Asymmetric Information and the Link Between Leverage and Mortgage Default

Christopher Hansman*

September 15, 2017

Abstract

Borrowers with large mortgages relative to their home values are more likely to default. This paper asks whether this correlation is due to moral hazard—larger balances causing borrowers to default—or adverse selection—ex-ante risky borrowers choosing larger loans. To separate these information asymmetries, I exploit a natural experiment resulting from (i) the unique contract structure of Option Adjustable Rate Mortgages and (ii) the unexpected divergence, during the 2008 crisis, of two financial indices used to determine interest rate adjustments for these loans. I find that moral hazard is responsible for 60-70 percent of the baseline correlation between leverage and default, but adverse selection explains the remaining 30-40 percent. I construct and calibrate a simple model of mortgage choice and default with asymmetric information to highlight the policy tradeoff informed by my estimates. I show that optimal regulation of mortgage leverage must weigh losses from defaults against under-provision of credit due to adverse selection.

*Columbia University. Email: c.hansman@columbia.edu. I am indebted to Pierre-André Chiappori, Jonas Hjort, Wojciech Kopczuk, and Bernard Salanié for their advice and guidance. I would additionally like to thank Eduardo Dávila, François Gerard, Arpit Gupta, Henrik Kleven, Rocco Macchiavello, Dan O'Flaherty, Tomasz Piskorski, Miikka Rokkanen, Christoph Rothe, Johannes Spinnewijn, Miguel Urquiola, Eric Verhoogen and members of the microeconomics colloquia at Columbia for their helpful comments and suggestions. This work has been supported by the National Science Foundation (SES-1559460) and by award # 98-16-02 from the Russell Sage Foundation. Any opinions expressed are mine alone and should not be construed as representing the opinions of the National Science Foundation or the Russell Sage Foundation.

1 Introduction

The historic rise in household debt in the early 2000s is central to many narratives of the financial crisis, and a key component of these accounts is the role of mortgage leverage.[1] Leading up to 2008, borrowers increasingly took on large mortgages relative to the values of their homes. As housing prices dropped, highly leveraged borrowers were the most likely to default.

This paper separates two potential explanations for the correlation between leverage and default. The first, sometimes called moral hazard,[2] is a causal effect: if housing prices fall, high balances may lead borrowers to default. The alternative is adverse selection: ex-ante riskier borrowers prefer larger loans. Despite a substantial theoretical literature examining these two classical information asymmetries in credit markets, distinguishing moral hazard from adverse selection remains a fundamental challenge for empirical work. Yet the distinction between the two is crucial for mortgage policy.[3] As I show, a policymaker that attributes the correlation solely to moral hazard will (i) *overestimate* the reduction in defaults generated by regulations on leverage and (ii) *underestimate* a significant source of welfare losses, as adverse selection entails safe borrowers taking inefficiently small loans to differentiate themselves from riskier types.

My research design exploits a natural experiment generated by the unique contract structure of Option Adjustable Rate Mortgages (Option ARMs), which traditionally have interest rate adjustments tied to either LIBOR or Treasury rates. The unexpected divergence of these two indices during the 2008 crisis caused borrowers who chose otherwise identical contracts to owe substantially different amounts ex-post. This variation allows me to (i) identify moral hazard effects by comparing borrowers with identical initial leverage choices and different realized balances, and (ii) document adverse selection by comparing borrowers with different initial leverage choices but the same realized balance.

The paper is organized in three parts. First, as a guide, I propose a model of asymmetric information in mortgage markets.[4] This model clarifies the sources of adverse selection and moral hazard and highlights the consequences for contracts in equilibrium. Next, I use the natural experiment described above to empirically disentangle adverse selection from moral hazard in a large sample of Option ARMs. Finally, I suggest and calibrate a simple structural model to quantify the contrast-

[1] The ratio of household debt to disposable personal income (DPI) peaked in late 2007 at 1.24, up from a historical average of 0.71. See the Financial Accounts of the United States. Mian and Sufi (2015) highlight the role of household debt in the financial crisis, Garriga and Schlagenhauf (2009), Corbae and Quintin (2015), Mayer, Pence and Sherlund (2009) and Campbell and Cocco (2015) discuss the role of mortgage leverage explicitly.

[2] This terminology in credit markets is used, for example, by Adams, Einav and Levin (2009).

[3] The empirical distinction also informs our understanding of the equilibrium form of mortgage contracts. Do lenders require down payments to ensure that borrowers repay ex-post (moral hazard) or to solve an ex-ante screening problem by sorting borrowers with different unobserved risk types (adverse selection)? Berger, Frame and Ioannidou (2011) provide recent empirical evidence on this question in credit markets.

[4] The model follows a substantial theoretical literature and draws particularly from Brueckner (2000) as well as Jaffee and Russell (1976).

ing welfare implications of each asymmetry. As an application of the model, I evaluate a policy that explicitly attempts to limit defaults by restricting leverage: a loan-to-value (LTV) cap.[5]

The theoretical framework, which begins the paper, clarifies the sources of asymmetric information in mortgage markets. In the model, moral hazard is the result of limited access to effective recourse.[6] In practice, mortgage lenders face severe constraints in recovering any loan balance beyond the value of the home itself. As a result, lenders are unable to write contracts that prevent borrowers from choosing to default when it is financially beneficial to do so. The default choice, in turn, may respond endogenously to the loan balance.

Adverse selection arises from heterogeneity across borrowers in their willingness to take advantage of the option to default. A large literature (e.g. Deng, Quigley and Van Order, 2000) suggests that there are significant differences across borrowers in this domain. Some borrowers walk away as soon as the home is worth less than the mortgage balance, whereas others choose not to default until the balance significantly outweighs the value of the home. If borrowers are privately aware that they are less likely to pay the loan back—a clear determinant of risk to the lender—they will prefer larger loans. This framework can be interpreted as a model of selection on (ex-post) moral hazard, as in Einav et al. (2013). Borrowers' demand for large loans is a function of their privately observed propensity to exercise the default option.

After presenting the model, I turn to disentangling moral hazard from adverse selection. The logic of the exercise comes in recognizing that, as in Karlan and Zinman (2009), the two give distinct empirical predictions. Adverse selection implies a positive correlation between the initial loan size and default, regardless of the balance the borrower actually faces. To identify this effect, the ideal experiment would reassign a large sample of borrowers who have endogenously chosen different loan sizes to identical contracts. With equal balances, any remaining correlation between default and the initial loan choice is attributable to adverse selection.

Conversely, moral hazard implies a positive correlation between a borrower's balance and default, regardless of the initial loan size. The ideal experiment would take a set of borrowers choosing identical loans and randomly assign each borrower a different balance. Any relationship between default and the randomly assigned balance identifies a moral hazard effect.

The natural experiment I utilize features the key property of both experiments: an exogenous change in borrowers' balances *after* the initial contract choice. I isolate changes in ex-post balances that result from plausibly exogenous difference-in-difference variation in monthly interest rates. The

[5]LTV caps limit the size of an initial loan relative to the home value, e.g. 90 percent. They are common worldwide and have recently been considered in the literature, for example by Gete and Reher (2015).

[6]While there is variation across states in recourse laws, the majority of the loans studied here are in states with limited or no recourse. For example, California, the most heavily represented state in my sample, does not allow deficiency judgments for owner-occupied homes (Pence, 2006).

variation itself comes as the result of a fine contract detail of all adjustable rate mortgages—the financial index used as a proxy for the cost of funds to the lender (typically a LIBOR or Treasury rate). While there was little reason for a borrower to prefer one index to another when taking a mortgage, the spread between the two increased significantly in late 2007. This led borrowers with otherwise identical loans to face a unique sequence of interest rates as a result of the index they chose and the origination month of their loan.

Two characteristics of Option ARMs translate this interest rate variation to changes *only* in borrowers' balances: fixed payment schedules and variable interest rates. For most other adjustable rate mortgages, the first order impact of an interest rate increase is not a change in the balance owed but rather a rise in the monthly payment. However, because payments are fixed for Option ARMs, at least in the first five years, excess interest accrual is absorbed directly into the loan balance. I use this variation to directly identify the causal effect of borrowers' balances on default—the moral hazard effect—and subsequently back out the role of adverse selection.

I find robust evidence that both moral hazard and adverse selection are present in the mortgage market. I estimate that moral hazard is responsible for 60-70 percent of the baseline correlation between leverage and default, while adverse selection is responsible for the remaining 30-40 percent. The moral hazard effect is directly policy relevant, quantifying how effective ex-post regulations that reduce balances are in preventing defaults.[7] My estimates imply, for example, that a 10-point reduction in a borrower's LTV 24 months after origination would reduce the average probability of default by over 4 percentage points. The policy implications of adverse selection are more difficult to determine. Ex-ante restrictions on mortgage contracts may have profoundly different impacts on equilibrium with and without adverse selection, but there is no standard framework to evaluate such regulations. Even the appropriate characterization of equilibrium in competitive contexts with adverse selection is controversial, and equilibria may fail to exist under conventional definitions.[8]

My final step is to calibrate and simulate a simple structural model to quantify the welfare implications of ex-ante regulation. To ensure the existence of equilibrium, I use the robust equilibrium concept recently proposed by Azevedo and Gottlieb (2016). I consider, as an example, the impact of a reduced LTV cap. I find that this policy is effective in limiting defaults, but the effect is smaller than a naive regulator—one who attributes the full correlation between leverage and default to moral hazard—would expect. Furthermore, for such a regulator, the presence of adverse selection generates significant unexpected welfare losses due to knock-on effects. While borrowers initially above the cap are mechanically forced to take smaller loans, the regulation propagates through the whole distribution: those below the cap also choose to take smaller loans in order to maintain a separation from

[7]e.g. the Home Affordable Refinance Program Principal Reduction Alternative (HAMP PRA).
[8]e.g. Rothschild-Stiglitz.

riskier types. Appropriately accounting for adverse selection, I estimate that default externalities on the order of $313,000 per default are necessary to make a reduction in the LTV cap from 100 to 90 welfare neutral. A naive regulator would underestimate this by 40 percent.

This paper's foremost contribution is to the growing empirical literature on asymmetric information in credit markets.[9] Complementing Karlan and Zinman (2009), a number of influential papers attempt to distinguish between adverse selection and moral hazard by exploiting ex-ante variation—experimental, regulatory, or institutional—in the set or shape of contracts offered. These include Ausubel (1999) and Agarwal, Chomsisengphet and Liu (2010) on the US credit card market; Adams, Einav and Levin (2009) and Einav, Jenkins and Levin (2012, 2013) on subprime auto loans; and Dobbie and Skiba (2013) on payday lending. However, separately identifying moral hazard effects in these contexts requires an assumption about why the relevant variation in ex-ante contracts does not also generate selection of borrowers on unobservables. To circumvent such assumptions, I isolate ex-post variation in the loan balance that is *unknown* to borrowers when selecting contracts.[10] Further, I propose and simulate a framework to evaluate policy in the presence of these asymmetries.

I also add to the papers above by studying the largest and arguably most important consumer debt market in the United States.[11] A small handful of empirical papers explicitly consider information asymmetries in mortgage markets, including Edelberg (2004), who uses structural assumptions to test for adverse selection and moral hazard in a broad class of consumer debts, and Ambrose, Conklin and Yoshida (2015) and Jiang, Nelson and Vytlacil (2014), who consider selection into and within low documentation mortgages.[12] Despite the significance of the mortgage market, the well-documented importance of screening in mortgage lending (e.g. Keys et al., 2010), and the quantity of theoretical work on information asymmetries (e.g. Brueckner, 2000; Dunn and Spatt, 1988; Stanton and Wallace, 1998; Harrison, Noordewier and Yavas, 2004; Chari and Jagannathan, 1989), attempts to cleanly separate adverse selection and moral hazard are relatively rare.

The estimated moral hazard effect directly contributes to the literature on the causes of mortgage

[9]This paper is also heavily indebted to broader empirical work on asymmetric information in insurance and other markets. Particularly Chiappori and Salanié (2000), Cardon and Hendel (2001), Finkelstein and Poterba (2004), Finkelstein, McGarry et al. (2006), Finkelstein and Poterba (2014), Hendren (2013), as well as recent work examining the welfare implications of information asymmetries such as Einav, Finkelstein and Cullen (2010), Einav et al. (2013), and Einav, Finkelstein and Schrimpf (2010).

[10]Conceptually, I build on Karlan and Zinman (2009), who experimentally generate ex-post variation in balances in an unsecured debt market. While this environment closely approximates the canonical Stiglitz and Weiss (1981) framework, the majority of lending is secured by some form of collateral. A large theoretical literature (e.g. Bester, 1985) shows that the use of collateral in credit contracts has significant implications for both welfare and the expression of adverse selection in equilibrium. By screening borrowers using contracts that differ along two dimensions—interest rates and collateral amounts—lenders can avoid the credit rationing that characterizes unsecured lending. In contrast to Karlan and Zinman (2009), I explicitly study the richer contract space of collateralized lending.

[11]Mortgage balances represented 68 percent of consumer debt in the first quarter of 2016. See the Federal Reserve Bank of New York's May 2016 Quarterly Report on Household Debt and Credit.

[12]There is also related work in the home equity lending market, in particular Agarwal et al. (2011), who explore dynamic relationships, and Agarwal, Chomsisengphet and Liu (2016), who follow the strategy of Adams, Einav and Levin (2009).

default, in which the role of home equity is a major concern. Vandell (1995) provides an overview of early research on borrowers' exercise of the default option. More recent work, including Bajari, Chu and Park (2008), Foote, Gerardi and Willen (2008), Elul et al. (2010), Bhutta, Shan and Dokko (2010), and Gerardi et al. (2015), has stressed the joint importance of triggers such as liquidity and job loss alongside home equity in mortgage default. However, the majority of this literature identifies the impact of home equity on default using variation that results from changes in local home prices. I provide a new source of borrower-level variation in home equity that avoids the potential for measurement error and other endogeneity concerns inherent to the use of home price variation.

The identification strategy complements a series of papers investigating the impacts of interest rate resets on delinquency and other outcomes for borrowers with adjustable rate mortgages. This includes Fuster and Willen (2012), Tracy and Wright (2012), Keys et al. (2014), Di Maggio, Kermani and Ramcharan (2014), and particularly Gupta (2016), who also utilizes the distinction between different indices. Because those papers examine more traditional adjustable rate mortgages, none are able to identify the impacts of loan liability on default, focusing instead on the liquidity impacts of monthly payment shocks that typically accompany rate resets. My primary innovation comes in developing a research design that cleanly translates interest rate resets into variation in borrowers' balances.

The paper is structured as follows: Section 2 lays out key definitions and presents a model of information asymmetries in the mortgage market. Section 3 provides background information on Option ARMs and the data used in the paper. Section 4 describes the contracts offered to borrowers and provides initial tests for information asymmetries. Sections 5 and 6 present the empirical strategy and results, respectively. Section 7 shows the results of simulations, and Section 8 concludes.

2 Definitions and a Model of Information Asymmetries

In this section, I define adverse selection and moral hazard as they pertain to the relationship between mortgage borrowers and lenders. I then discuss why we might expect information asymmetries to exist in mortgage markets, highlighting a particular sort of borrower-level heterogeneity—individual differences in willingness to default—that provides a source of adverse selection. I develop a simple model of mortgage choice and default incorporating this heterogeneity following Brueckner (2000) and show that it gives rise to a Spence-Mirrlees single crossing condition. Finally, I briefly outline the equilibrium implications of such a model, focusing on the potential for under-provision of credit due to adverse selection.

2.1 Definitions of Adverse Selection and Moral Hazard

The definitions of adverse selection and moral hazard that I specify follow largely from those used in Adams, Einav and Levin (2009):

(I) *Moral Hazard:* The mortgage market exhibits moral hazard if there is a causal relationship between the borrower's loan liability and default. That is, amongst homogeneous borrowers, those who face higher balances ex-post default more frequently.

(II) *Adverse Selection:* The mortgage market exhibits adverse selection if unobservably risky borrowers—those who are more likely to default with contract terms held equal—select higher leverage contracts.

Defining adverse selection in this way is fairly standard and adheres closely to the discussion in Chiappori and Salanié (2013) on insurance markets. Adverse selection exists if there is an exogenous correlation between a borrower's demand for leverage and the unobservable credit risk he poses to the lender. While there are a number of possible underlying models that could generate such a relationship, the equilibrium implications of the correlation are largely independent of the source, so I do not specify a mechanism in the baseline definition.

The way I define moral hazard is somewhat broader than usual. Typically, a credit market can be said to exhibit moral hazard if (i) the expected returns to the lender depend on some non-contractable action of the borrower and (ii) that action is itself influenced by the terms of the loan contract. If default is considered a strategic choice, my definition aligns with this traditional notion. Default itself can be thought of as the non-contractable action taken by the borrower. However, default is sometimes not an active choice. Borrowers may be insolvent or credit constrained to the extent that they are mechanically unable to make payments. While the empirical analysis that I conduct is explicitly designed to emphasize the strategic channel, the definition I use does not, in principal, rule out defaults due to a mechanical relationship between the loan balance and default. As in Adams, Einav and Levin (2009), whether the source is mechanical or strategic is not crucial for the policy implications I consider.

2.2 The Sources of Information Asymmetries in Mortgage Markets

In this subsection, I suggest potential sources of adverse selection and moral hazard in the mortgage market. While the definitions above are agnostic regarding mechanisms, understanding why we might expect these asymmetries to be present is helpful to frame further discussion.

Limited access to recourse for lenders provides an obvious explanation for the existence of moral hazard. The particular legal restrictions on contracts vary from state to state,[13] with some explicitly prohibiting lenders from recovering any excess balance from the borrower beyond the home itself in the event of default. However, even in states with laws that are favorable to lenders, deficiency judgments are relatively rare in practice (Pence, 2006). As a result, lenders cannot effectively contract against borrowers defaulting when their mortgages are underwater.

What is the source of heterogeneity that generates adverse selection? As a baseline, consider a simple model of mortgage default—often referred to as the frictionless option model—in which borrowers strategically default immediately if the value of their home drops below the value of the mortgage. Unless borrowers and lenders have different beliefs about housing prices, this model leaves little room for private information. All borrowers default according to a uniform rule.

However, a large literature suggests that borrowers do not default according to a frictionless option model (see Vandell, 1995, for a review). There is significant heterogeneity in willingness to exercise the default option (Deng, Quigley and Van Order, 2000), and a growing consensus that negative equity is a necessary but not sufficient condition for default (Bhutta, Shan and Dokko, 2010; Elul et al., 2010). Borrowers typically do not default until they owe more on their mortgage than the home is worth, and sometimes significantly more. Note that there is no need for a behavioral explanation for this phenomenon. There are real costs associated with default, including credit score reductions, moving costs, and social stigma.[14] These costs may differ in the population.

Heterogeneity in default costs provides a natural source of adverse selection. Borrowers who know that they are unlikely to repay will be less sensitive to the size of the mortgage balance. One way to think about this framework is as a model of selection on (ex-post) moral hazard, as in Einav et al. (2013). Lenders cannot contract on the hidden action of default, the costs of taking that action are heterogeneous in the population, and borrowers are privately informed of their costs.

2.3 The Model

To capture the intuition described above, I propose a two-period model of borrowers' leverage demand and default choice, following Brueckner (2000). Borrowers differ in a single dimension, which I refer to as the private default cost. This black box parameter represents all factors that influence the borrower's default decision at a given level of home equity. There are two primary takeaways. First, the distribution of private default costs in the population determines the magnitude of the moral hazard effect, i.e., the increase in defaults generated by a given change in the loan balance. Second, a

[13] California, for example, has laws that explicitly prevent the lender from recovering any balance from the borrower beyond the home itself in the case of default for owner-occupied homes with 1-4 units. Alternatively, Illinois allows deficiency judgments that can only be relieved in bankruptcy.

[14] Even for a given cost of defaulting, borrowers may be heterogeneous in access to liquidity

Spence-Mirrlees single crossing condition holds: borrowers with lower private default costs (i.e. risky borrowers) are relatively more willing to accept large balances.

In period 0, borrowers choose what portion of a risky housing purchase to finance. In period 1, the value of the home is realized, and borrowers choose whether to pay off their loan or to default. Mortgage contracts have two dimensions: the period 0 loan and the period 1 balance. I consider a non-recourse environment: in default, the borrower cedes the right to the home and is relieved of the loan balance.

Formally, let time be indexed by $t \in \{0, 1\}$ and borrowers be indexed by i. Borrowers must purchase a home with initial price H_0 and uncertain period 1 price H_1 distributed on support $[\underline{h}, \bar{h}]$ according to CDF $F(H_1)$. Lenders offer contracts of the form $\{L, B(L)\}$, where L is the value of the loan provided to the borrower in period 0, and $B(L)$ is the balance due on the loan in period 1.[15] In general $B(L)$ is increasing in L, that is, lenders demand higher balances for larger loans. A high leverage mortgage is one with a large L and correspondingly, a large $B(L)$.

Borrowers have per-period utility of consumption $u(\cdot)$, which is increasing and concave, receive income y_t in each period, which is not stochastic, and discount the future according to β. Each borrower i has privately known costs associated with defaulting, C_i, which captures the difference in dollar terms between defaulting and not defaulting.

Default Choice

In period 1, borrowers realize the value of their home and choose between repaying and defaulting. A borrower who repays retains the value of the home for net income $y_1 + H_1 - B$, while a borrower who defaults avoids paying the mortgage balance but incurs the default cost: $y_1 - C_i$. Borrowers choose to default when

$$H_1 - B < -C_i.$$

Borrowers with a low C_i are quicker to default, that is, for the same B they will default at higher home values.

This default rule demonstrates the importance of private default costs in determining the strength of the moral hazard effect. For a given C_i, the expected fraction of borrowers defaulting at balance B is $F(B - C_i)$, and the marginal effect of an increase in B is $f(B - C_i)$. The calculation becomes even more complicated with heterogeneity in C_i, as one must integrate over the set of borrowers at a given B.

[15] While other terms are often used to define mortgage contracts, these are usually equivalent to simple transformations of L and B in the two-period case. We could alternatively speak of the down payment $(H_0 - L)$, the interest rate ($\frac{B}{L} = 1 + r$), and the original loan-to-value $\frac{L}{H_0}$.

Contract Choice

In period 0, borrowers know C_i but face uncertainty about the period 1 home value. As a result, they choose $\{L, B\}$ to maximize

$$U(L,B;C_i) = u(y_0 - (H_0 - L)) + \beta\left[\int_{\underline{h}}^{B-C_i} u(y_1 - C_i)dF(H_1) + \int_{B-C_i}^{\bar{h}} u(y_1 + H_1 - B)dF(H_1)\right].$$

The term in brackets represents the expected period one utility, with the first term giving utility in the case of default and the second utility with repayment.

Note that the borrower's overall utility is increasing in the loan size L:

$$U_L(L,B;C_i) = u'(y_0 - H_0 + L) \geq 0.$$

Additionally, borrower utility is decreasing in the balance:

$$U_B(L,B;C_i) = -\beta \int_{B-C_i}^{\bar{h}} u'(y_1 + H_1 - B)dF(H_1) < 0.$$

This is likely unsurprising. Holding the balance fixed, borrowers prefer larger loans, and holding the loan size fixed, borrowers prefer a smaller balance.

There are a variety of reasons why borrowers prefer to take out large loans. In the presence of credit constraints, L provides a method of smoothing consumption over time, so borrowers can consume period 1 income and the expected gains from the home in period 0. However, even if it is possible to borrow at the risk free rate, borrowers still value mortgage loans because they provide a form of insurance against low realizations of H_1.[16] An increased B effectively allows borrowers to give up consumption when H_1 is high in exchange for sure consumption (in the form of L) even when H_1 is low.[17]

At actuarily fair prices, borrowers prefer to take advantage of the insurance provided by a mortgage. In a totally frictionless context,[18] borrowers will choose an extreme form of full insurance when offered a fair price. In particular, they will take out as large a loan as possible and default on the loan in all states of the world. While this may seem surprising, it is a standard result: a risk averse agent will be willing to sell a risky asset for its expected value.

Yet borrowers with different values of C_i do not value this insurance equally. In fact, a Spence-

[16] Assuming mortgage debt is non-recourse, but other debt cannot be forgiven.

[17] The mortgage literature refers to this as the put option contained in a mortgage: the borrower retains the right to sell the home to the bank in exchange for the balance on the mortgage.

[18] By totally frictionless, I mean a context with (i) borrowing and lending at the risk free rate, (ii) no default costs to the borrower, and (iii) lenders who can perfectly recover the home value after a default.

Mirrlees single crossing condition holds:

$$\frac{\partial(\frac{U_B}{U_L})}{\partial C_i} = \frac{-\beta u'(y_1 - C_i) f(B - C_i)}{u'(y_0 - H_0 + L)} < 0.$$

Because low C_i are more likely to default, all else equal, they are more likely to take advantage of the insurance provided by the mortgage. As a result, they are willing to accept smaller increases in the loan size L in exchange for the same increase in the balance B.

If borrowers with different levels of C_i, say $C_R < C_S$ (where R and S denote risky and safe borrowers), are offered the same menu of contracts, the single crossing condition constrains the set of contracts chosen. In particular, if these types buy contracts $\{L_R, B_R\}$ and $\{L_S, B_S\}$, respectively, it must be the case that $L_R \geq L_S$. Of course, for borrower C_S to be willing to accept a smaller loan, it must also be the case that $B_R \geq B_S$. Further, if C_R and C_S buy different contracts along one dimension, both inequalities must hold strictly.

2.4 Equilibrium Consequences of Single Crossing

In this subsection, I provide a brief graphical discussion of the consequences of single crossing on the equilibrium allocation of credit. The intuition is familiar from Rothschild and Stiglitz (1976) and numerous other works on screening. With two borrower types, single crossing makes pooled contracts unsustainable. Any contract that is sold to both risky and safe borrowers can be undercut by a contract that offers slightly less credit but only attracts safe borrowers. As a result, safe borrowers receive smaller loans in equilibrium (à la Rothschild-Stiglitz) than they would in a world of perfect information. This notion is analogous to the under-provision of insurance to safe types in insurance markets.

I first take a world with perfect competition amongst lenders and a regulatory limit of $L \leq H_0$ (i.e. an LTV cap of 100 percent). I consider borrowers who prefer loans at or above the LTV cap with fair prices.[19] Panels A and B of Figure I present the perfect information case. These figures show lenders zero profit curves (solid curves, labeled with $\pi = 0$) and borrowers' indifference curves (dashed lines, labeled with U) in the space of contracts (L, B).

Borrowers prefer contracts to the southeast: larger loans with smaller balances. Lenders unambiguously prefer contracts to the west: smaller initial loans. However, the net effect of balance increases is ambiguous for lenders. Profits rise in the case of repayment, but the probability of repayment falls. With perfect information, borrowers choose initial loans right at the LTV cap. This holds whether there is a single borrower type, as shown in Panel A of Figure I, or multiple types, as in Panel

[19] For large enough C_i or sufficient difference between y_0 and y_1 with borrowing constraints, borrowers may prefer smaller loans even at fair prices.

B. With two borrowers (risky R and safe S), the riskier type must simply pay a higher balance for the same loan, as shown by $B_R > B_S$ in Panel B.

Panels C and D of Figure I zoom in on the boxed portion of the graph shown in Panel B and highlight the complications posed by the single crossing property if lenders cannot offer different contracts to different borrowers. The balance necessary to give lenders zero expected profits on a pooled contract lies above the balance paid by safe borrowers with perfect information. Because the indifference curves of the risky type are steeper than those of the safe borrower—risky types are willing to accept a larger increase in the balance for the same increase in the initial loan—there is a region of contracts that are preferred by the safe borrower to the pooled contract but also make non-negative profits for lenders. In Panel C, this region is shaded in gray. This generates an opportunity for cream skimming. Given any pooled contract, lenders may offer a contract with a slightly smaller initial loan that attracts only the low risk types.

Panel D shows the form of a Rothschild-Stiglitz equilibrium in this context, should it exist. Risky borrowers end up with the same loan they would get in a world with perfect information, while safe borrowers are forced to take a smaller loan to distinguish themselves from riskier types. The welfare loss from adverse selection comes from safe borrowers receiving these inefficiently small loans ($L_S < H_0$) relative to the perfect information outcome.

3 Background and Data

In this section, I provide historical background on adjustable rate mortgages generally and the Option Adjustable Rate Mortgage in particular. I focus on the unique features of the Option ARM product that are key to my identification strategy. Because these loans feature fixed payment schedules and variable interest rates, changes in the benchmark financial indices used to determine interest rate adjustments translate directly to changes in borrowers' balances. I then describe the source of variation in interest rates I utilize in the empirical analysis, which is generated by differences between the financial indices used to determine rate adjustments. Finally, I discuss the characteristics of borrowers that chose Option ARMs relative to the larger population of mortgage borrowers and describe the data sources used in the empirical analysis.

3.1 Background on the Option ARM

Prior to the late 1970s, regulation effectively limited residential mortgage products in the United States to long term fixed rate loans, with set payments that remained constant over the amortization period. However, in 1978 the Federal Home Loan Bank Board began to allow federal savings and loan institutions to originate adjustable rate mortgages (ARMs) in California, and by the end of 1981, restrictions on adjustable rate products had been significantly relaxed nationwide. Originations of ARMs grew rapidly, representing as much as 68 percent of all new mortgages for certain months in the 1980s (Peek, 1990).

The industry largely settled on what are called Hybrid ARMs. These mortgages feature fixed interest rates and payments for a set initial period, often 5 or 7 years. After the initial period, interest rates begin to adjust according to market conditions, usually changing annually or semiannually. Monthly payments are designed to be fully amortizing, that is, calculated to exactly pay off the loan over the full term at current interest rates. As a result, payments change to keep pace with interest rates and may unexpectedly increase if interest rates rise.

According to lenders, the potential danger of these unexpected payment increases motivated the creation of the Option ARM.[20] Banks wanted a product that incorporated floating interest rates while protecting borrowers from sharp payment increases and mortgage holders from the associated default risk. The Option ARM is characterized by a series of features that reflect this desire:

(I) *Fixed minimum payment schedule:* Borrowers are offered a relatively low initial payment, often based on the fully amortizing payment for an extremely low "teaser" interest rate. For the first 5 years, this payment adjusts upward once yearly by a fixed amount, usually 7.5 percent.[21] After 5 years, the minimum payment adjusts to the fully amortizing amount. This schedule may be interrupted if the loan balance rises above a fixed proportion of the original home value, often 110 or 125 percent.

(II) *Monthly interest rate changes:* While interest rates for most ARMs adjust annually or semi-annually, Option ARMs update much more frequently, usually monthly. As in typical ARMs, new interest rates are calculated as the sum of a fixed component (referred to as the *margin*), which is determined at origination, and a financial index that proxies for the cost of funds to the lender (hereafter the *index*).

(III) *Negative amortization:* Oftentimes the minimum payment required in a given month will be lower than the amount of accrued interest. In these circumstances, Option ARMs allow for negative amortization, that is, allow the excess interest accrual to be incorporated into the balance. As a result, the loan balance will typically grow in the early years of the mortgage.

(IV) *Proposed Payment Options:* The name, Option ARM, refers to a menu of payment options offered to borrowers on monthly statements. In addition to the minimum payment, statements offer the

[20] See Golden West's history of the Option ARM, available at http://www.goldenwestworld.com/wp-content/uploads/history-of-the-option-arm-and-structural-features-of-the-gw-option-arm3.pdf.

[21] In theory, 7.5 percent is a cap, and the minimum payment might adjust by less if a 7.5 percent increase were to exceed the fully amortizing payment. In practice, the cap is nearly always binding.

possibility of an interest only payment, covering the entirety of the interest accrual, along with amortizing payments calculated according to 15- and 30-year schedules. These possibilities are suggestions. Only the minimum payment is binding, and the borrower may in principle make any payment between the options or in excess of the 15-year amortizing payment (sometimes subject to certain caps). In practice most borrowers make minimum payments every month.

For the purposes of the identification strategy, (I) and (II) are key. Because payments are fixed for the first 5 years,[22] borrowers' balances change as a function of realized interest rates. In the next subsection, I discuss these features in greater depth.

In the 1980s and 1990s, the Option ARM was primarily a niche product directed towards sophisticated borrowers. The flexibility of payments was intended to appeal to borrowers who expected their income to rise in the future or those with high income volatility. With the growth of a secondary market for non-traditional mortgages in the early 2000s, banks began to market Option ARMs as affordability products, allowing borrowers to purchase more expensive homes than they would be able to afford with a traditional mortgage. Borrowers might take out such loans with the intention of refinancing the mortgage or selling the home after several years, and thus never making payments much above the initial minimum. In the years leading up to the crisis, Option ARMs became a significant fraction of the market, representing approximately 9 percent of originations in 2006.[23]

As the crisis hit, borrowers with Option ARMs defaulted at high rates. In the sample studied here, 41 percent of borrowers were seriously delinquent (60 days past due) on their mortgages at some point within the first 5 years, and 33 percent wound up in foreclosure. The combination of high default rates and non-traditional features made Option ARMs a poster-child for excess in mortgage lending. Their role in the crisis has been highlighted by various media sources and policy makers—Ben Bernanke noted that "the availability of these alternative mortgage products proved to be quite important and, as many have recognized, is likely a key explanation of the housing bubble." Despite these criticisms, recent research has argued that these loans approximate the optimal mortgage contract (Piskorski and Tchistyi, 2010).

3.2 Diverging Indices and Interest Rate Resets

In addition to fixed payments and variable interest rates, the identification strategy relies on the financial indices used to determine interest rate adjustments. Interest rates for Option ARMs are typically

[22] All analyses performed here consider outcomes within the first 5 years. Appendix Figure A.I presents a sample balance and payment trajectory for an Option ARM to highlight these product features from origination through that period.

[23] See the 2008 Mortgage Market Statistical Annual.

tied to LIBOR or Treasury rates.[24]

Prior to the crisis, borrowers had little reason to prefer one index to another. Although there tended to be a spread between LIBOR and Treasury rates,[25] the two moved quite closely together, and any fixed difference could be accounted for in the margin. Furthermore, Bucks and Pence (2008) suggest that borrowers tended to be relatively uninformed about their contract terms. When asked what index their loan depended on, only 25 percent of borrowers responded with even plausibly correct indices, while 30 percent of borrowers simply answered that they did not know.

If borrowers were unaware of the distinction between indices, why did some end up with a Treasury index and others with LIBOR? Much of the variation comes as a result of the lender. Appendix Table A.I shows the proportion of LIBOR-indexed loans for the top originators in the sample. Most originators appear to specialize in either LIBOR or Treasury indices, and some offer only Treasury rates. A similar pattern can be seen among servicers, although slightly less pronounced. According to Gupta (2016), differences across lenders are often a function not of the borrowers they lend to, but rather their intentions on the secondary market.

Panel A of Figure II shows the spread between the 1-year CMT and 1-year LIBOR. While there were fluctuations in the years preceding the crisis, the difference was contained in a relatively narrow band. However, in mid-2007, Treasury rates began to fall and the spread increased, eventually peaking at over 3 percentage points in late 2008 following the Lehman Brothers bankruptcy filing and news of the AIG bailout. As a result, borrowers taking out similar loans prior to the crisis faced substantially different interest rates when their loans reset.

Panel B of Figure II displays this phenomenon for a large sample of adjustable rate mortgages, including traditional hybrid ARMs. The black line shows the average difference in interest rates between resetting loans indexed to LIBOR versus Treasury for each month. There is a noticeable spike in the in-sample difference in early 2009, corresponding to the late 2008 spike in Panel A.[26] The red line in Panel B shows the in sample difference in default for resetting loans indexed to LIBOR versus Treasury. In sync with the spike in relative interest rates, there was a sharp spike in relative defaults in early 2009. In the month that LIBOR-indexed loans experienced the most severe difference in interest rates, they also exhibited the most severe difference in defaults. This figure demonstrates the basic idea behind the identification strategy. Borrowers face different interest rates depending on whether their loan is indexed to LIBOR or Treasury and resultantly default at different rates.

But how do differences in interest rates cleanly translate into differences in loan balances? This is where the unique features of the Option ARM come into play. For a traditional adjustable rate

[24]Treasury rates are usually the 1-year Constant Maturity Treasury (CMT) or the 12-month Moving Treasury Average (MTA). Typically LIBOR refers to the 3-month LIBOR.

[25]For example, the spread between 1-year CMT and 1-year LIBOR was generally below 50 basis points.

[26]A lagged value of the index is typically used.

mortgage, a change in the interest rate also changes the monthly payment, which adjusts to ensure that payments are fully amortizing at the new rate. These payment shocks are thought to be the first order link between interest rate changes and default (see, e.g., Fuster and Willen, 2012). Alternatively, for Option ARMs, the required monthly payment is fixed for the initial period. As a result, changes in interest rates have no direct impact on monthly obligations. Because the mortgage must account for changes in the interest rate somehow, any additional interest accrual is incorporated directly into the balance. This means that for Option ARMs, the divergence between Treasury and LIBOR rates caused borrowers with otherwise identical loans to have sizable differences in loan balances ex-post. Appendix Figure A.II provides a stylized example of this pattern. Consider two identical $100,000 loans at origination, one of which faces a high realization of interest rates, while the other faces a low realization. Two years into the loan, the two borrowers will still have the same monthly payment, but the first borrower may owe thousands of dollars more than the second.

The impact of a LIBOR or Treasury index on the loan balance is not uniform across the sample period. Each origination month for each index generates a unique path of interest rates and a unique balance trajectory. Figure III demonstrates this difference-in-difference variation in borrowers' balances. The plot shows the loan balance over time for four sample $100,000 loans: one LIBOR- and one Treasury-indexed loan originated in January 2005, and one of each originated in January 2007.[27] Each of the four shows a distinct balance trajectory.

3.3 Data

The data on Option ARMs used in this paper are taken from a loan-level panel of privately securitized mortgages provided by Moody's Analytics (formerly provided by Blackbox Logic), representing over 90 percent of non-agency residential mortgage backed securities. These data provide detailed information about loans at origination, including borrower information, property characteristics, and contract terms. They also include dynamic information on monthly payments, loan balances, modifications, delinquency, and foreclosure. I focus on a sample of around 500,000 Option ARMs originated between 2004 and 2007, tied to either LIBOR or Treasury rates.

3.3.1 Summary Statistics: Balance Across indices

Table I shows summary statistics for the primary sample studied here, divided between loans indexed to Treasury and those indexed to LIBOR. Note that Treasury is the dominant index, representing approximately 90 percent of loans. Despite this, the majority of variables are reasonably balanced

[27] These figures are based on simulated loans with a margin of 3.5 for both samples, based on the 3-month LIBOR and 12-month MTA respectively.

across the two groups. Borrowers have fairly high FICO credit scores for both indices, with an average of 706 for Treasury loans and 714 for LIBOR loans. Furthermore, the majority of loans are low or no documentation—79 percent of Treasury loans and 77 percent for LIBOR. The original leverage choice, summarized by the original loan-to-value ratio (LTV), is also quite similar across the two indices, at 76.6 for Treasury and 77 for LIBOR.[28] Nearly all loans are subject to some form of prepayment penalties, and the majority of both Treasury and LIBOR loans are for primary residences.[29] There is a difference in the average margin for each—3.21 for Treasury loans versus 2.85 for LIBOR—but this gap reflects the baseline spread between the indices themselves.

The four most common states for both indices are California, Florida, Arizona, and Nevada. While Treasury loans are slightly more concentrated in California, the overall geographic patterns are similar across states. Appendix Figure A.III shows the relatively uniform density: between 5 and 15 percent of loans are indexed to LIBOR in nearly all states. The largest difference between the two samples is in the timing of origination. LIBOR loans are significantly more concentrated in 2004, while Treasury loans are more heavily represented in 2005 and 2006. This pattern is also reflected in the slightly higher average balances for Treasury loans. Overall, the observable details in both groups are reasonably balanced.

3.3.2 In Comparison to the Broader Market

While the unique characteristics of the Option ARM may have attracted a certain sample of the population, the growth of the product was not the result of an inflow of observably low quality borrowers. Option ARM borrowers have relatively good credit scores. The average FICO score in the sample studied here is over 700, and a negligible number of borrowers have scores below 620.[30] In this observable dimension, borrowers with Option ARMs reflect the general pool of borrowers rather than some particularly subprime subset.[31]

The geographic patterns of Option ARM originations also reflect the broader mortgage market. As in the sample of Option ARMs, the top two states for mortgage lending are California and Florida, representing 24 percent and 9 percent of all originations, respectively. Arizona (3.5 percent) and Nevada (1.7 percent) are also prominent nationally. Furthermore, these states all experienced significant growth relative to the market as a whole in the years leading up to the crisis.[32]

[28] I restrict the sample to loans with original LTVs between 50 and 100.

[29] The stated level of owner occupancy is likely an overstatement due to false reporting by investors (Piskorski, Seru and Witkin, 2013).

[30] The average credit score in the US is below 690, while the average among conforming loans purchased by Freddie Mac is 723 (Frame, Lehnert and Prescott, 2008). 620 is a common threshold to identify subprime borrowers.

[31] Amromin et al. (2011) find that borrowers with complex mortgages tend to be sophisticated, with high incomes and credit scores relative to the subprime population.

[32] All gained as a proportion of the market between 1996 and 2006: California by 25 percent, Florida by 60 percent, Arizona

The initial leverage choices of borrowers with Option ARMs are not out of line with the market as a whole. The average original LTV for Option ARMs, close to 77, is slightly larger than conforming loans purchased by Fannie Mae or Freddie Mac but below the average LTV for subprime adjustable rate mortgages.[33] The average initial loan size for Option ARMs is also larger than that of conforming loans,[34] although still below the conforming loan limit.

One peculiarity distinguishing Option ARMs from conforming loans is the rarity of income verification. Given low payments, protections against payment increases, and generally favorable expectations about housing prices, lenders were relatively unconcerned about borrowers' ability to meet their monthly obligations. This was especially true given that most loans were made to borrowers with high credit scores.[35] This led to the prevalence of low or no documentation loans—nearly 80 percent in this sample. For these loans, borrowers provide little or no formal evidence of sufficient income to meet monthly payments, often simply stating income with no verification. In the market as a whole in 2007, low or no documentation loans represented only 9 percent of outstanding loans. However, nearly 80 percent of Alt-A securitizations in 2006 were low or no documentation, mirroring the pattern in this sample (Financial Crisis Inquiry Commission, 2011).

4 Positive Correlation Tests

Before attempting to disentangle adverse selection from moral hazard, I confirm the motivating relationship of the paper: a positive correlation between leverage and default. I first clarify that borrowers do not choose an amount of leverage in isolation. Instead, they choose a contract that entails both the loan size and the interest rate as a pair. I then explicitly conduct positive correlation tests following Chiappori and Salanié (2000) to demonstrate the existence of asymmetric information. More simply, I show that a positive correlation exists between loan size and default and that this correlation persists conditional on the relevant information available to the bank at the time of contracting.

4.1 The Contract Menu

Mortgage contracts often have many features, but the key choice analyzed here is the tradeoff between the amount of leverage—summarized by the original LTV—and the interest rate.[36] A borrower may

by 55 percent, and Nevada by 44 percent. All statistics presented here are available in the 2006 Mortgage Market Statistical Annual.

[33] The average original LTV for Fannie Mae and Freddie Mac in 2007 was 72 and 71, respectively, according to Frame, Lehnert and Prescott (2008). The average LTV for subprime adjustable rate mortgages was over 80 as early as 2004 (Chomsisengphet, Pennington-Cross et al., 2006).

[34] The average conforming loan size was $236,400 for purchases and $233,800 for refinances in 2006 (Avery, Brevoort and Canner, 2007).

[35] See http://www.mortgagevox.com/meltdown/option-arms.html

[36] Just as in the model, borrowers choose an initial loan L and Balance B.

select a high LTV contract that requires a high interest rate or a low LTV contract with a lower rate. This menu of different interest rates and leverage pairs represents two dimensions of what Geanakoplos (2014) calls the credit surface. The particular slope and curvature of the surface depends on the current economic climate, the borrower's credit score, and other observable characteristics. For a given lender, contracts are usually summarized in a rate sheet, a series of guidelines indicating the required rates for different mortgage products, features, and borrower characteristics.[37]

Two features of the contract menu are suggestive of the existence of information asymmetries. The first is the very presence of multiple contract options for a given borrower, a standard property of markets with separating equilibria. Of course, these options might reflect market segmentation based on unobserved heterogeneity in preferences that is irrelevant to the lender. The second is that the relationship between leverage and interest rate is increasing: larger loans are offered at a higher unitary price. This shows that lenders account for the increase in default risk that comes when observationally equivalent borrowers take larger loans. Figure IV shows this increasing relationship in the data, plotting the empirical schedule between leverage (original LTV) and the margin—the fixed portion of the interest rate—for Option ARMs. I show the median margin offered within each 5-point bin of LTV from 50 to 90.[38] A clear pattern emerges: borrowers who choose higher original LTVs are also choosing higher interest rates.

4.2 A Positive Correlation Between Leverage and Default

Figure V shows the raw positive correlation between leverage and default. For Option ARMs, borrowers who choose large original LTVs are also more likely to default within the first 60 months of the loan. The relationship is nearly monotonic: those with original LTVs close to 50 default less that 10 percent of the time, while those with original LTVs near 90 default more than 50 percent of the time. The black line represents a local linear smoothing through the raw data, while black circles show the proportion of defaults for each 1-point point bin of original LTV.

By itself this raw correlation is not conclusive evidence of an information asymmetry. In principle, the correlation could be driven by selection on the basis of characteristics that are observed by the lender and appropriately priced into the contract. As a result, a crucial feature of the tests for information asymmetries suggested by Chiappori and Salanié (2000) are comprehensive controls for the lender's information set. If *observationally equivalent* borrowers who select larger loans default at higher rates, it can only be because (i) those borrowers are more likely to default on the basis of some

[37] Customarily, lenders specify explicit margin increases associated with each 5-point bin of LTV. For eligible mortgages delivered to Fannie Mae, these increases are explicitly codified. Appendix Table A.II shows an example of pricing adjustments necessary for mortgages purchased by Fannie Mae.

[38] This figure presents raw data, unconditional on the borrower's credit score or other characteristics, and hence may not represent the actual contract menu offered to any specific borrower.

unobservable (adverse selection), (ii) the larger loans actually *cause* more defaults (moral hazard), or (iii) some combination of the two.

Figure VI shows that the positive correlation between original LTV and default holds conditional on the information available to the lender, providing an affirmative test for the existence of asymmetric information. Each dot in the plot shows the coefficient on original LTV from OLS regressions of a binary indicator of default by 60 months on original LTV, successively controlling for more comprehensive subsets of the information available to the lender. The leftmost coefficient, from a regression with no controls, displays the raw relationship: a 1-point increase in the borrower's original LTV is associated with an approximately 1.2 percentage point increase in defaults. The coefficient labeled *full controls* shows the relationship with the most comprehensive set of controls available. The inclusion of *full controls* drops the coefficient slightly, from just under 1.2 to just under 1.0. It remains large and significant. Note that only the inclusion of month fixed effects causes a meaningful drop, while controlling for the loan purpose and home value actually causes the coefficient to increase. The left-hand side of Panel A in Table II exhibits the information in Figure VI in table form.

While I would ideally be able to condition exactly on the information available to the lender, the Moody's data is missing two features typically known by mortgage originators. These are the borrower's income as well as soft personal information about a borrower's risk that is not recorded. Fortunately, because of the preponderance of low or no documentation Option ARMs, lenders also did not have information about borrowers' income for the majority of loans in the sample. The fourth column of Panel A displays results limiting the sample only to these loans. The coefficient remains significant, and in fact increases slightly. The problem of soft information is more difficult to deal with. However, any missing soft information is likely to bias the coefficient on the original LTV towards 0. If lenders are aware that borrowers are bad credit risks in some way that is unobservable in the Moody's data, they will be less likely to offer desirable terms for high leverage contracts.

The right-hand side of Table II shows that whether one considers the borrower's leverage or interest rate as the defining feature of the contract is not crucial. There is a robust positive correlation between the borrower's margin and default. Finally, Panel B shows that the positive correlation holds consistently when looking at defaults by 12, 24, 36, or 48 months.

The final two coefficients shown in Figure VI preview the remainder of the empirical analysis. The first, labeled *Ex-post LTV,* shows the coefficient on original LTV when controlling not just for the information of the lender, but also the (imputed) LTV at 60 months. The drop in the coefficient when including *Ex-post LTV* roughly represents the moral hazard effect. This is the portion of the correlation that is due not to selection, but to the incentives to default provided by the loan liability. The residual coefficient on original LTV represents selection. The final plotted coefficient, labeled *Option Value*, repeats this exercise using a more flexible set of controls in addition to the *Ex-post LTV*, including 6 months of leads and lags in interest rates and zip code level home prices. Note that there is not a significant drop in the coefficient when including this more flexible specification.

5 Empirical Strategy

The primary specification to separate adverse selection from moral hazard is a standard binary model of default. In simple terms, I regress an indicator for default on the borrower's current negative equity and original leverage. I argue that this single equation model is sufficient to capture the effects of both information asymmetries. Appropriately instrumented, the impact of current negative equity on default captures the moral hazard effect, while the effect of original leverage captures adverse selection.[39]

More specifically, for borrower i in MSA j at loan age t, I consider models of the form:

$$D_{ijt+1} = \mathbb{1}\{\alpha E_{ijt} + \gamma L_i + x'_{ijt}\beta + \omega_{m_i} + \delta_{index_i} + \zeta_j + u_{ijt} > 0\}. \tag{1}$$

Where E_{ijt} is instrumented by a function of the borrower's index choice and origination month:

$$E_{ijt} = f_t(m_i, index_i) + \lambda_{m_i} + \mu_{index_i} + z'_i\pi_t + \phi_j + e_{ijt}. \tag{2}$$

Here, D_{ijt+1} is a measure of default by time $t+1$. E_{ijt} is a measure of the borrower's negative equity, measured either as the current difference between a borrower's balance and the value of the home or as the current LTV. L_i is the borrower's initial leverage choice, measured as the original LTV. The vector x_{ijt} contains all time varying covariates relevant to the default decision as well as z_i, the set of all borrower characteristics and loan features known to the bank at the time of contracting. I include fixed effects for the origination month of the loan (ω_{m_i} or λ_{m_i}), the index choice (δ_{index_i} or μ_{index_i}), and the borrower's MSA (ζ_j or ϕ_j). Standard errors are clustered at the MSA level. Most specifications additionally allow for state specific time trends. I estimate the equation separately at different t, so do not include loan age effects. I interpret $\gamma > 0$ as evidence of adverse selection and $\alpha > 0$ as evidence of moral hazard.

In the next subsection, I discuss the challenges of consistently estimating α and γ that necessitate an IV strategy and specify the form of $f_t(m_i, index_i)$. I then justify the use of this single equation model, showing that Equation 1 can be derived by collapsing a more comprehensive model that explicitly

[39] In my basic specifications I use current negative equity to estimate the moral hazard effect, rather than explicitly using the borrower's balance. Negative equity captures any changes in the borrower's balance, but also captures the incentives to default driven by changes in housing prices.

specifies the borrower's demand for leverage alongside the default choice. Doing so also clarifies the interpretation of α as the moral hazard effect and γ as the adverse selection effect. Finally, I propose a secondary strategy to jointly estimate leverage demand alongside the default choice. While more complex, this approach allows me to recover fundamental parameters more directly relevant for the simulation exercise performed in the next section.

5.1 Identification

The basic challenge in separately identifying α and γ—the effects of current equity and initial leverage on default—is the mechanical relationship between L_i and E_{ijt}. In the absence of other differences, borrowers with identical L_i will tend to have identical E_{ijt}. For borrowers consistently making minimum payments, there are only two factors that might cause those with identical L_i to have different E_{ijt}: differences in home prices or differences in interest rates that lead to different balances.

Unfortunately, shocks to home prices may, in general, be correlated with u_{ijt}. For example, a local labor market shock may influence both home prices and, separately, the borrower's probability of default. Additionally, because home prices can never be observed directly but rather must be inferred from the sale prices of surrounding homes, E_{ijt} is measured with error. Similarly, variation across time in interest rates is likely correlated with macro conditions, while cross-sectional variation potentially reflects borrowers' endogenous contract choices. Isolating exogenous variation in E_{ijt} is non-trivial but necessary to accurately estimate α and γ.

I focus on plausibly exogenous variation in E_{ijt} that comes from the interaction of the borrower's index and the origination month of the loan. Each {Index Type, Origination Month} pair generates a unique trajectory of interest rates for a borrower. Utilizing this difference-in-difference variation allows me to control for any origination month-specific cohort effects or trends in the macro-economy, while also accounting for any fixed differences between borrowers with different indices.[40]

Equation 2 shows the basic framework for isolating this variation in E_{ijt}. The function $f_t(m_i, index_i)$ is effectively a set of instruments for E_{ijt}. These instruments capture changes in E_{ijt} that result from the interaction between the origination month m_i and the index, but are distinct from fixed month and index effects λ_{m_i} and μ_{index_i}.

Developing an instrument involves choosing a functional form for $f_t(m_i, index_i)$. In what follows, I focus primarily on a specification that exploits all possible variation in the interaction between the month of origin and the index, that is:

$$f_t(m_i, index_i) = \lambda_{m_i} \times \mu_{index_i}.$$

[40] The difference-in-difference framework allows me to control for fixed lender characteristics, even for originators or servicers who exclusively feature one of the two indices.

In words, I use a full set of fixed effects for every possible {Index Type, Origination Month} pair as instruments for the borrower's home equity. This specification has the advantage of limiting assumptions about functional forms and provides a large number of instruments.

However, because this large set of instruments does not provide an easily interpretable first stage and may suffer from problems associated with many weak instruments, I also consider a secondary option. The basic idea behind this exercise is to produce a strong predictor of a borrower's balance using only the origination month and index. To do so, I mechanically calculate the full balance trajectory for a sample loan for each index type and origination month. The sample loan sets all potentially endogenous terms, which vary for any given loan, to standard values.[41] As a result, the instrument captures the variation in the balance that is driven by the interest rate realizations while excluding any variation due to endogenous contract choices. I refer to the instrument developed using this calculation as the "simulated" instrument.

5.2 Leverage Demand and Default Choices

In this section, I show that Equation 1 can be derived from a more explicit model of the borrower's leverage and default choices. I begin with the default rule suggested by the theoretical model in Section 2. A borrower defaults if the value of the home (H) falls far enough below the balance (B) to justify incurring any default costs C_{ijt}. C_{ijt} is a reduced form parameter that captures any observable and unobservable factors that influence the borrower's decision to default at a given level of home equity. The default condition is then:

$$D_{ijt+1} = \mathbb{1}\{B_{ijt} - H_{ijt} > C_{ijt}\}.$$

A slight relabeling generates a condition that resembles Equation 1 above. First, let $E_{ijt} = B_{ijt} - H_{ijt}$, a measure of the borrower's negative equity. E_{ijt} is large when the borrower owes much more than the home is worth. Next, decompose C_{ijt} into its observable and unobservable components, where $-C_{it} = \sigma_\varepsilon(x'_{ijt}\beta + \omega_{m_i} + \zeta_j + \delta_{index_i} + \varepsilon_{ijt})$ and only $\varepsilon_{ijt} \sim N(0,1)$ is unobservable. Defining $\alpha = \frac{1}{\sigma_\varepsilon}$, we can write the default condition as:

$$D_{ijt+1} = \mathbb{1}\{\alpha E_{ijt} + x'_{ijt}\beta + \omega_{m_i} + \zeta_j + \delta_{index_i} + \varepsilon_{ijt} > 0\} \tag{3}$$

While the borrower's contract choice is the result of a complex maximization problem, I abstract from this structure and specify a linear demand model for leverage. Letting L_i represent the original LTV

41 A margin of 3.5, an initial loan size of $400,000, and a minimum payment based on a 1.75 percent teaser rate.

chosen by borrower i:

$$L_i = z_i'\psi + \theta_{m_i} + \eta_{index_i} + v_i. \tag{4}$$

Within this framework, moral hazard and adverse selection have straightforward empirical predictions:

(I) *Moral Hazard:* $\alpha > 0$ provides evidence of a moral hazard effect, where α quantifies the impact of the borrower's equity on default.

(II) *Adverse Selection*: $\rho = Corr(v_i, \varepsilon_{ijt}) > 0$ provides evidence of adverse selection. Borrowers who choose higher than average L_{ij} based on unobservables (large v_i) are more likely to default holding home equity constant (large ε_{ijt}).

In the next subsection, I describe an approach to estimating Equations 3 and 4 jointly. However, to arrive at Equation 1, I collapse leverage demand and the default decision based on the correlation between v_i and ε_{ijt}. In particular, I write $\varepsilon_{ijt} = \gamma v_i + u_{ijt}$, where $\gamma > 0$ holds if the two are positively correlated, that is, if there is adverse selection.[42] Replacing v_i using Equation 4 gives $\varepsilon_{ijt} = \gamma(L_i - z_i'\psi - \theta_{m_i} - \eta_{index_i})$. Replacing ε_{ijt} in Equation 3, collapsing month and index fixed effects, and absorbing z_i into x_{ijt} gives Equation 1.

Joint Model

Although the collapsed model in Equation 1 is more straightforward, there are also benefits to jointly estimating the leverage demand and default choice. Given the need to instrument for E_{ijt}, doing so actually involves estimating three equations:

$$\begin{aligned} D_{ijt+1} &= \mathbb{1}\{\alpha E_{ijt} + x_{ijt}'\beta + \omega_m + \zeta_j + \delta_{index} + \varepsilon_{ijt} > 0\} \\ L_i &= z_i'\gamma + \theta_m + \eta_{index} + v_i \\ E_{ijt} &= f(m_i, index_i) + \lambda_m + \mu_{index} + z_i'\pi_t + \phi_{jt} + e_{ijt}. \end{aligned}$$

I impose a parametric structure on the errors. In particular:

$$\begin{pmatrix} \varepsilon_{imt} \\ v_i \\ e_{it} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 1 & & \\ \rho_{\varepsilon v}\sigma_v & \sigma_v^2 & \\ \rho_{\varepsilon e}\sigma_e & \rho_{ve}\sigma_v\sigma_e & \sigma_e^2 \end{bmatrix} \right).$$

Again, I estimate cross-sectionally at different t and hence make no assumption about the evolution of errors over time. This specification allows a relatively straightforward estimation. I effectively employ

[42] In the normal case, $\gamma = \rho\frac{\sigma_\varepsilon}{\sigma_v}$.

a control function approach following Blundell and Powell (2004), incorporating an additional linear equation.

The benefit of this approach is that I am able to recover a few parameters that provide a basis for simulation in Section 2. Perhaps most importantly, I directly recover $\rho_{\varepsilon v}$, the correlation between ε_{ijt} and v_i. This correlation determines the strength of the adverse selection effect. Furthermore, under the normality assumption, I am able to recover the underlying distribution of the default costs C_{ijt} for any individual i. Recalling that $-C_{ijt} = \sigma_\varepsilon(x'_{ijt}\beta + \omega_{m_i} + \zeta_j + \delta_{index_i} + \varepsilon_{ijt})$, we have:

$$C_{ijt}|x_{ijt}, \omega_{m_i}, \zeta_j, \delta_{index_i} \sim N\left(\frac{-(x'_{ijt}\beta + \omega_{m_i} + \zeta_j + \delta_{index_i})}{\alpha}, \frac{1}{\alpha}\right).$$

The distribution of C_{ijt} characterizes the moral hazard effect, while $\rho_{\varepsilon v}$ summarizes the degree to which borrowers' knowledge of their place in this distribution impacts leverage demand.

6 Results

In this section, I describe the central empirical results of the paper. I begin by defining a few variables used in the analysis. I then describe first stage results: the instruments I use are correlated with borrowers' balances, are directly predictive of default in a reduced form, and do not predict borrower characteristics such as credit scores. I next turn to the results of the primary, single equation model of default. I find strong evidence of both moral hazard and adverse selection, which hold across numerous robustness checks. I explore heterogeneity in these results by state recourse status and initial loan features. Finally, I discuss my estimation of the joint model of leverage demand and default choice, which provides parameters that directly inform the simulations presented in the next section.

6.1 Definitions of Key Variables

The empirical analysis revolves around three variables: default D_{ijt+1}, original leverage L_i, and current equity E_{ijt}. Here, I discuss the definitions of each as used below:

(I) **Default (D_{ijt+1}):** The standard definition for default used here is a borrower being 60 or more days past due on monthly payments. Typically, D_{ijt+1} measures the outcome of default between years t and $t+1$. However, when explicitly stated, D_{ijt+1} may also refer to default at any point between loan origination and $t+1$.

(II) **Original Leverage (L_i):** Original leverage is measured as original loan-to-value in percentage terms. While the CLTV (combined loan-to-value), which incorporates any second liens, is sometimes used as a measure of leverage, my focus here is on the leverage contained in a particular contract. I control for the presence of any observable additional liens in all specifications.

(III) **Current Equity (E_{ijt}):** I use two alternative measures of E_{ijt} throughout the analysis. The first is the borrower's *negative equity*. This is defined as the current balance on the loan less the value of the home. The second is the borrower's current *loan-to-value*, the ratio of the current balance to the current value of the home. Both of these measures grow as the borrower's balance increases and fall if the price of the home increases. As home values are generally only recorded when houses are sold, I follow the literature and impute the current home value based on local home price indices.[43]

Unfortunately, I do not observe E_{ijt} for borrowers who exit the sample prior to time t. This prevents me from using the full sample for specifications that incorporate current home equity. However, given the contract terms for a mortgage—the margin, initial monthly payment, initial balance, and index—predicting the balance up to the first delinquency or partial prepayment is a straightforward mechanical calculation. Using just loan terms at origination and interest rates, I am able to predict the observed values of E_{ijt} with a high degree of accuracy ($R^2 > 0.95$). In regressions that incorporate the full sample, I use these imputed values of E_{ijt}, which are available for all borrowers, rather than the observed values.

6.2 First Stage

The plots in Figure VII highlight a few features of the instruments for E_{ijt} used in the main analysis. Because the primary specification uses a large number of fixed effects, and hence does not provide an easy to interpret first stage, I use the simulated instrument described above—a single variable—to produce these figures. This variable allows me to address three points.

First, do borrowers actually have significantly higher E_{ijt} when the instrument suggests balances should be high? Panel A of Figure VII shows that this is the case. The plot presents the coefficient on the simulated instrument from the simplest possible specification for considering relevance: a regression of E_{ijt} on the instrument, controlling for origination month and index fixed effects. When the simulated instrument is high, borrowers' E_{ijt} are high. This pattern holds across the first several years of the loan, although the size of the correlation declines over time.

Second, Panel B shows that borrowers also default more when the instrument is high. This is a reduced form and shows coefficients from an identical exercise to that in Panel A, replacing E_{ijt} with default D_{ijt}. Third, Panel C shows evidence of instrument exogeneity. Despite predicting borrowers' balances and defaults, the instrument is not correlated with FICO credit scores, a key measure of borrowers' creditworthiness.

[43] I use Zillow's zip code level home price index, available at http://www.zillow.com/research/data/.

Table III shows a more formal first stage. The coefficients shown are analogous to those in Panel *A* of Figure VII, utilizing the simulated instrument. However, because the primary specifications use the full set of $\lambda_m \times \mu_{index}$ fixed effects, I also include F-statistics calculated using the full set of fixed effects alongside those from the simulated instruments. These tables largely mirror the information shown in the figure. At 24 months, predictive power is strong, with F-statistics suggesting that the instruments are relevant in both the fixed effects and simulated instrument specifications (although the F-statistic drops below 10 when using fixed effects to predict negative equity). However, by 48 months, as the sample diminishes, the instruments lose their predictiveness. This is especially so for the simulated instrument, which has effectively no predictive power by 48 months when controls are included. The fixed effects specification, while weak, retains some relevance.[44]

6.3 Main Results: Single Equation Model

The primary specifications attempt to isolate adverse selection and moral hazard following Equation 1. In the main tables, I use linear probability models and a standard instrumental variables approach. In the appendix, I show probit estimates, accounting for the endogeneity of E_{ijt} following Blundell and Powell (2004).

The main tables are structured to show three versions of each specification of interest: baseline, OLS, and IV. The first is a reference and shows the baseline relationship between original LTV and default, including relevant controls but excluding any measure of current equity. For OLS regressions, I add a measure of E_{ijt} to the baseline regression but do not account for endogeneity in E_{ijt}. Finally, in the IV regressions I explicitly instrument for E_{ijt} with the full set of $\lambda_m \times \mu_{index}$ fixed effects. The coefficient on E_{ijt} gives the moral hazard effect, while the coefficient on original LTV gives the adverse selection effects. Comparing the IV regressions to the baseline regressions gives a sense of the role of moral hazard in the overall correlation between leverage and default.

Table IV presents the primary set of specifications, showing a cross-section of borrowers 24 months after origination.[45] This table includes E_{ijt} defined as both current negative equity (Panel A) and current LTV (Panel B). In different specifications I include two levels of controls: a basic set with only origination month and index fixed effects, and a comprehensive set, including MSA fixed effects, flexible controls for the original FICO credit score (dummies for each 20-point bin), state-level time trends, loan originator and servicer fixed effects, and controls for documentation, loan purpose, occupancy,

[44]Because the current loan-to-value or negative equity at 24 or 48 months is only observed for loans that actually survive to those points, the first stage regressions in Panel A are necessarily conducted on a selected sample. Panel B shows identical specifications to those in Panel A but replaces the observed values of E_{ijt} with imputed values, which allows the use of the full sample. While the strength of the instruments is not substantially better at 24 months for either the current LTV or home equity, there are significant improvements at 48 months. While the F-statistics only exceed 10 without covariates, both versions of the instrument are relevant (if weak) with the full sample.

[45]Other loan ages are shown in the appendix.

property type, prepayment penalties, private mortgage insurance, second liens, and the original home value.

The left-hand side of Panel A shows that there is strong evidence of both adverse selection and moral hazard when defining E_{ijt} as current negative equity and including only basic controls. The estimated moral hazard effect in the IV specification suggests that a $100,000 increase in negative equity increases the one year default probability by 4.5 percentage points (17 percent). The estimated adverse selection effect suggests that a 10-point increase in the borrower's original LTV is associated with a 3.3 percentage point (12.5 percent) increase in the one year default probability. The OLS results are quite similar, showing slightly larger moral hazard effects and a slightly larger role for adverse selection. Comparing the role of original LTV in the IV regression (0.331) to that in the baseline estimate (0.586) implies that adverse selection is responsible for more than half of the baseline correlation. However, it is crucial not to over-interpret the coefficient on original LTV with this limited set of controls. Without controlling for information available at loan origination, this result pools true selection on unobservables with lenders' steering of riskier borrowers towards smaller loans.

The right-hand side of Panel A includes the full set of controls and shows (i) a larger moral hazard effect and (ii) an adverse selection effect that is similar in levels but smaller as a fraction of the baseline estimate. The estimated moral hazard effect implies that a $100,000 increase in negative equity increases the one year default probability by 8.9 percentage points (33.5 percent). The estimated adverse selection effect shows that, all else equal, borrowers who choose 10-point larger initial LTVs are 2.6 percentage points (10 percent) more likely to default between 24 and 36 months. However, including the full set of controls also leads to a significant increase in the baseline correlation between original LTV and default. As a result, adverse selection accounts for approximately 36 percent of the baseline correlation with appropriate controls. This leaves moral hazard responsible for the remaining 64 percent.

Panel B repeats the exercise from Panel A but defines E_{ijt} as current loan-to-value. With full controls, the effects are quite similar to those found in Panel A. Adverse selection is responsible for 32 percent of the baseline correlation between original LTV and default, while the remaining 68 percent is due to moral hazard. These estimates imply that borrowers that choose 10-point higher original LTVs are 2.3 percentage points more likely to default between 24 and 36 months, all else equal, while a 10-point increase in current LTV at 24 months increases the probability of default by just over 4 percentage points. Furthermore, specifications without controls highlight the potential complications of ignoring the information available to the bank. The OLS and IV show negative (although insignificant) coefficients on the original LTV when controlling for the current LTV.

6.4 Heterogeneity in Results from the Single Equation Model

Table V considers how the results in Table IV change across three relevant subgroups: (i) in states with full recourse versus those with limited recourse, (ii) for borrowers providing full documentation versus those providing limited or no documentation, and (iii) for home purchases versus refinances. In each panel, I show the baseline relationship between original LTV and default for the relevant subgroup, then IV regressions with E_{ijt} defined first as negative equity and next as current LTV. All specifications include the full set of controls.

6.4.1 State Recourse Status

The most notable difference between states with full versus limited recourse[46] is in the strength of the estimated moral hazard effect. Both categories show a significant baseline correlation between original LTV and default. However, the impact of E_{ijt} on default—defined either as current negative equity or LTV—is large and statistically significant in limited recourse states, and near zero in full recourse states. This pattern is intuitive: in states where borrowers are responsible for the loan balance even in default, the marginal incentive to default generated by an increase in the current balance is low. Perhaps more surprising is that both types of states show strong evidence of adverse selection across OLS and IV specifications. In both cases, original LTV is strongly associated with default, controlling for current incentives to default. It should be noted that the sample size is much smaller in full recourse states, and the estimates are correspondingly less precise.

6.4.2 Documentation

Dividing borrowers by documentation provided, shown in Panel B of Table V, suggests that income verification may be an important factor in screening borrowers. The results for the low or no documentation sample largely match the full sample. In contrast, in the sample providing full documentation, the entirety of the raw correlation between leverage and default is explained by moral hazard. The optimistic view of this result is that documentation solves the adverse selection problem: the additional information on income allows lenders to distinguish an individual's riskiness before offering a set of contracts. However, because I do not observe income, I am also not perfectly able to control for the information set of the lender in the full documentation sample. As a result, the coefficient on original LTV in the full documentation sample pools an adverse selection effect with any steering of borrowers by lenders on the basis of income.

[46] By state recourse status, I refer to a state's provisions regarding a lender's ability to recover any balance that exceeds the value of the home in the case of default. I categorize states as full or limited recourse on the basis of that in Rao and Walsh (2009), with full recourse referring to states with strong provisions regarding deficiency judgments and limited recourse referring to those with mixed, weak, or nonexistent provisions.

6.4.3 Purchases vs. Refinances

The differences between those purchasing homes versus those refinancing existing mortgages, shown in Panel C of Table V, are less severe than those in Panels A and B. While the baseline correlation between original LTV and default is higher in the refinance sample, both show comparable moral hazard effects: a 10 point increase in the current LTV causes an average of just over 4 percent more defaults within a year on average in both samples. However, the estimated adverse selection effects are slightly smaller in the purchase sample and only significant when E_{ijt} is defined as current negative equity.

6.5 Robustness for the Single Equation Model

The appendix includes a number of tables intended to serve as robustness checks to Table IV and to provide alternative estimates of interest. Here, I briefly discuss these exercises.

6.5.1 Loans at 48 Months

The results for loans at 48 are similar to those at 24 months, if somewhat muted. Appendix Table A.III presents identical regressions to those in Table IV, except with current E_{ijt} defined at 48 months and the dependent variable defined to be default between 48 and 60 months. With full controls, the baseline relationship between original LTV and default is somewhat lower than at 24 months, and the proportion of the correlation due to adverse selection somewhat higher (greater than 50 percent in the IV specifications). Further, the estimated moral hazard effects are smaller, and insignificant when E_{ijt} is defined to be current LTV. Given the weakness of the instrument at 48 months, these estimates should be interpreted cautiously, but they largely support the results found in Table IV. Results at other cross-sections are similar.

6.5.2 Cumulative Default Probabilities

The regressions in Table IV take an indicator for default within one year as the dependent variable. Doing so poses two potential issues. First, considering the default probability between 24 and 36 months limits the sample to borrowers who are still active at 24 months. This generates a potential source of bias, as borrowers who default or prepay in the early years of the loan may differ from the larger population, or may be responding endogenously to new knowledge of their anticipated future balance. Second, lenders may be more concerned with whether a borrower defaults at all, rather than a borrower's hazard rate, particularly with loans that feature negative amortization.

To address these issues, Appendix Table A.IV considers the impact of the original LTV and current

E_{ijt} on cumulative default outcomes in the full sample. This approach avoids sample selection issues, but requires a slight reinterpretation of the treatment. The moral hazard effect no longer captures a response to the realized balance but rather the borrower's response to the anticipated balance trajectory. Furthermore, because E_{ijt} is not observed directly for those defaulting prior to t, I use the imputed version of the E_{ijt}, based on original contract terms and realized interest rates.

I first estimate specifications meant to mimic those in Table IV, this time utilizing the outcome of cumulative default by 36 months. These are shown on the left-hand side of Appendix Table A.IV. I include imputed E_{ijt} measured at 36 months. For these estimates the baseline relationship between the original LTV and default is higher than in Table IV. However, the portion owing to adverse selection—approximately 17 percent when E_{ijt} is defined as current negative equity, and 29 percent when defined as current LTV—is somewhat lower. Regardless, there is strong evidence that both moral hazard and adverse selection are present.

As a more robust test of the adverse selection effect, the right-hand side of Appendix Table A.IV considers the outcome of default by 60 months. For these regressions, I include a comprehensive set of controls for E_{ijt}, not just at a given point in time, but across the life of the loan. These controls are meant to account for the full impact of the non-linear loan trajectory throughout the first 60 months. Even controlling for the full trajectory of E_{ijt}, the initial leverage choice is strongly predictive of default. In these specifications, adverse selection remains responsible for approximately 30 percent of the baseline relationship between original LTV and default.

6.5.3 Alternate Functional Forms

A potential concern is that the observed effect of original LTV on default when controlling for E_{ijt} does not truly reflect selection, but rather some more complicated functional form relating E_{ijt} to default that is not captured by a linear specification. Appendix Table A.V examines whether there is still evidence of adverse selection across three more complex specifications: (i) including a cubic specification in current E_{ijt}, (ii) controlling for current and past minimum payments and interest rates, and (iii) interacting E_{ijt} with covariates.[47] The estimated adverse selection effect is persistent across all specifications.

6.5.4 Further Robustness

Appendix Tables A.VI and A.VII explore further robustness. The results are robust to (i) probit and control function specifications, which are potentially more realistic than the linear probability model,

[47] In column 5 of Appendix Table A.V, the OLS specification, I fully interact E_{ijt} with all covariates. However, because I do not have sufficient instruments to do so in an IV specification, in column 6 I simply interact E_{ijt} with two covariates: the borrower's credit score and whether the loan was to purchase a home or refinance an existing mortgage.

(ii) the use of the simulated instrument rather than the full set of fixed effects, and (iii) alternative definitions of default, ranging from mild (30 days past due) to extreme (foreclosure).

6.6 Joint Model

The final step of my empirical analysis is to estimate a joint model of leverage demand alongside the default choice. Doing so allows me to recover parameters that more directly relate to the model developed in Section 2 and that can be used to inform the simulations developed in the next section. Because of the increased computational complexity of this estimation, I slightly reduce the richness of included controls, e.g. substituting MSA fixed effects with state fixed effects. I estimate the model separately at 24, 36, and 48 months, and again define E_{ijt} as both current LTV and home equity. These estimates are presented in Table VI and qualitatively align with estimates from the single equation model.

The primary benefit is in providing estimates of three parameters: (i) $\rho_{\varepsilon}v$, the correlation between the errors in the leverage and default choices, where a positive value indicates adverse selection, (ii) $\sigma_{\varepsilon} = \frac{1}{\alpha}$, the standard deviation of the default error in units of E_{ijt}, where a positive and significant value of α indicates moral hazard (corresponding to a finite positive value of σ_{ε}), and (iii) the mean of borrowers' default costs, conditional on observables. This can also be interpreted as the default threshold, that is, the level of E_{ijt} above which the average borrower (with a given set of observables) defaults. While the estimates at 24 and 36 months show strong evidence of both adverse selection and moral hazard—a positive ρ and α—the estimates at 48 months are less precise.

The first and fourth columns display the estimated parameters at 24 months, with E_{ijt} defined in terms of negative equity and current LTV, respectively. The estimated threshold for default in the first column (at average values of observables) is just under \$100,000, meaning that a borrower will not default until the balance on their loan is \$100,000 above what the home is worth. The standard deviation of the default error in this specification is approximately \$190,000. Similarly, the fourth column suggests that the average borrower must owe 1.34 times what the home is worth before defaulting. The standard deviation of unobserved default costs in the population is just over 50 percent of what the home is worth: $\sigma_{\varepsilon} = 0.55$. Finally, the correlation between unobserved default costs and the unobserved portion of the original leverage (original LTV) choice—which measures adverse selection—is significant, and just under 0.07. I use these estimates to parameterize the model in the next section.

7 Simulations and Welfare Analysis

In this section, I consider the policy implications of separately accounting for adverse selection and moral hazard. The estimated moral hazard effect—the causal effect of a change in home equity on the probability of default—provides precisely the relevant parameter for considering the effectiveness of ex-post policies that reduce loan balances in preventing defaults. Understanding the impact of ex-ante regulations is more challenging. I consider the case of an LTV cap and argue that ignoring the role of adverse selection leads policymakers to (i) *overestimate* the reduction in defaults generated by a reduction in the LTV cap and (ii) *underestimate* the welfare loss generated by borrowers taking smaller mortgages. I consider a slightly expanded version of the model suggested in Section 2 and use the equilibrium concept proposed by Azevedo and Gottlieb (2016) to address the challenges of evaluating counterfactual policies in competitive markets with adverse selection.

7.1 The Impact of Home Equity and Ex-Post Balance Writedowns

The estimated moral hazard effect has direct policy relevance. It captures the causal effect of a change in home equity on the probability of default, which is necessary to predict the effectiveness of ex-post principal writedowns in preventing mortgage defaults. The estimates in Table IV suggest that, for the sample studied here, a 10 percentage point reduction in all borrowers' LTV at 24 months would have reduced defaults within a year by just over 15 percent. Relative to the literature, these estimates are on the large side,but not outside of normal bounds. For example, Bajari, Chu and Park (2008) find that a 25-point increase in the current LTV is necessary to generate a 15 percent increase in the default probability. Alternatively, Elul et al. (2010), find that borrowers with increasing CLTV from between 100 and 110 to between 110 and 120 raises the quarterly default hazard by about 30 percent of the mean.

After the crisis, policies of this form were enacted, for example the Home Affordable Mortgage Refinance Program Principal Reduction Alternative (HAMP PRA). Scharlemann and Shore (2016) use a kink in the schedule for HAMP PRA to analyze the effectiveness of the regulation. They estimate that principal writedowns— balance reductions of 28 percent on average—reduced the quarterly delinquency hazard by 18 percent (from 3.8 percent per quarter to 3.1 percent). However, their study examines only those who participated in the program (and hence were already delinquent), while my estimates consider the full population of active borrowers.

7.2 A Model to Evaluate Ex-Ante Regulations

Understanding the effects of ex-ante regulations on welfare requires the specification of a model of borrower and lender behavior, and an equilibrium concept. I begin with the model, which is a minor expansion of the one presented in Section 2.

7.2.1 Consumer Preferences

Given a contract $\{L_k, B(L_k)\}$, I characterize the observed portion of a borrower's ex-ante utility exactly as in Section 2:

$$U_i(L_k) = u(y_0 - (H_0 - L_k)) + \beta \Bigg[\underbrace{\int_{\underline{h}}^{B(L_k)-C_i} u(y_1 - C_i) dF(H_1)}_{\text{Default}} + \underbrace{\int_{B(L_k)-C_i}^{\bar{h}} u(y_1 + H_1 - B(L_k)) dF(H_1)}_{\text{Repayment}} \Bigg].$$

As in the theoretical model, the only source of heterogeneity in U_i is C_i, the borrower's private costs of default. However, in practice, borrowers choose mortgages on the basis of a number of factors beyond just their default costs. Recall that the estimated correlation between the leverage choice and a borrower's private default costs was only 0.07. In a richly specified model, initial mortgage choice might also be a function of heterogeneity in borrower's income, preferences (e.g. risk aversion or intertemporal elasticity of substitution), or period 0 knowledge of future C_i.

I abstract from these details and consider a simplified model in which borrowers' utility for a contract with a particular leverage choice is characterized by an observed portion, as defined above, and an independent, idiosyncratic error ϵ_{iL}:

$$V_i(L_k) = U_i(L_k) + \epsilon_{iL}.$$

This error captures, in a reduced form way, all factors that influence borrowers with the same C_i to choose different contracts. When the variance of ϵ_{iL} is high, there is a weak relationship between C_i and the chosen L. When the variance is low, the correlation increases.

It is convenient to specify ϵ_{iL} to be type 1 extreme value, in which case a borrower's choice probability for a given L can be written as:

$$P_{ik} = \frac{e^{\gamma U_i(L_k)}}{\sum_{k'} e^{\gamma U_i(L_{k'})}},$$

where γ is a viscosity parameter determined by the variance of ϵ_{iL}. Of course, this specification imposes a standard independence of irrelevant alternatives (IIA) assumption, which may not hold in a more sophisticated model of heterogeneity across borrowers.

7.2.2 Lender Profits

With these choice probabilities in hand, computing lender profits is straightforward. I assume lenders are able to recover a fraction $\delta \leq 1$ of what the home is worth in the case of default. The expected profits of a lender selling contract $\{L_k, B(L_k)\}$ to borrower i with private default cost C_i are:

$$\pi(L_k, B(L_k); C_i) = -L_k + \frac{1}{1+r_f}\Bigg[\underbrace{\int_{\underline{h}}^{B(L_k)-C_i} \delta H_1 dF(H_1)}_{\text{Default}} + \underbrace{\int_{B(L_k)-C_i}^{\overline{h}} B(L_k) dF(H_1)}_{\text{Repayment}}\Bigg].$$

The expected profits of a lender are the profits for each individual i, multiplied by the probability that i chooses contract k, integrated over the distribution of C_i (specified here as $G(C)$):

$$\Pi_k = \int P_{ik}\pi(L_k, B(L_k); C_i) dG(C).$$

7.2.3 Equilibrium Concept

There is no clear consensus on the appropriate definition of equilibrium in competitive markets with adverse selection (Chiappori and Salanié, 2013). Furthermore, because equilibria often fail to exist under standard concepts, e.g. Rothschild-Stiglitz, evaluating the counterfactual implications of policy can be difficult. However, a recent development by Azevedo and Gottlieb (2016) characterizes an equilibrium concept that is both robust—an equilibrium always exists—and straightforward to implement in a variety of applications. Equilibria of this form satisfy three requirements: (i) consumers optimize over the available set of contracts, (ii) lenders make zero profits on each contract, and (iii) there is free entry, in the sense that the equilibrium is robust to small perturbations, as defined formally in Azevedo and Gottlieb (2016).

For the purposes of simulation, utilizing this equilibrium concept is straightforward. I calculate equilibrium in what Azevedo and Gottlieb (2016) call a perturbation. I propose a fixed set of contracts (in the example presented, every integer LTV between 50 and 100). I then consider a mass of uniformly distributed behavioral borrowers equal to 1 percent of the population, who always choose a given contract. Behavioral borrowers pay back the loan in all states of the world and, as a result are costless to the lender. I use a fixed point algorithm to determine equilibrium. In each iteration, consumers choose optimally taking prices as given, and interest rates are adjusted up or down for profitable or unprofitable contracts. Convergence is achieved when profits across all contracts fall below a predefined threshold. The existence of behavioral borrowers is crucial for convergence to intuitive equilibria. Because behavioral borrowers are costless, the interest rate on any contract that is only purchased by these types is reduced until either (i) a risky borrower is indifferent between the contract and his current choice or (ii) the interest rate reaches the risk free rate. This rules out equilibria with contracts that have arbitrarily high prices and only make zero profits because they are not chosen.

7.2.4 Calibration

I calibrate three features of the simulation to the estimates from Table VI. I define the mean and variance of the private costs of default based on those estimated in Column 4 of VI. Furthermore, I choose γ, or equivalently the variance of ϵ_{iL}, so that the correlation between borrowers' choice of L and C_i in Regime I below matches the estimated $\rho_{\varepsilon v}$ in Column 4. All other parameters are set based on the data when possible and explicitly described in the bottom panel of Table VII. For the purposes of the simulation, I assume that borrowers have exponential utility, with CARA coefficient a.

7.3 Welfare Implications of an LTV Cap

I consider the implications of a decreased LTV cap, that is, a limit on the initial loan provided by lenders. This can be thought of as roughly the mirror image of a standard policy in insurance markets: a mandated minimum level of coverage. I evaluate three policy regimes:

(I) **LTV Cap of 100**: In the first regime, lenders do not observe C_i, and equilibrium is as discussed above, with all loans making zero profits. The set of potential contracts contains all original LTVs between 50 and 100.

(II) **LTV Cap of 90 (No Supply Response)**: The second regime presents a naive view of the impact of an LTV cap of 90, ignoring the impacts of adverse selection. This regime evaluates the choices made by borrowers if an LTV cap of 90 were implemented but lenders did not otherwise adjust their contracts. As a result, lenders may make positive or negative profits under this regime.

(III) **LTV Cap of 90 (With Supply Response)**: The final regime considers the equilibrium allocation of credit when lenders are able to endogenously adjust contracts in response to a change in the LTV cap.

7.3.1 A Naive Evaluation of an LTV Cap: No Supply Response

I first consider a comparison of Regimes I and II, which can be thought of as the anticipated response to an LTV cap for a naive policymaker. For these purposes, I consider a naive regulator to be one who understands borrower preferences and can anticipate the contracts borrowers will choose from any given set, but who disregards adverse selection. Such a policy maker believes that the proportion of defaults for a given contract does not depend on the population purchasing that contract, and hence that there will be no supply response to a change in the LTV cap. The intuition behind this comparison is demonstrated by the dark and light gray bars in Figure VIII. This figure shows results with an exaggerated degree of adverse selection, to better present the patterns across the three regimes, while Table VII presents numbers based on simulations calibrated to the empirical results.

The black bars illustrate the allocation of original LTV under Regime I and exhibit a basic pattern of adverse selection. While borrowers would prefer initial loans with LTVs of 100 in a world with perfect information, the clustering of the riskiest borrowers raises the interest rate of a 100 LTV loan significantly. As a result, safe borrowers take smaller loans to distinguish themselves from risky types and avoid paying inflated interest rates.

Under the naive view, the only borrowers impacted by the regulation are those initially choosing LTVs above 90. The borrowers who choose contracts with original LTVs below 90 in Regime I will continue to do so, while the majority of those choosing original LTVs above 90 will bunch close to the LTV cap, creating the large mass of borrowers captured by the gray bars in Figure VIII.[48] Furthermore, the naive view will expect a significant reduction in defaults generated by the regulation. Because it assumes no heterogeneity across borrowers in default propensities, it also expects that those that choose an LTV of 90 under Regime II will default at the same rate as borrowers choosing an LTV of 90 under Regime I.

Columns 1 and 2 of Table VII compare Regimes I and II. There is indeed a reduction in loan size, from \$270,055 to \$246,265, and a corresponding reduction in average interest rates from 8.6 to 7.5 percent. This corresponds to a welfare loss of just over \$8,000. Under the naive view, the expected number of defaults is significantly larger than the true reduction, even without a supply response. The naive view suggests that an LTV cap of 90 would cut the proportion of defaults by 35 percent, from 0.12 to 0.078. Appropriately accounting for the risk of the borrowers initially allocated above 90 reveals the true reduction to be just 24 percent.

7.3.2 Allowing a Supply Response

In addition to overstating the reduction in defaults generated by the regulation, the naive view understates the reduction in mortgage size generated by knock-on effects of the regulation. Reducing the LTV cap does indeed force some risky borrowers to decrease their LTV to 90. However, as a result, the interest rates on 90 LTV loans must also rise. Correspondingly, some borrowers who previously chose LTVs of 90 will choose slightly smaller loans, thereby leading lenders to increase interest rates on those smaller loans, causing further knock-on effects. In the presence of adverse selection, lever-

[48] Because borrowers have a random component ϵ_{iL} of their preference for contracts, and because of the IIA assumption, borrowers who initially chose LTVs above 90 will not strictly choose contracts at 90. Rather, they will distribute their choices across remaining loans such that the relative choice probabilities are the same before and after the regulation.

age can be seen as a sorting device. Eliminating high LTV loans does not eliminate the incentive of borrowers to distinguish themselves, but instead forces them to do so over a smaller range of loans.

The leftward shift of the white bars in Figure VIII relative to the light gray bars demonstrates the additional reduction in mortgage size due to knock-on effects. In the calibrated simulations of Regime III, shown in the third column of Table VII, the knock-on effects cause an additional reduction in loan size of more than \$250 on average. Furthermore, because lenders in Regime III appropriately account for the reallocation of risky borrowers, the interest rates of all contracts rise from 7.5 percent on average to 7.9 percent. As a result, the average borrower has a final balance that is nearly \$600 larger under Regime III than Regime II. The reduction in loan size and increased borrower balance combine to generate a welfare loss that is \$617 larger per borrower in Regime III as compared to Regime II.

Optimal regulation involves balancing reductions in defaults with the welfare loss that results from borrowers taking smaller loans. In the simulations provided here, a naive regulator overstates the number of defaults by 11 percent and underestimates the welfare loss due to reductions in loan size by 7.5 percent. The naive estimates suggest that for this regulation to be welfare neutral, default externalities on the order of \$194,000 per default are necessary. When accounting for adverse selection, much larger default externalities are necessary to justify the regulation, on the order of \$313,000 per default.

8 Conclusion

In this paper, I empirically separate moral hazard from adverse selection in the mortgage market. I begin by developing a theoretical framework to highlight the sources of information asymmetries. In the model, moral hazard exists as the result of limited recourse: lenders cannot contract against borrowers choosing to default when it is in their ex-post interest to do so. Adverse selection, on the other hand, results from borrower heterogeneity in willingness to default. Borrowers differ in access to liquidity, value of future credit access, attachment to the home, and many other factors that influence the default choice. If borrowers know about this heterogeneity when choosing mortgage contracts, riskier borrowers will tend to prefer larger loans.

The primary empirical contribution comes in separating adverse selection from moral hazard. I do so by exploiting a natural experiment resulting from two features of Option ARMs: fixed payments and variable interest rates. Because monthly payments do not change, the balances borrowers owe are a direct function of market interest rates. This creates a distinction between borrowers' initial leverage choices and the balances they owe ex-post. To isolate plausibly exogenous variation in balances, I focus on difference-in-difference variation in interest rates that comes as the result of the financial index used to determine rate adjustments. Because of the unexpected divergence between LIBOR and Treasury rates during the crisis, borrowers experienced substantially different balances as a function of the loan's index and origination month.

This variation in borrowers' balances allows me to construct a series of instruments to identify the causal effect of home equity on default—the moral hazard effect—and subsequently to back out the role of adverse selection. I find significant evidence of both information asymmetries. Moral hazard is responsible for 60-70 percent of the baseline correlation between leverage and default, while adverse selection is responsible for the remaining 30-40 percent. The estimated moral hazard effect at 24 months suggests that a policy that reduced all borrowers' loan-to-values at 24 months would have reduced defaults by over 8 percent.

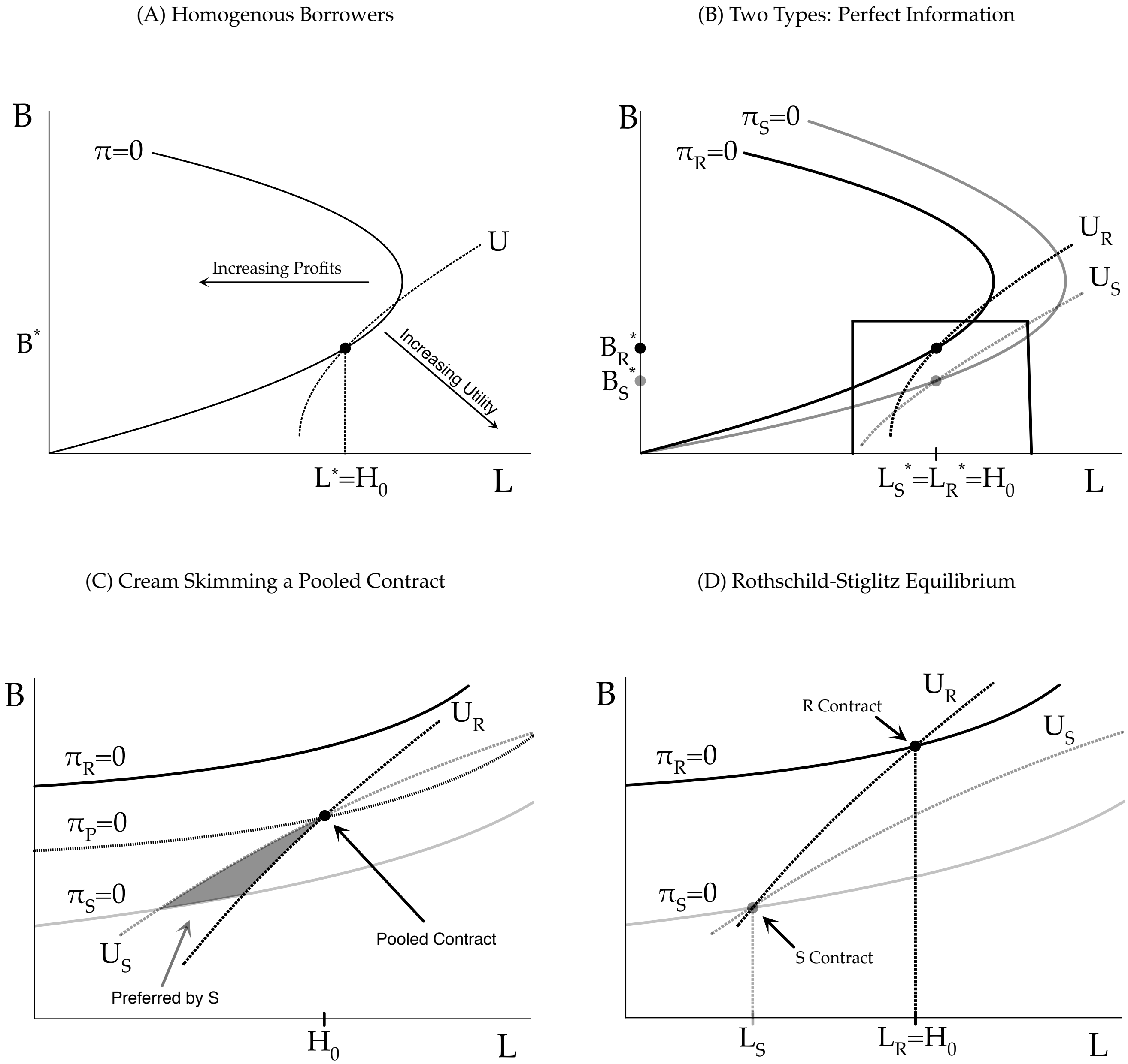
The main welfare implications of adverse selection come in its impact on equilibrium. As in standard insurance models, adverse selection imposes an externality on low risk borrowers, who must take smaller loans than they would in a world with perfect information in order to distinguish themselves from riskier types. The final contribution of this paper is to construct and simulate a model of competitive equilibrium to consider the consequences of this externality for policy. Because even defining competitive equilibrium is a notorious challenge in the presence of adverse selection, I use the robust equilibrium concept recently developed by Azevedo and Gottlieb (2016).

I evaluate the impact of a reduction in an LTV cap, a common policy aimed at reducing defaults by limiting borrowers' initial leverage. I find that a naive policymaker who does not account for adverse selection will significantly *overestimate* the number of defaults prevented by a reduced LTV cap and significantly *underestimate* the welfare losses generated by borrowers' taking smaller loans. The effects of the cap propagate through the distribution. Risky borrowers are forced to take smaller loans, but safer borrowers choose to do so as well in order to differentiate themselves. I estimate that externalities on the order of $313,000 per default are necessary to make a reduction in the LTV cap from 100 to 90 welfare neutral.

This paper separates adverse selection from moral hazard in a particular segment of the mortgage market and examines a single policy. However, the relative role of these information asymmetries is relevant to some of the most important policy questions for the market as a whole. There is significant debate over a number of core mortgage regulations in the US, including the mortgage interest tax deduction and the role of the GSEs. Some argue that the potential magnitude of positive externalities from homeownership is insufficient to justify the current level of intervention in the mortgage market. The existence of adverse selection provides an additional rationale for intervention. Even in the absence of positive homeownership externalities, policies that encourage borrowers to take on larger loans may be welfare enhancing in the presence of adverse selection.

Along these lines, understanding the importance of information asymmetries may help to explain when and why some segments of the mortgage market break down. For some observable portions of the population, mortgage credit is effectively unavailable. If this is due to moral hazard, there is little room for welfare improving intervention. Defaults may simply be so high in those populations that no interest rate is profitable for borrowers, even in the absence of adverse selection. However, if these markets are unravelling due to adverse selection, there may indeed be place for regulation. While this paper provides only a first step, fully understanding the relative roles of adverse selection and moral hazard is key to determining the effectiveness of a broad class of mortgage policies.

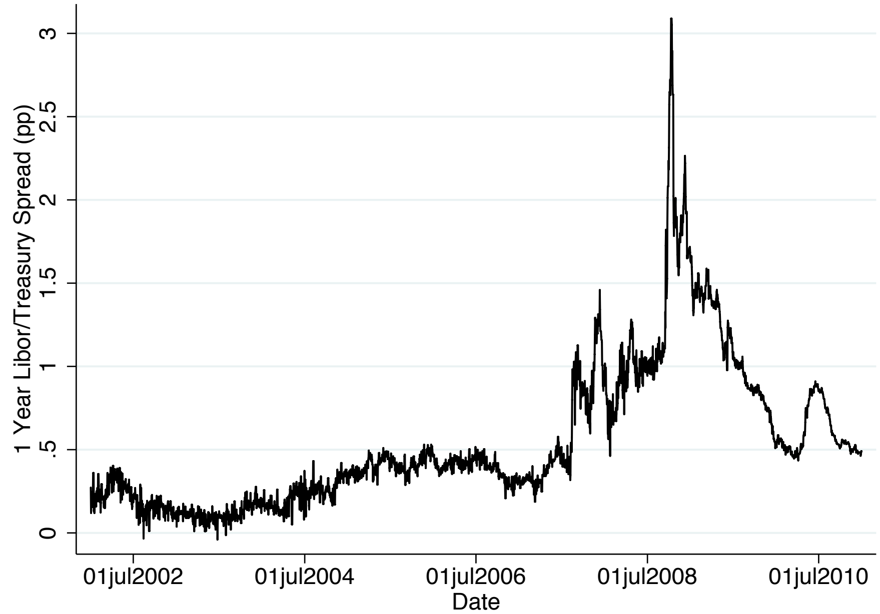
FIGURE I
THEORY: EQUILIBRIUM CONTRACTS WITH AN LTV CAP OF 100



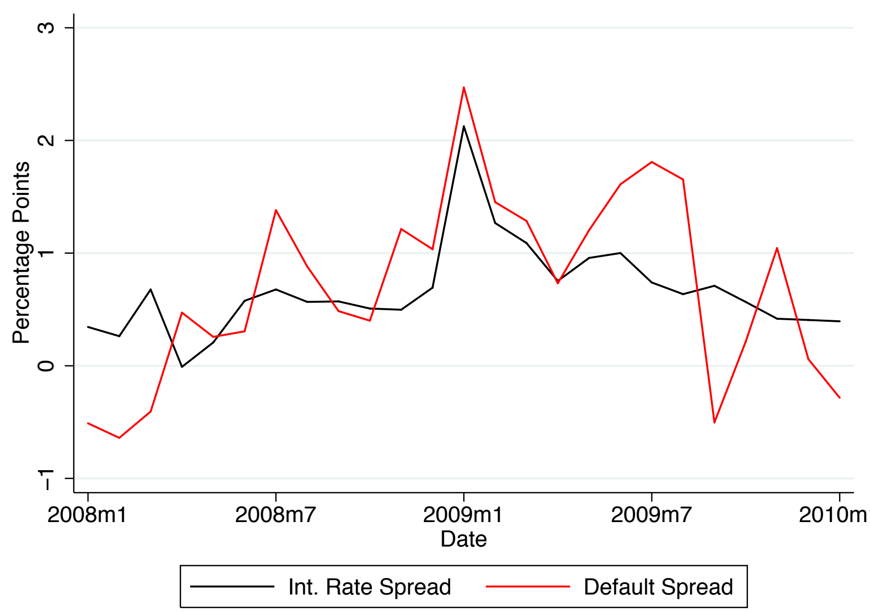
Each graph shows the contract space for stylized two period mortgage loans. Balance is shown on the y-axis, the initial loan is shown on the x-axis. Borrowers' indifference curves are marked with U, and lenders' zero profit lines are marked with Π. Borrower utility increases with contracts offered to the southeast, lender profits increase with contracts offered to the west. Panel A displays the equilibrium contract with a single borrower type and an LTV cap of H_0. Panel B shows the equilibrium contract with a risky (R) and safe (S) borrower, and perfect information. Both borrowers receive loans of H_0, but the riskier type pays a higher interest rate. Panel C zooms in on the boxed area of Panel B, and shows the difficulty of sustaining a pooled contract with single crossing when lenders cannot observe risk types. The shaded region shows profitable contracts preferred by the borrower to the pooled contract. Panel D shows the form of a Rothschild-Stiglitz equilibrium, should it exist. The safe borrower takes a smaller loan, $L_S < L_R$, than they would with perfect information.

FIGURE II

PANEL A: SPREAD BETWEEN LIBOR AND TREASURY INCREASED DRAMATICALLY DURING CRISIS

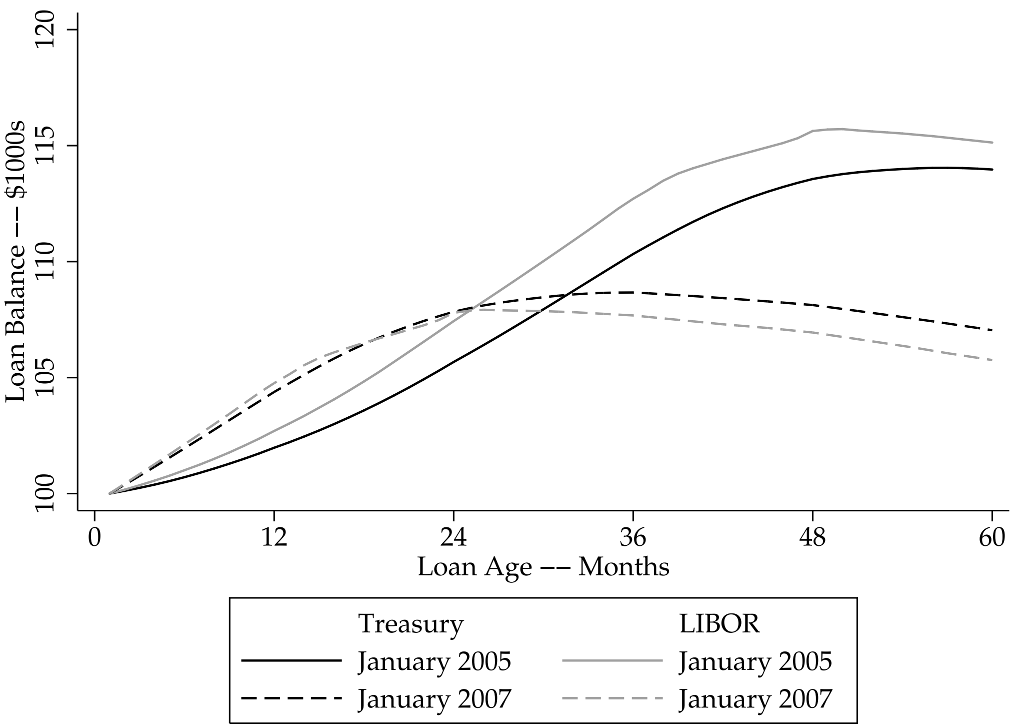


PANEL B: SPREAD IS MIRRORED IN INTEREST RATES AND DEFAULT PATTERNS FOR ARMS



Top panel shows spread between 1-year LIBOR and 1-year Constant Maturity Treasury (CMT) between 2002 and 2010. The black line in the bottom panel shows the difference in (reset) rates between LIBOR-indexed loans and Treasury-indexed loans resetting in the corresponding month. The lighter line shows the difference in the one year default probability between LIBOR and Treasury indexed loans resetting in that month. A large sample of adjustable rate mortgages of different types are included. These figures are recreations of Figure IV, Panel B and Figure V, Panel A of Gupta (2016).

FIGURE III
INDEX × ORIGINATION MONTH GENERATES DIFFERENCE-IN-DIFFERENCE VARIATION IN BALANCE TRAJECTORIES FOR OPTION ARMS



Simulated balance trajectories for $100,000 LIBOR- and Treasury-indexed loans originated in January 2005 or January 2007. Trajectories assume margin of 3.5 percent and initial payment based on 1.75 percent teaser. Treasury refers to 1-year MTA, LIBOR refers to 3-month duration.

FIGURE IV
INTEREST RATE SCHEDULE IS INCREASING IN ORIGINAL LOAN-TO-VALUE

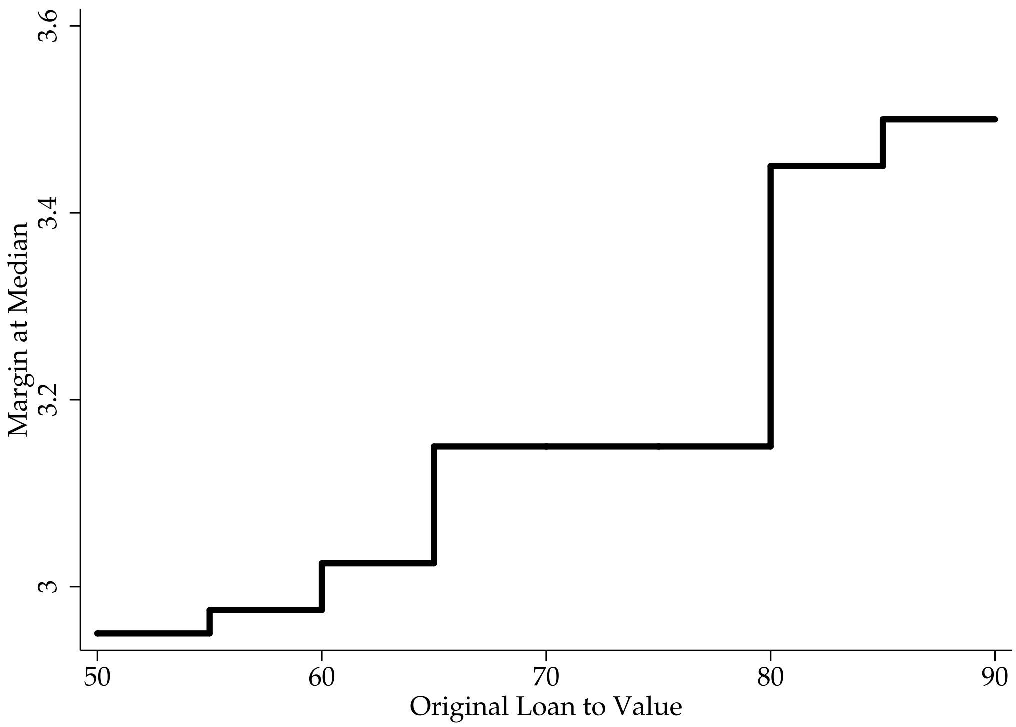
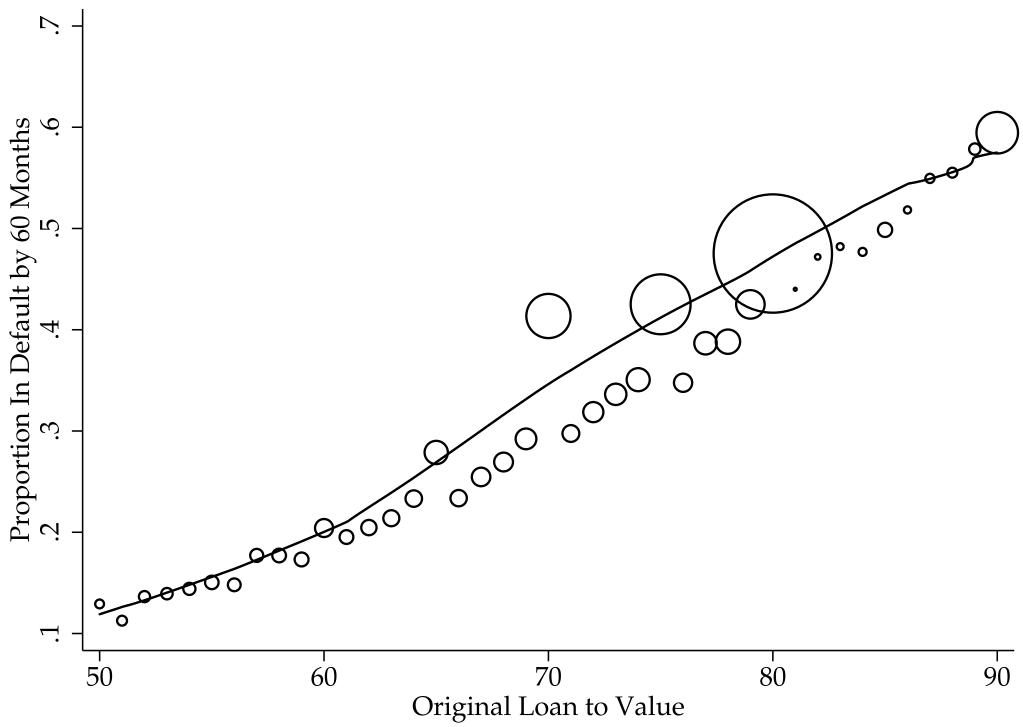


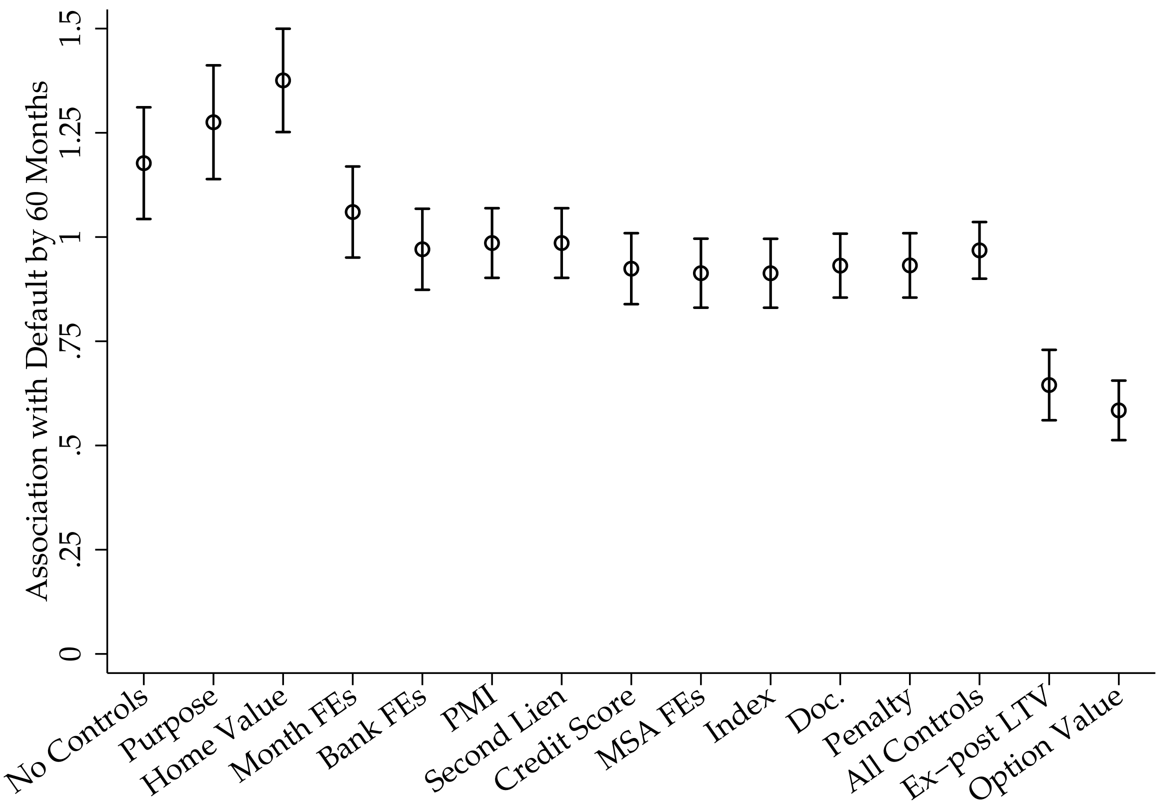
Figure displays the median margin, the fixed portion of each borrower's interest rate, for each 5 point loan-to-value bin. Full sample of Option ARMs is included.

FIGURE V
ORIGINAL LTV IS POSITIVELY CORRELATED WITH DEFAULT WITHIN 60 MONTHS



Hollow dots show the average proportion of loans defaulting within the first 60 months for each 1-point bin of original loan-to-value. Size of dots is proportional to number of borrowers within each bin. Default is defined as 60 or more days past due. The solid line shows a local linear smoothing of the raw data. Full sample of Option ARMs is included.

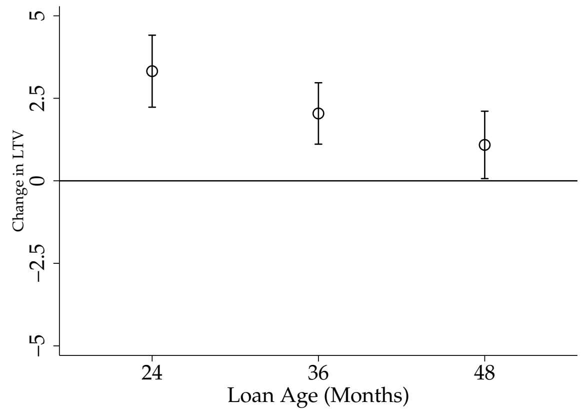
FIGURE VI
CORRELATION BETWEEN ORIGINAL LTV AND DEFAULT HOLDS CONDITIONAL ON INFORMATION AVAILABLE TO LENDER



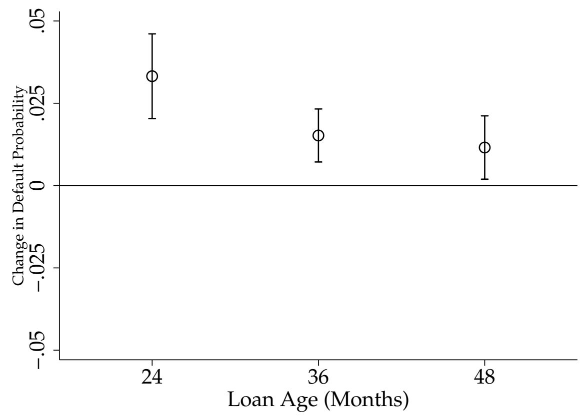
Results from OLS regressions of default within the first 60 months on original loan-to-value. Circles show coefficients on original loan-to-value with 95 percent confidence intervals based on standard errors clustered at the MSA level. The leftmost coefficient includes no controls, and each step to the left increases the set of controls included. Purpose refers to dummies for purchase, refinance, or cash out refinance. Bank FEs include originator and servicer fixed effects. Credit score refers to dummies for each 20-point bin of original FICO, with an additional category for missing values. Index is a dummy for LIBOR. Penalty is equal to one if the loan features a prepayment penalty. Full controls additionally includes a dummy for single family home and investor vs. owner occupant. Ex-post LTV refers to the imputed loan-to-value at 60 months based on initial contract terms. Option value provides a more flexible specification of mortgage value, including six leads and lags of home prices and interest rates at 60 months. Full sample of Option ARMs is included.

FIGURE VII
REGRESSIONS OF DEFAULT AND CREDIT SCORE ON INSTRUMENT FOR HOME EQUITY

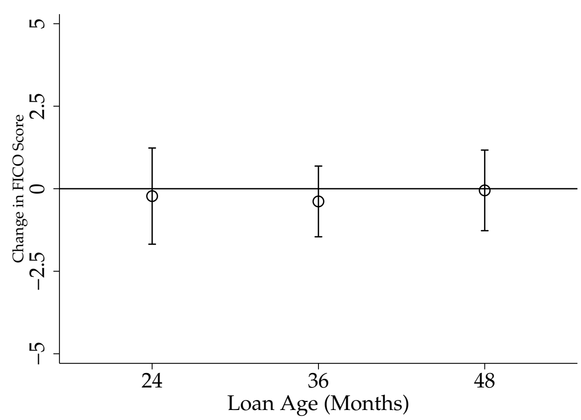
PANEL A: INSTRUMENT RELEVANCE—INSTRUMENT PREDICTS CURRENT LOAN-TO-VALUE



PANEL B: REDUCED FORM—INSTRUMENT PREDICTS DEFAULT ACROSS LIFECYCLE

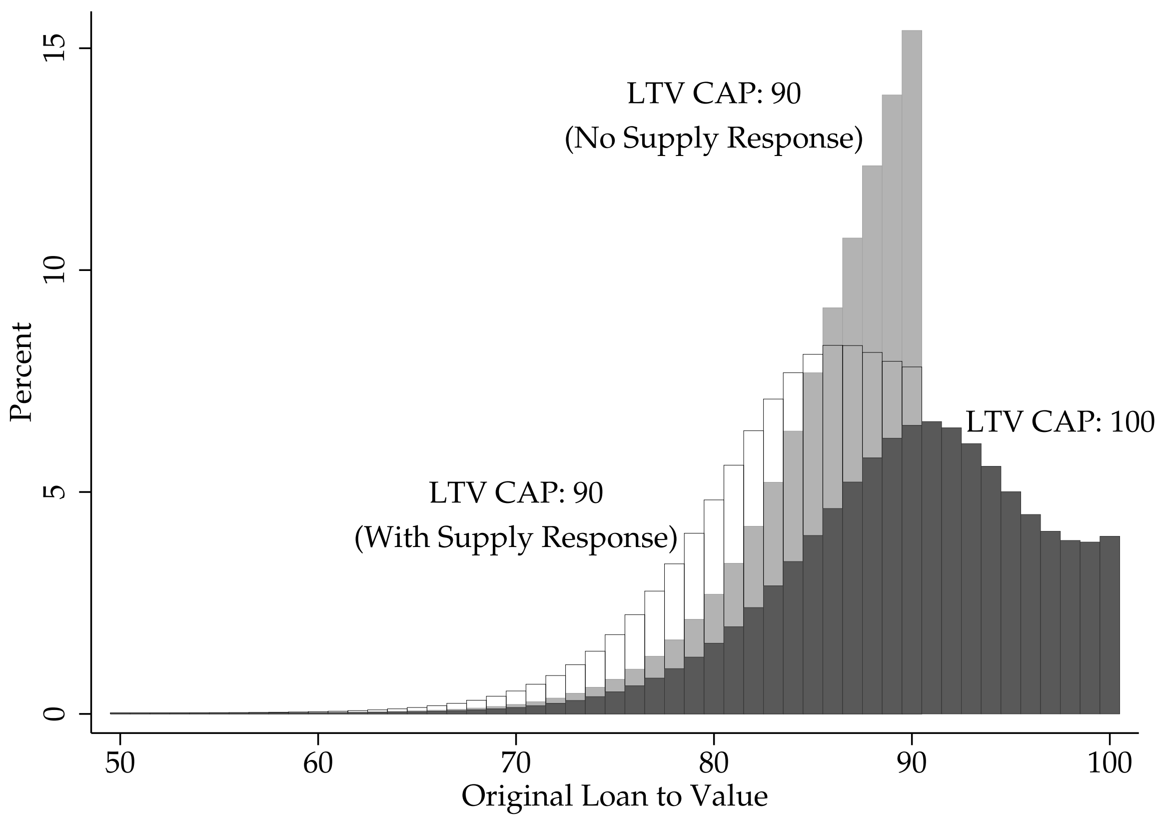


PANEL C: INSTRUMENT EXOGENEITY—INSTRUMENT DOES NOT PREDICT CREDIT SCORE



OLS regressions of outcomes on simulated instrument including origination month and index fixed effects. The simulated instrument is the mechanical calculation of balance based upon the borrowers' index choice and origination month. Margin is fixed to 3.5 for all borrowers, original loan to $100,000 and initial monthly payment is based on 1.75 percent teaser rate. Top panel shows the outcome of default within one year at 24, 36, and 48 months, where default is defined as being 60 or more days past due. Bottom panel shows the outcome of borrowers FICO scores for those remaining at 24, 36 and 48 months.

FIGURE VIII
EFFECT OF LTV CAP OF 90 ON LEVERAGE: WITH AND WITHOUT SUPPLY RESPONSE



Bars show simulated proportion of borrowers choosing each original LTV under three regimes. The dark gray bars show equilibrium LTV choices at an LTV cap of 100, the light gray show borrowers' LTV choices after a reduction in the LTV cap to 90, but allowing no changes in contracts below 90. White bars show equilibrium LTV choices with an LTV cap of 90 allowing for the supply response. Figure is based on exaggerated level of adverse selection. Table VII shows appropriately calibrated results.

Table I
Summary Statistics: Balance Across Indicies

	Treasury		Libor	
	Mean	SD	Mean	SD
FICO Score	706.1	45.9	713.8	45.1
Original Balance	370.5	264.4	346.1	282.1
Loan for Purchase	0.33		0.42	
No/Low Documentation	0.79		0.77	
Primary Residence	0.77		0.68	
Condo, Co-op or Multifamily	0.14		0.16	
Prepayment Penalty	.99		0.94	
Margin	3.21	0.53	2.85	0.51
Original LTV	76.6	8.40	77.0	8.30
State:				
- California	0.46		0.35	
- Florida	0.14		0.16	
- Arizona	0.043		0.040	
- Nevada	0.037		0.054	
Origination Year:				
- 2004	0.081		0.31	
- 2005	0.41		0.35	
- 2006	0.43		0.24	
- 2007	0.082		0.089	
Observations	490132		45199	

Summary statistics for full sample of Option ARMs. Treasury refers to loans indexed to Treasury rates, LIBOR refers to those indexed to LIBOR.

Table II
Positive Correlation Tests: Original Leverage and Interest Rate Predict Default

	Panel A: Association Between Contract Terms and Delinquency by 60 Months							
	No Controls	Month FEs	Full Controls	Low/No Doc	No Controls	Month FEs	Full Controls	Low/No Doc
Original Loan-to-Value	1.177***	0.929***	0.992***	1.122***				
	(0.068)	(0.064)	(0.037)	(0.040)				
Margin					0.222***	0.113***	0.068***	0.071***
					(0.007)	(0.005)	(0.004)	(0.005)
Mean of Dep. Var.	0.42	0.42	0.42	0.47	0.42	0.42	0.42	0.47
Raw Correlation	0.20	0.20	0.20	0.22	0.24	0.24	0.24	0.24
χ^2 Test Statistic	22585	15096	11018	1257	30362	7680	414	618
N	534909	534909	534909	419645	534909	534909	534909	419645
Origination Month FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Index FEs	No	No	Yes	Yes	No	No	Yes	Yes
MSA FEs	No	No	Yes	Yes	No	No	Yes	Yes
Full Controls	No	No	Yes	Yes	No	No	Yes	Yes
	Panel B: Association Between Contract Terms and Delinquency by Year							
	12 Months	24 Months	36 Months	48 Months	12 Months	24 Months	36 Months	48 Months
Original Loan-to-Value	0.097***	0.452***	0.803***	0.956***				
	(0.007)	(0.023)	(0.030)	(0.035)				
Margin					0.013***	0.047***	0.074***	0.076***
					(0.001)	(0.002)	(0.004)	(0.004)
Mean of Dep. Var.	0.035	0.15	0.29	0.38	0.035	0.15	0.29	0.38
Raw Correlation	0.06	0.15	0.19	0.20	0.10	0.21	0.26	0.25
χ^2 Test Statistic	175	1663	1290	1878	172	1008	821	974
N	534909	534909	534909	534909	534909	534909	534909	534909
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Results from OLS regressions of the binary outcome of default on borrowers' original loan-to-value or margin and various controls. Default is defined as 60 or more days past due. Full controls refers to fixed effects for index type, documentation, the loans purpose and occupancy, the existence of prepayment penalties and private mortgage insurance, and single family homes. I also include indicators for each 20-point bin of borrowers' FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. I allow individual state time trends. Raw correlation refers to the pairwise correlation between original loan-to-value or margin and default. The χ^2 test statistic is the test suggested in Chiappori and Salanié (2000), and tests for a correlation between the generalized residual from probit versions of the specifications presented here, and the residual from an OLS regression of original loan-to-value or margin on the listed covariates. Standard errors are clustered at the MSA level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE III
FIRST STAGE: INSTRUMENTS PREDICT REALIZED NEGATIVE EQUITY AND LOAN-TO-VALUE

	Panel A: OLS Regressions of Observed Loan-to-Value and Home Equity on Simulated Instruments							
	24 Months		48 Months		24 Months		48 Months	
Simulated Home Equity ($100,000s)	1.345*** (0.214)	0.795*** (0.197)	0.314** (0.152)	−0.001 (0.108)				
Simulated Loan-To-Value					3.559*** (0.621)	1.736*** (0.471)	1.076** (0.521)	0.090 (0.459)
F (Simulated Instrument)	39.5	16.3	4.3	0.0	32.8	13.6	4.3	0.0
F (Fixed Effects)	10.3	7.1	6.0	3.8	5.0	11.3	4.0	4.6
N	265134	265134	107917	107917	268364	268364	108987	108987
Orig. Month/Index FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	No	Yes	No	Yes	No	Yes	No	Yes
Full Controls	No	Yes	No	Yes	No	Yes	No	Yes
	Panel B: OLS Regressions of Imputed Loan-to-Value and Home Equity on Simulated Instruments							
	24 Months		48 Months		24 Months		48 Months	
Simulated Home Equity ($100,000s)	1.349*** (0.226)	0.339** (0.164)	0.593*** (0.133)	0.273*** (0.105)				
Simulated Loan-To-Value					2.632*** (0.463)	0.545 (0.363)	1.835*** (0.426)	0.287 (0.329)
F (Simulated Instrument)	35.7	4.2	19.8	6.8	32.3	2.3	18.5	0.8
F (Fixed Effects)	11.3	6.8	8.5	5.4	8.1	8.5	7.2	8.3
N	443600	443600	443600	443600	443600	443600	443600	443600
Orig. Month/Index FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	No	Yes	No	Yes	No	Yes	No	Yes
Full Controls	No	Yes	No	Yes	No	Yes	No	Yes

First stage regressions of measures of borrower equity on instruments for equity based on borrowers' index and month of origin. Displayed in each column is the coefficient from regressing borrowers' true equity value, measured either as the loan-to-value ratio (in percentage terms) or as the level of negative equity in $100,000s,on the simulated instrument for that equity. The simulated instrument is calculated using the borrowers true index and origination month to determine interest rates but fixing all other loan terms to standard values: a margin of 3.5%, an initial minimum payment based upon a 1.75% teaser rate, home price appreciation equal to the national average, and the assumption that the borrower always makes minimum payments. With these terms, home equity can be calculated mechanically. The F(simulated instrument) is the F-statistic from this regression, while F(Fixed effects) is the F-statistic from regressions that include the full set of interactions between origination month and index type as instruments. Full controls refers to fixed effects for index type, documentation, the loans purpose and occupancy, the existence of prepayment penalties and private mortgage insurance, and single family homes. I also include indicators for each 20-point bin of borrowers' FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. I allow individual state time trends. Standard errors are clustered at the MSA level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE IV
SEPARATING ADVERSE SELECTION AND MORAL HAZARD:
THE IMPACT OF ORIGINAL AND CURRENT LEVERAGE ON 1 YEAR DEFAULT PROBABILITIES

	Panel A: OLS and IV Regressions at 24 Months Including Current Negative Equity					
	Baseline	OLS	IV	Baseline	OLS	IV
Original Loan-to-Value	0.586***	0.252***	0.331***	0.721***	0.407***	0.260***
	(0.046)	(0.054)	(0.118)	(0.026)	(0.037)	(0.047)
Current Negative Equity in $100,000s		0.059***	0.045**		0.061***	0.089***
		(0.006)	(0.021)		(0.005)	(0.010)
Mean of Dep. Var	0.264	0.264	0.264	0.264	0.264	0.264
N	265134	265134	265134	265134	265134	265134
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	No	No	No	Yes	Yes	Yes
Full Controls	No	No	No	Yes	Yes	Yes
	Panel B: OLS and IV Regressions at 24 Months Including Current Loan-to-Value					
	Baseline	OLS	IV	Baseline	OLS	IV
Original Loan-to-Value	0.586***	−0.059	−0.241	0.721***	0.244***	0.229***
	(0.046)	(0.056)	(0.234)	(0.026)	(0.053)	(0.050)
Current Loan-to-Value		0.573***	0.735***		0.402***	0.415***
		(0.029)	(0.212)		(0.037)	(0.041)
Mean of Dep. Var	0.264	0.264	0.264	0.264	0.264	0.264
N	265134	265134	265134	265134	265134	265134
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	No	No	No	Yes	Yes	Yes
Full Controls	No	No	No	Yes	Yes	Yes

OLS and IV regressions of default between 24 and 36 months on borrowers original loan-to-value and current equity at 24 months, defined as either the level of negative equity in $100,000s (Panel A), or current loan-to-value (Panel B). Default is defined as 60 or more days past due. Baseline refers to OLS regressions omitting current equity. IV regressions include the full set of interactions between index and origination month as instruments for current equity. Full controls refers to fixed effects for index type, documentation, the loans purpose and occupancy, the existence of prepayment penalties and private mortgage insurance, and single family homes. I also include indicators for each 20-point bin of borrowers' FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. I allow individual state time trends. Standard errors are clustered at the MSA level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE V
HETEROGENEITY IN THE IMPACT OF ORIGINAL AND CURRENT LEVERAGE ON 1 YEAR DEFAULT PROBABILITY

	Panel A: IV Regressions at 24 Months by State Recourse Status					
	Full Recourse			Limited Recourse		
Original Loan-to-Value	0.546***	0.530***	0.494***	0.741***	0.276***	0.239***
	(0.044)	(0.116)	(0.179)	(0.025)	(0.045)	(0.057)
Current Negative Equity in $100,000s		0.003			0.089***	
		(0.034)			(0.009)	
Current Loan-to-Value			0.044			0.421***
			(0.189)			(0.044)
Mean of Dep. Var	0.198	0.198	0.198	0.272	0.272	0.272
N	28565	28565	28565	236569	236569	236569
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	Yes	Yes	Yes	Yes	Yes	Yes
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes
	Panel B: IV Regressions at 24 Months by Original Documentation					
	No/Low Documentation			Full Documentation		
Original Loan-to-Value	0.766***	0.355***	0.336***	0.492***	−0.017	−0.095*
	(0.026)	(0.056)	(0.058)	(0.031)	(0.067)	(0.057)
Current Negative Equity in $100,000s		0.077***			0.113***	
		(0.010)			(0.015)	
Current Loan-to-Value			0.356***			0.538***
			(0.045)			(0.052)
Mean of Dep. Var	0.286	0.286	0.286	0.166	0.166	0.166
N	215366	215366	215366	49768	49768	49768
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	Yes	Yes	Yes	Yes	Yes	Yes
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes
	Panel C: IV Regressions at 24 Months by Loan Purpose					
	Purchase			Refinance		
Original Loan-to-Value	0.587***	0.119**	0.076	0.762***	0.313***	0.277***
	(0.033)	(0.056)	(0.050)	(0.026)	(0.050)	(0.059)
Current Negative Equity in $100,000s		0.101***			0.084***	
		(0.013)			(0.009)	
Current Loan-to-Value			0.452***			0.405***
			(0.042)			(0.048)
Mean of Dep. Var	0.252	0.252	0.252	0.270	0.270	0.270
N	93226	93226	93226	171908	171908	171908
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	Yes	Yes	Yes	Yes	Yes	Yes
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes

OLS and IV regressions of default between 24 and 36 months on borrowers original loan-to-value and current equity at 24 months, defined as either the level of negative equity in $100,000s (or current loan-to-value). Baseline refers to OLS regressions omitting current equity, all other specifications are IV regressions including the full set interactions between index and origination month as instruments for current equity. States are categorized as full recourse if they are considered to have strong provisions regarding deficiency judgments in Rao and Walsh (2009). Full controls refers to fixed effects for index type, documentation, the loans purpose and occupancy, the existence of prepayment penalties and private mortgage insurance, and single family homes. I also include indicators for each 20-point bin of borrowers' FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. I allow individual state time trends. Standard errors are clustered at the MSA level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE VI
JOINT ESTIMATES OF THE IMPACT OF ORIGINAL AND CURRENT LEVERAGE ON 1 YEAR DEFAULT PROBABILITIES AND LEVERAGE DEMAND

	Negative Equity			Loan-to-Value		
	24 Months	36 Months	48 Months	24 Months	36 Months	48 Months
Current Negative Equity in $100,000s	0.529*** (0.048)	0.371*** (0.058)	0.232* (0.136)			
Current Loan-To-Value				1.811*** (0.194)	1.046*** (0.199)	0.420 (0.358)
ρ: Correlation of Errors in Default and Leverage Choice	0.036** (0.017)	0.048** (0.020)	0.050 (0.044)	0.067*** (0.016)	0.071*** (0.019)	0.078** (0.037)
Default Threshold	0.906	1.707	3.822	1.338	1.681	3.144
S.D. of Default Error	1.890	2.697	4.313	0.552	0.956	2.378
N	263177	162103	106921	263177	162103	106921
Origination Quarter FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes

Estimates from joint model of leverage demand and default choice. Table shows coefficient on current equity at 24, 36 and 48 months, where current equity is defined as either the level of negative equity in $100,000s (or current loan-to-value). ρ displays the estimated correlation between the errors in the leverage and default equations, capturing adverse selection. Also shown are the default threshold for a borrower at the mean covariate level, and the standard deviation of the error in the default choice in units of current equity. Full controls refers to fixed effects for index type, documentation, the loans purpose and occupancy, the existence of prepayment penalties and private mortgage insurance, and single family homes. I also include indicators for each 20-point bin of borrowers' FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. I allow individual state time trends. Standard errors are clustered at the MSA level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE VII
SIMULATION RESULTS: THE IMPACT OF A REDUCTION IN THE LTV CAP FROM 100 TO 90

	Col. 1: LTV Cap of 100	Col. 2: LTV Cap of 90 (No Supply Response)	Col. 3: LTV Cap of 90 (With Supply Response)
Average Loan Amount	\$270,055	\$246,265	\$246,002
Average Interest Rate	8.6%	7.5%	7.9%
Average Balance	\$293,359	\$264,863	\$265,476
Defaults	12.0%	9.1%	9.2%
Naive Defaults	-	7.8%	-
Welfare Loss (CV Rel. to Col. 1)	-	\$8,135	\$8,752
		Parameters	
	Initial Price: $H_0 = \$300000$	Final Value: $H \sim N(500000, 100000)$	Proportion Behavioral: 1%
	CARA Coefficient: a=0.000002	Viscosity: $\gamma = \frac{10}{504}$	$\beta = \frac{1}{1+r_f} = 0.95$
	Prop. Recovered in Default: $\delta = 0.9$	$C_i \sim N(90,000, 190,000)$	Borrowers: $N = 1000$

Simulations from structural model described in Section 7. CARA utility assumed, 1000 simulated borrowers, with 1% behavioral. Viscosity set to match estimated $\rho = 0.067$. The first column shows equilibrium outcomes with an LTV cap of 100. The second column shows borrower responses to the removal of all contracts with initial LTV between 90 and 100, holding fixed all contracts with initial LTV less than 90. Naive defaults refer to expected defaults calculated ignoring borrower heterogeneity and extrapolating from default probabilities at each LTV with an LTV cap of 100. The third column shows equilibrium outcomes with an LTV cap of 90. CV calculated based on expected utility prior to realization of EV1 error.

References

Adams, William, Liran Einav, and Jonathan Levin. 2009. "Liquidity constraints and imperfect information in subprime lending" *American Economic Review*, 99(1): 49–84.

Agarwal, Sumit, Brent W Ambrose, Souphala Chomsisengphet, and Chunlin Liu. 2011. "The role of soft information in a dynamic contract setting: Evidence from the home equity credit market" *Journal of Money, Credit and Banking*, 43(4): 633–655.

Agarwal, Sumit, Souphala Chomsisengphet, and Chunlin Liu. 2010. "The importance of adverse selection in the credit card market: Evidence from randomized trials of credit card solicitations" *Journal of Money, Credit and Banking*, 42(4): 743–754.

Agarwal, Sumit, Souphala Chomsisengphet, and Chunlin Liu. 2016. "An Empirical Analysis of Information Asymmetry in Home Equity Lending" *Journal of Financial Services Research*, 49(1): 101–119.

Ambrose, Brent W, James Conklin, and Jiro Yoshida. 2015. "Reputation and exaggeration: Adverse selection and private information in the mortgage market" Working paper.

Amromin, Gene, Jennifer Huang, Clemens Sialm, and Edward Zhong. 2011. "Complex mortgages" National Bureau of Economic Research.

Ausubel, Lawrence M. 1999. "Adverse selection in the credit card market" working paper, University of Maryland.

Avery, Robert B, Kenneth P Brevoort, and Glenn B Canner. 2007. "2006 HMDA Data, The" *Fed. Res. Bull. A73*, 93.

Azevedo, Eduardo M, and Daniel Gottlieb. 2016. "Perfect Competition in Markets with Adverse Selection" *Econometrica*, Forthcoming.

Bajari, Patrick, Chenghuan Sean Chu, and Minjung Park. 2008. "An empirical model of subprime mortgage default from 2000 to 2007" National Bureau of Economic Research.

Berger, Allen N, W Scott Frame, and Vasso Ioannidou. 2011. "Tests of ex ante versus ex post theories of collateral using private and public information" *Journal of Financial Economics*, 100(1): 85–97.

Bester, Helmut. 1985. "Screening vs. rationing in credit markets with imperfect information" *The American Economic Review*, 75(4): 850–855.

Bhutta, Neil, Hui Shan, and Jane Dokko. 2010. "The depth of negative equity and mortgage default decisions"

Blundell, Richard W, and James L Powell. 2004. "Endogeneity in semiparametric binary response models" *The Review of Economic Studies*, 71(3): 655–679.

Brueckner, Jan K. 2000. "Mortgage default with asymmetric information" *The Journal of Real Estate Finance and Economics*, 20(3): 251–274.

Bucks, Brian, and Karen Pence. 2008. "Do borrowers know their mortgage terms?" *Journal of urban Economics*, 64(2): 218–233.

Campbell, John Y, and Joao F Cocco. 2015. "A model of mortgage default" *The Journal of Finance*, 70(4): 1495–1554.

Cardon, James H, and Igal Hendel. 2001. "Asymmetric information in health insurance: evidence from the National Medical Expenditure Survey" *RAND Journal of Economics*, 408–427.

Chari, Varadarajan V, and Ravi Jagannathan. 1989. "Adverse selection in a model of real estate lending" *The Journal of Finance*, 44(2): 499–508.

Chiappori, Pierre-André, and Bernard Salanié. 2000. "Testing for asymmetric information in insurance markets" *Journal of Political Economy*, 108(1): 56–78.

Chiappori, Pierre-André, and Bernard Salanié. 2013. "Asymmetric information in insurance markets: Predictions and tests" In *Handbook of insurance*. 397–422. Springer.

Chomsisengphet, Souphala, Anthony Pennington-Cross, et al. 2006. "The evolution of the subprime mortgage market" *Federal Reserve Bank of St. Louis Review*, , (Jan): 31–56.

Corbae, Dean, and Erwan Quintin. 2015. "Leverage and the Foreclosure Crisis" *Journal of Political Economy*, 123(1): 1–65.

Deng, Yongheng, John M Quigley, and Robert Van Order. 2000. "Mortgage terminations, heterogeneity and the exercise of mortgage options" *Econometrica*, 68(2): 275–307.

Di Maggio, Marco, Amir Kermani, and Rodney Ramcharan. 2014. "Monetary policy pass-through: Household consumption and voluntary deleveraging" *Columbia Business School Research Paper*, , (14-24).

Dobbie, Will, and Paige Marta Skiba. 2013. "Information asymmetries in consumer credit markets: Evidence from payday lending" *American Economic Journal: Applied Economics*, 5(4): 256–282.

Dunn, Kenneth B, and Chester S Spatt. 1988. "Private information and incentives: Implications for mortgage contract terms and pricing" *The Journal of Real Estate Finance and Economics*, 1(1): 47–60.

Edelberg, Wendy. 2004. "Testing for adverse selection and moral hazard in consumer loan markets"

Einav, Liran, Amy Finkelstein, and Mark R Cullen. 2010. "Estimating welfare in insurance markets using variation in prices" *The Quarterly Journal of Economics*, 125(3): 877–921.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf. 2010. "Optimal mandates and the welfare cost of asymmetric information: Evidence from the uk annuity market" *Econometrica*, 78(3): 1031–1092.

Einav, Liran, Amy Finkelstein, Stephen P Ryan, Paul Schrimpf, and Mark R Cullen. 2013. "Selection on Moral Hazard in Health Insurance" *The American Economic Review*, 103(1): 178–219.

Einav, Liran, Mark Jenkins, and Jonathan Levin. 2012. "Contract pricing in consumer credit markets" *Econometrica*, 80(4): 1387–1432.

Einav, Liran, Mark Jenkins, and Jonathan Levin. 2013. "The impact of credit scoring on consumer lending" *The RAND Journal of Economics*, 44(2): 249–274.

Elul, Ronel, Nicholas S Souleles, Souphala Chomsisengphet, Dennis Glennon, and Robert Hunt. 2010. "What" Triggers" Mortgage Default?" *The American Economic Review*, 490–494.

Financial Crisis Inquiry Commission. 2011. *The financial crisis inquiry report: Final report of the national commission on the causes of the financial and economic crisis in the United States* Public Affairs.

Finkelstein, Amy, and James Poterba. 2004. "Adverse selection in insurance markets: Policyholder evidence from the UK annuity market" *Journal of Political Economy*, 112(1): 183–208.

Finkelstein, Amy, and James Poterba. 2014. "Testing for Asymmetric Information Using Unused Observables in Insurance Markets: Evidence from the U.K. Annuity Market" *Journal of Risk and Insurance*, 81(4): 709–734.

Finkelstein, Amy, Kathleen McGarry, et al. 2006. "Multiple dimensions of private information: evidence from the long-term care insurance market" *American Economic Review*, 96(4): 938–958.

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

Frame, Scott, Andreas Lehnert, and Ned Prescott. 2008. "A Snapshot of Mortgage Conditions with an Emphasis on Subprime Mortgage Performance" *Federal Reserve's Home Mortgage Initiatives, Federal Reserve Bank of Richmond, Richmond*.

Fuster, Andreas, and Paul S Willen. 2012. "Payment size, negative equity, and mortgage default" Staff Report, Federal Reserve Bank of New York.

Garriga, Carlos, and Don E Schlagenhauf. 2009. "Home equity, foreclosures, and bailouts" *Manuscript, Federal Reserve Bank of St. Louis*.

Geanakoplos, John. 2014. "Leverage, default, and forgiveness: Lessons from the American and European crises" *Journal of Macroeconomics*, 39: 313–333.

Gerardi, Kristopher, Kyle F Herkenhoff, Lee E Ohanian, and Paul S Willen. 2015. "Can't Pay or Won't Pay? Unemployment, Negative Equity, and Strategic Default" National Bureau of Economic Research.

Gete, Pedro, and Michael Reher. 2015. "Two extensive margins of credit and loan-to-value policies" *Journal of Money, Credit, and Banking, Forthcoming*.

Gupta, Arpit. 2016. "Foreclosure Contagion and the Neighborhood Spillover Effects of Mortgage Defaults"

Harrison, David M, Thomas G Noordewier, and Abdullah Yavas. 2004. "Do riskier borrowers borrow more?" *Real Estate Economics*, 32(3): 385–411.

Hendren, Nathaniel. 2013. "Private information and insurance rejections" *Econometrica*, 81(5): 1713–1762.

Jaffee, Dwight M, and Thomas Russell. 1976. "Imperfect information, uncertainty, and credit rationing" *The Quarterly Journal of Economics*, 651–666.

Jiang, Wei, Ashlyn Aiko Nelson, and Edward Vytlacil. 2014. "Liar's loan? Effects of origination channel and information falsification on mortgage delinquency" *Review of Economics and Statistics*, 96(1): 1–18.

Karlan, Dean, and Jonathan Zinman. 2009. "Observing unobservables: Identifying information asymmetries with a consumer credit field experiment" *Econometrica*, 77(6): 1993–2008.

Keys, Benjamin J, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig. 2010. "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans" *The Quarterly Journal of Economics*, 125(1): 307–362.

Keys, Benjamin J, Tomasz Piskorski, Amit Seru, and Vincent Yao. 2014. "Mortgage rates, household balance sheets, and the real economy" National Bureau of Economic Research.

Mayer, Christopher, Karen Pence, and Shane M Sherlund. 2009. "The rise in mortgage defaults" *The Journal of Economic Perspectives*, 23(1): 27–50.

Mian, Atif, and Amir Sufi. 2015. *House of debt: How they (and you) caused the Great Recession, and how we can prevent it from happening again* University of Chicago Press.

Peek, Joe. 1990. "A call to ARMs: adjustable rate mortgages in the 1980s" *New England Economic Review*, , (Mar): 47–61.

Pence, Karen M. 2006. "Foreclosing on opportunity: State laws and mortgage credit" *Review of Economics and Statistics*, 88(1): 177–182.

Piskorski, Tomasz, Amit Seru, and James Witkin. 2013. "Asset Quality Misrepresentation by Financial Intermediaries: Evidence from RMBS Market" National Bureau of Economic Research, Inc.

Piskorski, Tomasz, and Alexei Tchistyi. 2010. "Optimal Mortgage Design" *Review of Financial Studies*, 23(8): 3098–3140.

Rao, John, and Geoff Walsh. 2009. *Foreclosing a dream: State laws deprive homeowners of basic protections* National Consumer Law Center.

Rothschild, Michael, and Joseph Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information" *The Quarterly Journal of Economics*, 90(4): 629–649.

Scharlemann, Therese C., and Stephen H. Shore. 2016. "The Effect of Negative Equity on Mortgage Default: Evidence From HAMP's Principal Reduction Alternative" *Review of Financial Studies*.

Stanton, Richard, and Nancy Wallace. 1998. "Mortgage choice: what's the point?" *Real estate economics*, 26(2): 173–205.

Stiglitz, Joseph E, and Andrew Weiss. 1981. "Credit rationing in markets with imperfect information" *American Economic Review*, 393–410.

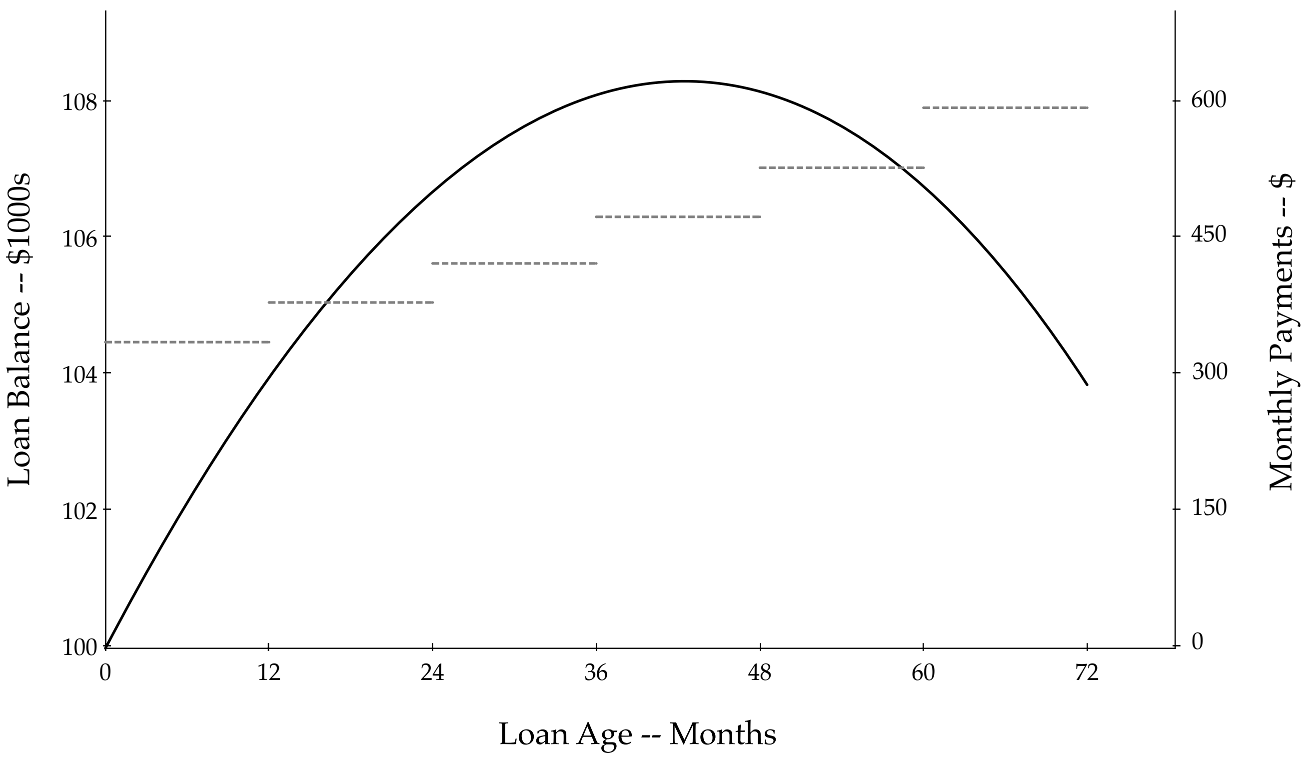
Tracy, Joseph S, and Joshua Wright. 2012. "Payment changes and default risk: The impact of refinancing on expected credit losses" *FRB of New York Staff Report*, , (562).

Vandell, Kerry D. 1995. "How ruthless is mortgage default? A review and synthesis of the evidence" *Journal of Housing Research*, 6(2): 245.

A Appendicies

A.1 Appendix Tables and Figures

FIGURE A.I
STYLIZED MONTHLY PAYMENT AND BALANCE TRAJECTORY FOR OPTION ARM



The solid line shows the balance trajectory for a stylized Option ARM with an initial loan of $100,000. The balance is initially increasing, demonstrating negative amortization. Monthly payments, shown by the dashed lines, increase by 7.5% per year regardless of balance. As payments grow, the balance begins to decrease, as shown by the parabolic shape of the balance trajectory. At 5 years the monthly payment jumps to the fully amortizing amount.

FIGURE A.II
STYLIZED EXAMPLE OF IMPACT OF INTEREST RATE VARIATION ON OPTION ARM BALANCE

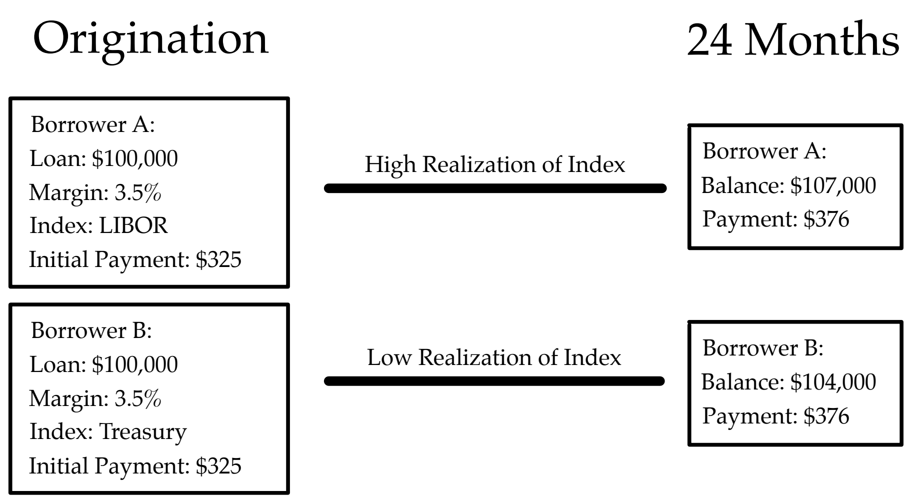
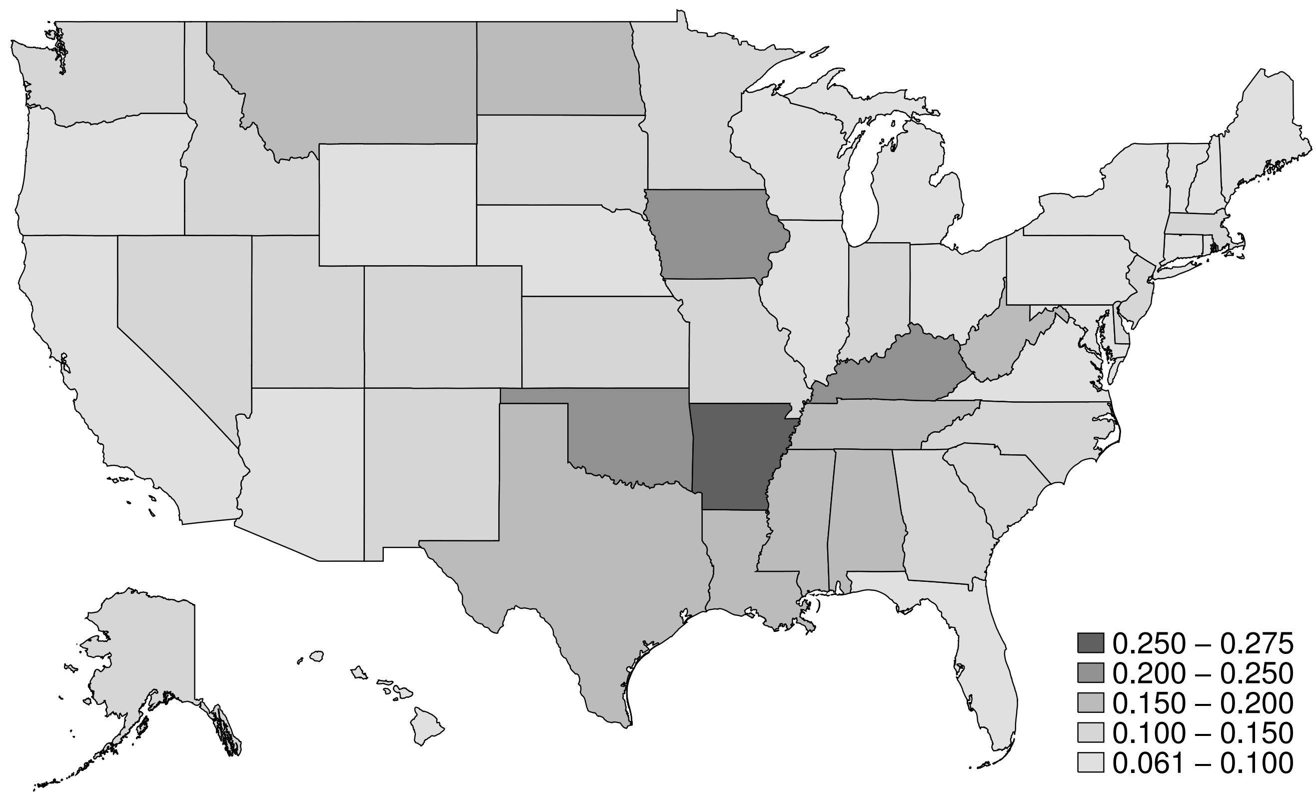


FIGURE A.III
UNIFORM DENSITY OF LIBOR INDEXED OPTION ARMS ACROSS STATES



Plot shows number of LIBOR indexed Option ARMs as a proportion of all LIBOR- or Treasury-indexed Option ARMs. The minimum is 6.1 percent, while the max is 27.5 percent. In the majority of states, between 5 and 15 percent of Option ARMs are indexed to LIBOR.

TABLE A.I
FRACTION OF LIBOR-INDEXED LOANS BY LENDER

Originator	Percent of Loans Indexed to LIBOR
American Home Mortgage	$< 1\%$
Bank United	$< 1\%$
Bank of America	85%
Countrywide	3%
Downey	0%
EMC	0%
Greenpoint	91%
IndyMac	$< 1\%$
MortgageIT	5%
Residential Funding	9%
Servicer	**Percent of Loans Indexed to LIBOR**
American Home Mortgage	$< 1\%$
Bank of America	10%
Central Mortgage	1%
Countrywide	15%
EMC	7%
IndyMac	$< 1\%$
JP Morgan Chase	2%
Nationstar	31%
Ocwen	2%
Washington Mutual	$< 1\%$

Table displays percent of LIBOR-indexed loans for the top 10 originators and servicer in the sample. Servicer is available for 99 percent of loans, while originator is only available for 27 percent of loans.

TABLE A.II
FANNIE MAE LOAN-LEVEL PRICING ADJUSTMENTS

Table 1: All Eligible Mortgages (excluding MCM) – LLPA by Credit Score/LTV Ratio									
Representative Credit Score	**LTV Range**								
	Applicable for all mortgages with terms greater than 15 years								
	≤ 60.00%	**60.01 – 70.00%**	**70.01 – 75.00%**	**75.01 – 80.00%**	**80.01 – 85.00%**	**85.01 – 90.00%**	**90.01 – 95.00%**	**95.01 – 97.00%**	**SFC**
≥ 740	0.000%	0.250%	0.250%	0.500%	0.250%	0.250%	0.250%	0.750%	N/A
720 – 739	0.000%	0.250%	0.500%	0.750%	0.500%	0.500%	0.500%	1.000%	N/A
700 – 719	0.000%	0.500%	1.000%	1.250%	1.000%	1.000%	1.000%	1.500%	N/A
680 – 699	0.000%	0.500%	1.250%	1.750%	1.500%	1.250%	1.250%	1.500%	N/A
660 – 679	0.000%	1.000%	2.250%	2.750%	2.750%	2.250%	2.250%	2.250%	N/A
640 – 659	0.500%	1.250%	2.750%	3.000%	3.250%	2.750%	2.750%	2.750%	N/A
620 – 639	0.500%	1.500%	3.000%	3.000%	3.250%	3.250%	3.250%	3.500%	N/A
< 620[(1)]	0.500%	1.500%	3.000%	3.000%	3.250%	3.250%	3.250%	3.750%	N/A

Loan-level interest rate increases necessary for different categories of original LTV and credit scores for loans delivered to Fannie Mae.

Table A.III
Impact of Original and Current Leverage on 1 Year Default Probability at 48 Months

	Panel A: OLS and IV Regressions Including Current Negative Equity					
	Baseline	OLS	IV	Baseline	OLS	IV
Original Loan-to-Value	0.283***	−0.005	−0.136	0.452***	0.187***	0.231***
	(0.038)	(0.025)	(0.219)	(0.028)	(0.021)	(0.074)
Current Negative Equity in $100,000s		0.073***	0.106**		0.054***	0.045***
		(0.004)	(0.053)		(0.002)	(0.015)
Mean of Dep. Var	0.213	0.213	0.213	0.213	0.213	0.213
N	107917	107917	107917	107917	107917	107917
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	No	No	No	Yes	Yes	Yes
Full Controls	No	No	No	Yes	Yes	Yes

	Panel B: OLS and IV Regressions Including Current Loan-to-Value					
	Baseline	OLS	IV	Baseline	OLS	IV
Original Loan-to-Value	0.283***	0.027	−0.067	0.452***	0.165***	0.371***
	(0.038)	(0.028)	(0.160)	(0.028)	(0.026)	(0.068)
Current Loan-to-Value		0.195***	0.266**		0.180***	0.050
		(0.008)	(0.122)		(0.010)	(0.038)
Mean of Dep. Var	0.213	0.213	0.213	0.213	0.213	0.213
N	107917	107917	107917	107917	107917	107917
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	No	No	No	Yes	Yes	Yes
Full Controls	No	No	No	Yes	Yes	Yes

OLS and IV regressions of default between 48 and 60 months on borrowers' original loan-to-value and current equity at 24 months, defined as either the level of negative equity in $100,000s (Panel A), or current loan-to-value (Panel B). Default is defined as 60 or more days past due. Baseline refers to OLS regressions omitting current equity. IV regressions include the full set of interactions between index and origination month as instruments for current equity. Full controls refers to fixed effects for index type, documentation, the loans purpose and occupancy, the existence of prepayment penalties and private mortgage insurance, and single family homes. I also include indicators for each 20-point bin of borrowers FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. I allow individual state time trends. Standard errors are clustered at the MSA level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE A.IV
IMPACT OF ORIGINAL AND CURRENT LEVERAGE ON CUMULATIVE DEFAULT PROBABILITIES

	Panel A: Current Negative Equity					
	36 Months			60 Months: Controlling for Trajectory		
	Baseline	OLS	IV	Baseline	OLS	IV
Original Loan-to-Value	0.857***	0.712***	0.148**	1.041***	0.892***	0.338***
	(0.030)	(0.034)	(0.064)	(0.037)	(0.032)	(0.127)
Imputed Negative Equity at 36 Months in $100,000s		0.026***	0.128***		0.004	−0.157*
		(0.004)	(0.016)		(0.005)	(0.081)
Imputed Negative Equity at 24 Months in $100,000s					−0.006*	0.174***
					(0.004)	(0.050)
Imputed Negative Equity at 48 Months in $100,000s					0.012	0.174**
					(0.008)	(0.089)
Imputed Negative Equity at 60 Months in $100,000s					0.016**	−0.060
					(0.007)	(0.054)
Mean of Dep. Var	0.310	0.310	0.310	0.454	0.454	0.454
N	443600	443600	443600	443600	443600	443600
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	No	No	No	Yes	Yes	Yes
Full Controls	No	No	No	Yes	Yes	Yes
	Panel B: Current Loan-to-Value					
	36 Months			60 Months: Controlling for Trajectory		
	Baseline	OLS	IV	Baseline	OLS	IV
Original Loan-to-Value	0.857***	0.509***	0.255***	1.041***	0.583***	0.299**
	(0.030)	(0.037)	(0.066)	(0.037)	(0.049)	(0.124)
Imputed Loan-to-Value at 36 Months		0.231***	0.402***		0.006	−0.341
		(0.016)	(0.043)		(0.022)	(0.228)
Imputed Loan-to-Value at 24 Months					0.217***	0.806***
					(0.048)	(0.187)
Imputed Loan-to-Value at 48 Months					0.088***	−0.109
					(0.032)	(0.356)
Imputed Loan-to-Value at 60 Months					0.019	0.263
					(0.024)	(0.292)
Mean of Dep. Var	0.310	0.310	0.310	0.454	0.454	0.454
N	443600	443600	443600	443600	443600	443600
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	No	No	No	Yes	Yes	Yes
Full Controls	No	No	No	Yes	Yes	Yes

The left columns show OLS and IV regressions of default by 36 months on borrowers' original loan-to-value and imputed current equity at 36 months, defined as either the level of negative equity in $100,000s (Panel A) or current loan-to-value (Panel B) at 36 months. Right columns show OLS and IV regressions of default by 60 months on borrowers' original loan-to-value and imputed current equity at 24 months, 36 months, 48 months and 60 months, defined as either the level of negative equity in $100,000s (Panel A) or current loan-to-value (Panel B). Default is defined as 60 or more days past due. Baseline refers to OLS regressions omitting current equity. IV regressions include the full set of interactions between index and origination month as instruments for current equity. Full controls refers to fixed effects for index type, documentation, the loans purpose and occupancy, the existence of prepayment penalties and private mortgage insurance, and single family homes. I also include indicators for each 20-point bin of borrowers FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. I allow individual state time trends. Standard errors are clustered at the MSA level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE A.V
IMPACT OF ORIGINAL LEVERAGE WITH FLEXIBLE CONTROLS FOR CURRENT LEVERAGE AND TIME-VARYING COVARIATES

	Panel A: Current Negative Equity					
	Cubic in Neg. Equity		Current Rates and Payments		Neg. Equity × Covariates	
	OLS	IV	OLS	IV	OLS	IV
Original Loan-to-Value	0.332***	0.296***	0.339***	0.204***	0.372***	0.309***
	(0.041)	(0.053)	(0.035)	(0.050)	(0.040)	(0.055)
Current Negative Equity in $100,000s	0.107***	0.146***	0.061***	0.089***	0.413**	−0.082
	(0.008)	(0.015)	(0.005)	(0.010)	(0.171)	(0.064)
Current Negative Equity^2	0.016***	−0.016				
	(0.002)	(0.019)				
Current Negative Equity^3	0.001***	−0.010***				
	(0.000)	(0.003)				
Minimum Payment in $			0.000***	0.000***		
			(0.000)	(0.000)		
Interest Rate			0.047***	0.047***		
			(0.003)	(0.003)		
Current Negative Equity × Fico Score						−0.000
						(0.000)
Current Negative Equity × Purchase						0.086***
						(0.028)
Mean of Dep. Var	0.264	0.264	0.275	0.275	0.264	0.264
N	265134	265134	240189	240189	265134	265134
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	No	No	No	Yes	Yes	Yes
Full Controls	No	No	No	Yes	Yes	Yes

	Panel B: Current Loan-to-Value					
	Cubic in LTV		Current Rates and Payments		LTV × Covariates	
	OLS	IV	OLS	IV	OLS	IV
Original Loan-to-Value	0.168***	0.175***	0.157***	0.127**	0.241***	0.229***
	(0.055)	(0.060)	(0.054)	(0.058)	(0.054)	(0.070)
Current Loan-to-Value	−0.542***	0.556***	0.389***	0.414***	−0.033	0.308***
	(0.144)	(0.067)	(0.038)	(0.043)	(0.751)	(0.049)
Current Loan-to-Value^2	1.128***	0.000				
	(0.176)	(1.866)				
Current Loan-to-Value^3	−0.391***	−0.038				
	(0.059)	(0.170)				
Minimum Payment in $			0.000***	0.000***		
			(0.000)	(0.000)		
Interest Rate			0.043***	0.043***		
			(0.003)	(0.003)		
Current Loan-to-Value × Fico Score						0.000
						(0.001)
Current Loan-to-Value × Purchase						0.001
						(0.189)
Mean of Dep. Var	0.264	0.264	0.275	0.275	0.264	0.264
N	265134	265134	240189	240189	265134	265134
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	No	No	No	Yes	Yes	Yes
Full Controls	No	No	No	Yes	Yes	Yes

OLS and IV regressions of default between 24 and 36 months on borrowers' original loan-to-value and current equity at 24 months, defined as either the level of negative equity in $100,000s (Panel A), or current loan-to-value (Panel B). The first two columns include a cubic in current equity. The third and fourth columns include current and original minimum payments, as well as the current interest rate. The 5th column interacts current equity with all control variables in an OLS specification. The final column includes current equity interacted with each borrowers FICO score and an indicator equal to one if the loan was used to purchase a home. Default is defined as 60 or more days past due. IV regressions include the full set of interactions between index and origination month as instruments for all terms including current equity. Full controls refers to fixed effects for index type, documentation, the loans purpose and occupancy, the existence of prepayment penalties and private mortgage insurance, and single family homes. I also include indicators for each 20-point bin of borrowers FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. I allow individual state time trends. Standard errors are clustered at the MSA level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE A.VI
IMPACT OF ORIGINAL AND CURRENT LEVERAGE ON 1 YEAR DEFAULT PROBABILITY: PROBIT ESTIMATES AND ALTERNATIVE INSTRUMENTS

	Panel A: Probit and Control Function				
	Baseline	Probit	Control Function	Probit	Control Function
Original Loan-to-Value	3.282***	2.028***	0.556**	1.602***	1.047***
	(0.100)	(0.134)	(0.260)	(0.169)	(0.296)
Current Negative Equity in $100,000s		0.251***	0.529***		
		(0.020)	(0.051)		
Current Loan-to-Value				1.347***	1.808***
				(0.111)	(0.216)
Mean of Dep. Var	0.264	0.264	0.264	0.264	0.264
N	265128	265128	265128	265128	265128
Origination Month FEs	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes
Full Controls	Yes	Yes	Yes	Yes	Yes
	Panel B: OLS and IV Incorporating Simulated Instrument				
	Baseline	OLS	IV	OLS	IV
Original Loan-to-Value	0.721***	−0.641*	0.256***	−0.541	0.229***
	(0.026)	(0.339)	(0.047)	(0.348)	(0.050)
Current Negative Equity in $100,000s		0.216***	0.090***		
		(0.060)	(0.010)		
Current Loan-to-Value				1.002***	0.415***
				(0.314)	(0.041)
Mean of Dep. Var	0.264	0.264	0.264	0.264	0.264
N	265134	265134	265134	265134	265134
Origination Month FEs	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes
Full Controls	Yes	Yes	Yes	Yes	Yes

Top panel shows probit and control function specifications of default between 24 and 36 months on borrowers' original loan-to-value and current equity at 24 months, defined as either the level of negative equity in $100,000s or current loan-to-value. Default is defined as 60 or more days past due. Baseline refers to probit regressions omitting current equity. Control Function regressions include the full set of interactions between index and origination month as instruments for current equity, and are estimated following Blundell and Powell (2004). The bottom panel includes OLS regressions as in Table **??**, but uses the simulated instrument. Full controls refers to fixed effects for index type, documentation, the loans purpose and occupancy, the existence of prepayment penalties and private mortgage insurance, and single family homes. I also include indicators for each 20-point bin of borrowers FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. I allow individual state time trends. Standard errors are clustered at the MSA level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE A.VII

IMPACT OF ORIGINAL AND CURRENT LEVERAGE ON ONE YEAR DELINQUENCY, DEFAULT, AND FORECLOSURE RATES

	30 Days Past Due			90 Days Past Due			Foreclosure		
	Baseline	Equity	Loan-to-Value	Baseline	Equity	Loan-to-Value	Baseline	Equity	Loan-to-Value
Original Loan-to-Value	0.688***	0.379***	0.211***	0.710***	0.407***	0.256***	0.494***	0.315***	0.213***
	(0.026)	(0.034)	(0.049)	(0.026)	(0.037)	(0.051)	(0.019)	(0.029)	(0.038)
Current Negative Equity in $100,000s		0.060***			0.059***			0.035***	
		(0.005)			(0.005)			(0.004)	
Current Loan-to-Value			0.404***			0.381***			0.234***
			(0.034)			(0.035)			(0.022)
Mean of Dep. Var	0.304	0.304	0.304	0.248	0.248	0.248	0.177	0.177	0.177
N	213535	213535	213535	278761	278761	278761	294636	294636	294636
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

OLS and IV regressions of default between 24 and 36 months on borrowers' original loan-to-value and current equity at 24 months, defined as either the level of negative equity in $100,000s (Panel A), or current loan-to-value (Panel B). Default is defined as 30/90 or more days past due or as foreclosure. Baseline refers to OLS regressions omitting current equity. IV regressions include the full set of interactions between index and origination month as instruments for current equity. Full controls refers to fixed effects for index type, documentation, the loans purpose and occupancy, the existence of prepayment penalties and private mortgage insurance, and single family homes. I also include indicators for each 20-point bin of borrowers FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. I allow individual state time trends. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.