (2011) note, "the fair market value restriction is likely present because the lender is often the only bidder at the foreclosure sale... In the absence of such a restriction, the lender could profit from a foreclosure by placing an artificially low bid."

them.[7] This is consistent with the empirical evidence in Ghent and Kudlyak (2011), who find that — while recourse lowers the probability that an underwater homeowner defaults — it does not lower default rates generally, either in aggregate *or* conditional on loan and borrower characteristics.

A surprising implication of the model is that recourse actually *encourages* homeowners with small amounts of positive equity to default. Under recourse, an abovewater homeowner who does not default this period loses the opportunity to default costlessly if she becomes underwater after the house depreciates. For homeowners with little equity, this is an important concern, so recourse encourages them to default. This finding emphasizes that structural models of default intended for policy analysis should match the LTV distribution of defaulters. It also offers an explanation for the puzzling finding in Ghent and Kudlyak (2011) that, conditional on observable characteristics, borrowers in recourse states pay *higher* interest rates.

Finally, the model has implications of interest beyond the mortgage literature. In particular, it cleanly identifies the elasticity of substitution between consumption and housing, through a simple, intuitive, and entirely new mechanism.[8] Homeowners (almost by definition) enjoy a high level of housing consumption. For high values of this elasticity, they can therefore tolerate large reductions in nonhousing expenditures without defaulting. However, the data requires that the model generate many homeowners that default with considerable equity at a large financial cost. This within-cell variation identifies a low value for this elasticity of .37 that is well in line with cross-sectional estimates.[9]

There is an extensive literature on mortgage default. Papers on "strategic" default model default as a put option on a financial asset, e.g. Kau et al. (1994), Mitman (2012), Jeske et al. (2013), and Davis and Quintin (2014). Papers on "double trigger" default, in which a liquidity shock and negative equity together induce default, include Deng et al. (2000), Foote et al. (2008), Haughwout and Okah (2009), Gerardi et al. (2009), Elul et al. (2010), Garriga and Schlagenhauf (2010), Bhutta et al. (2010), Foote et al. (2010), Quintin (2012), Chen et al. (2013), Gerardi et al. (2013), Guren and McQuade (2013), Herkenhoff and Ohanian (2013), Khandani et al. (2013), Chen et al. (2013), Laufer (2013), Corbae and Quintin (2014), Hatchondo et al. (2014), Campbell and Cocco (2014), Chatterjee and Eyingungor (2014), Li et al. (2014b), and Schelkle (2014). Individually, each of these

---

[7]Since the distribution of equity among homeowners before and during the mortgage crisis was very different than during normal times, my findings do *not* imply that recourse had no effect during the mortgage crisis.

[8]Because of how I model family size in the utility function, in my model this elasticity should only be interpreted as the elasticity of substitution between housing and nonhousing expenditures for agents with high levels of expenditures. At lower levels of expenditures, the elasticity will be lower.

[9]Many papers in macroeconomics and finance, e.g. Piazzesi et al. (2007), Davis and Ortalo-Magné (2011), and Halket and Vasudev (2014), argue that this elasticity is close to 1 since aggregate expenditure shares on housing have generally been steady over time. These models typically ignore cross-sectional data. Studies that use cross-sectional data, e.g. Hanushek and Quigley (1980), Flavin and Nakagawa (2008), Stokey (2009), and Li et al. (2014a), estimate a low elasticity (between .15 and .65), because people spend more on housing in places where housing is expensive.

papers provides valuable insights on the causes and consequences of mortgage default. As a whole, the literature overemphasizes the role of negative equity in mortgage default.

The rest of this paper proceeds as follows. Section 2 presents new empirical evidence from the PSID that divorce and family size, along with unemployment and negative equity, are important predictors of mortgage default. Section 3 constructs the structural model, which is estimated in Section 4. Section 5 discusses results, and Section 6 concludes. The appendix provides more information on the distribution of LTVs among defaulting homeowners.

2   Mortgage Distress in the PSID

In this section I show that divorce and family size are powerful predictors of mortgage distress. In my baseline specification, unemployment of the household head increases the probability of mortgage distress by 25.9%, while divorce increases it by 42.2%. Households are 10.6% more likely to be distressed if they are underwater, while distress probability grows by 4.2% for each individual in the household. These estimates imply that an underwater household with two members, and an above water household with five, have similar distress probabilities.

Other papers study the effect of divorce on default. For example, Deng et al. (2000) find that divorce increases the probability of default by about four times as much as unemployment does. Their estimates of the effects of divorce and unemployment are both highly statistically significant, even though they proxy for them using regional divorce and unemployment rates, an approach that is subject to the attenuation bias discussed by Gyourko and Tracy (2013).

The only other paper I am aware of that examines the effect of divorce on default with individual-level data, Gerardi et al. (2013), finds that divorce has no significant effect on default. Though there are several major differences between their approach and mine, two are critical to my results.

First, the sample in Gerardi et al. (2013) consists of mortgagors in 2009. Unfortunately, this sample excludes households who had already lost their home to foreclosure. Foreclosure (especially completed foreclosure) is often considered the worst of all potential outcomes for a delinquent mortgage, so it is important to account for. I use the PSID's retrospective questions on foreclosure to include foreclosed homeowners in my sample. I also use the 2011 PSID to distinguish between delinquent mortgagors in 2009 who would not enter foreclosure, and those that would.

Second, Gerardi et al. (2013) consider divorce that occurs before default is measured, which I also find to have no significant effect on default risk. However, concurrent divorce (observed because of the PSID's panel dimension) is a powerful predictor of mortgage distress.[10] These findings raise

---

[10]The much stronger estimated effect of concurrent divorce may be due to the (potentially contentious) division of assets that occurs during divorce, which is an important force in my model. It may also be due to a selection effect.

serious concerns about the endogeneity of divorce, so I instrument for divorce with the number of years a couple has lived together. Instrumenting in this way has virtually no effect on the size or significance of my estimates.

To the best of my knowledge, this paper is the first to study family size as a predictor of default.

2.1 Data

Data come from the Panel Survey of Income Dynamics (PSID), a longitudinal survey that has followed families and their offshoots since 1968. The PSID's panel dimension and detailed demographic information make it an attractive dataset to study the determinants of mortgage default.

However, there are two major problems with the PSID. First, like many other datasets, the PSID only has data on mortgage default from the recent housing boom (when unemployment and negative equity were rare and mild) and subsequent mortgage crisis (when unemployment and negative equity were widespread and severe). How well results from this time period generalize to more normal times is an open question. Therefore estimates of the effects of unemployment and negative equity should be treated with caution, though there is less reason to be wary of the results for divorce and family size.

Second, PSID data from the mortgage crisis understates both mortgage delinquency rates and the right tail of the LTV distribution of defaulters. Gerardi et al. (2013) argue this may be because the PSID undersampled subprime borrowers. Since I am interested in studying default in "normal" times, this is not a major drawback for this paper.[11] Researchers interested in the mortgage crisis or subprime lending, however, should keep this in mind.

The PSID was annual until 1997, and biannual after; the analysis in this paper uses all available PSID waves. The main sample of the PSID was designed to be nationally representative, but the Survey of Economic Opportunity (SEO) subsample oversamples low-income households. Since the PSID is a small survey, I include the SEO data in my analysis. This means the data is not nationally representative, so family-level weights provided by the PSID are used whenever necessary. These weights are intended to make the sample nationally representative within but not necessarily across years, so I weight all years equally.

I keep only data from households with heads between the ages of 21 and 67. To remove outliers, observations with total family income, wealth, or any category of expenditures in the bottom or top percentile are dropped. 3.96% of of the sample is dropped as outliers. Tax liabilities are estimated using the NBER's TAXSIM program, and subtracted from household income.

Households that undergo divorce without defaulting immediately may have particularly high income or liquid wealth that allow them to make their mortgage payments despite the divorce.

[11]See the appendix for more information on the distribution of LTVs among defaulters before and during the crisis.

The final sample consists of 2,827 families in 2009 that reported either: (i) currently having a mortgage, or (ii) having a home foreclosed since 2002. The PSID did not ask any questions about the performance of a mortgage until 2009. In 2009, they asked a series of questions on the current state of respondents' mortgages. They also asked, retrospectively, about foreclosures since 2002.

My dependent variable is based on these questions. Therefore I do not estimate the unconditional probability that a mortgage ends in default. Rather, I estimate the conditional probability that a mortgage is in one of several states by 2009, given that it lasted until at least 2002 and was not prepaid.

2.2   Empirical Strategy

Exploiting the panel nature of the PSID, I control for "ex-ante" observables, like education, at the time the mortgage is originated. I also control for "ex-post" variables like unemployment, which are measured after the mortgage is originated and may be regarded as shocks. Controlling for "ex-ante" variables is critical, because other drivers of mortgage default may be correlated with ex-post variables; for example, underwater mortgagors may be less educated than the general population of mortgagors. Failing to control for education would therefore overstate the strength of negative equity as a default trigger. Indeed, in Section 5, I provide evidence that traditional estimates of the effect of equity on default are significantly biased, in part because they fail to adequately control for ex-ante household-level variables.

I estimate an ordered probit model to take advantage of the PSID's detailed questions on mortgage distress. The observed outcome variable, Y, is therefore an ordered categorical measure of mortgage distress, with values:

0. The household is not, and does not anticipate becoming, behind on mortgage payments.
1. The household is "somewhat" or "very" likely to fall behind on payments.
2. The household is one month behind in mortgage payments.
3. The household is two or more months behind in mortgage payments.
4. The household is in foreclosure, or will be by 2011.
5. The household has lost its home to foreclosure, or will before 2011.

The ordered probit model I estimate is:

$$Y^* = \beta_0 + \beta_{ante}X^{ante} + \beta_{post}X^{post} + \epsilon \tag{1}$$

$$Y = \begin{cases} 0 & : Y^* < c_1 \\ 1 & : c_1 \leq Y^* < c_2 \\ 2 & : c_2 \leq Y^* < c_3 \\ 3 & : c_3 \leq Y^* < c_4 \\ 4 & : c_4 \leq Y^* < c_5 \\ 5 & : Y^* \geq c_5 \end{cases}$$

The goal is to estimate β_{post}, to evaluate the magnitude of different drivers of mortgage default. The ex-ante control variables (continuous variables are in logs) are:[12]

1. Age

2. Income: Average historical total family income, net of taxes and transfers.

3. Wealth: Total household net worth, excluding home equity.

4. Home value

5. Loan-to-Value Ratio: Whether above 80, and if so, every point above 80.

6. Whether the household has held debt before, and if so, how much (to proxy for credit score).[13]

7. Gender

8. Race

9. Marital status

10. Dummy for whether the head is white collar or not.

11. Dummies for whether the head and spouse are self-employed or not.

12. Education levels of the household head and spouse.

13. Dummy for whether the home is in a CSA metropolitan area or not.

14. Years elapsed between mortgage origination and 2009.

The "ex-post" variables in X^{post} (e.g. unemployment) are measured the last year the mortgage is observed, i.e. *after* mortgage origination. The only exception is concurrent divorce, which in the basline results is measured immediately after the last year the mortgage is observed.

The main variables of interest in Equation (1) are divorce and family size, so it is important to understand how they may be correlated with the error term . Family size is likely to be positively correlated with income and wealth growth; Black et al. (2013) find that children are a normal good,

---

[12]Including interest rates in the ex-ante controls does not significantly change the results, but does require dropping over 150 observations due to missing data. Since homeowners who cannot report the interest rate on their mortgage are unlikely to be a random sample, I exclude interest rates from the baseline results.

[13]This is not an ideal proxy, but it is the best available in the PSID. The coefficients are generally statistically significant and of the expected sign.

conditional on observables. The literature also typically finds that income or wealth shocks increase observed fertility, e.g. Black et al. (2013) and Lovenheim and Mumford (2013). Controlling directly for ex-post income increases the estimated effect of family size on mortgage distress. Therefore, my estimates are probably a lower bound on the causal effect of family size.

Correlation between divorce and   is more of a concern, since its direction is theoretically ambiguous. First, divorce may be correlated with omitted or mismeasured variables that also predict default, though the empirical evidence for this argument is decidedly mixed. In fact, there is some evidence that divorce hazard increases with wealth and income. For example, Farnham et al. (2011) find that declining house prices "lock" people in to marriage, while Hellerstein et al. (2013) find that divorce is pro-cyclical. The basic idea in both papers is that, since divorce is costly, a couple may wait for financial circumstances to improve before divorcing. However, there is also evidence that in some cases divorce hazard may decrease with wealth and income. Weiss and Willis (1997) find that positive shocks to male income reduce divorce hazard, although positive shocks to female income increase it. Charles and Stephens (2004) find that job displacement or layoffs increase divorce hazard, but disability and plant closings do not, perhaps because only the former are negative signals of a spouse's earning ability. These contrasting findings may explain why I find no evidence that unobserved or mismeasured income or wealth are a major concern. Controlling for observed ex-post income and home value does not affect the size or significance of the estimated effect of divorce on default.

Divorce and   will also be correlated if default causes (or prevents) divorce. Again, the direction is not clear. Losing a home may eliminate an important reason to maintain a marriage, and so it may encourage divorce. However, default may also discourage divorce, by denying couples the financial ability to live separately. Time-series evidence favors the latter view; in the midst of the recent foreclosure crisis, divorce rates hit a 40-year low.[14]

Still, correlation between divorce and   is a concern, so I instrument for divorce with the number of years a couple has been observed together. This is an excellent predictor of divorce, with an estimated p-value well below 1%. Moreover, there is no reason to expect it to predict default, except through its effect on divorce, since I control directly for the age of the household head and the number of years the couple has spent in the house. Instrumenting for divorce in this way reinforces suggestive evidence presented earlier that divorce and   are not strongly correlated; it has virtually no effect on my estimate of the effect of divorce on default.

Finally, I also include marital status in the list of ex-post controls. This is important to estimate the causal effect of family size; unmarried individuals have smaller families, but are more likely to

---

[14]http://www.cbsnews.com/videos/divorce-rate-up-15-since-the-great-recession/

default than married couples. It is also useful to disentangle the short- and long-term effects of divorce, since divorcing individuals become single. This long-run effect may be quantitatively more important, even if it is smaller in magnitude, since it lasts until remarriage.

2.3 Results

As is standard for probit models, I present Average Marginal Effects (AME) instead of coefficient estimates. I show the AME on the probability that $Y > 0$, i.e. the probability of some form of mortgage distress. For example, Specification III in Table 1 below shows that concurrent divorce increases the estimated probability that a mortgage is distressed by 42.2 percent, on average.

Results Table 1 presents the results on the determinants of mortgage distress. Specification I considers divorce that occurs before default is measured. Specification II does not include retrospective or 2011 information on foreclosures. Specification III is the baseline specification. To test whether unobserved income or house value shocks may bias estimates of the effects of divorce, Specification IV includes observed income and house value as controls. To more directly control for the endogeneity of divorce, Specification V instruments for divorce using the number of years the couple has lived together. Instrumenting in this way does not change the size or significance of the divorce coefficient.

The primary result of Table 1 is the size and significance of the coefficients for divorce and family size. It is also interesting to note that house price appreciation has virtually no effect on default risk; this may be because so many homeowners extracted home equity during the boom years, as noted by Laufer (2013) and Chen et al. (2013). Interestingly, controlling for income has almost no effect on the estimated effect of unemployment, which Gerardi et al. (2013) also find.

The result that recent (rather than concurrent) divorce has no significant effect on default risk is fragile. Recent divorce is obviously highly correlated with marital status, so removing marital status from the list of controls increases the estimated effect of recent divorce. But recent divorce also greatly lowers the probability of a concurrent divorce; simultaneously controlling for both types of divorce increases the estimated effect of recent divorce to statistical significance in some specifications. In every specification, however, concurrent divorce is a much more powerful predictor of default than recent divorce.

3 Model

This section builds a realistic model of housing, mortgages, and families over the lifecycle, which draws from the housing literature (Li et al. (2014a), Bajari et al. (2013)), mortgage literature

Table 1: **Ex-Post Determinants of Default**

	Specification				
	(I)	(II)	(III)	(IV)	(V)
Unemployed	25.8***	29.9***	25.9***	24.2***	25.2***
	(3.9)	(4.4)	(3.9)	(3.8)	(3.8)
Income				-8.0***	
				(1.7)	
Underwater	10.1***	14.4***	10.6***	11.4***	10.2***
	(2.4)	(2.2)	(2.3)	(2.4)	(2.2)
Home Value				0.2	
				(2.3)	
Divorce	5.9	1.8	42.2***	43.6***	41.7***
	(4.8)	(5.1)	(6.1)	(5.9)	(9.4)
Family Size	4.3***	2.7***	4.2***	4.6***	4.0***
	(.6)	(.6)	(.6)	(.6)	(.6)
Single Male	0.4	-0.9	0.1	-1.1	0.1
	(5.3)	(4.7)	(5.0)	(4.8)	(5.1)
Single Female	5.3	6.9	8.8*	7.6*	8.4*
	(4.6)	(4.3)	(4.5)	(4.4)	(4.6)
Divorced/Widowed Male	10.4	12.7	12.1	10.2	12.4
	(8.8)	(9.7)	(8.7)	(9.0)	(9.0)
Divorced/Widowed Female	8.9**	5.9	11.5**	7.5*	11.8**
	(4.5)	(4.4)	(4.5)	(4.5)	(4.7)
Ex-Ante Controls?	Yes	Yes	Yes	Yes	Yes
Concurrent Divorce?	No	Yes	Yes	Yes	Yes
Completed Foreclosures?	Yes	No	Yes	Yes	Yes
Instrument for Divorce?	No	No	No	No	Yes
Observations	2,822	2,646	2,821	2,796	2,821
Pseudo R^2	.1197	.1487	.1392	.1451	.1392

Notes: Average marginal effects on the probability that $Y > 0$ (i.e. mortgage distress) are reported. *, **, and *** denote statistical significance at the 10, 5, and 1 percent levels, respectively. Standard errors are in parentheses.

(Corbae and Quintin (2014), Laufer (2013), Chatterjee and Eyingungor (2014), Jeske et al. (2013)), and divorce literature (Voena (2012), Fernández and Wong (2013)). I explicitly model the problems of both homeowners and renters, and allow for non-convex adjustment costs to housing. This is the first structural model to include divorce and family size as default triggers.

I depart from the "strategic" and "double trigger" paradigms of default and assume that, because of search frictions in the housing market (as discussed in Wheaton (1990), Krainer (2001), Head et al. (2014), and Guren (2014)), houses sales occur at the end of the period. Hence, negative equity is not a necessary condition for default. This minor and realistic assumption allows the model to match the LTV distribution of defaulters with standard preference parameters and realistic income and family processes.

In the real world, homeowners also refinance their mortgage (or use a home equity line of credit ("HELOC")) to extract equity. Before the credit boom of the 2000s, it was typically difficult or impossible for homeowners to extract equity beyond an LTV of 80. Empirically and in my model, the

majority of defaulters have LTVs above 80, and so would not be able to refinance or tap a HELOC to avoid default. Therefore, I do not include refinancing or HELOCs in my model. Why some homeowners with LTVs below 80 default rather than refinance is an interesting question for future research. In the model, these high equity defaults are driven almost entirely by the "disastrous" income shock, which is fairly persistent. Hence, minimum Loan-To-Income (LTI) requirements and refinancing costs are probably part of the explanation.

3.1 Environment

The baseline environment is a standard model of consumption, housing, and portfolio choice over the lifecycle, with similarities to Li et al. (2014a) and Bajari et al. (2013). Onto this baseline, I add mortgages and marriage and divorce, which I introduce in turn.

Baseline Time is discrete. Consumers receive an exogenous, stochastic income flow $\{y_t\}$, from period $t = 0$ until retirement at T_{retire}. After retirement, they receive an exogenous, nonstochastic income until certain death in period T. Consumers value consumption c, and discount the future at rate β.

Agents can purchase liquid, risk-free assets a, which earn a rate of return R. Agents can borrow against a fraction ξ of their minimum income next period $\underline{y}'$, so the borrowing constraint is given by:

$$a' \geq -\xi \underline{y}' \tag{2}$$

Agents care about the flow value of where they live. Non-homeowners can choose to spend an arbitrary amount on rent r, or buy a house h, which yields a service flow κh and incurs a maintance cost ζh. Since $\kappa > \zeta$, one benefit of owning a home is its imputed rental value.

To reflect the economies of scale in a household, consumption and housing are both deflated by effective family size e, which is a concave function of actual family size. Moreover, a household must achieve minimum levels of consumption and housing services, $\underline{c}$ and $\underline{r}$, per effective family member. Therefore, utility over consumption and housing is given by $u(\frac{c}{e} - \underline{c}, \frac{r}{e} - \underline{r})$ for renters, and $u(\frac{c}{e} - \underline{c}, \frac{\kappa h}{e} - \underline{r})$ for homeowners.

Houses can be sold, incurring a proportional cost ϕh that represents both broker fees and moving costs. The proceeds from selling a house of size h are therefore $(1 - \phi)h$.

The focus of this paper is mortgage default with equity. In a frictionless model, homeowners with equity would sell their houses and repay their mortgages to extract this equity. Since, as I show, the LTV distribution of defaulting homeowners is a critical moment to match, I assume that frictions

in the housing market prevent agents from selling their home until the end of the period. This is equivalent to assuming that, though agents receive proceeds from house sales right away, they must save the proceeds until the next period. Therefore the borrowing limit for sellers is:

$$a' \geq -\xi \underline{y}' + (1-\phi)h \tag{3}$$

Finally, house values move stochastically over time. Let $\mathcal{H}(h)$ denote the distribution of h', given h.

Mortgages Home buyers can finance their purchase with a mortgage. To keep the state space manageable, I assume that home buyers obtain fixed-rate mortgages that last until the terminal period. Agents' mortgages can therefore be summarized by the constant mortgage payment p.

Thus, the budget constraint for agents who are not adjusting their housing stock is:

$$R^{-1}a' + c + r + p + \zeta h = a + y \tag{4}$$

Mortgage payments are offered to the mortgage lender in exchange for a loan of size L. L is chosen by risk-neutral mortgage lenders, who charge interest rate R_b, discount the future at rate β_B[15], observe the borrower's state and choice variables, and choose L to maximize their profits subject to perfect competition. These conditions determine the function $L = L(a, h', p', m, y)$.

House purchases, unlike sales, are processed immediately. Therefore the budget constraint for buyers is:

$$R^{-1}a' + c + \zeta h' + h' = a + y + L(a, h', p', m, y) * I(d = 0) \tag{5}$$

Home sellers must repay their mortgage. This involves buying back the nominal sequence of payments, $\{p_t, p_{t+1}, ..., p_T\}$ at the interest rate R_b, which implies that the cost of repaying a mortgage of constant payment p at time t is:

$$\Pi(p, t) = p \frac{(1 - (R_b^{-1})^{T-t-1})}{(1 - R_b^{-1})} \tag{6}$$

Therefore the budget constraint for sellers is:

$$R^{-1}a' + c + \zeta h + \Pi(p, t) = a + y + (1-\phi)h \tag{7}$$

Agents may default to avoid paying their mortgage or maintaining their home. Banks foreclose

---

[15]In line with the rest of the literature, I assume that $\beta_b R_b = 1$.

14

on defaulting homeowners at the end of the period, and sell the house at a proportional discount χ.[16] If the proceeds from the sale exceed the outstanding mortgage balance, the excess is returned to the defaulter. A defaulter therefore receives proceeds of $\max\{0, (1-\chi)h - \Pi(p,t)\}$, and has as a budget constraint:

$$R^{-1}a' + c = a + y + \max\{0, (1-\chi)h - \Pi(p,t)\} \tag{8}$$

Defaulters, like sellers, do not receive the proceeds (if any) until the end of the period. Therefore, like sellers, they are in effect constrained to save an amount at least equal to the proceeds. This means their borrowing limit is:

$$a' \geq -\xi \underline{y}' + R\max\{0, (1-\chi)h - \Pi(p,t)\} \tag{9}$$

After defaulting, agents are excluded from the mortgage market with a default flag, $d = 1$. Every period there is a probability π that d exogenously switches from 1 to 0. To reduce the number of state variables for homeowners (and following Chatterjee and Eyingungor (2014)), agents with a default flag lose it after buying a home (if they can afford to without a mortgage).

Marriage & Divorce: Marriage and divorce are exogenous, stochastic processes.[17] Denote marital status by m; agents in my model can be married couples ($m = 1$), single males ($m = 2$), or single females ($m = 3$). Each are given their own income and family size processes. Marriage probabilities are gender- and age-specific, while transition probabilities out of marriage are age-specific. As in Fernández and Wong (2013), the continuation value of married couples is taken to be the weighted average of the individual continuation values.

An important question is how to model the transition of assets through marriage. Love (2010) assumes that people marry people with the median wealth to income ratio. While tractable, this assumption does not respect the assortativeness of marriage. Fernández and Wong (2013) assume that singles know the characteristics of the person they will marry, if they marry; this is a flexible and useful assumption, but it comes at the expense of an additional state variable. Voena (2012) makes the simpler assumption that agents marry other agents with identical asset positions, which is computationally very tractable but does not respect gender differences in wealth, as Fernández and Wong (2013) point out.

I make an assumption that respects gender differences in wealth and the assortativeness of

---

[16]χ also includes the cost of the foreclosure process itself.

[17]In reality, divorce is endogenous. The assumption that it is exogenous in my model will bias my results if divorce is induced by other processes that also induce default, e.g. negative income shocks or declining house prices, or if default itself induces divorce. Section 2 provides theoretical and empirical evidence that this is not a major concern.

marriage, but does not introduce another state variable. Let $\Lambda(t)$ denote the ratio of median male wealth to median female wealth at age t. I assume that for every dollar in liquid assets and home equity that a woman brings to a marriage, a man brings $\Lambda(t)$ dollars. Since the housing space is discrete (and a couple must choose which house to live in), I assume that the newlywed couple stays in the house of the current agent.

An even more difficult question is what happens to assets after divorce. There is little quantitative research on the topic, even though a large proportion of assets may be lost in divorce because of litigation and the difficulty of splitting indivisible assets, like houses. The majority of economic models with divorce, including Voena (2012) and Fernández and Wong (2013), assume that it does not destroy any assets at all. Cubeddu and Ríos-Rull (2003) allow for a proportion of assets to be lost in divorce, but do not take a stand on the size of this proportion. Love (2010) assumes that 10 percent of assets are lost in divorce, but acknowledges that this choice does not have strong empirical support. The issue is further complicated by the fact that almost all previous papers with assets and divorce have only one asset class, instead of the three considered here.

My assumptions are intended to provide a reasonable lower bound on the amount of wealth lost in divorce. Following Love (2010), I assume that assets are split equally in divorce.[18] Therefore, a divorcing couple with no home and assets a produces two individuals each with assets $\frac{a}{2}$.

Figure 2: The Divorce Game
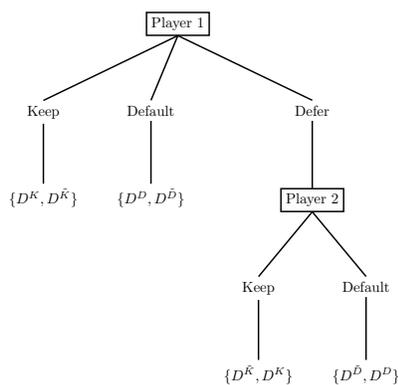
Notes: D^K and D^D denote the value functions of divorcing agents who keep the home and default on it, respectively. $D^{\bar{K}}$ and $D^{\bar{D}}$ denote the value functions of divorcing agents whose spouses keep the home and default on it, respectively.

Divorce for couples with a home is more complex. First, they split their liquid assets a equally,

[18]Wardle and Nolan (2011) write that during divorce "in most states, the court is charged to make an "equitable" distribution of the marital or community property between the spouses. The most common approach (codified or expressed as a judicial presumption in a few states, but the de facto unexpressed reality in most) is to make a roughly equal division of the property between the spouses... [Despite exceptions], equal division of property acquired during marriage is the common rule and justified under the notion that the marriage is, among other things, an economic partnership..." (pp. 123-124).

1

and each agent is awarded the home in a court settlement with probability $\frac{1}{2}$. The outcome of the divorce is then given by the unique Subgame Perfect Nash Equilibrium of the two-stage game depicted in Figure 2. The agent awarded the home ("Player 1") chooses whether to keep the home, default on it, or defer to Player 2. If Player 1 chooses to defer, Player 2 chooses whether to keep the home or default on it. If either agent keeps the home (whether or not they sell it), they must pay the other half the value of their home equity (if positive) at the end of the period. If either player defaults on the mortgage, they pay the other half the proceeds (if any) at the end of the period, and both agents are given a default flag.[19]

Finally, widowhood is also allowed in the model, but is much simpler and involves no change in assets. It is modeled simply as a transition to the income and family size processes for singles.

3.2 Value Functions

Banks   Banks are risk-neutral and discount the future at rate β_B. Let M_t denote the value to a bank of a mortgage owed by a homeowner with state $(a, h, p; y, m)$. When a bank extends a mortgage of size L to a home buyer, its return on the transaction is $\beta_b E(M_{t+1}(\tilde{a}', \tilde{h}', p'; y', m')) - L$. I assume that the mortgage market is perfectly competitive, which implies that at the time of origination L solves:

$$L = R_B^{-1} E(M_{t+1}(\tilde{a}', \tilde{h}', p'; y', m')) \tag{10}$$

Let $D_t(a, h, p; y, m)$ denote the decision rule of the agent regarding the mortgage, with $D_t = 0$ for payment, $D_t = 1$ for prepayment, and $D_t = 2$ for default. Then $M_t(a, h, p; y, m)$ satisfies the Bellman-like recursion:

$$M_t(a, h, p; y, m) = \begin{cases} p + \beta_b E(M_{t+1}(\tilde{a}', \tilde{h}', p' = p; y', m')) & \text{if } D(a, h, y, m, p, t) = 0, \\ \Pi(p, t) & \text{if } D(a, h, y, m, p, t) = 1, \\ \min\{\Pi(p, t), (1 - \chi_D)h\} & \text{if } D(a, h, y, m, p, t) = 2, \end{cases}$$

This equation for M_t, together with the equilibrium condition in the mortgage market (10), implicitly defines $L(a', h', p', m, t)$.

---

[19]Note that the model does not allow divorcing agents to refinance the mortgage to keep the home, which could potentially cause the model to overestimate the effect of divorce on default probability. However, in Section 4, I show that the model captures almost exactly the proportion of defaults that are associated with divorce, so this is not a major concern.

Non-homeowners Consider the problem of an agent who does not own a home. Denote his value function by V_t^R, and note that it is a function of assets a, default flag d, income y, and marital status m.

The agent has two options. First, he can continue to rent a home, in which case his Bellman equation is:

$$V_t^{RR}(a, d; y, m) = \max_{c,r,a'} u(\frac{c}{e}, \frac{r}{e}) + \beta E_{\tilde{a}',d',y',m'} V_{t+1}^R(\tilde{a}', d'; y', m')$$

subject to the borrowing limit (2), and the budget constraint for non-adjusters (4).

The agent can also buy a home, by obtaining a mortgage of size L with payments p' in which case his Bellman equation is:

$$V_t^{RB}(a, d; y, m) = \max_{c,a',h',L} u(\frac{c}{e}, \frac{\kappa h'}{e}) + \beta E_{\tilde{a}',\tilde{h}',y',m'} V_{t+1}^H(\tilde{a}', \tilde{h}', p'; y', m')$$

subject to the borrowing limit (2), the budget constraint for adjusters (4), and the equilibrium condition for p', (10).

Finally, an agent who does not own a home chooses to keep renting or to buy optimally, and so:

$$V_t^R(a, d; y, m) = \max\{V_t^{RR}, V_t^{RB}\}$$

Homeowners Now, consider the problem of an agent who does have a home. Denote his value function by V_t^H, and note that it is a function of assets a, house h, mortgage payment p, income y and marital status m.

The agent has three options: (1) stay in the home, (2) sell the home, or (3) default on the mortgage.[20]

If the agent stays in his home, his Bellman equation is:

$$V_t^{HH}(a, h, p; y, m) = \max_{c,a'} u(\frac{c}{e}, \frac{\kappa h}{e}) + \beta E_{\tilde{a}',\tilde{h}',y',m'} V_{t+1}^H(\tilde{a}', \tilde{h}, p' = \delta p; y', m')$$

subject to the borrowing limit (2) and the budget constraint for non-adjusters (4).

If the agent sells his home, his Bellman equation is:

---

[20]Note that, if house sales were processed immediately, the household would have a fourth choice of selling the home and buying a new one.

$$V_t^{HR}(a, h, p; y, m) = \max_{c, a'} u(\frac{c}{e}, \frac{\kappa h}{e}) + \beta E_{\tilde{a}', y', m'} V_{t+1}^R(\tilde{a}', d' = 0; y', m')$$

subject to the borrowing limit for sellers (3) and the budget constraint for sellers (7).

If the agent defaults on his mortgage, his Bellman equation is:

$$V_t^{HD}(a, h, p; y, m) = \max_{c, r, a'} u(\frac{c}{e}, \frac{r}{e}) + \beta E_{\tilde{a}', y', m'} V_{t+1}^R(\tilde{a}', d' = 1; y', m')$$

subject to the borrowing limit for defaulters (9), and the budget constraint for defaulters (8).

Finally, an agent who owns a home chooses to stay, sell, or default optimally, and so:

$$V_t^H(a, h, p; y, m) = \max\{V_t^{HH}, V_t^{HS}, V_t^{HD}\}$$

4 Data and Estimation

4.1 Estimation

I estimate the parameters of the model by the Simulated Method of Moments (SMM). This two-step procedure, developed by Pakes and Pollard (1989) and Duffie and Singleton (1993), is now a standard tool to estimate the parameters of structural models without closed form solutions.

In the first step, I choose parameter values that are standard in the literature, or that can be estimated from the data without the use of a structural model. In the second step, I take the parameters from the first stage as given, and estimate the remaining parameters by minimizing the distance between empirical moments and model output.

The details of the two steps are described in turn.

4.1.1 First Stage

The values for parameters that can be estimated without the use of the structural model are described below.

Initial Distribution of State Variables: At the beginning of life, agents receive the marital status, nonhousing wealth, and housing wealth of a randomly-drawn 23-year old from the PSID. The persistent income state is drawn from the stationary distribution. Renters begin with no default flag.

19

Bequest Function: The bequest function B is taken from De Nardi (2004). Note that, upon death, an agent with assets a, housing h, and mortgage payment p leaves behind wealth worth $w = a + (1 - \phi)h - \Pi(p, t)$. Let $c^*(w)$ and $r^*(w)$ denote optimal consumption and rent, respectively, from the one-period renter's problem with cash-on-hand w. Then I set:

$$B_t(l, m) = \upsilon_1 u_t(c^*(w + \upsilon_2), r^*(w + \upsilon_2), m)$$

The parameter υ_1 controls the strength of the bequest motive. In theory, it could be directly estimated since it should be identified by the lifecycle wealth profile. However, in practice it does not seem to be well-identified in my model, with large changes in υ_1 having no effect on the wealth profile early in life (which is driven by housing), and only a small effect on the wealth profile near retirement. Therefore, I set $\upsilon_1 = 1$, following De Nardi (2004). υ_2 controls the extent to which bequests are luxury goods. In De Nardi (2004), υ_2 is identified by the fraction of households who leave no bequests. Since my model is not designed to match this moment, I set $\upsilon_2 = 11.6$, following De Nardi (2004).

Debt and Liquid Assets: I set the proportion of labor income that can be borrowed against, ξ, to .2, following Heathcote et al. (2010). The real interest rate on liquid assets is set to -1.48%, following Kaplan and Violante (2012).

Housing and Mortgages: The real appreciation rate of home values is 0, matching the rate in Shiller (2008) for 1987-2000, and the value used in Li et al. (2014a) .

The flow value of housing, κ, is set to 7.5%, while maintenance costs ζ are set to 2%. Both of these values are standard in the literature, e.g. Li & Yao (2007), Li et al. (2014a). The proportional cost of selling a home, $\phi = .06$, is also standard, e.g. Bajari et al. (2013).

For computational tractability, I follow Mitman (2012) in assuming that shocks to a home also affect the flow utility derived from it. Therefore, I need to estimate the standard deviation of log house value growth (as opposed to log house price growth), which is not easy to do. Papers that differentiate between the quality of housing (which enters the utility function), and its price (which does not) typically assume that the quality of housing does not change (after homeowners pay maintenance costs), but that the standard deviation of log house price growth is about .1, e.g. Cocco (2004) and Li et al. (2014a). Granziera and Kozicki (2012) estimate that a model where house prices are driven purely by fundamentals has $\frac{5}{23}$ of the standard deviation of empirical log house price growth. Therefore, I set the standard deviation of log house value growth equal to $(\frac{5}{23})(.1) \approx .0217$.

Assuming a 25% tax bracket, the median real after-tax interest rate on mortgages in the PSID is approximately 3.66%. Therefore I set the interest rate on mortgages, R_B, to 3.66%.

Recall that χ denotes the deadweight loss of foreclosure, as a fraction of the value of the foreclosed home. In the literature, this is typically interpreted as the discount on foreclosed homes and calibrated accordingly, often to around $\chi = .2$. However, there are other costs of foreclosure that it is important to account for, including processing and maintenance costs and legal fees. Therefore I set $\chi = .255$, which Qi and Yang (2009) estimate to be the fraction of a loan lost after foreclosure on a homeowner with no equity.

The per-period probability ν that an agent with a default flag loses it is set to .25 following Chatterjee and Eyingungor (2014).

Utility: There are minimum acceptable levels of consumption and housing services, denoted $\underline{c}$ and $\underline{r}$, respectively. Denote discretionary consumption by $\tilde{c} = \frac{c}{e} - \underline{c}$, and discretionary housing services by $\tilde{r} = \frac{r}{e} - \underline{r}$.

The utility function is CES between discretionary consumption and discretionary housing, and CRRA over time:

$$u(c_t, h_t, e_t) = \frac{(\omega(\tilde{c}_t)^{\frac{\theta-1}{\theta}} + (1-\omega)(\tilde{h}_t)^{\frac{\theta-1}{\theta}})^{\frac{(1-\gamma)(\theta)}{\theta-1}}}{1-\gamma} \tag{11}$$

The ratio $\frac{\underline{r}}{\underline{r}+\underline{c}}$ is directly identified in the data, by the observation that families in the PSID below the poverty level spend, on average, 33% of total expenditures on rent. Therefore, I set $\frac{\underline{r}}{\underline{r}+\underline{c}} = .33$. The sum of these parameters, $\underline{x} = \underline{c} + \underline{r}$, is estimated using the structural model in the next section. Given $\underline{x}$ and $\frac{\underline{r}}{\underline{r}+\underline{c}}$, trivial algebra gives $\underline{c}$ and $\underline{r}$.

The coefficient of relative risk aversion, γ, is set to the standard value of 2. θ is estimated using the structural model. Given a value of θ, ω is directly identified in the data. In the limit (as income increases and $\underline{x}$ becomes irrelevant), only ω and θ determine the fraction of expenditures that renters in the model spend on rent. Therefore, given θ, I set ω to match the fact that high-income PSID respondents spend on average 23% of total expenditures on rent.

Mortality: Mortality data comes from the most recently available National Longitudinal Mortality Survey (NLMS), from 2008. The age- and gender-specific mortality rate is assumed to be the proportion of NLMS respondents of that age and gender who were observed to die at a given age.

Demographics: Households begin life at age 23, retire at 65, and die with certainty at 80.

The effective family size e (i.e. the consumption and housing deflator) is calculated for each

observation as in Li et al. (2012), who follow Scholz et al (2006). If N_c is the number of children in a household and I_m is an indicator for a household headed by a married couple, then the deflator for a household is $(1 + I_m + .7 * N_c)^{.7}$ In the model, I set e to the median deflator value in the PSID, conditional on age and marital status.

Age-specific divorce and widowhood probabilities (for married couples), and age- and gender-specific marriage probabilities (for singles) are taken from the PSID.

Gender wealth ratios (for determining assets after marriage) are taken from the data, as the ratio of average male wealth to average female wealth.

Income: Following much of the lifecycle literature, I assume that labor income follows a deterministic trend but is subject to transitory and persistent idiosyncratic shocks. Specifically,

$$\log(y_t) = g_t + z_t + \epsilon_t \qquad (12)$$

where g_t is the deterministic component of income and ϵ_t is the transitory shock. The persistent shock z_t follows the AR(1) process,

$$z_t = \rho z_{t-1} + \eta_t$$

ϵ_t and η_t are normal random variables with mean 0 and variances σ_ϵ^2 and σ_η^2, respectively. I estimate the parameters of the process (ρ, σ_ϵ^2, and σ_η^2) by Minimum Distance Estimation, the method originally proposed by Chamberlain (1984) and since used by Guvenen (2009), Kaplan and Violante (2010), Fernandez and Wong (2011), and many others. Table 2 presents the results.

Table 2: *Estimated Parameters of the Income Process.* Standard errors are in parentheses.

$\hat{\rho}$	$\hat{\sigma_\eta^2}$	$\hat{\sigma_\epsilon^2}$
.965	.008	.056
(.018)	(.003)	(.010)

These estimates are well in line with others in the literature, though they are all somewhat low. There are several potential explanations for my low estimates, as my sample differs in some details from others in the literature, most notably the inclusion of households headed by single males and especially single females. It is possible, for example, that singles face less income uncertainty than married couples, though this seems unlikely given the risk sharing that is often assumed to occur between married couples. An explanation I prefer is described in detail in Guvenen (2009); income profiles (absent any shocks) may in fact be heterogeneous, while the estimation procedure restricts

them not to be. This forces the estimation to treat heterogeneity as shocks.[21] Since my estimation procedure allows for profiles to differ by household type and race, it allows for more heterogeneity than usual, which in turn may reduce the portion of income variance that is attributed to shocks.[22]

This paper emphasizes the role of divorce and family size in triggering mortgage default, so it is important not to underestimate the role of income shocks. Guvenen et al. (2014) note that the standard lognormal income process cannot match the negative skewness and high kurtosis of income seen in the data. This is a serious concern for this paper, since large negative income shocks are precisely the ones likely to trigger default. One common approach in the mortgage literature is to explicitly introduce an unemployment shock. However, there are other large negative income shocks besides unemployment, like disability, so accounting only for unemployment will understate the probability of such an event. Therefore, I include an i.i.d. "disaster" shock (which may include unemployment) in the income process, calibrated to match the 3.3% probability in the PSID of a working-age non-divorcing family reporting at least a 50% drop in income.[23] The size of this shock is calibrated to match the fact that, empirically, the median household who reports such a shock reports a 61.67% drop in income. Finally, I set the per period probability of escaping this state to 65.26%, which is the probability that a household reports at least a 20% increase in income following such a shock.[24].

Age- and marital status-specific income profiles are taken as the empirical median. Income after retirement follows Li et al. (2014a). Specifically, after retirement, households receive a fraction of their pre-retirement persistent income. I estimate the average value of this fraction by household type in the PSID, to get values of .876 for married couples, .923 for single males, and .816 for single females.

4.1.2 Second Stage

Here, I estimate the remaining structural parameters β, θ, and $\underline{x}$ by minimizing the distance between empirical moments and model output. Recall that β is the discount factor, θ is the elasticity of substitution between "discretionary" housing and nonhousing consumption, and $\underline{x}$ is the necessary level of total consumption per effective family member.

The LTV distribution of defaulting homeowners embeds a great deal of information, including

---

[21]Heterogeneity in growth rates will be interpreted as very persistent idiosyncratic shocks. As detailed in Guvenen (2009), failing to account for this heterogeneity will bias $\hat{\rho}$ upwards. The direction of the bias for $\hat{\sigma}_\eta^2$ is not clear, and depends on the relative variance of profile heterogeneity and true persistent shocks. This may explain while accounting for heterogeneity moves Guvenen's $\hat{\sigma}_\eta^2$ upwards and mine downwards.

[22]Allowing for only one profile, common to households of all types and both races, changes the estimates to $\hat{\rho} = 0.969$, $\hat{\sigma}_\epsilon^2 = 0.0718$, $\hat{\sigma}_\eta^2 = 0.011$, which are all much closer to typical values in the literature.

[23]In the PSID, the probability of experiencing such a shock seems almost constant over the working lifecycle, which is what I assume.

[24]This disaster shock is similar in spirit to the "disastrous" income shock in Cocco et al. (2005)

what the causes of default are and what the effects of policies (like recourse) could be, and is a high priority target in my estimation. Surprisingly, no other paper I know of targets it.[25] My other two targets are more standard: (1) the median wealth profile by household type,[26] and (2) a foreclosure rate of .5%.[27]

I solve and simulate the model for 50,000 agents over a Halton grid of 1,000 different combinations of parameter values.

I estimate a value for $\beta = .958$, which is actually on the high side of standard. This demonstrates that it is not simply a very low discount rate that generates abovewater default in my model. The values for $\theta = .374$ and $\underline{x} = 6.19$ are mainly what drive abovewater default. Abovewater default requires that reductions in nonhousing expenditures be unpleasant for homeowners, who (since they are in a house they bought) enjoy a comparatively high level of housing consumption. For high values of θ, this high level of housing consumption will allow agents to tolerate large reductions in nonhousing expenditures without defaulting. Hence, high levels of abovewater default identify a low value of θ.

5   Results

Model Fit   This subsection descibes the fit of the model to targeted and untargeted moments in the data.

The primary goal of this paper is to understand why so many homeowners default with so much equity. Therefore the aggregate foreclosure rate and the LTV distribution of defaulters are high-priority moments to match. The estimated model generates a foreclosure rate of .49%, which is almost identical to the targeted foreclosure rate of .5%. Figure 3 plots the LTV distribution of defaulters in the model and in the data. Overall, the model does quite well.

The estimation also targeted median wealth profiles by household type. Figure 4 plots these profiles, both in the estimated model and in the data. The model does well reproducing the wealth profile of married couples, although it slightly underpredicts the wealth of older married couples. This may be because the model lacks a high-return illiquid asset, such as retirement accounts. However, the model slightly overpredicts the wealth of single males and females.
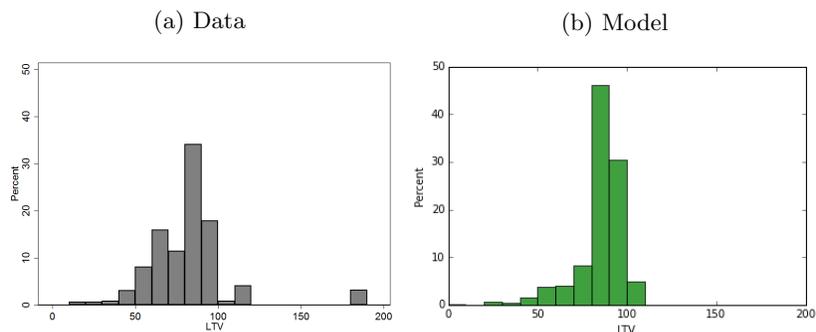
The homeownership rate was not targeted in estimation, although it is closely related to wealth.

[25]Laufer (2013) targets default hazard rates by LTV. In his model, houses can be sold instantly, so the high hazard rates of abovewater homeowners identifies counterfactually high selling costs. See the appendix for more details.

[26]Kaplan and Violante (2012) note that, because the wealth distribution is right-skewed, economists face a tradeoff in targeting the median or average wealth profile. Like Kaplan and Violante (2012), I target the median wealth profile.
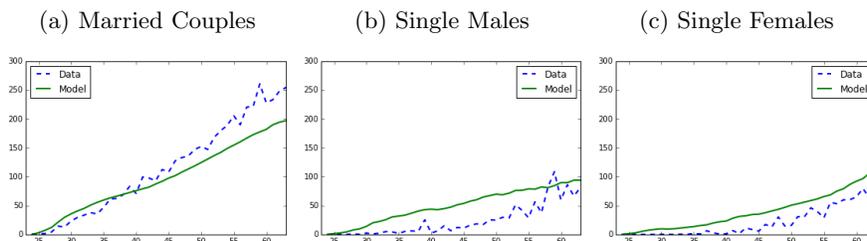
[27]Jeske et al. (2013) report that, from 2000-2006 roughly .4% of mortgages actually resulted in liquidation, which is the consequence of default in both their model and mine. House price growth was unusually strong in 2000-2006, so they view this .4% as a lower bound on the true figure, and target .5% instead. This is the number I target as well.

Figure 3: Distribution of Defaulters' LTVs in the Data and Model

(a) Data

(b) Model

Figure 4: Median Wealth over the Lifecycle

(a) Married Couples

(b) Single Males

(c) Single Females

Notes: Solid green lines are model output. Dotted blue lines are from the PSID.

Figure 5 plots homeownership rates, both in the estimated model and in the data. The model does an excellent job reproducing the homeownership profile of married couples, and only slightly overstates the homeownership rate of single females. However, the model cannot explain why so few single males own homes.

Figure 5: Homeownership Rates over the Lifecycle

(a) Married Couples

(b) Single Males

(c) Single Females
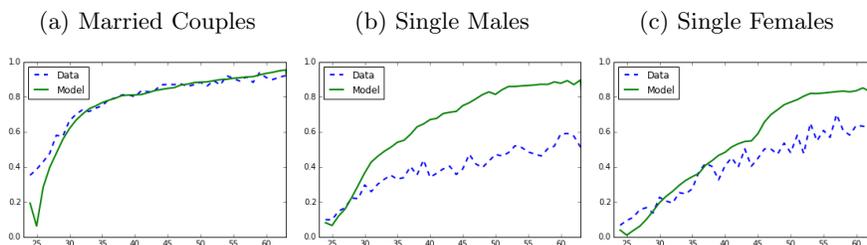
Notes: Solid green lines are model output. Dotted blue lines are from the PSID.

The estimation does not target the LTV of mortgagors. However, given this paper's focus on mortgage default and the LTV of defaulting homeowners, this is an important moment to match. Figure 6 plots the median LTV of mortgagors over the lifecycle and shows that, in general, the model

does quite well.

Figure 6: Median LTV of Mortgagors over the Lifecycle
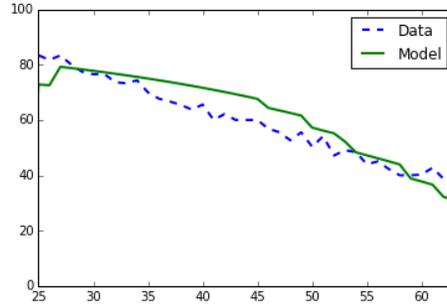
Notes: The Loan-to-Value Ratio is defined as 100 times the ratio of total mortgage debt outstanding to house value. Data are from the PSID.

To check that the model does not overstate the importance of divorce as a driver of default, it is useful to compare divorce rates of foreclosed homeowners in the data and the model. In the model, 26.7% of defaults occur in the same period as a divorce (and roughly 1% the period before or after). In the data, 9.8% of foreclosures immediately follow a divorce, while 18.3% of foreclosures involve a concurrent divorce. Hence, the importance of divorce as a driver of default is virtually identical in the model and the data. Moreover, in the data less than 20% of defaulters have negative equity. Therefore, in both the estimated model and the data, divorce drives more defaults than negative equity does.

The Causes of Default   One point of this paper is that the LTV distribution of defaulters embeds important information on the causes of mortgage default. Figure 7 highlights the default triggers associated with particular defaults in the model. Note that either divorce or a "disastrous" income shock are virtually required for homeowners with LTVs below 90 to default. Hence, models that do not match the LTV distribution of defaulters, but do match the aggregate default rate, almost certainly underestimate the magnitude of the liquidity shocks facing defaulters.

The importance of lifecycle effects are also apparent in Figure 7. Divorce and "disaster" drive comparable amounts of default between LTVs of 80 and 90. However, divorce hazard falls over the lifecycle, while the probability of a "diastrous" income shock does not. Hence older agents, who have higher LTVs, rarely get divorced, so divorce causes few defaults below LTVs of 70. These high-equity defaults are driven almost entirely by the "disastrous" income shock.

Another way to study the importance of default triggers in the model is to shut them off. In particular, divorce and family size as default triggers are unique to this paper, and were added to allow the model to match the aggregate default rate and the LTV distribution of defaulters. Re-estimating the model, setting the probability of divorce to 0 every period and minimum expenditures
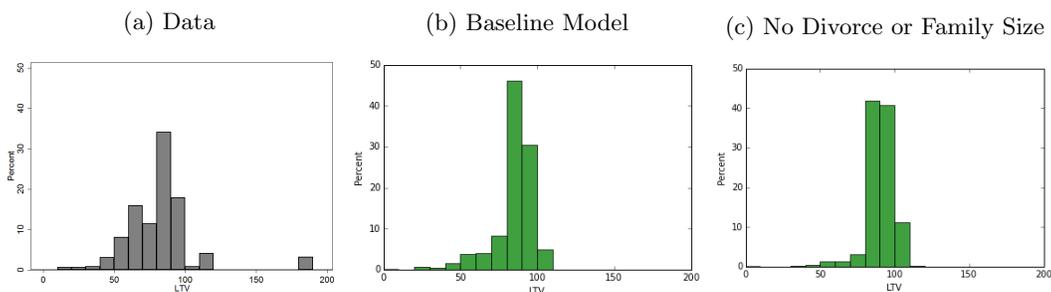
26

Figure 7: Default Triggers by LTV
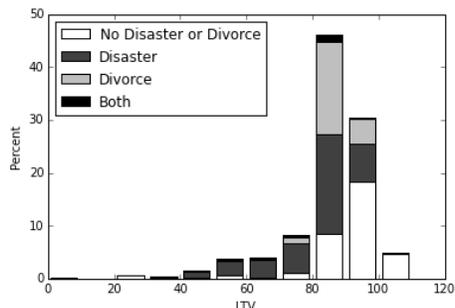
Notes: The Loan-to-Value Ratio is defined as 100 times the ratio of total mortgage debt outstanding to house value. "Disaster" denotes defaulting agents who are in the "disastrous" income state. "Divorce" denotes defaulting agents who received an exogenous divorce shock at the beginning of the period.

$\underline{x} = 0$, yields almost identical parameter estimates $\beta = .957$ and $\theta = .330$. However, it changes the output of the model radically. Figure 8 plots the LTV distribution of defaulters in the re-estimated model without divorce or family size against the same distribution in the baseline model and in the data.

Figure 8: Distribution of Defaulters' LTVs in the Data and Model

(a) Data   (b) Baseline Model   (c) No Divorce or Family Size

Notes: The Loan-to-Value Ratio is defined as 100 times the ratio of total mortgage debt outstanding to house value. Data are from the 1998 & 2001 SCF. In the model without divorce or family size, the probability of divorce is set to 0 in all periods, and necessary expenditures per family member $\underline{x}$ are set to 0.

Note that divorce and family size are critical in allowing the model to match the LTV distribution of defaulters. They are also critical in allowing the model to match the aggregate foreclosure rate. The targeted foreclosure rate is .50%; in the baseline model the foreclosure rate is .49%, while in the re-estimated model without divorce or family size it collapses to .31%.

Recourse   The model has so far assumed that mortgages are "non-recourse", i.e. they are secured only by the house. However, many states ("recourse" states) allow lenders to seize other assets of underwater defaulters, though they differ considerably in how practical this is. For a detailed discussion of recourse laws by state, see Ghent and Kudlyak (2011). In this subsection, I examine the effect of adding recourse to the model.

Empirically, Clauretie (1987) and Ghent and Kudlyak (2011) find that recourse has no effect on aggregate default rates. This may appear to suggest that recourse has no effect on individual default probabilities. However, several "double trigger" models of default, e.g. Quintin (2012), Laufer (2013), Hatchondo et al. (2014), and Corbae and Quintin (2014), note the potential for selection bias. Specifically, if recourse reduces individual default probabilities, it should also lower individual interest rates. In turn, this should expand the pool of homeowners to include borrowers who previously could not obtain or afford a mortgage. Since these less creditworthy borrowers are precisely the most likely to default, the effect of recourse on aggregate default rates in these models is theoretically ambiguous. However, they agree that recourse should lower default probabilities and interest rates, conditional on loan and borrower characteristics. In contrast, Mitman (2012) constructs a model of "strategic" default that also includes a bankrupty decision. In his model, recourse has almost no effect on the default rates of underwater homeowners, since they can always declare bankrupty to discharge themselves of deficiency judgments.

Interestingly, Ghent and Kudlyak (2011) find that empirically recourse does not lower default rates *conditional* on loan and borrower characteristics, and in fact *increases* conditional interest rates, even though it *does* lower the default rates of underwater homeowners. In this subsection, I add recourse to my model to examine whether it can explain these findings. I do this by assuming that, if underwater homeowners default, lenders seize any other assets they have at the time of foreclosure to cover the difference between the mortgage balance and the value of the home. I do not assume that any assets are exempt from recourse, that lenders must pay litigation costs to seize assets, or that defaulters can declare bankruptcy to discharge their debts to lenders. Therefore, it is likely that I exaggerate the effectiveness of recourse.

My first result is standard. I find that recourse acts as a powerful deterrent to underwater defaulters by raising the costs of default, which echoes the theoretical findings of Quintin (2012), Laufer (2013), Hatchondo et al. (2014), and Corbae and Quintin (2014), as well as the empirical results in Ghent and Kudlyak (2011). Figure 9 illustrates this result, by plotting the default decision rules of underwater homeowners as a function of cash on hand.

However, recourse *encourages* "effectively" underwater households to default in the model because it lowers the expected value of maintaining the mortgage. If the house depreciates, "effectively" underwater households will become actually underwater, and in recourse states will face the unpleasant choice of either maintaining an underwater mortgage or defaulting and losing assets to recourse. Therefore, under recourse, "effectively" underwater households have less incentive to avoid defaulting this period, and so are *more* likely to default. Figure 10 illustrates this result, by plotting the default decision rules of "effectively" underwater homeowners as a function of cash on hand.

Figure 9: Effect of Recourse on Underwater Default
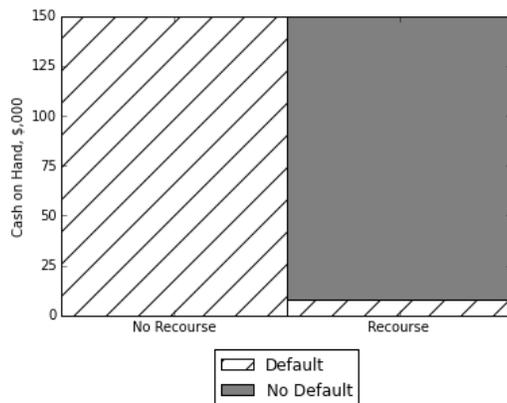
Notes: Decision rules shown for married agents with an LTV of 116 in the "disastrous"
income state, in a house worth $176,000 in the first period.

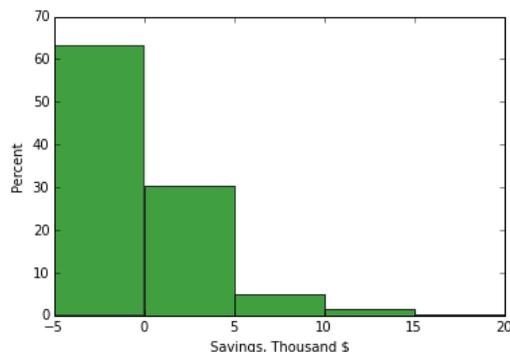Figure 10: Effect of Recourse on "Effectively" Underwater Default
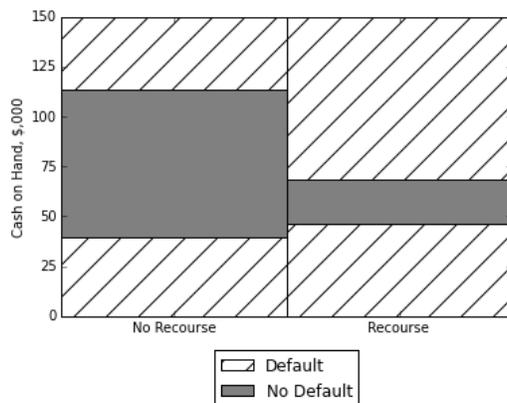
Notes: Decision rules shown for married agents with an LTV of 100 in the "disastrous"
income state, in a house worth $176,000 in the first period.

Finally, recourse has no effect on the decisions of abovewater defaulters, for two reasons. First, abovewater homeowners are not subject to recourse, since their house is worth more than the oustanding mortgage balance. Second, abovewater defaulters have very few assets to seize. The distribution of assets brought in to the next period, for defaulters in the no-recourse regime with LTVs below 90, is shown in Figure 11. Over 60% of these defaulters save no assets whatsoever; another 30% save less than $5,000. Without any legal costs of recourse or asset exemptions (which is what I assume), these agents could respond to expanded recourse simply by consuming these savings and still defaulting. With a small fixed legal cost of pursuing recourse (or asset exemption) of $5,000, over 90% of abovewater defaulters would be entirely unaffected, even if recourse were expanded to include them, because they have so little to seize. This is highly intuitive, since severe

liquidity constraints drive abovewater default.

Figure 11: Savings Distribution of Abovewater Defaulters

Introducing recourse to the model has virtually no effect on any aggregate outcome of interest.
It lowers foreclosure rates from .486% to .484%. The aggregate homeownership rate increases by
.05%, and interest rates decrease by less than one-tenth of one basis point.[28] This result is intuitive.
Recourse does nothing to address the liquidity constraints that drive the vast majority of defaults.
Recourse does prevent underwater homeowners with assets from defaulting, but these make up only
a small fraction of the population of mortgagors. Therefore, in aggregate, recourse is ineffective
at discouraging default, expanding homeownership, or lowering interest rates, consistent with the
empirical results in Ghent and Kudlyak (2011).

Effect of Equity on Default   Many empirical papers, e.g. Deng et al. (2000), Foote et al. (2008),
Gerardi et al. (2009), and Elul et al. (2010), show that default hazard rates fall with equity. These
findings have motivated many policy proposals, e.g. LTV caps on new or refinancing homeowners.
Hence, it is critical that estimates of the causal effect of equity on default hazard be precise.

Unfortunately, virtually all estimates in the literature come from loan-level datasets, and so
cannot control for household-level characteristics — such as income, wealth, or divorce hazard —
that also predict default. This may be problematic, since these characteristics are likely to be
correlated with equity. These papers typically can control for the LTV at loan origination, and
hazard models by construction control for the age of the mortgage, so the extent of the omitted
variable bias is not obvious.

This subsection attempts to estimate the size of this bias, using simulated data from the model.

---

[28]As in most structural models with mortgages, interest rates are not a state variable in my model. However, one
can still calculate the effective interest rate on the loan home buyers receive. For a given stream of promised future
mortgage payments, the smaller the loan, the higher the effective interest rate.

First, as is standard in the literature, I estimate a Cox proportional hazard model of default controlling only for the LTV at origination.[29] I then estimate a "household" hazard model of default that also controls for log income, a dummy for the "disastrous" income state, log wealth (if greater than \$1,000), marital status, divorce, and family size. So that changes in default hazard, rather than levels, are clear, I then normalize hazard rates so that they equal 1 at an LTV of 80. Relative hazard rates from these two models, as a function of LTV, are shown in Figure 12a.

Figure 12: Relative Hazard Rates, with and without Household Controls

(a) LTV at origination

(b) LTV at origination dummies
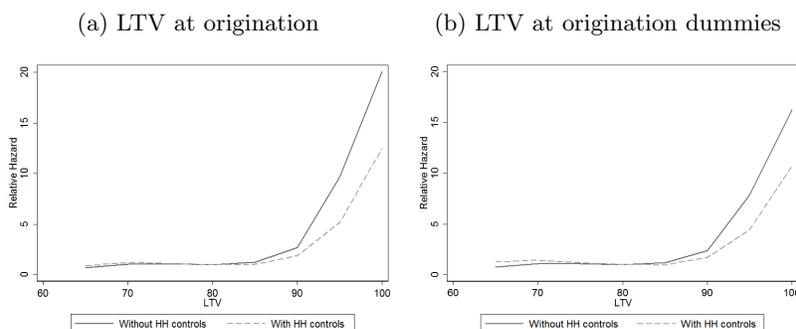
Roughly, Figure 12a shows that the default hazard rate at an LTV of 100 is 20.0 times higher than the default hazard rate at an LTV of 80, without controlling for household-level characteristics. However, controlling for these characteristics, the default hazard rate at an LTV of 100 is only 12.5 times as high as the hazard rate at 80. Hence, in the model, almost 40% of the estimated fall in default hazard from an LTV of 100 to 80 is due to omitted variables, e.g. income, wealth, and divorce risk (through age), that are correlated with equity but not caused by it.[30]

Much, but not all, of this correlation comes from borrower selection. Younger, poorer households choose higher LTV loans and are more likely to default. Figure 12a follows the literature standard by controlling for a linear term in LTV at origination. However, the effect of LTV at origination on default hazard (induced by covariates) is not linear. Figure 12b presents results from an almost identical hazard model that controls for bins of LTV at origination, rather than a simple linear term.[31] The default hazard rate in the baseline specification is considerably lower, though the default hazard rate in the model with household level controls is virtually unchanged. This is because non-linear terms in LTV at origination serve as a reasonable proxy for household-level characteristics.

[29]Empirical papers typically also control for other aggregate variables, such as interest rates and regional unemployment rates, that are not relevant in my structural model.

[30]Empirical results from the PSID, available upon request, are even stronger, but may not be credible. Proportional hazard models require fairly large samples, while the sample in the PSID is small.

[31]The bins are less than 80, by 5s up to 100, and above 100.

While these results using simulated data from the model are illuminating, they do not substitute for empirical work using high-quality, household level data. In particular, because the model does not distinguish between the price of its house and its value to the homeowner, the baseline model overestimates the decline in default hazard from LTVs of 100 to 80 relative to empirical work using property-level data, e.g. Foote et al. (2008). Estimating this hazard using household-level data should be a priority for future empirical work. If household-level data are not available, then non-linear terms in LTV at origination (and other observables, such as FICO score) should mitigate the bias induced by omitted variables, but may not eliminate it.

It is interesting to note that the model *overestimates* the causal effect of equity on default, while matching the empirical fact that the vast majority of defaulters have positive equity. This demonstrates that evidence that underwater households are particularly likely to default, e.g. Foote et al. (2008), should not be interpreted as evidence that most defaulters are underwater.

6 Conclusion

The existing mortgage default literature focuses almost exclusively on the "strategic " and "double trigger" paradigms of default, in which negative effective equity is a necessary condition for default. This paper provides empirical evidence that, in fact, the vast majority of defaulters are abovewater and that many defaults are driven by divorce and family size, rather than negative equity. This paper also constructs a structural model that matches the LTV distribution of defaulters with realistic parameters and income and family processes, demonstrating that abovewater default is entirely consistent with theory.

Abovewater default has profound implications. It provides clear identification of the elasticity of substitution between consumping and housing using only aggregate moments, which until now has been elusive. It suggests that loan- or property-level estimates of the effects of equity on default hazard should not be interpreted as entirely causal. It also provides a theoretical explanation for the empirical results in Ghent and Kudlyak (2011), who show that lender recourse — a policy that is often assumed to be effective in preventing defaults and expanding credit access — has almost no effect on individual default risk or interest rates, even though it prevents underwater default. Together, these lessons should provide useful guidance for future policy, as the effects of the recent housing crisis fade and the proportion of mortgagors who are underwater returns to normal (i.e. low) levels.[32]

---

[32]Zillow reports that 31.4% of mortgagors were underwater in the first quarter of 2012. By the second quarter of 2014, only 17% of mortgagors were underwater. This number is expected to fall further. See http://www.zillow.com/research/2014-q2-negative-equity-report-7465/

References

Ambrose, Brent W. and Charles A. Capone (1998) "Modeling the Conditional Probability of Foreclosure in the Context of Single-Family Mortgage Default Resolutions," *Real Estate Economics*, Vol. 26(3), pp. 391–429.

Bajari, Patrick, Phoebe Chan, Dirk Krueger, and Daniel Miller (2013) "A Dynamic Model of Housing Demand: Estimation and Policy Implications," *International Economic Review*, Vol. 54(2), pp. 409–442.

Bhutta, Neil, Jane Dokko, and Hui Shan (2010) "The Depth of Negative Equity and Mortgage Default Decisions," *Federal Reserve Board Finance and Economics Discussion Series*.

Black, Dan A., Natalia Kolesnikova, Seth G. Sanders, and Lowell J. Taylor (2013) "Are Children "Normal"?" *The Review of Economics and Statistics*, Vol. 95(1), pp. 21–33.

Campbell, John Y. and João F. Cocco (2014) "A Model of Mortgage Default," *Working Paper*.

Charles, Kerwin Kofi and Melvin Stephens (2004) "Job Displacement, Disability, and Divorce," *Journal of Labor Economics*, Vol. 22(2), pp. 489–522.

Chatterjee, Satyajit and Burcu Eyingungor (2014) "A Quantitative Analysis of the US Housing and Mortgage Markets and the Foreclosure Crisis," *Working Paper*.

Chen, Hui, Michael Michaux, and Nikolai Roussanov (2013) "Houses as ATMs? Mortgage Refinancing and Macroeconomic Uncertainty," *NBER Working Paper 19421*.

Clauretie, Terrence M. (1987) "The Impact of Interstate Foreclosure Cost Differences and the Value of Mortgages on Default Rates," *Real Estate Economics*, Vol. 15(3), pp. 152–167.

Cocco, João F. (2004) "Portfolio Choice in the Presence of Housing," *The Review of Financial Studies*, Vol. 18(2), pp. 535–567.

Cocco, João, Francisco J. Gomes, and Pascal J. Maenhout (2005) "Consumption and Portfolio Choice over the Life Cycle," *The Review of Financial Studies*, Vol. 18(2), pp. 491–533.

Corbae, Dean and Erwan Quintin (2014) "Leverage and the Foreclosure Crisis," *Forthcoming, Journal of Political Economy*.

Cubeddu, Luis and José-Víctor Ríos-Rull (2003) "Families as Shocks," *Journal of the European Economic Association*, Vol. 1(2/3), pp. 671–682.

Davis, Morris A. and François Ortalo-Magné (2011) "Household Expenditures, Wages, Rents," *Review of Economic Dynamics*, Vol. 14, pp. 248–261.

Davis, Morris A. and Erwan Quintin (2014) "Default when Current House Prices are Uncertain," *Working Paper*.

De Nardi, Mariacristina (2004) "Wealth Inequality and Intergenerational Links," *Review of Economic Studies*, Vol. 71, pp. 743–768.

Deng, Yongheng, John M. Quigley, and Robert Van Order (2000) "Mortgage Terminations, Heterogeneity, and the Excercise of Mortgage Options," *Econometrica*, Vol. 68(2), pp. 275–307.

Duffie, Darrell and Kenneth J. Singleton (1993) "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica*, Vol. 61(4), pp. 929–952.

Elul, Ron, Nicholas S. Souleles, Souphala Chomsisengphet, Dennis Glennon, and Robert Hunt (2010) "What "Triggers" Mortgage Default?" *The American Economic Review, Papers and Proceedings*, Vol. 100(2), pp. 490–494.

Farnham, Martin, Lucie Schmidt, and Purvi Sevak (2011) "House Prices and Marital Stability," *American Economic Review: Papers & Proceedings*, Vol. 101:3, pp. 615–619.

Fernández, Raquel and Joyce Cheng Wong (2013) "Divorce Risk, Wages, and Working Wives: A Quantitative Life-Cycle Analysis of Female Labor Force Participation," *Mimeo*.

Flavin, Marjorie and Shinobu Nakagawa (2008) "A Model of Housing in the Presence of Adjustment Costs: A Structural Interpretation of Habit Persistence," *The American Economic Review*, Vol. 98(1), pp. 474–495.

Foote, Chistopher, Kristopher Gerardi, Lorenz Goette, and Paul Willen (2010) "Reducing Foreclosures: No Easy Answers," in Daron Acegmoglu, Kenneth Rogoff, and Michael Woodford eds. *NBER Macroeconomics Annual 2009, Volume 24*: University of Chicago Press.

Foote, Christopher L., Kristopher Gerardi, and Paul. S. Willen (2008) "Negative Equity and Foreclosure: Theory and Evidence," *Journal of Urban Economics*, Vol. 64, pp. 234–245.

Garriga, Carlos and Don E. Schlagenhauf (2010) "Home Equity, Foreclosures, and Bail-out Programs during the Subprime Crisis," *Working paper*.

Gerardi, Kristopher, Kyle F. Herkenhoff, Lee E. Ohanian, and Paul S. Willen (2013) "Unemployment, Negative Equity, and Strategic Default," *Federal Reserve Bank of Atlanta Working Paper Series*.

Gerardi, Kristopher, Adam Hale Shapiro, and Paul S. Willen (2009) "Decomposing the Foreclosure Crisis: House Price Depreciation versus Bad Underwriting," *Federal Reserve Bank of Atlanta Working Paper*.

Ghent, Andra C. and Marianna Kudlyak (2011) "Recourse and Residential Mortgage Default: Evidence from US States," *The Review of Financial Studies*, Vol. 24(9), pp. 3139 – 3186.

Granziera, Eleonora and Sharon Kozicki (2012) "House Price Dynamics: Fundamentals and Expectations," *Bank of Canada Working Paper*, Vol. 12.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2013) "The Determinants of Attitudes Toward Strategic Default on Mortgages," *The Journal of Finance*, Vol. 68(4), pp. 1473–1515.

Guren, Adam M. (2014) "The Causes and Consequences of House Price Momentum," *Working Paper*.

Guren, Adam and Tim McQuade (2013) "How Do Foreclosures Exacerbate Housing Downturns?" *Working Paper*.

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song (2014) "What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk?" *Working Paper*.

Gyourko, Joseph and Joseph Tracy (2013) "Unemployment and Unobserved Credit Risk in the FHA Single Family Mortgage Insurance Fund," *NBER Working Paper 18880*.

Halket, Jonathan and Santhanagopalan Vasudev (2014) "Saving Up or Settling Down: Home Ownership Over the Life Cycle," *Review of Economic Dynamics*, Vol. 17, pp. 345–366.

Hanushek, Eric. A. and John M. Quigley (1980) "What is the Price Elasticity of Housing Demand?" *The Review of Economics and Statistics*, Vol. 62(3), pp. 449–454.

Hatchondo, Juan Carlos, Leonardo Martinez, and Juan M. Sánchez (2014) "Mortgage Defaults and Prudential Regulations in a Standard Incomplete Markets Model," *Working Paper*.

Haughwout, Andrew F. and Ebiere Okah (2009) "Below the Line: Estimates of Negative Equity Among Nonprime Mortgage Borrowers," *FRBNY Economic Policy Review*.

Head, Allen, Huw Lloyd-Ellis, and Hongfei Sun (2014) "Search, Liquidity, and the Dynamics of House Prices and Construction," *American Economic Review*, Vol. 104(4), pp. 1172–1210.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante (2010) "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of Political Economy*, Vol. 118(4), pp. 681–722.

Hellerstein, Judith K., Melinda Sandler Morrill, and Ben Zou (2013) "Business Cycles and Divorce: Evidence from Microdata," *Economics Letters*, Vol. 118, pp. 68–70.

Herkenhoff, Kyle F. and Lee E. Ohanian (2013) "Foreclosure Delay and U.S. Unemployment," *Working Paper*.

Jeske, Karsten, Dirk Krueger, and Kurt Mitman (2013) "Housing, Mortgage Bailout Guarantees and the Macro Economy," *Journal of Monetary Economics*, Vol. 60, pp. 917–935.

Kaplan, Greg and Gianluca Violante (2012) "A Model of the Consumption Response to Fiscal Stimulus Payments," *Mimeo*.

Kau, James B., Donald C. Keenan, and Taweon Kim (1994) "Default Probabilities for Mortgages," *Journal of Urban Economics*, Vol. 35, pp. 278–296.

Khandani, Amir E., Andrew W. Lo, and Robert C. Merton (2013) "Systemic Risk and the Refinancing Ratchet Effect," *Journal of Financial Economics*, Vol. 108, pp. 29–45.

Kiel, Katherine A. and Jeffrey E. Zabel (1999) "The Accuracy of Owner-Provided House Values: The 1978 - 1991 American Housing Survey," *Real Estate Economics*, Vol. 27(2), pp. 263–298.

Krainer, John (2001) "A Theory of Liquidity in Residential Real Estate Markets," *Journal of Urban Economics*, Vol. 49, pp. 32–53.

Laufer, Steven (2013) "Equity Extraction and Mortgage Default," *Finance and Economics Discussion Series*.

Li, Wenli, Haiyong Liu, Fang Yang, and Rui Yao (2014a) "Housing over Time and over the Life Cycle: A Structural Estimation," *Working Paper*.

Li, Wenli, Costas Meghir, and Florian Oswald (2014b) "Consumer Bankruptcy and Mortgage Default," *Working Paper*.

Love, David A. (2010) "The Effects of Marital Status and Children on Savings and Portfolio Choice," *The Review of Financial Studies*, Vol. 23(1), pp. 385–432.

Lovenheim, Michael F. and Kevin J. Mumford (2013) "Do Family Wealth Shocks Affect Fertility Choices? Evidence from the Housing Market," *The Review of Economics and Statistics*, Vol. 95(2), pp. 464–475.

Mitman, Kurt (2012) "Macroeconomic Effects of Bankruptcy & Foreclosure Policies," *Working Paper*.

Pakes, Ariel and David Pollard (1989) "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, Vol. 57(5), pp. 1027–1057.

Pennington-Cross, Anthony (2003) "Subprime & Prime Mortgages: Loss Distributions," *OFHEO Working Paper 03-1*.

———— (2010) "The Duration of Foreclosures in the Subprime Mortgage Market: A Competing Risks Model with Mixing," *Journal of Real Estate and Financial Economics*, Vol. 40, pp. 109–129.

Piazzesi, Monika, Martin Schneider, and Selale Tuzel (2007) "Housing, Consumption, and Asset Pricing," *Journal of Financial Economics*, Vol. 83, pp. 531–569.

Qi, Min and Xiaolong Yang (2009) "Loss Given Default of High Loan-to-Value Residential Mortgages," *Journal of Banking and Finance*, Vol. 33, pp. 788–799.

Quintin, Erwan (2012) "More Punishment, Less Default?" *Annals of Finance*, Vol. 8, pp. 427–454.

Schelkle, Thomas (2014) "Mortgage Default During the U.S. Mortgage Crisis," *University of Cologne Working Paper Series in Economics*, Vol. 72.

Shiller, Robert (2008) "Understanding Recent Trends in House Prices and Homeownership," *Housing, Housing Finance and Monetary Policy, Jackson Hole Conference Series*, pp. 85–123.

Stokey, Nancy L. (2009) "Moving Costs, Nondurable Consumption, and Portfolio Choice," *Journal of Economic Theory*, Vol. 144, pp. 2419–2439.

Voena, Alessandra (2012) "Yours, Mine, and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples?" *Working Paper*.

Wardle, Lynn and Laurence Nolan (2011) *Family Law in the USA*: Kluwer Law International.

Weiss, Yoram and Robert J. Willis (1997) "Match Quality, New Information, and Marital Dissolution," *Journal of Labor Economics*, Vol. 15(1), pp. S293–S329.

Wheaton, William C. (1990) "Vacancy, Search, and Prices in a Housing Market Matching Model," *Journal of Political Economy*, Vol. 98(6), pp. 1270–1292.

7 Appendix

This appendix examines the distribution of LTVs among defaulters (which I denote $\mathcal{D}$) from a number of sources and time periods. The evidence suggests that around 80 to 90% of defaulters have positive

equity in "normal" times, while 30 to 50% of defaulters were abovewater during the recent mortgage crisis. However, there are concerns with every dataset considered, and these estimates should only be viewed as approximations. A key priority for future empirical work should be to examine how $\mathcal{D}$ has evolved over time.

The assumption that negative (effective) equity is a necessary condition for default is ubiquitous in the literature. This assumption is often justified with evidence that underwater homeowners are particularly likely to default, which has been shown by many papers including Deng et al. (2000), Foote et al. (2008), Gerardi et al. (2009), Haughwout and Okah (2009), Elul et al. (2010), and Gerardi et al. (2013). However, many of these same papers show that abovewater homeowners default with positive probability, e.g. Foote et al. (2008), Gerardi et al. (2009), Haughwout and Okah (2009), Elul et al. (2010), and Gerardi et al. (2013). This means that negative equity is not a necessary condition for default. Indeed, if most homeowners are abovewater, then most defaulters may be abovewater, even though abovewater homeowners are less likely to default individually.[33] This is the case in normal times.

Homeowners with small amounts of equity may still be "effectively underwater", i.e. they could lose money by selling their home after accounting for transaction costs. Estimates of these transaction costs are typically near 6% (Bajari et al. (2013)), and can be as as high as 15% (Li et al. (2014a)) of the house value. However, estimates near 15% come from structural models of housing demand (not default) that need to explain low levels of mobility in the data. These high estimates are understood to include psychic costs of selling in addition to actual financial costs, which Li et al. (2014a) place at 10% of the house value. Regardless, households with LTVs as low as 50 are observed to default with an empirically significant probability, so transaction costs alone cannot explain abovewater default. This is discussed further below.

Conversely, unconstrained homeowners may prefer selling their home to defaulting even if this entails a small financial cost, since there are many nonfinancial costs associated with default. Defaulters often lose meaningful access to credit markets for years. They may have other assets seized by the mortgage lender, if they are underwater in a recourse state. There is also evidence of significant psychic costs of default; in survey data, Guiso et al. (2013) find that 82% of respondents believe that strategic default is morally wrong. Therefore, the LTV at which a homeowner would financially "break even" by choosing to sell rather than default is a lower bound of the LTV at which they optimally choose not to default, even in a model where house sales are instantenous.

Regardless of these concerns, formulating effective policy requires a good understanding of

---

[33]For example, Elul et al. (2010) report that, in their panel sample — which is chosen specifically to maximize the incidence of negative equity — less than 4% of homeowners are underwater. Therefore abovewater homeowners outnumber underwater homeowners in their sample by more than 24 to 1.

the distribution $\mathcal{D}$. For example, homeowners with LTVs of 150 and 98 may both be effectively underwater. However, as shown in Section 5, recourse can discourage the first from defaulting while encouraging the second.

Data before 2002 Evidence from before the housing boom (and subsequent bust) of the 2000s strongly suggests that most defaulters had positive equity. Ambrose and Capone (1998) estimate that, in their sample of defaulted FHA mortgages from 1988 through 1994, 90.3% of foreclosed homeowners had positive equity. Using Freddie Mac data from 1976-1992, Deng et al. (2000) estimate that defaulters have a 9.18% probability of being underwater, on average. Pennington-Cross (2003) finds that, in a sample of Fannie Mae and Freddie Mac foreclosed homeowners from 1995-1999, 99.7% had positive equity while 38.1% had LTVs below 80. Unfortunately, the datasets in these papers are loan-level. Therefore their LTV estimates are based only on the first lien against a property, and so they will seriously understimate the combined LTV of homeowners with other liens. However, in the 1998 and 2001 SCF, less than 20% of defaulting homeowners have either a second mortgage or a home equity line of credit. Hence, these papers are still strong evidence that (in "normal" times) most defaulters have positive equity.

However, since the SCF has data on second and third liens, its data may be more informative regarding the distribution $\mathcal{D}$. I replicate Figure 1 here, which also includes the LTV distribution of mortgagors for comparison. There is an intuitive explanation for why abovewater homeowners constitute the majority of defaulters: they constitute the vast majority of mortgagors.

Figure 13: Distribution of Loan-to-Value Ratios of Mortgagors and Defaulters
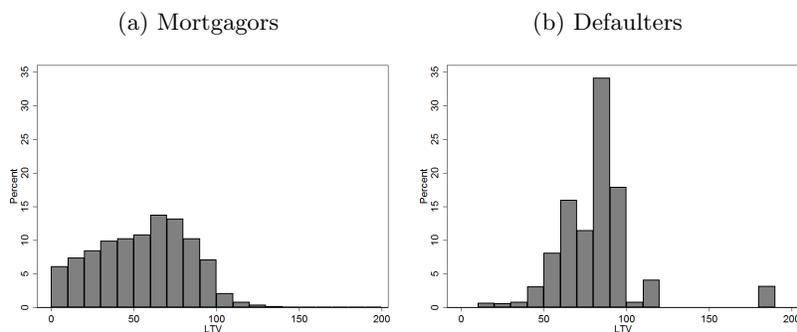
(a) Mortgagors　　　　　　　　　(b) Defaulters

Notes: The Loan-to-Value Ratio is defined as 100 times the ratio of total mortgage debt outstanding to house value. Data are from the 1998 & 2001 SCF. The SCF does not ask how many months respondents are behind on mortgage payments, so defaulters are assumed to be households who: (1) were two months behind on debt payments within the last year, and (2) are currently behind schedule on their mortgage. Defaulters with LTVs below 10 or above 250 are dropped as outliers.

There are concerns with SCF data. House values are self-reported in the SCF (and other surveys, like the PSID). They will be biased upwards if home owners systematically overestimate the value of their homes. Research typically finds that, on average, homeowners overestimate the value of

their home by 5-10% (e.g. Kiel and Zabel (1999)). This bias is important to keep in mind, but the homeowner's valuation of the home is the input into her decision to default on it. The market value of the home only matters if homeowners interact with the market, which they may do either by trying to sell the home or by refinancing the mortgage. But if homeowners interact with the market, they are likely to learn the market value of their home. Most importantly, as discussed earlier the estimates in Ambrose and Capone (1998), Deng et al. (2000), and Pennington-Cross (2003) are from loan-level datasets and are very similar. Therefore it does not seem that this concern is quantitatively significant.

Another concern is that the distribution of LTVs among defaulters will differ from the distribution of LTVs among foreclosed homeowners, if homeowners become more likely to transition from default to completed foreclosure as LTV increases. Available evidence suggests that this concern is not as important as it might seem.[34] Ambrose and Capone (1998) find that, in their sample, 31.2% of defaulters with positive equity lose their home to foreclosure, while 42.7% of defaulters with negative equity do. The sample in Pennington-Cross (2003) consists of foreclosed homeowners. Further evidence from the mortgage crisis, presented below, confirms this view.

Data from the Mortgage Crisis   Much more data on mortgages and default is available since the beginning of the mortgage crisis. I present some of this data here.

Due to the surge in mortgage credit before the crisis, as documented by Laufer (2013) and Chen et al. (2013), and the subsequent collapse in house prices, far more homeowners and defaulters had negative equity during the crisis than in earlier periods. Figure 14 compares the distribution $\mathcal{D}$ in the 1998-2001 SCF with the same distribution in 2010. Though a much higher proportion of defaulters had negative equity in 2010 than in 1998 and 2001, many — perhaps most — still had positive equity.

Again, one concern with cross-sectional SCF data is that it only covers defaulters, who may or may not lose their home to foreclosure. However, differences in hazard rates into completed foreclosure do not seem to be a first-order concern for defaulters with LTVs above 60.[35] In a sample of seriously delinquent subprime loans, Pennington-Cross (2010) shows that the hazard rate into completed foreclosure is virtually constant between LTVs of 80 and 110. In his sample, even homeowners with LTVs between 50 and 60 have over one quarter the hazard rate into completed foreclosure of underwater homeowners. The PSID (which only has default data beginning in 2008) and the 2007-2009 SCF panel include data on homeowners who were delinquent on their mortgage at the time of the interview, and who subsequently lost their home to foreclosure. The LTV distribution

---

[34]One potential explanation is that mortgage lenders have greater incentives to avoid foreclosure as LTV increases.
[35]However, it does seem that defaulters with very low LTVs (below 50) rarely lose their home to foreclosure.
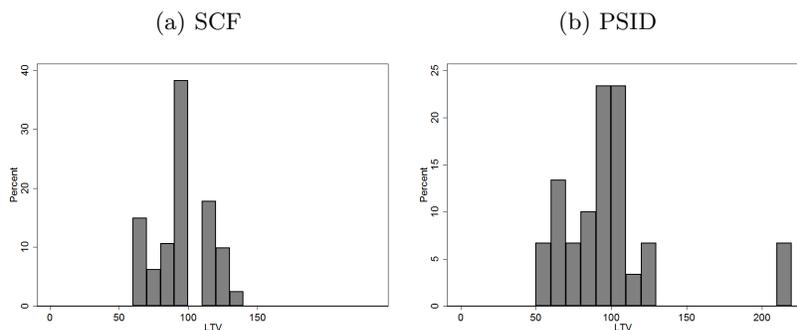
Figure 14: Distribution of Loan-to-Value Ratios of Defaulters Over Time
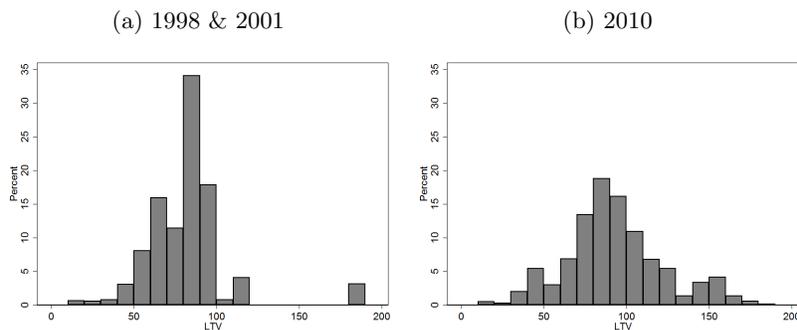
(a) 1998 & 2001   (b) 2010

Notes: The Loan-to-Value Ratio is defined as 100 times the ratio of total mortgage debt outstanding to house value. Data are from the SCF. The SCF does not ask how many months respondents are behind on mortgage payments, so defaulters are assumed to be households who: (1) were two months behind on debt payments within the last year, and (2) are currently behind schedule on their mortgage.

of these homeowners is shown in Figure 15. Figure 15 draws from a very small sample, consisting of just 8 families in the SCF and 26 families in the PSID, and should be treated with caution. Still, it is entirely consistent with the findings in Pennington-Cross (2010).

Figure 15: Distribution of Loan-to-Value Ratios of Foreclosed Homeowners

(a) SCF   (b) PSID

Notes: The Loan-to-Value Ratio is defined as 100 times the ratio of total mortgage debt outstanding to house value. Data shown are from homeowners who were delinquent on their mortgage at the time of the first survey (2007 in the SCF, 2008 in the PSID), and had lost their home to foreclosure by the following survey (2009 in the SCF, 2010 in the PSID). To maximize sample size, the PSID sample also includes homeowners who were current on their mortgage during the 2008 survey, but thought they were "very likely" to fall behind in the future, and subsequently lost their home to foreclosure.

Thus, all available evidence from the PSID and SCF suggests that, even at the height of the mortgage crisis, over half of defaulters had positive equity. Still, given the concerns about these datasets described above, this is probably an upper bound on the true figure.

A useful lower bound comes from Haughwout and Okah (2009). In their sample of subprime or Alt-A ("nonprime") securitized mortgages in December 2008, they estimate that 31% of homeowners that are in foreclosure, or have been foreclosed on, have positive equity. Since nonprime mortgagors are more likely to have negative equity than prime mortgagors, I view this likely to be a lower bound on the true fraction of abovewater foreclosed homeowners in December 2008.

In a paper on the recent mortgage crisis, Laufer (2013) (whose measure of default is the beginning of foreclosure) notes the high amount of abovewater default in his data, and writes: "In the data, there is a considerable amount of default among homeowners with positive equity. Homeowners with an LTV ratio less than .75 default at a rate of 0.14% per quarter, which is more than half the total default rate during the early part of the sample."

Finally, Zillow predicts individual property values, and they have mortgage data from TransUnion that covers all liens against a property. Their property-level data is not publicly available, and their equity reports begin only in 2013. However, they report that in the first quarter of 2014, 1.4% of mortgagors were both seriously delinquent (90 days delinquent) and had negative equity.[36] The New York Federal Reserve reports that in the same quarter, 3.7% of mortgagors were seriously delinquent.[37] Hence, 62% of seriously delinquent mortgagors were abovewater in the first quarter of 2014. This suggests that the proportion of defaulters who are abovewater is returning to its high pre-crisis level.

---

[36]http://www.zillow.com/research/2014-q1-negative-equity-report-6937/
[37]http://www.newyorkfed.org/householdcredit/2014-q1/data/pdf/HHDC_2014Q1.pdf