SELECTION, LEVERAGE, AND DEFAULT IN THE MORTGAGE MARKET*

Arpit Gupta† **Christopher Hansman**‡

September 15, 2018

Abstract

We ask whether the correlation between mortgage leverage and default is due to moral hazard—the causal effect of leverage—or adverse selection: ex-ante risky borrowers choosing larger loans. We separate these information asymmetries using a natural experiment resulting from (i) the unique contract structure of Option Adjustable-Rate Mortgages and (ii) the unexpected 2008 divergence of the indices that determine interest rate adjustments. Moral hazard is responsible for 60-70% of the correlation, while adverse selection explains 30-40%. We construct and calibrate a simple model to show that optimal regulation of leverage must weigh defaults against under-provision of credit due to adverse selection.

JEL classification: D14, G21, D82

*This paper subsumes and extends two previously circulated papers, "Asymmetric Information in the Adjustable Rate Mortgage Market" and "Asymmetric Information and the Link Between Leverage and Mortgage Default." We would like to thank Pierre-André Chiappori, Eduardo Dávila, François Gerard, Jonas Hjort, Wojciech Kopczuk, Henrik Kleven, Rocco Macchiavello, Dan O'Flaherty, Tomasz Piskorski, Miikka Rokkanen, Christoph Rothe, Bernard Salanié, Johannes Spinnewijn, Miguel Urquiola, Stijn Van Nieuwerburgh, Eric Verhoogen, Albert Zevelev and participants of many seminars. 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 should not be construed as representing the opinions of the National Science Foundation or the Russell Sage Foundation.

†NYU Stern, Email: arpit.gupta@stern.nyu.edu

‡Imperial College London, Email: c.hansman@imperial.ac.uk (Corresponding Author)

1 Introduction

In the years since the financial crisis, household leverage—particularly vis-a-vis mortgage debt—has come sharply into relief as a critical element of financial stability (e.g. Geanakoplos and Pedersen, 2012; Mian and Sufi, 2015; Korinek and Simsek, 2016). These observations have motivated an increased empirical focus on the *determinants* and *regulation* of mortgage leverage (e.g. Bailey et al., 2017; Gete and Reher, 2016; DeFusco, Johnson and Mondragon, 2017). Meanwhile, a long theoretical literature has emphasized the profound implications of adverse selection in determining *equilibrium* leverage (e.g. Bester, 1985). Despite this work, empirical research has struggled to connect asymmetric information to household leverage.

In this paper we develop a strategy to explicitly test for adverse selection in the mortgage market and trace out implications of our findings for macroprudential regulation. The key issue researchers face in identifying selection lies in disentangling two potential explanation for an intuitive stylized fact: highly leveraged borrowers tend to default more frequently. This correlation could reflect adverse selection: ex-ante "risky" borrowers choose high leverage mortgages. However, it could alternatively reflect the causal effect of leverage: when housing prices fall, a lack of equity may directly lead high leverage borrowers to default. This second channel is often referred to as moral hazard.[1] Typically, these two mechanisms are observationally equivalent in the data.

Our research design exploits a natural experiment generated by the unique contract structure of Option Adjustable-Rate Mortgages (Option ARMs). These loans typically have interest rate adjustments tied to either LIBOR or Treasury rates. Though the choice between LIBOR and Treasury was not salient prior to the financial crisis, the unexpected divergence of these two financial indexes in 2008 caused borrowers who chose otherwise identical contracts to owe substantially different amounts ex-post. This variation allows us to (i) identify moral hazard by comparing borrowers with identical initial leverage choices but different realized home equity, and (ii) identify adverse selection by comparing borrowers with different initial leverage choices but the same realized home equity.

Disentangling these mechanisms is crucial for understanding how the *supply* of leverage—whether a lender is willing to offer a given leverage ratio, and if so, at what interest rate—is determined in equilibrium. Naturally, a lender must adjust rates on higher leverage loans to account for any causal impact on the default probability. However, adverse selection imposes further, distinct, constraints. A lender may have to radically increase rates on (or refuse to offer) a loan product that attracts only the riskiest borrowers. Consequently, the distinction has direct implications for the efficacy of macroprudential policy. As our analysis shows, a policymaker who ignores the presence of adverse selection will both *overestimate* the reduction in defaults generated by regulations on household leverage and

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

underestimate the corresponding reduction in borrowing.

The paper begins with a simple conceptual framework to clarify the sources of asymmetric information in mortgage markets. The presence of moral hazard is a standard feature of many models of mortgage default (e.g. Campbell and Cocco, 2015). Because lenders face limited access to effective recourse, home equity (or a lack thereof) is a key driver of a borrower's choice to default. However, a large literature (e.g. Deng, Quigley and Van Order, 2000) has documented notable heterogeneity in this relationship across borrowers. Some "risky" borrowers default as soon as the home is worth less than than the mortgage balance, whereas others will not to default until the balance significantly outweighs the value of the home. Adverse selection arises precisely when this heterogeneity is correlated with *demand* for leverage—that is, when those who are quickest to default (for unobservable reasons) also prefer larger loans.[2]

We then turn to our primary empirical exercise: disentangling adverse selection from moral hazard. The logic of our approach comes in recognizing that the two give distinct empirical predictions, as in Karlan and Zinman (2009). Adverse selection implies a correlation between default and a borrowers initial leverage choice—independent of the home equity a borrower actually faces. Conversely, moral hazard implies a positive correlation between default and a borrowers realized home equity—regardless of initial leverage. As a result, it is possible to separately identify the two by isolating exogenous variation in equity that is distinct from a borrower's leverage choice.

In our context, the divergence of LIBOR and Treasury rates generate exactly this sort of variation. This is due to two features of Option ARMs, which have *fixed minimum payments* in the early years of the loan, but *interest rates that update monthly*. As a result, fluctuations in interest rates generate meaningful changes in the balances borrowers will ultimately owe, without impacting the amount they must pay in the short term.[3] An identical borrower will receive different path of interest rates—and hence a different realized level of home equity—depending on (i) the loan's origination month and (ii) whether it is indexed to a LIBOR or Treasury rate. Our analysis instruments for a borrower's home equity with the interaction between the loan's financial index and origination month, effectively a form of difference-in-difference variation. Using this strategy, we 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. Our estimates imply, for example, that a 10-point reduction in a borrower's loan-to-value ratio two years after origination would reduce the average probability of default by over 4 percentage points.

The fixed (and unusually small) minimum payments for Option ARMs facilitate our analysis by

[2]This framing is analogous to a model of selection on ex-post moral hazard, as in Einav et al. (2013).

[3]While a standard adjustable-rate product might adjust payments to adjust to account for interest accrual, for Option ARMs any excess accrual is absorbed into the balance.

allowing us to set aside the role of monthly payments in the default decision. However, doing so may limit our ability to extend our insights to the broader market, where borrowers short term liquidity concerns may play an important role in the relationship between leverage and default. Borrowers may default because they chose to not—or cannot—make their monthly payments, rather than because of a lack of home equity in isolation. Unobserved sensitivity to monthly payments may also influence or correlate with borrowers demand for leverage, a form of selection.

To address the role of monthly payments—which allows us to extrapolate to mortgage markets more generally—we utilize a similar identification strategy in a large sample of more traditional adjustable-rate mortgages. For these loans, the divergence of LIBOR and Treasury rates led to sharp changes in monthly payments for similar borrowers. We use this variation to estimate the impact of payments on default. We find that the payment shock caused by an additional 100 basis point rise in interest rates—sustained over a year—leads to a 3.3 percent increase in default rates.[4] We then combine our estimates on monthly payments and home equity to provide a suggestive decomposition of the correlation between leverage and default in the market as a whole. On a broad, fairly representative sample of fixed-rate mortgages, we find that the association between an additional 10-LTV points and default (0.22 percentage points) is 32% due to moral hazard, 18% due to adverse selection, and 49% due to monthly payments.

Following our empirical analysis, our final step is to consider the implications of our findings from a regulatory perspective. Our estimates for the impacts of changes in home equity and monthly payments are directly policy relevant, quantifying the effectiveness of ex-post policies that relieve borrowers' balances or reduce payment sizes, respectively. However, quantifying the impact of ex-ante regulation in the presence of adverse selection is more challenging. While it is clear that restrictions on mortgage contracts may have profoundly different impacts on equilibrium with and without adverse selection, there is no standard framework to evaluate such policies. Even the appropriate characterization of equilibrium in competitive contexts with adverse selection is controversial, and equilibria may fail to exist under conventional definitions.[5]

To assess ex-ante regulations, we develop a stylized structural model and calibrate it using our estimates. To ensure the existence of equilibrium, we use the robust equilibrium concept recently proposed by Azevedo and Gottlieb (2017). We consider, as an example, the impact of a loan-to-value cap—a widely used policy tool intended to reduce mortgage defaults by restricting leverage. We find that this policy is indeed effective in limiting defaults, but to a lesser extent than a naive regulator who ignores the presence of adverse selection would expect. Furthermore, adverse selection generates

[4]While this result is consistent with a broader literature examining the role of ARM reset shocks, we also find evidence of selection on borrower's sensitivity to these payment shocks: borrowers who are prone to defaulting in response to higher monthly payments tend to choose higher leverage mortgages.

[5]e.g. Rothschild-Stiglitz.

unexpected welfare losses due to *knock on effects.* While the regulation only mechanically forces the risky borrowers initially above the cap to take smaller loans, its effects ultimately propagate through the entire distribution. Safer borrowers initially below the cap also choose to take less leverage in order to maintain separation from riskier types and obtain suitably low interest rates. In equilibrium, interest rates rise across the whole loan-to-value distribution, and all borrowers choose smaller loans. Appropriately accounting for adverse selection, we 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 is most closely related to the growing empirical literature on asymmetric information in credit markets.[6] Our primary contribution to this work comes in identifying a form of ex-post variation in the size of a borrower's loan that is unknown to borrowers when selecting contracts. 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; Hertzberg, Liberman and Paravisini (2018) on online consumer credit; 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. Furthermore, we do so in the largest and arguably most important consumer debt market in the US,[7] in which there has been relatively little work on disentangling these asymmetries.[8]

Our second contribution comes in connecting the work on information asymmetries to broader debates on the role of macroprudential regulations on household leverage. As noted in Cerutti, Claessens and Laeven (2017), macroprudential caps (along the dimensions of loan-to-value or debt-to-income) are in place in over 41 countries as of 2014. A growing literature explores the ways in which this prudential regulation impacts credit markets, and evaluates their effectiveness. These include: Greenwald (2018), Gete and Reher (2016), DeFusco, Johnson and Mondragon (2017), and Behn, Haselmann and Vig (2017). We provide the first empirically grounded evidence highlighting the com-

[6]We are 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).

[7]For example, 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.

[8]One early exception is Edelberg (2004), who uses structural assumptions to test for adverse selection and moral hazard in a broad class of consumer debts, including mortgages. Ambrose, Conklin and Yoshida (2015) and Jiang, Nelson and Vytlacil (2014), consider selection into and within low documentation mortgages. There is also related work in the home equity lending market, in particular Agarwal et al. (2011) and Agarwal, Chomsisengphet and Liu (2016).

plications that adverse selection poses for regulations of this form.

Our estimates of the causal impacts of home equity and monthly payments also each contribute to the literature on mortgage default in their own right. Our key innovation is our use of borrower-level cross-sectional variation in interest rates, which is otherwise impossible in most contexts. Vandell (1995) provides an overview of early research on the role of home equity in borrower's 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. Our borrower-level variation in home equity avoids the potential for measurement error and other endogeneity concerns inherent to the use of home price variation. A key exception is Ganong and Noel (2017), which uses experimental evidence from the HAMP program based on a sample of loan modification applicants. We complement their work by focusing on less deeply underwater borrowers, and—in contrast to their findings—find strong evidence that borrowers balances causally drive default decisions. Our estimates on monthly payments—and our identification strategy—complement 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 (2017), Tracy and Wright (2012), Di Maggio et al. (2017), and Gupta (2017).

The paper is structured as follows: Section 2 lays out key definitions and provides a framework describing information asymmetries in the mortgage market. Section 3 provides background information on ARMs and the data used in the paper. Sections 4 presents the empirical strategy. Section 5 present results on the Option ARM sample, while Section 6 uses the 5/1 ARM sample to isolate the role of payment shocks. Section 7 shows the results of simulations, and Section 8 concludes. In the Appendix, Section B describes a model formalizing the intuition in Section 2, and Section C contains additional robustness tests.

2 Definitions and Sources of Information Asymmetries

In this section we define adverse selection and moral hazard as they pertain to the relationship between mortgage borrowers and lenders. We 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. Finally, we consider the implications of these asymmetries for mortgage market equilibrium.

2.1 Definitions of Adverse Selection and Moral Hazard

The definitions of adverse selection and moral hazard that we 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 borrowers loan liability and default. That is, amongst homogenous 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 we do not specify a mechanism in the baseline definition.

Our definition of 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, our definition aligns with this traditional notion. Default itself can be thought of as the non-contractable action taken by the borrower. However, default may not be an active choice in certain circumstances. Borrowers may be insolvent or credit constrained to the extent that they are mechanically unable to make payments. While the discussion below primarily highlights the strategic channel, the definition we use does not, in principal, rule out defaults due to a mechanical relationship between the loan balance and default. While this distinction is generally not critical for the policy implications we consider, we examine the importance of liquidity concerns in more detail in Section 6.

2.2 Sources of Information Asymmetries in Mortgage Markets

In this subsection, we discuss sources of adverse selection and moral hazard in the mortgage market. While the definition above is agnostic regarding mechanisms, understanding why these information asymmetries might be present is helpful to frame further discussion.

Moral hazard results naturally from the limited liability that is implicit in a mortgage contract. The particular legal restrictions on contracts vary from state to state, with some explicitly prohibiting lenders from recovering any excess balance from the borrower beyond the home itself in the event of default.[9] 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. For any given borrower, a larger balance increases the probability of owing more than the home is worth, which in turn increases the probability of default.

What is the source of heterogeneity in borrower risk that leads to 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 borrower's have individual (and private) insights into the future path of real estate prices, this model leaves little room for differences in risk. Borrowers would 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 necessary need for a behavioral explanation for this phenomenon. There are real costs associated with default, including credit score reductions, moving costs, and social stigma. These costs may vary in the population, generating differences in the risk of default across borrowers.

This notion can be shown—in a stylized manner—via a simple default rule. Consider a borrower i who owes a loan balance B for a house with value H.[10] The borrower must choose between repaying the loan and receiving the value of the home, or defaulting, which entails both losing the home and incurring any costs of default C_i (e.g. moving costs or credit score reductions). This gives rise to the following rule: borrowers default if

$$\underbrace{H - B}_{\text{Value of Repaying}} < \underbrace{-C_i}_{\text{Value of Defaulting}}$$

The key feature for our purposes is simply that C_i may be different for different individuals i. Some borrowers may have very low C_i and default anytime $H < B$, whereas others may have relatively high

[9] 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.

[10] H might also be thought of as the option value of continuing to service the mortgage.

C_i. For a given distribution of housing values ex-ante—and a fixed balance—borrowers with lower C_i will default with a higher probability. Further, a portion of this cost is likely to be unobservable to lenders.

Adverse selection arises when this heterogeneity in default risk is correlated with a borrower's demand for leverage when selecting a mortgage.[11] In other words, if riskier borrowers—e.g. those with unobservably low values of C_i—prefer higher leverage mortgages.

In classic models of adverse selection, this correlation exists because borrowers have private information about their risk when choosing a loan.[12] For example, if they have some private signal regarding the value of C_i. A borrower who knows ex-ante that he is relatively unlikely to repay will prefer to put less money down—i.e. to choose a higher leverage loan—and will be willing to accept an increased interest rate to do so. In Appendix B, we outline a simple two period version of such a model, and show that a Spence-Mirrlees single crossing condition holds: risky borrowers are relatively more willing to accept large balances (i.e. higher interest rates) in exchange for more leverage.[13]

However, a correlation between default risk and leverage demand might occur even if borrowers are not so forward looking as to take into account their default probability when selecting a mortgage. For example, a borrower who has limited access to cash for a down payment when selecting a mortgage may also find it disproportionally costly to raise the funds to repay the loan. The key feature is simply that borrowers are heterogeneous in their risk, and that this may correlate with leverage demand. For the purposes of our empirical exercises, the precise mechanism underlying this correlation is not crucial, although we are required to make assumptions when conducting policy simulations in Section 7.

2.3 Implications of Adverse Selection for Market Equilibrium

Mortgage contracts may differ in many dimensions, but a key tradeoff is generally between the leverage ratio—equivalently, the size of the down payment—and the interest rate. Because lenders may offer these different interest rate–leverage pairs, the theoretical implications of adverse selection are familiar from the literature on collateralized lending following Bester (1985) and more general work on screening (e.g. Rothschild and Stiglitz, 1976).[14] While the appropriate definition of equilibrium is controversial, and different forms of selection and market power can have different consequences, a few standard insights are suggested by the theory.

[11] More formally, if the *unobservable* portion of this heterogeneity is correlated with leverage demand.

[12] See, e.g. Brueckner (2000) and Bester (1985).

[13] In other words, borrowers with low costs of default place a higher value on implicit hedge against home price reductions provided by the option to default. 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 action of default, the willingness to take that action varies in the population, and borrowers are privately informed of their own willingness.

[14] This is in contrast to the consequences in environments with fixed (or no) collateral following Stiglitz and Weiss (1981).

A first implication is that, in competitive contexts, leverage acts as a screening device. For a given home, any borrower would prefer to take a high leverage mortgage at fair interest rates. However, because safer borrowers value leverage less they are willing to take smaller mortgages in order to differentiate themselves and avoid paying the inflated interest rates charged for the high leverage contracts that attract riskier types. As a result, lenders offer a menu of contracts with low leverage, low interest rate contracts to attract safer borrowers, and high leverage, high interest rate contracts to attract risky borrowers. Faced with such a menu, safe borrowers wind up choosing sub-optimally low leverage compared to a world with full information. This generates a welfare loss analogous to the under-provision of insurance in the healthcare context. In Appendix B we demonstrate these insights in a simple model.

A second implication is that any regulation that restricts (or expands) the set of contracts lenders may offer will have consequences for the form of all contracts in equilibrium— even those that are not directly impacted. This is because, in a world with adverse selection, it matters exactly *which* borrowers choose each contract: the same leverage ratio might require a much higher interest rate if it attracts a riskier pool of individuals. Accordingly, a regulation which forces even a small number of borrowers to choose a different contract may cause interest rates to adjust on all loans and ultimately prompt initially unaffected borrowers to choose different contracts as well.

These implications suggest two observable properties of equilibrium in a mortgage market with adverse selection:

- Multiple contracts offered to a given borrower featuring an increasing relationship between leverage and the interest rate. Appendix Figure A.I shows that this holds empirically in the market we study. The figure plots the upward slope between borrowers initial leverage (Loan-to-Value ratio) and the median interest rate.[15]
- A positive correlation exists between leverage and default, conditional on all observables. In other words, observationally equivalent borrowers who select higher leverage mortgages default at higher rates. This is the credit market analogue of the positive correlation test suggested in Chiappori and Salanié (2000). Figure I confirms that an initial version of this test holds in our setting as well: there is an unconditional correlation between leverage and the probability of default in our context. In Section 5 we confirm that the relationship also holds conditional on observables available to the lender.

While adverse selection generates these predictions, it is possible that both might also be present in a world with no selection. In particular, if moral hazard is present—and borrowers have idiosyncratic

[15] In particular, the figure displays the *margin* as specified in the loan contract: the fixed mark-up over some floating financial index.

preferences for leverage that cause some to prefer more or less leverage—equilibria might feature both of these observable properties even in the absence of any unobserved heterogeneity in borrower risk. In the remainder of the paper, we develop an empirical strategy to explicitly test for the presence of adverse selection.

3 Background and Data

In this section, we provide historical background on Adjustable-Rate Mortgages (ARMs) generally and the Option Adjustable-Rate Mortgage in particular. We focus on the unique features of the Option ARM product that are key to our 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. We then discuss the natural experiment we exploit: the divergence of the two most common financial indices used for Option ARMs. Finally, we 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 ARMs

Though ARMs were commonplace prior to the Great Depression, U.S. regulation effectively limited residential mortgage products to long-term fixed-rate loans until the late 1970s. 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. Driven by high interest rates 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.

Payments on ARMs are determined by an annuity calculation taking into account scheduled amortization over the loan term (typically 30 years, but 15 years is common as well), as well as a composite (or market) interest rate to account for lender repayment. This composite rate is the sum of a fixed margin component (priced at loan origination), as well as a fluctuating index component tied to a market index. Variation in relative spreads on this market index result in within-month variation in interest rates and payments made by borrowers, which we exploit in our analysis.

According to lenders, the potential danger of these unexpected payment increases motivated the creation of the Option ARM.[16] 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.[17] 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 margin component and an index term to proxy for the cost of funds to the lender.

(III) *Negative amortization:* The minimum payment required in a given month will be often 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 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 the 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,[18] borrower's balances change as a function of realized interest rates. In the next

[16] 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.

[17] 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.

[18] All analyses performed here consider outcomes within the first 5 years. Appendix Figure A.II presents a sample balance

subsection, we 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.[19]

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 (see Amromin et al. (2011)). Their role in the crisis has been highlighted by various media sources and policymakers—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 in favor of these products, both suggesting that they approximate the optimal mortgage contract when borrowers have stochastic income (Piskorski and Tchistyi, 2010)[20] and noting potential benefits from a macroprudential perspective (Campbell, Cocco and Clara, 2017).

3.2 Diverging Indices and Interest Rate Resets

Our identification strategy exploits the divergence of the two most common financial indices used to determine interest rate adjustments for Option ARMs. Interest rates for ARMs are typically tied to LIBOR or Treasury rates.[21]

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,[22] the two indices 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

and payment trajectory for an Option ARM to highlight these product features from origination through that period.

[19]See the 2008 Mortgage Market Statistical Annual.

[20]See also Guren, Krishnamurthy and McQuade (2018), who confirm the findings Piskorski and Tchistyi (2010), but suggest that the Option ARM may perform particularly poorly in crises.

[21]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.

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

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 of Option ARMs. Most originators appear to specialize in either LIBOR or Treasury indices, and some offer only Treasury rates. A similar—if slightly less pronounced—pattern can be seen among servicers. According to Gupta (2017), differences across lenders are often a function not of the borrowers they lend to, but rather their intentions on the secondary market.

The interest rate divergence between the two indices is illustrated in Panel A of Figure II, which shows the spread between the 1-year Constant Maturity Treasury (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 the impact of this phenomenon for a large sample of adjustable-rate mortgages, including the 5/1 ARMs we study in Section 6. 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. 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. While this figure does not capture our primary strategy precisely, it illustrates the basic concept: the divergence of LIBOR and Treasury indices generated meaningful differences in the interest rates borrowers faced, with subsequent consequences for default patterns.

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 product, a change in the interest rate also changes the monthly payment, which adjusts to ensure that the payments are fully amortizing at the new rate. Alternatively, because the minimum payments are set in the initial period for Option ARMs, 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.III 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 borrower's 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.[23] Each of the four shows a distinct balance trajectory.

3.3 Data

The data on ARMs and 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. We focus on a sample of around 500,000 Option ARMs originated between 2004 and 2007, tied to either LIBOR or Treasury rates. We limit our sample to loans with initial loan-to-value ratios between 50 and 100.

3.4 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. For Option ARMs highlighted in Panel A, note that Treasury is the dominant index, representing approximately 90 percent of loans. Despite this, the majority of variables are reasonably balanced 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 origination loan-to-value ratio (LTV), is also quite similar across the two indices, at 76.6 for Treasury and 77 for LIBOR. Nearly all loans are

[23] 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.

subject to some form of prepayment penalties, and the majority of both Treasury and LIBOR loans are for primary residences.[24] 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.IV 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.

3.5 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.[25] In this observable dimension, borrowers with Option ARMs reflect the general pool of borrowers rather than some particularly subprime subset.[26]

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.[27]

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.[28] The average initial loan size for Option ARMs is also larger than that of conforming

[24] The stated level of owner occupancy is likely an overstatement due to false reporting by unreported investors (Piskorski, Seru and Witkin, 2015).

[25] 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.

[26] 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.

[27] All gained as a proportion of the market between 1996 and 2006: California by 25 percent, Florida by 60 percent, Arizona by 55 percent, and Nevada by 44 percent. All statistics presented here are available in the 2006 Mortgage Market Statistical Annual.

[28] The average original LTV for Fannie Mae and Freddie Mac in 2007 was 72 and 71, respectively, according to Frame, Lehnert

loans,[29] 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 borrower's ability to meet their monthly obligations. This was especially true given that most loans were made to borrowers with high credit scores.[30] 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 Empirical Strategy

In this section we describe a simple empirical model of borrower's leverage and default choices. We then clarify the distinct empirical predictions of both moral hazard and adverse selection this context. Finally, we show how the model translates to a series of estimating equations and describe the strategy we use to separately identify adverse selection and moral hazard.

4.1 An Empirical Model of Leverage Demand and Default

To specify a model of borrower default, we begin with the rule discussed in Section 2. At a given point in time t, a borrower i defaults if their equity in the home—the difference between the home price (H_{it}) and the loan balance (B_{it})—drops below some borrower specific cost of default: $H_{it} - B_{it} < C_{it}$. For our purposes, it is convenient to define the borrowers *Negative Equity* as the difference between the balance owed and the value of the home ($E_{ijt} = B_{it} - H_{it}$). Rearranging, and decomposing the individual specific costs into observable and unobservable components, we may rewrite this as:[31]

$$D_{i,t+1} = \mathbb{1}\{\alpha E_{it} + x_i'\beta + \varepsilon_{it} > 0\} \tag{1}$$

where $D_{i,t+1} = 1$ if borrower i defaults between period t and $t+1$.

and Prescott (2008). The average LTV for subprime adjustable-rate mortgages was over 80 as early as 2004 (Chomsisengphet, Pennington-Cross et al., 2006).

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

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

[31] We decompose $C_{it} = x_i'\frac{\beta}{\alpha} + \frac{1}{\alpha}\varepsilon_{it}$, where x is a set of observable characteristics available to the lender at origination, and $\frac{1}{\alpha}\varepsilon_{it}$ is the unobservable portion of the borrower's default costs. α is then the precision of the default costs in the population.

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

$$L_i = x_i'\psi + v_i. \tag{2}$$

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

(I) *Moral Hazard:* $\alpha > 0$ provides evidence of a moral hazard effect. α quantifies the causal impact of the borrower's equity on default. α can also be interpreted as the precision of unobserved default costs in the population.

(II) *Adverse Selection*: $\rho = Corr(v_i, \varepsilon_{it}) > 0$ 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_{it}).

4.2 Estimating Equations

In practice, our primary specifications collapse equations 1 and 2 into a single reduced form specification.[32] For borrower i in MSA j at loan age t, we estimate:

$$D_{it+1} = \mathbb{1}\{\alpha E_{it} + \gamma L_i + x_i'\beta + \omega_{m(i)} + \delta_{index_i} + \zeta_{j(i)} + u_{it} > 0\}. \tag{3}$$

Here, D_{jt+1} is a measure of default between year t and $t+1$. 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_i contains the set of all borrower characteristics and loan features known to the bank at the time of contracting. We include fixed effects for the origination month of the loan ($\omega_{m(i)}$), the index choice (δ_{index_i}), and the borrower's MSA ($\zeta_j(i)$). Standard errors are clustered at the MSA level. Most specifications additionally allow for state specific time trends. Importantly, we separately estimate the above cross-sectionally at different loan ages, so do not include loan age effects (or time effects separate from origination month) in our specifications.

In this reduced form specification, we interpret $\gamma > 0$ as evidence of adverse selection and $\alpha > 0$ as evidence of moral hazard. For robustness, and to provide parameters that are more directly

[32] In particular, we write $\varepsilon_{it} = \gamma v_i + u_{it}$, where $\gamma > 0$ holds if v_i and ε_{it} are positively correlated, that is, if there is adverse selection (In the normal case, $\gamma = \rho \frac{\sigma_\varepsilon}{\sigma_v}$). Replacing v_i using Equation 2 gives

$$\varepsilon_{it} = \gamma(L_i - x_i'\psi - \theta_{m_i} - \eta_{index_i}).$$

Replacing ε_{it} in Equation 1 and collapsing month and index fixed effects gives Equation 3.

relevant for our simulations in Section 7, we also explicitly estimate a joint model of Equations 1 and 2 (alongside Equation 4, described below). This joint specification allows us to directly recover the correlation ρ, as well as the distribution of the costs of default in the population.[33]

4.3 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_{it}. In the absence of other differences, borrowers with identical L_i will tend to have identical E_{it}. For borrowers consistently making minimum payments, there are only two factors that might cause those with identical L_i to have different E_{it}: 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_{it}. 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_{it} is measured with error. Similarly, variation across time in interest rates is likely correlated with macro conditions, while cross-sectional variation potentially reflects borrower's endogenous contract choices. Isolating exogenous variation in E_{it} is non-trivial but necessary to accurately estimate α and γ.

To address the biases in the estimation of moral hazard and adverse selection, we introduce an instrumental variable approach that utilizes the divergence of Libor and Treasury rates. In our analysis, E_{ijt} is instrumented by a function of the borrower's index choice and origination month:

$$E_{it} = f_t(m_i, index_i) + \lambda_{m(i)} + \mu_{index_i} + x_i'\pi_t + \phi_{j(i)} + e_{it}. \tag{4}$$

[33]To estimate Equations 1, 2 and 4 jointly, we impose a parametric structure on the errors. In particular:

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

Again, we 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. We effectively employ a control function approach following Blundell and Powell (2004), incorporating an additional linear equation. The benefit of this approach is that we are able to recover a few parameters that provide a basis for simulation in Section 2. Perhaps most importantly, we 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, we are able to recover the underlying distribution of the default costs C_{ijt} for any individual i. Recalling that $-C_{it} = \sigma_\varepsilon(x_i'\beta + \omega_{m(i)} + \zeta_{j(i)} + \delta_{index_i} + \varepsilon_{it})$, we have:

$$C_{it} | x_i, \omega_{m(i)}, \zeta_{j(i)}, \delta_{index_i} \sim N\left(\frac{-(x_i'\beta + \omega_{m(i)} + \zeta_{j(i)} + \delta_{index_i})}{\alpha}, \frac{1}{\alpha} \right).$$

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

As described in Section 3, we 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 for this product. Each {Index Type, Origination Month} pair generates a unique trajectory of interest rates for a borrower. Utilizing this difference-in-difference variation allows us 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.[34]

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

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

In words, we 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 we also use a second, complementary, option. The basic idea behind this option is to produce a strong instrument for a borrower's balance using only the origination month and index. To do so, we 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.[35] 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. We refer to the instrument developed using this calculation as the "simulated" instrument.

5 Results

In this section, we describe the central empirical results of the paper. We begin by defining a few variables used in the analysis. Next, we describe first stage results: the instruments we use are correlated with borrower's balances and payments, are directly predictive of default in a reduced form, and do not predict borrower characteristics such as credit scores. We then turn to the results of the primary, single equation model of default. We find strong evidence of both moral hazard and adverse selection, which hold across numerous robustness checks. Finally, we discuss the estimation of the

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

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

joint model of leverage demand and default choice, which provides parameters that directly inform the simulations presented in Section 7.

5.1 Definitions of Key Variables

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

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

(III) **Current Equity (E_{it}):** we use two alternative measures of E_{it} 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, we follow the literature and impute the current home value based on local home price indices.[36]

Unfortunately, we do not observe E_{it} for borrowers who exit the sample prior to time t. This prevents us 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, we are able to predict the observed values of E_{it} with a high degree of accuracy ($R^2 > 0.95$). In regressions that incorporate the full sample, we use these imputed values of E_{it}, which are available for all borrowers, rather than the observed values.

5.2 First Stage

The plots in Figure IV highlight a few features of the instruments for E_{it} used in the main analysis. Because the primary specification uses a large number of fixed effects, and hence does not provide

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

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_{it} when the instrument suggests balances should be high? Panel A of Figure IV 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_{it} on the instrument, controlling for origination month and index fixed effects. When the simulated instrument is high, borrowers' E_{it} 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_{it} with default E_{it}. 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.

Table II 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, we also include F-statistics calculated using the full set of fixed effects. 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).[37]

5.3 Main Results: Single Equation Model

The primary specifications attempt to isolate adverse selection and moral hazard following Equation 3. In the main tables, we use linear probability models and a standard instrumental variables approach. In the Appendix, we 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, we add a measure of E_{ijt} to the baseline regression but do not account for endogeneity in E_{ijt}. Finally, in the IV regressions we 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

[37]Because the current loan-to-value or negative equity at 24 months is only observed for loans that actually survive to those points, the first stage regressions in the first four columns of Panel A are necessarily conducted on a selected sample. Columns (5)-(8) of panel A replaces the observed values of E_{it} with imputed values, which allows the use of the full sample. While the F-statistics only exceed 10 without covariates, both versions of the instrument are relevant (if weak) with the full sample.

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 III presents the primary set of specifications for Option ARMs, showing a cross-section of borrowers 24 months after origination.[38] This table includes E_{ijt} defined as both current negative equity (Panel A) and current LTV (Panel B). In different specifications we 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, 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 con-

[38]Cross-sections at other loan ages are shown in the Appendix.

trols, 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.

5.4 Robustness for the Single Equation Model

Appendix C discusess a number of tables intended to serve as robustness checks to Table III and to provide alternative estimates of interest. Appendix Table A.II considers how the results in Table III change across three relevant subgroups. The basic results are as follows: we find stronger moral hazard effects in states with limited recourse vs. those with full recourse, stronger selection effects in no/low documentation loans versus full documentation loans, and minor differences in purchase vs. refinance mortgages.

To show that the effects we find are not specific to default at 24 months, Appendix Table A.III presents identical regressions to those in Table III, except with current E_{ijt} defined at 48 months and the dependent variable defined to be default between 48 and 60 months. The results are similar to those at 24 months, albeit somewhat muted.

To deal with the potential for differential survival to 24 months, Appendix Table A.IV considers the impact of the original LTV and current E_{ijt} on cumulative default outcomes, that is, on default at any point up to 36 months. These estimates show strong evidence of both moral hazard and adverse selection. Details of the approach and estimation are outlined in Appendix C.

Appendix Tables A.V A.VI and A.VII explore further robustness. The results are robust to (i) alternative functional forms of E_{ijt}, (ii) probit and control function specifications, which are potentially more realistic than the linear probability model, (iii) the use of the simulated instrument rather than the full set of fixed effects, and (iv) alternative definitions of default, ranging from mild (30 days past due) to extreme (foreclosure).

5.5 Joint Model

The final step of our empirical analysis is to estimate a joint model of leverage demand alongside the default choice. Doing so allows us 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, we slightly reduce the richness of included controls, e.g. substituting MSA fixed effects with state fixed effects. We 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 IV 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 borrower's 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. We use these estimates to parameterize the model in Section 7.

6 Isolating Payment Shocks

The preceding results decompose adverse selection and moral hazard using a sample of Option ARMs. The particular features of the product enabled us to largely ignore the role of monthly payments in our analysis. However, outside of the special case of the Option ARM, borrowers with high leverage mortgages may default for a third reason beyond adverse selection or the causal impact of home equity. In particularly, there is potential for a payment channel: borrowers with high leverage must make larger monthly payments to service their debts.[39]

To account for this possibility, and understand the full set of mechanisms connecting leverage

[39] The causal role of mortgage payments on default has also been analyzed by a growing literature, including Ganong and Noel (2017) and Di Maggio et al. (2017).

to mortgage default, we conduct a complimentary empirical approach on a sample of more traditional adjustable rate mortgages: 5/1 arms. For these products, differences in indexes generated sharp changes in borrowers monthly payments.

In this section, we lay out an identification strategy to estimate the impact of monthly payments on default in this sample. We then describe the results of this approach. Finally, we use our estimates, combined with those in the previous section, to provide a back-of-the-envelope decomposition of the overall correlation between leverage and default into three channels: (i) adverse selection, (ii) moral hazard, and (iii) the causal effect of higher payments.

6.1 Identification on the ARM Sample

In contrast to Option ARMs, 5/1 arms feature fixed interest rates for the initial 5 years. However, at 5 years, borrowers receive a new interest rate, referred to as the *reset rate*, based on some underlying financial index. To amortize the loan, payments change alongside these interest rate changes. As a result, borrowers with similar loans, but with different indexes, may face sharply different monthly payments.

To capture variation in monthly payments driven by this difference, we conduct a strategy similar to the one outlined above. For borrower i at loan age t we estimate:

$$D_{it+1} = \mathbb{1}\{\eta r^o_{it} + \tau r^r_{it} + x'_i\beta + \omega_{m(i)} + \delta_{index_i} + \zeta_{j(i)} + u_{it} > 0\}. \quad (5)$$

Where D_{it+1} continues to refer to delinquency between age t and $t+1$. We include two new variables of interest: r^o_{it} refers to the *original interest rate* of the borrower. That is, the interest rate the borrower faces in the first five years. r^r_{it} refers to the reset rate. Our primary coefficient of interest is τ the causal impact of the reset rate on default.

As before, a difference-in-difference approach provides an instrument. However, in this case, the object of instrumentation is the interest rate upon reset:

$$r^r_{it} = m_i \times index_i + \lambda_{m(i)} + \mu_{index_i} + x'_i\pi_t + \phi_{j(i)} + e_{it}. \quad (6)$$

In measuring the reset interest rate, we consider both the composite rate—including both the index component as well as a margin term—as well as the actual monthly payment made by the borrower.

6.2 Estimating Payment Shocks

To illustrate the validity of the instrument as applied to 5/1 ARMs, recall Figure II, which demonstrates instrument relevance. Panel B of this figure shows the in-sample variation in interest rates paid by borrowers resetting on different months across borrowers linked to LIBOR and Treasury indices; the resulting interest rates apply to these borrowers for payments made for the following 12 month period. Overlaid on the interest rate spread is the default spread—the in sample variation in mortgage default by borrowers linked to LIBOR and Treasury loans. The correlation between these two spreads indicates that when LIBOR-linked borrowers face higher monthly payments than Treasury-linked borrowers due to movements in the underlying index; they also default at greater rates in the following 12 month period.

Appendix Figure A.VI suggests that endogeneity is not driving this result. This specification regresses the instrumented post-reset interest rate against several borrower characteristics. We find no evidence that borrower exposure to higher index-driven rates is related to other ex-ante borrower characteristics, suggesting that we can view payment shocks as exogenous in our context.

Table V uses the index-driven interest rate instrument to causally identify the role of payment shocks on loan default in our 5/1 ARM sample. We focus on borrowers in the year after the first loan reset, taking place after five year of initial payments have been made by borrowers. We exploit the fact that borrowers in the post-reset window will face different payment shocks due to cross-sectional variation in interest rates across different indices. The first column of Panel A shows results from an OLS regression of mortgage default on interest rates in the year after default, controlling for the fixed effect of the index choice itself, as well as the origination month and servicer. The second column shows the IV version of the same regression, based on the strategy described above. Column 3 of this table also adds a control for origination loan-to-value. Columns 4-6 repeat the analysis in columns 1-3, but include a variety of other standard controls (origination loan characteristics, origination credit score). We focus on column 6, which we view as our most comprehensive specification. This estimate suggests that a 100 basis point increase in instrumented interest rates results in a 3.3 percentage point increase on loan default, relative to a baseline of 6.8% in the year after reset. This sizable estimate suggests a substantial sensitivity of borrower default to monthly payment shocks, and so suggests that a component of the relationship between borrower leverage and default more generally may be accounted for the by the fact that greater leverage mechanically induces higher monthly payments.

To further explore the heterogeneity in this result, in Panel B of this table we explore the sensitivity of interest rate shocks on mortgage default by subsetting on different levels of origination leverage. By doing so, we examine how sensitive are borrowers to shocks to payments across different levels of ex-ante leverage. Columns 4–6 of this table instrument for the post-reset rate using our interaction of index rate and month of origination. We control, as before, for a variety of controls, including loan level covariates, as well as the MSA, index type, servicer, and origination month. We also control, in this panel, for the Current CLTV, to try to isolate the role of origination leverage separate from its impact on ex post leverage outcomes.

We find in this specification that borrowers exhibit substantially greater sensitivity to payment shocks when they have also selected into greater degrees of leverage. Because we are controlling for ex post CLTV, this is not just because borrowers who post smaller down payments are more likely to be in negative equity; but instead it suggests that borrowers are selected along sensitivity to payment shocks. Borrowers with greater leverage may by unobservably risky to lenders not just because of a propensity to default at a higher level of home equity, but also because they may be more sensitive to payment shocks. This suggests the potential for multidimensional heterogeneity in borrower risk, an important avenue for further work.

6.3 Decomposing the Channels Linking Leverage and Mortgage Default

To (i) bring together the three channels connecting leverage to mortgage default estimated in this and previous sections, and (ii) provide a suggestive view of the importance of our results for the broader market, we perform a simple, back of the envelope decomposition of the correlation between leverage and default. To do so, we use a sample of Fixed Rate Mortgages (FRMs) purchased by Fannie Mae and Freddie Mac representing over 11 million loans. We select these loans because they provide a picture of perhaps the most standard US mortgage: the long-term, fixed rate agency loan.

We begin by establishing a baseline correlation between leverage and default in this sample. This estimate is show in the first column of Panel C in Table V. This shows results from regression of default on borrowers original LTV. Our results suggest that a 10 point rise in origination LTV is associated with an additional 0.22 percentage point increase in the yearly mortgage default rate (relative to a mean of 1.43%). This relation suggests a strong leverage-default association among a broader sample of fixed-rate loans.

To understand the role of payment shocks in this correlation, we first must understand the correlation between leverage and the interest rate. In column 2, we regress borrowers interest rate on the original LTV. (here, we abstract from the additional payment generated by the excess loan balance itself). This specification suggests that a 10-point rise in LTV results in higher interest rates of about 3.33 basis points to a borrower.

With these estimates in hand, we turn to decomposing the correlation into three channels. As a starting point, we take the estimated 0.22 percentage point increase in the default rate associated with a 10-point rise in LTV. First, to translate the effect of higher payments induced by these interest rates on default rates, we use the estimates in column 6 of Panel A of V. These results suggest this rise in interest rates can be expected to induce an additional 0.10 percentage point increase in defaults (the 0.033 increase in interest rate × the 3.3percentage point coefficient on the role of interest rates on default).

To isolate the channels which drive this association, we first isolate the role of payment shocks. To do so, we regress in column (2) of Panel C in this table origination LTV against interest rate. This specification suggests that a 10-point rise in LTV results in higher interest rates of about 3.33 basis points to a borrower. To translate the effect of higher payments induced by these interest rates on default rates, we use Panel A of this table estimated on adjustable-rate loans. Our results in column (6) of that table would suggest that this rise in interest rates can be expected to induce an additional 0.0010 increase in defaults (the 0.033 increase in the interest rate × the 3.3 percentage point relationship between interest rates on default).

The mechanical payment component, therefore, can account for just under 50% $\approx 0.0010/0.0022$ of the increase in defaults due to greater origination leverage. The remainder of this increase is not mechanically tied to the payment, and instead arises from the roles of adverse selection and moral hazard. Of the non-payment shock component, we use our estimates in Table III, as discussed in section 5.3. Those estimates, based on Option ARM borrowers for which the payment component could be ignored, suggest that the residual leverage-default relationship can be decomposed into 36% selection (0.0399 percentage points) and 64% moral hazard (0.071 percentage points).

Taken together, our results suggest that the original 0.22 percentage point increase in default can be decomposed in to a component which includes the role of greater monthly payments (0.10 percentage points, or 49% of the original leverage-default association); a component accounting for moral hazard (0.071 percentage points, or 32% of the default link), and another dimension related to adverse selection (0.0399 percentage points, or 18%).

To provide some additional heterogeneity in analysis of these results, we also examine two additional subsets of our fixed-rate data. First, we analyze fixed-rate loans originated in years 2004–2007, to better match the years for which use used adjustable-rate loan data. On this sample, we estimate that the leverage-default association is 0.0027—of which we estimate that monthly payments can account for 54% of the association, adverse selection 16%, and moral hazard 29%.[40]

Of course, these estimates are subject to significant caveats, both based on extrapolating from different samples, and the generality of our decomposition exercise. However, the exercise allows for at least a suggestive look into the relative importance of adverse selection, moral hazard, and

[40]We also estimate in columns (5)-(6) the relationship between leverage and default for loans originated from 2004–2007, restricted to the first five years after origination, to also better match our main ARM samples. We estimate a decomposition of 47% (payments), 19% (adverse selection), and 33% (moral hazard) on this sample.

payments.

7 Simulations and Welfare Analysis

In this section, we highlight the reasons why an appropriate estimation of moral hazard and adverse selection has implications for public policy. After briefly discussing the relevance of our reduced form estimates, we turn to analyzing a particular policy—an LTV cap—in a simulation framework. To address the challenges of evaluating counterfactual policies in competitive markets with adverse selection, we use the equilibrium concept proposed by use the equilibrium concept proposed by Azevedo and Gottlieb (2017).

We view an LTV cap as a fairly representative example of the types of macroprudential regulations that have been implemented in various countries since the recession. In addition to concerns about asset price booms and busts, these policies are commonly motivated by moral hazard concerns—in hopes that limiting consumer leverage directly limits the scope for ex-post default. In our analysis, we quantify the value of these policies in lowering default rates, and identify precisely the appropriate default externalities that policymakers need to envision in order to justify such policies. The key innovation in our analysis, however, is the introduction of adverse selection. We 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.

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 III 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 our 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. We begin with the model:

7.2.1 Consumer Preferences

We consider a two period model. In the first, consumers have a fixed home to purchase and choose a mortgage. In the second, a stochastic ex-post home value is realized. Borrowers then either default or repay the balance on the loan. Borrowers face a menu of contracts that each have two elements: a loan size/leverage L_k and a balance $B(L_k)$.

Given a contract $\{L_k, B(L_k)\}$ and a distribution of home prices, we characterize the observed portion of a borrower's ex-ante based on the model in Appendix 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.

We 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. We 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 (2017) 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 (2017).

For the purposes of simulation, utilizing this equilibrium concept is straightforward. We calculate equilibrium in what Azevedo and Gottlieb (2017) call a perturbation. We propose a fixed set of contracts (in the example presented, every integer LTV between 50 and 100). We 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. We 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

We calibrate three features of the simulation to the estimates from Table IV. We define the mean and variance of the private costs of default based on those estimated in Column 4 of IV. Furthermore, we 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 VI. For the purposes of the simulation, we assume that borrowers have exponential utility, with CARA coefficient a.

7.3 Welfare Implications of an LTV Cap

We 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. We 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

We 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, we 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 policymaker 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 V. This figure shows results with an exaggerated degree of adverse selection, to better present the patterns across the three regimes, while Table VI 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 V.[41] 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 VI 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.

[41]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.

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, leverage 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 V 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 VI, 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, we empirically separate moral hazard from adverse selection in the mortgage market. We 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 lies in separating adverse selection from moral hazard. We 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 borrower's initial leverage choices and the balances they owe ex-post. To isolate plausibly exogenous variation in balances, we 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 borrower's balances allows us 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. We 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 borrower's loan-to-values at 24 months would have reduced defaults by over 8 percent.

We then translate what these estimates of moral hazard and adverse selection imply for macroprudential policy. 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, we use the robust equilibrium concept recently developed by Azevedo and Gottlieb (2017).

We evaluate the impact of a reduction in an LTV cap, a common policy aimed at reducing defaults by limiting borrower's initial leverage. We 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 borrower's 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. Relative prices adjust for all borrowers, even those far from the binding leverage cap. We 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.

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. 2017. "Perfect Competition in Markets with Adverse Selection" *Econometrica*, 85(1): 67–105.

Bailey, Michael, Eduardo Dávila, Johannes Stroebel, and Theresa Kuchler. 2017. "House Price Beliefs and Leverage Choice" *Working Paper*.

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

Behn, Markus, Rainer Haselmann, and Vikrant Vig. 2017. "The Limits of Model-Based Regulation" *SAFE Working Paper No. 75*.

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, João Cocco, and Nuno Clara. 2017. "Structuring Mortgages for Macroeconomic Stability" *Working Paper*.

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.

Cerutti, Eugenio, Stijn Claessens, and Luc Laeven. 2017. "The use and effectiveness of macroprudential policies: New evidence" *Journal of Financial Stability*, 28(C): 203–224.

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.

DeFusco, Anthony, Stephanie Johnson, and John Mondragon. 2017. "Regulating Household Leverage" *Working paper*.

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, Benjaim Keys, Tomsz Piskorski, Rodney Ramcharan, Amit Seru, and Vincent Yao. 2017. "Interest Rate Pass-Through: Mortgage Rates, Household Consumption, and Voluntary Deleveraging" *American Economic Review*, 107(11): 3550–88.

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.

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

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. 2017. "Payment Size, Negative Equity, and Mortgage Default" Staff Report, Federal Reserve Bank of New York 4.

Ganong, Peter, and Pascal Noel. 2017. "The Effect of Debt on Default and Consumption: Evidence from Housing Policy in the Great Recession" *Working Paper*.

Geanakoplos, John, and Lasse Heje Pedersen. 2012. "Monitoring leverage" In *Risk Topography: Systemic Risk and Macro Modeling*. 113–127. University of Chicago Press.

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. 2016. "Two extensive margins of credit and loan-to-value policies" *Journal of Money, Credit, and Banking*, 48(7): 1397–1438.

Greenwald, Dan. 2018. "The Mortgage Credit Channel of Macroeconomic Transmission" *Working Paper*.

Gupta, Arpit. 2017. "Foreclosure Contagion and the Neighborhood Spillover Effects of Mortgage Defaults" *Journal of Finance*, Forthcoming.

Guren, Adam M., Arvind Krishnamurthy, and Timothy J. McQuade. 2018. "Mortgage Design in an Equilibrium Model of the Housing Market" *NBER Working Paper No. 24446*.

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

Hertzberg, Andrew, Andres Liberman, and Daniel Paravisini. 2018. "Screening on Loan Terms: Evidence from Maturity Choice in Consumer Credit" *The Review of Financial Studies, forthcoming*.

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.

Korinek, Anton, and Alp Simsek. 2016. "Liquidity trap and excessive leverage" *American Economic Review*, 106(3): 699–738.

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. 2015. "Asset Quality Misrepresentation by Financial Intermediaries: Evidence from RMBS Market" 6.

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.

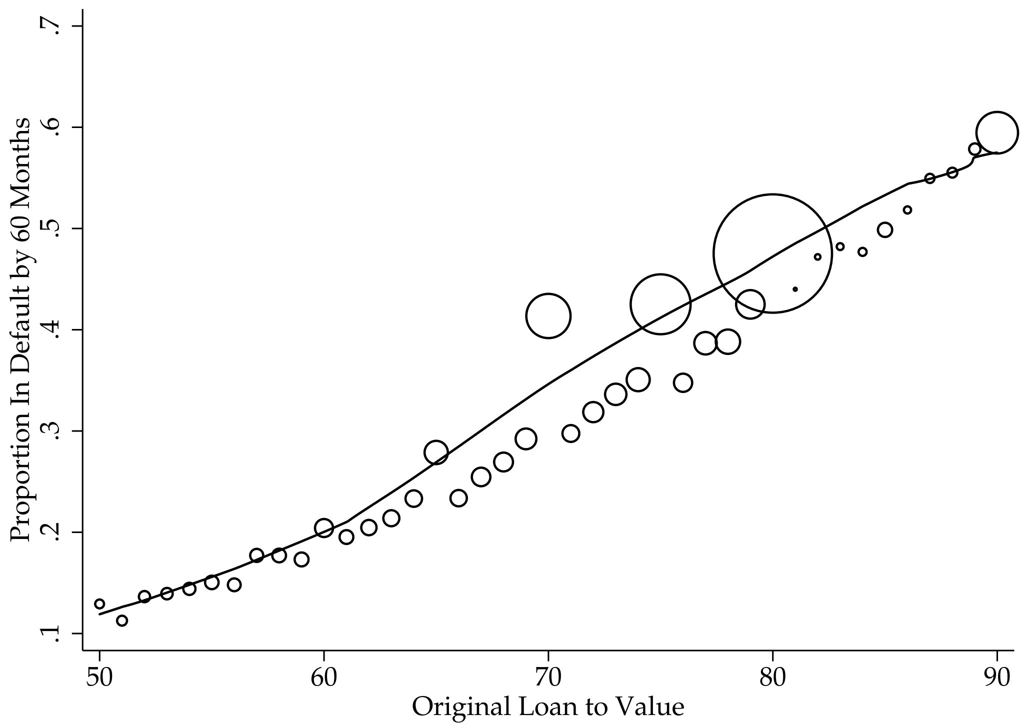
Scharlemann, Therese C., and Stephen H. Shore. 2016. "The Effect of Negative Equity on Mortgage Default: Evidence From HAMP's Principal Reduction Alternative" *Review of Financial Studies*.

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.

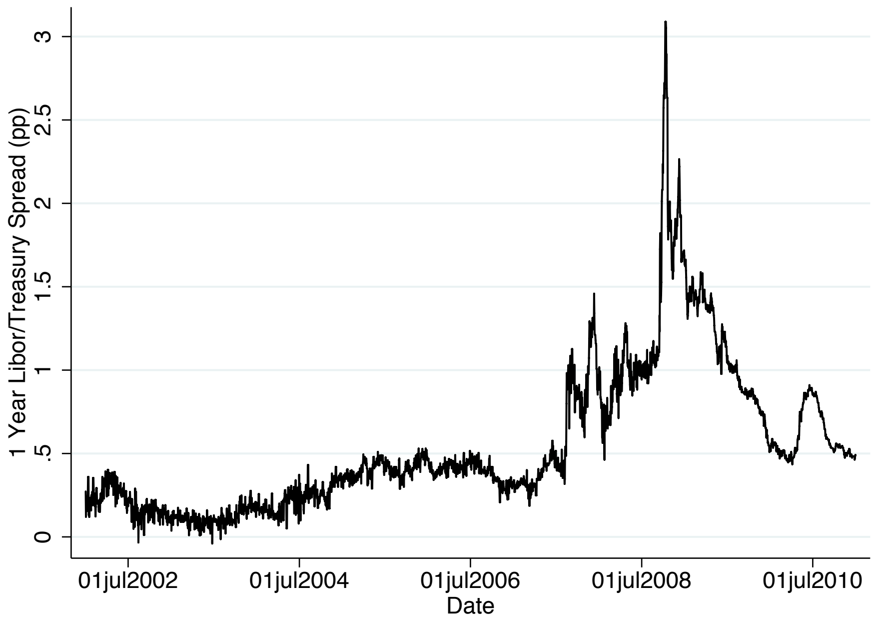
FIGURE I
ORIGINAL LTV IS POSITIVELY CORRELATED WITH DEFAULT WITHIN 60 MONTHS



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. The full sample of Option ARMs is included.

FIGURE II

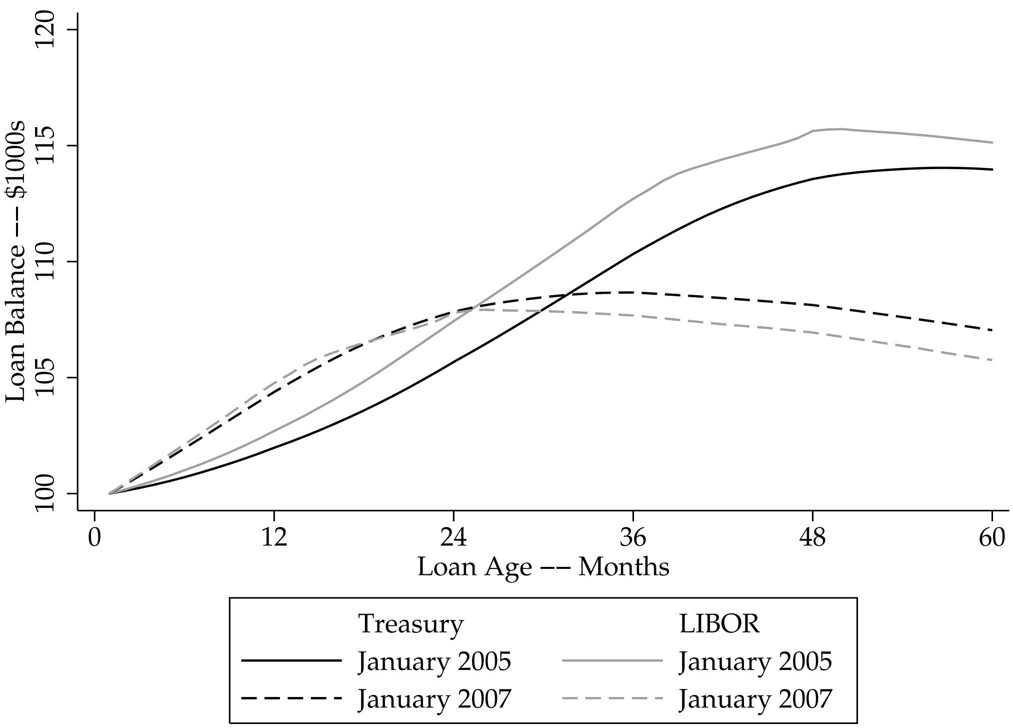
PANEL A: SPREAD BETWEEN LIBOR AND TREASURY INCREASED DRAMATICALLY DURING CRISIS



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 5/1 ARM loans resetting in that month.

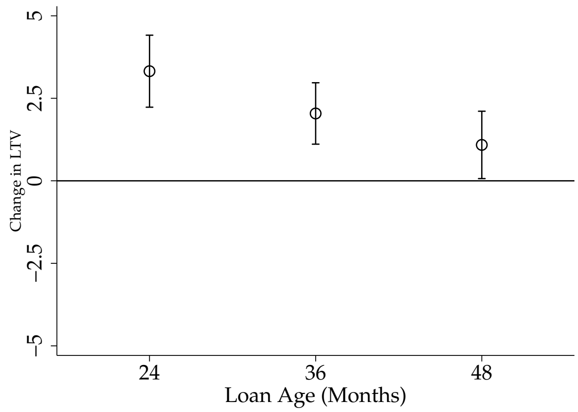
FIGURE III
INDEX × ORIGINATION MONTH GENERATES DIFFERENCE-IN-DIFFERENCE VARIATION IN BALANCE TRAJECTORIES FOR OPTION ARMS



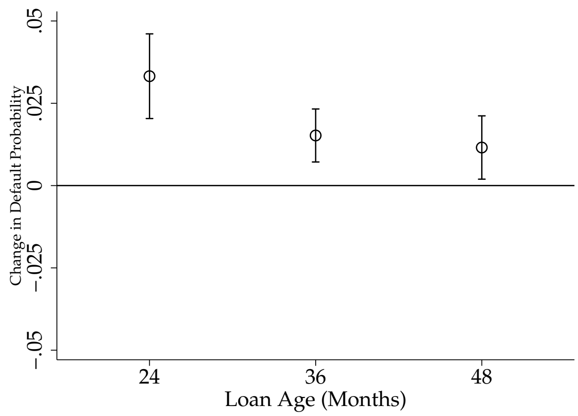
Simulated balance trajectories for $100,000 LIBOR- and Treasury-indexed loans originated in January 2005 or January 2007. Trajectories assume margin of 3.5 percent and initial payment based on 1.75 percent teaser. Treasury refers to 1-year MTA, LIBOR refers to 3-month duration.

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

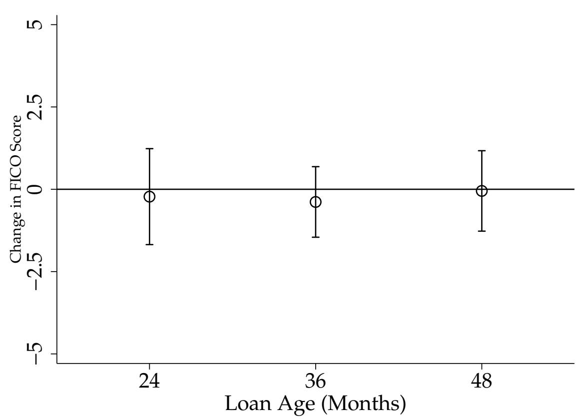
PANEL A: INSTRUMENT RELEVANCE—INSTRUMENT PREDICTS CURRENT LOAN-TO-VALUE



PANEL B: REDUCED FORM—INSTRUMENT PREDICTS DEFAULT ACROSS LIFECYCLE

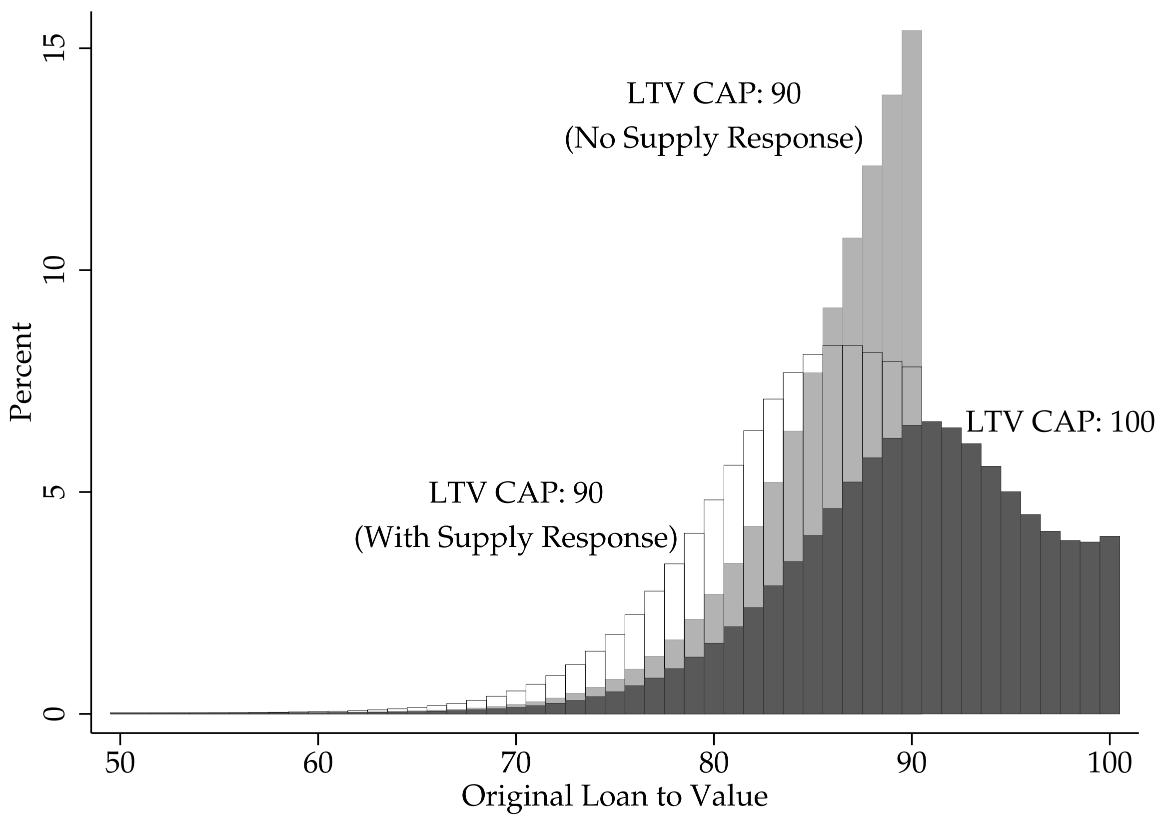


PANEL C: INSTRUMENT EXOGENEITY—INSTRUMENT DOES NOT PREDICT CREDIT SCORE



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 V
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 VI shows appropriately calibrated results.

Tables

TABLE I
SUMMARY STATISTICS: BALANCE ACROSS INDICES

	Panel A: Option ARM Sample			
	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	0.99		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	490,132		45,199	

	Panel B: 5/1 ARM Sample			
	Treasury		LIBOR	
	Mean	SD	Mean	SD
FICO Score	728.7	50.8	718.1	52.3
Origination Balance	483.6	302.7	439.0	313.6
Loan for Purchase	0.24		0.20	
No/Low Documentation	0.52		0.70	
Primary Residence	0.87		0.84	
Prepayment Penalty	1		1.00	
Margin	2.77	0.31	2.36	0.39
Original LTV	74.3	9.02	76.2	8.37
State:				
- California	0.48		0.40	
- Florida	0.063		0.093	
- Arizona	0.032		0.041	
- Nevada	0.021		0.029	
Origination Year:				
- 2002	0.029		0.0072	
- 2003	0.16		0.058	
- 2004	0.25		0.16	
- 2005	0.37		0.25	
- 2006	0.12		0.34	
- 2007	0.051		0.17	
Observations	109,214		344,594	

Summary statistics for full sample of Option ARMs (Panel A) and 5/1 ARMs (Panel B). Treasury refers to loans indexed to Treasury rates, LIBOR refers to those indexed to LIBOR.

TABLE II
FIRST STAGE: INSTRUMENTS PREDICT REALIZED NEGATIVE EQUITY, LOAN-TO-VALUE, INTEREST RATES, AND PAYMENTS

	Panel A: Option ARMs—Observed/Imputed Equity on Simulated Instruments							
	Conditional on Survival to 24 Months				Full Sample (Imputed Current Equity)			
	Home Equity		Loan-to-Value		Home Equity		Loan-to-Value	
Simulated Home Equity	1.345***	0.795***			1.349***	0.339**		
($100,000s)	(0.214)	(0.197)			(0.226)	(0.164)		
Simulated Loan-To-Value			3.559***	1.736***			2.632***	0.545
			(0.621)	(0.471)			(0.463)	(0.363)
F (Simulated Instrument)	39.5	16.3	32.8	13.6	35.7	4.2	32.3	2.3
F (Fixed Effects)	10.3	7.1	5.0	11.3	11.3	6.8	8.1	8.5
N	265134	265134	268364	268364	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 borrower's index and month of origin. Displayed in each column is the coefficient from regressing borrower's 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. We also include indicators for each 20-point bin of borrower's FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. We allow individual state time trends. Standard errors are clustered at the MSA level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE III
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. We also include indicators for each 20-point bin of borrower's FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. We allow individual state time trends. Standard errors are clustered at the MSA level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE IV
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. We also include indicators for each 20-point bin of borrower's FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. We allow individual state time trends. Standard errors are clustered at the MSA level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE V
IDENTIFYING THE IMPACT OF PAYMENTS: INTEREST RATE SHOCKS AND DEFAULT

	Panel A: Causal Effect of Reset Interest Rates on One Year Default Rate					
	OLS	IV	IV	OLS	IV	IV
Reset Rate	0.014***	0.037***	0.037***	0.014***	0.033***	0.033***
	(0.003)	(0.007)	(0.007)	(0.003)	(0.006)	(0.005)
Original Loan-to-Value			0.180***			0.208***
			(0.014)			(0.017)
Mean of Dep. Var.	0.068	0.068	0.068	0.068	0.068	0.068
N	107438	107438	107438	107438	107420	107420
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes
Servicer	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	No	No	No	Yes	Yes	Yes
Full Controls	No	No	No	Yes	Yes	Yes
Current CLTV	No	No	No	No	No	No
	Panel B: Selection—Relative Interest Rate Sensitivity by Original LTV					
	OLS			IV		
	Orig. LTV< 80	Orig. LTV= 80	Orig. LTV> 80	Orig. LTV< 80	Orig. LTV= 80	Orig. LTV> 80
Reset Rate	0.003	0.022***	0.006	0.011*	0.046***	0.096**
	(0.003)	(0.006)	(0.009)	(0.006)	(0.009)	(0.048)
Mean of Dep. Var.	0.050	0.083	0.12	0.050	0.083	0.12
N	45892	44819	4476	45867	44799	4440
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index FEs	Yes	Yes	Yes	Yes	Yes	Yes
Servicer	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	Yes	Yes	Yes	Yes	Yes	Yes
Full Controls	Yes	Yes	Yes	Yes	Yes	Yes
Current CLTV	Yes	Yes	Yes	Yes	Yes	Yes
	Panel C: Decomposition—Effects on Fixed-Rate Sample					
	Full Sample		Origination Years 2004–2007		Origination Years 2004–2007, Observed Five Years	
	Delinquency	Interest Rate	Delinquency	Interest Rate	Delinquency	Interest Rate
Original Loan-to-Value	0.022***	0.333***	0.027***	0.446***	0.031***	0.447***
	(0.001)	(0.008)	(0.001)	(0.011)	(0.002)	(0.010)
Mean of Dep. Var.	0.0143	5.66	0.027	6.10	0.029	6.12
N	80704510	80704490	18208510	18208510	13369247	13369247
Origination Month FEs	Yes	Yes	Yes	Yes	Yes	Yes

Panels A and B display results from OLS and IV regressions with the binary outcome of default between 60 and 72 months for borrowers with 5/1 ARMs. Default is defined as 60 or more days past due. Panel A includes regressions of default on monthly payments or reset rates, as well as contract terms at origination (the original interest rate and margin). IV specifications instrument for monthly payments or reset rates with a difference-in-difference interaction of origination month and index type. Panels B OLS and IV regression of interest rate on default separately for borrowers with origination LTV below 80, at 80, and above 80. Full controls refers to fixed effects for index type, documentation, the loans purpose and occupancy, the existence of prepayment penalties and private mortgage insurance. We also include indicators for each 20-point bin of borrower's FICO credit scores, loan originator and servicer fixed effects, and controls for 5-point bins of current CLTV. Standard errors are clustered at the MSA level. Panel C displays results from the fixed-rate sample of loans securitized by Fannie and Freddie. Column (1) regresses origination LTV against yearly default rates, while column (2) regresses original LTV against origination interest rate. Columns (3)-(4) subset on loans originated from 2004–2007, to better match the ARM sample, while columns (5)-(6) restrict on origination years as well as subset on the first five years after origination. FRM specifications control in addition only for origination month, and standard errors are clustered at the 3-digit ZIP level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE VI
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.

Appendicies

A Appendix Tables and Figures

Figure A.I
Interest Rate Schedule is Increasing in Original Loan-to-Value

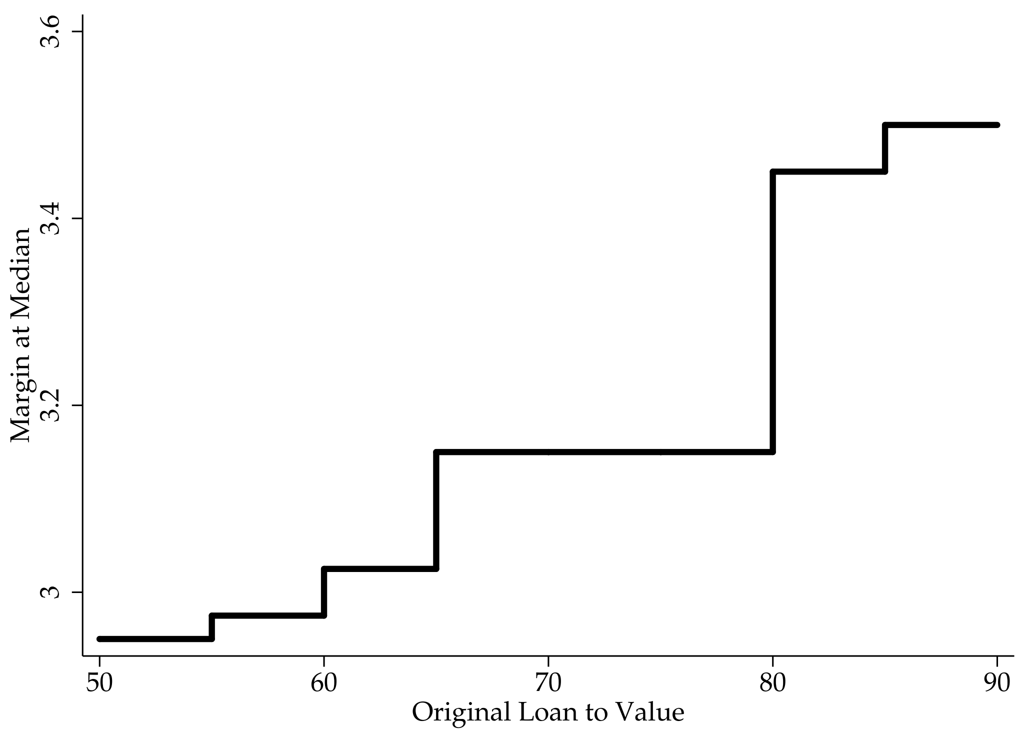
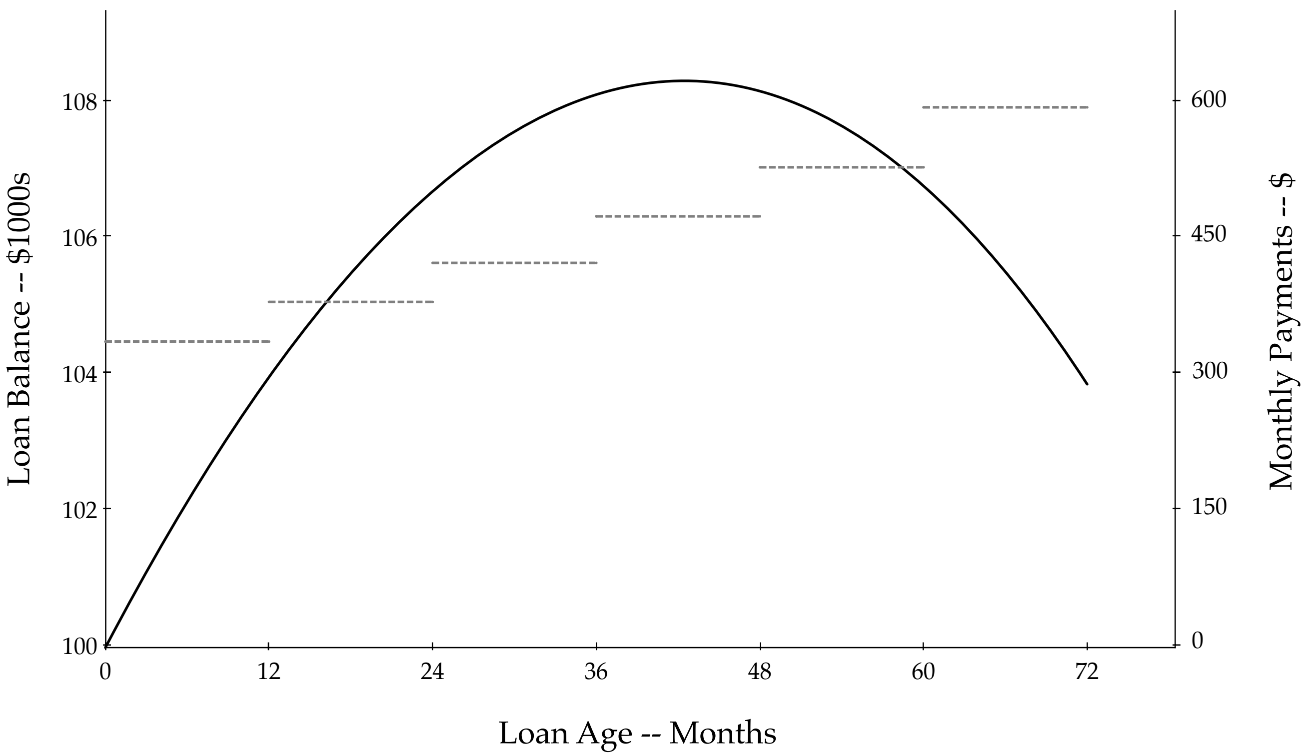


Figure displays the median margin, the fixed portion of each borrower's interest rate, for each 5 point loan-to-value bin. Full sample of Option ARMs is included.

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

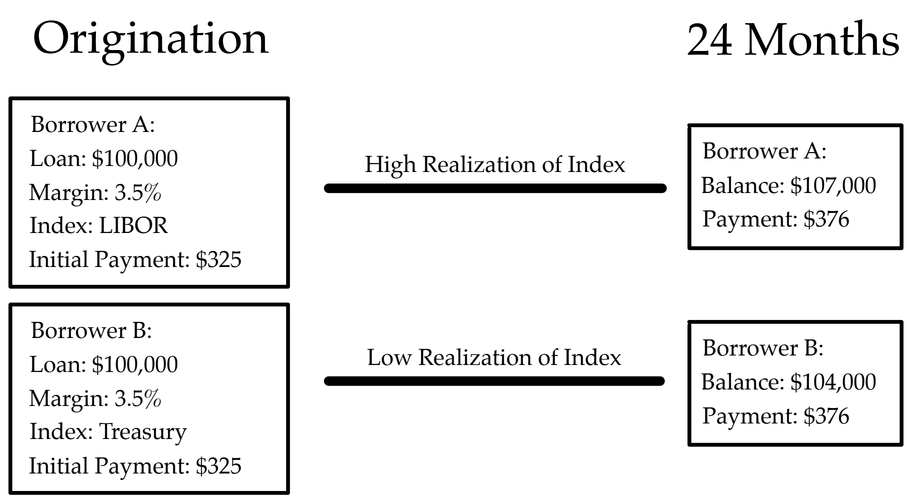
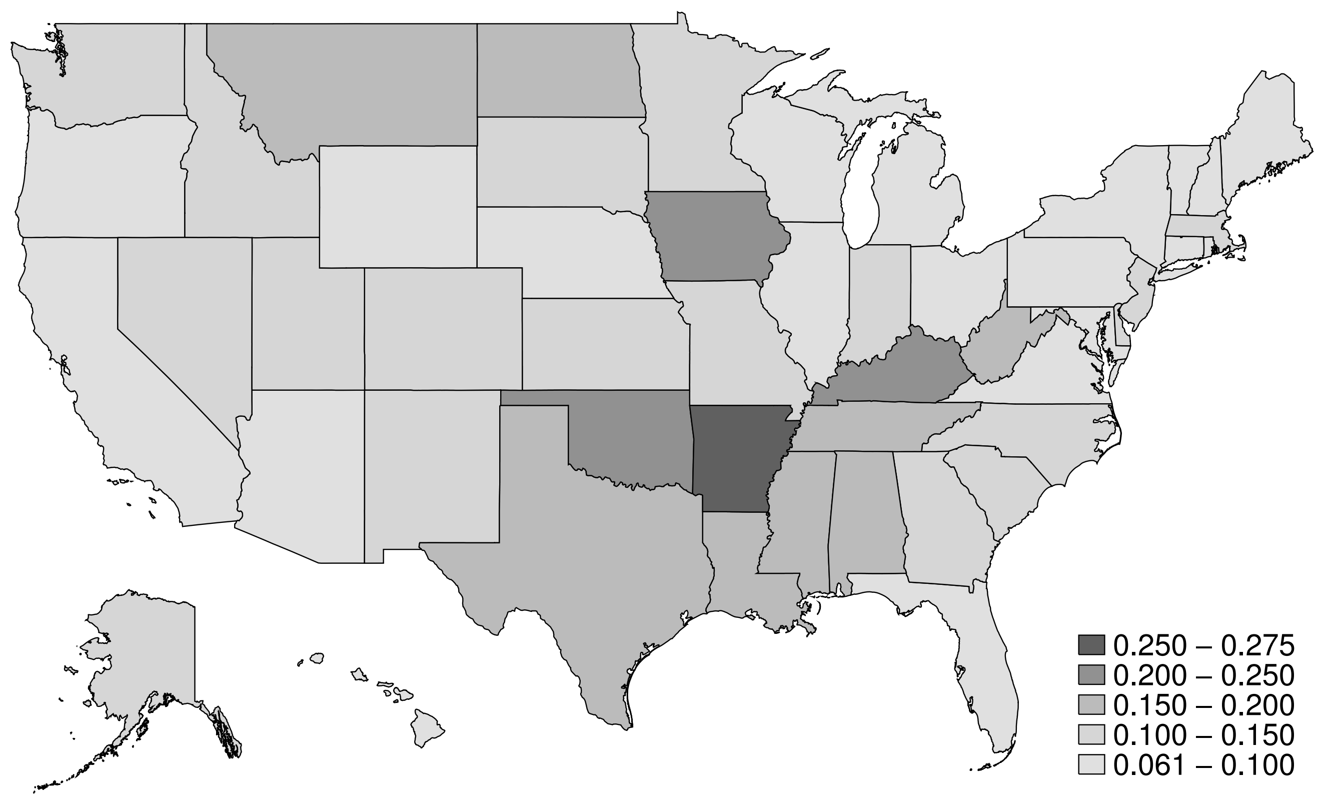
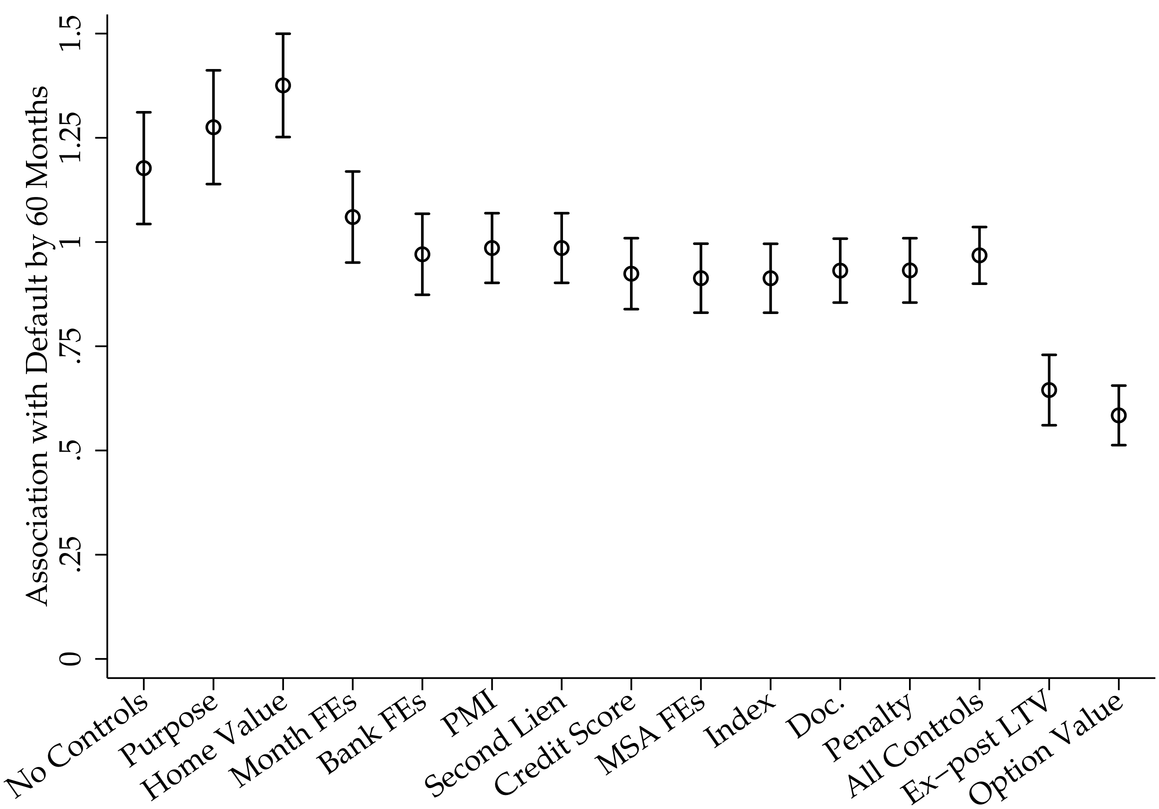


FIGURE A.IV
UNIFORM DENSITY OF LIBOR INDEXED OPTION ARMS ACROSS STATES



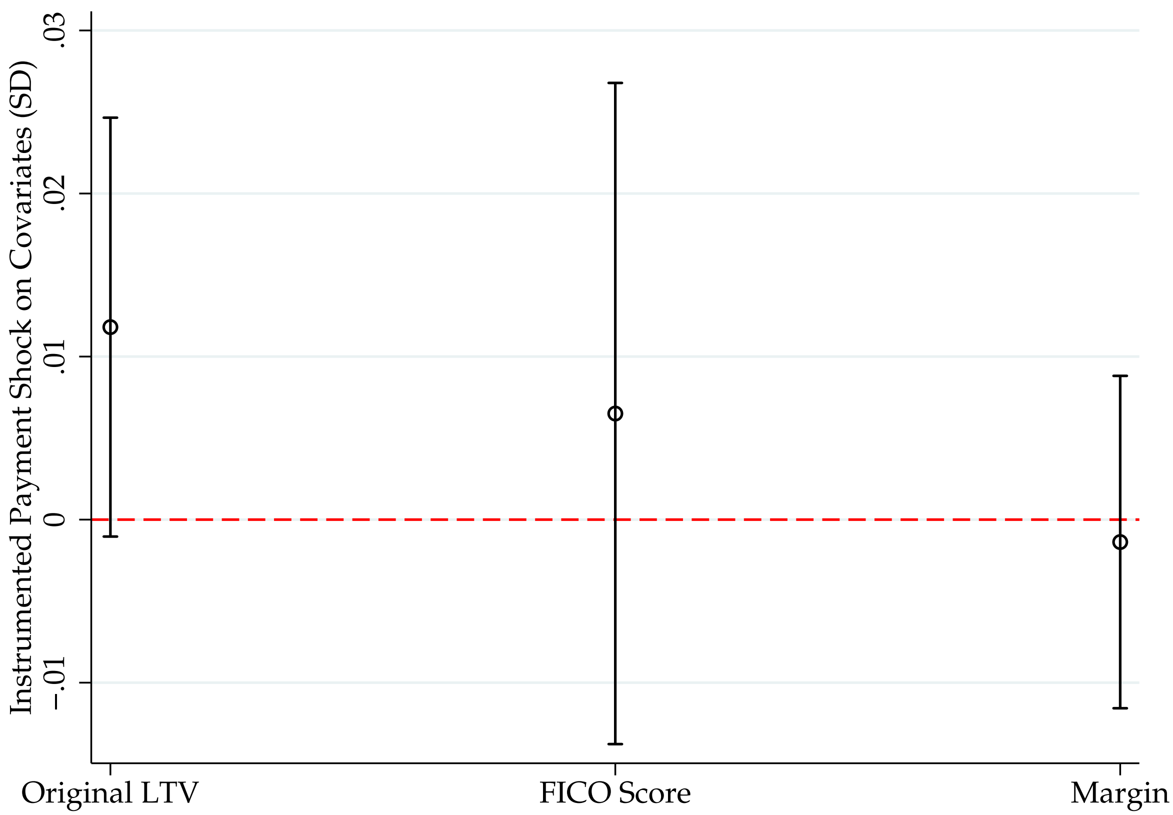
Plot shows number of LIBOR indexed Option ARMs as a proportion of all LIBOR- or Treasury-indexed Option ARMs. The minimum is 6.1 percent, while the max is 27.5 percent. In the majority of states, between 5 and 15 percent of Option ARMs are indexed to LIBOR.

FIGURE A.V
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 A.VI
EXOGENEITY OF INSTRUMENT ON 5/1 ARMS



Results from IV regressions of interest rates on loan covariates. Month of origination-by-index fixed effects are used to instrument for loan-level interest rates after initial reset of 5/1 ARMs. The standard set of covariates are included, but in each specification one coefficient is left out. The first line illustrates the effect of a one percent increase in instrumented interest rates on the standard deviation of original LTV; the second shows the instrumented effect on a standard deviation of origination FICO score; the third show the predicted instrumented effect on the standard deviation of initial margin.

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
HETEROGENEITY IN THE IMPACT OF ORIGINAL AND CURRENT LEVERAGE ON 1 YEAR DEFAULT PROBABILITY (OPTION ARM)

	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. We also include indicators for each 20-point bin of borrowers' FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. We allow individual state time trends. Standard errors are clustered at the MSA level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

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. We also include indicators for each 20-point bin of borrowers FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. We 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. We also include indicators for each 20-point bin of borrowers FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. We 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. We also include indicators for each 20-point bin of borrowers FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. We 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 III, 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. We also include indicators for each 20-point bin of borrowers FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. We 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. We also include indicators for each 20-point bin of borrowers FICO credit scores, loan originator and servicer fixed effects, and controls for second liens. We allow individual state time trends. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE A.VIII
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.

B Model

We propose a two-period model of borrower's leverage demand and default choice, following Brueckner (2000). Borrowers differ in a single dimension, which we 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 from the model. 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 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. We 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.[42] 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

$$H_1 - B < -C_i.$$

[42] 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 $\left(\frac{B}{L} = 1 + r\right)$, and the original loan-to-value $\frac{L}{H_0}$.

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.

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

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

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 result is 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.[43] 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.[44]

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

[44] 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.

At actuarily fair prices, borrowers prefer to take advantage of the insurance provided by a mortgage. In a totally frictionless context,[45] 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-Mirrlees single crossing condition holds:

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

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.

[45]By totally frictionless, we 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.

C Robustness

In this appendix, we describe a series of robustness checks intended to supplement the primary analysis.

Table A.II considers how the results in Table III 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, we 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.

C.0.3 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.

C.0.4 Documentation

Dividing borrowers by documentation provided, shown in Panel B of Table A.II, 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 we do not observe income, we are also not perfectly able to control for the information set of the lender in the full documentation sample. As a result, the

[46]By state recourse status, we 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. We 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.

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.

C.0.5 Purchases vs. Refinances

The differences between those purchasing homes versus those refinancing existing mortgages, shown in Panel C of Table A.II, 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.

C.0.6 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 III, 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 III. Results at other cross-sections are similar.

C.0.7 Cumulative Default Probabilities

The regressions in Table III 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, we 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 III, this time utilizing the outcome of cumulative default by 36 months. These are shown on the left-hand side of Appendix Table A.IV. We 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 III. 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, we 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.

C.0.8 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.

C.0.9 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, we fully interact E_{ijt} with all covariates. However, because we do not have sufficient instruments to do so in an IV specification, in column 6 we 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).