SELECTION, LEVERAGE, AND DEFAULT IN THE MORTGAGE MARKET*

Arpit Gupta† **Christopher Hansman**‡

December 2, 2019

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 indexes that determine interest rate adjustments. Moral hazard is responsible for 40% of the correlation, while adverse selection explains 60%. We construct and calibrate a simple model to show that optimal regulation of leverage must weigh default-prevention against market distortions 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, Joshua Coven, Eduardo Dávila, François Gerard, Jonas Hjort, Wojciech Kopczuk, Henrik Kleven, Rocco Macchiavello, Dan O'Flaherty, Tomasz Piskorski, Tarun Ramadorai, Miikka Rokkanen, Christoph Rothe, Bernard Salanié, Johannes Spinnewijn, Miguel Urquiola, Stijn Van Nieuwerburgh, Eric Verhoogen, Albert Zevelev, and participants of many seminars. We are particularly grateful to João Cocco and Felipe Severino for valuable discussions. 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)

Imperial College Business School

Imperial means
Intelligent Business

December 2nd 2019

This disclosure statement pertains to the submission of my paper "**Selection, Leverage, and Default in the Mortgage Market**," joint work with Arpit Gupta (NYU Stern).

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. Neither the NSF nor the Russell Sage Foundation had any input into our analysis or conclusions.

I declare that I have no conflicts of interest or commitment and have no financial or non-financial interests to disclose in relationship to this project.

Sincerely,

Christopher Hansman
Assistant Professor
Imperial College Business School

Dr Christopher Hansman
Assistant Professor of Financial Economics

Imperial College Business School
Imperial College London
South Kensington Campus
London SW7 2AZ, United Kingdom

T: +44 (0)20 7594 1044
E: c.hansman@imperial.ac.uk
W: imperial.ac.uk/people/c.hansman

NYU STERN
Center for
Real Estate
Finance Research

Arpit Gupta
Assistant Professor

Leonard N. Stern School of Business
Kaufman Management Center
44 West 4th Street, Suite 9-53
New York, NY 10012
P: 212 998 0305

arpit.gupta@stern.nyu.edu

December 2, 2019

This disclosure statement pertains to the submission of my paper "**Selection, Leverage, and Default in the Mortgage Market**," joint work with Christopher Hansman (Imperial College Business School).

I declare that I have no conflicts of interest or commitment and have no financial or non-financial interests to disclose in relationship to this project.

Sincerely,

Arpit Gupta (NYU Stern)
arpit.gupta@stern.nyu.edu

1 Introduction

In the years since the crisis, mortgage leverage has come into focus as a critical element of financial stability and a key target for macroprudential policy.[1] At the core of this concern is the correlation between leverage and default: when housing prices fall, highly leveraged homeowners are the least likely to meet their obligations. While intuitive, this correlation has two potential explanations. The first, sometimes called moral hazard, is a causal effect: leverage itself increases the probability of default.[2] The second is adverse selection: ex-ante risky borrowers simply prefer high-leverage mortgages. Despite a substantial theoretical literature examining these two classical information asymmetries in credit markets, differentiating between them remains a major challenge for empirical work.[3]

In this paper, we separate adverse selection from moral hazard in the mortgage market. To do so, we rely on a natural experiment generated by the unique contract structure of Option Adjustable-Rate Mortgages (Option ARMs). These loans have interest rate adjustments tied to pre-specified indexes, most commonly a LIBOR or Treasury rate. While the choice between LIBOR and Treasury was not salient prior to the financial crisis, the unexpected divergence of the two in 2008 caused borrowers who chose otherwise identical contracts to owe substantially different amounts ex-post. This variation allows us to (i) isolate moral hazard by comparing borrowers with the same initial leverage choices but different realized home equity, and (ii) identify adverse selection by comparing borrowers with different initial leverage choices but similar realized home equity. We find that adverse selection is responsible for 60 percent of the baseline correlation between leverage and default, while the causal effect—moral hazard—explains the remaining 40 percent.

The presence of adverse selection has fundamental consequences for the quantity and price of leverage in equilibrium, and resultantly for macroprudential policy. The theoretical literature following Bester (1985) suggests that—in the presence of adverse selection—leverage must act as a screening device to differentiate riskier types. As a result, even minor changes in policy or market fundamentals may significantly reshape the set of contracts lenders offer, with implications that differ sharply from a world without selection. For instance, our analysis shows that a regulator who ignores adverse selection will *overestimate* the reduction in defaults generated by macroprudential restrictions on household leverage and *underestimate* the corresponding costs to borrowers, who face higher interest rates and choose smaller loans in equilibrium.

Similarly, the causal relationship is of direct interest to policymakers, practitioners, and academics. Clear estimates of this elasticity are necessary for effective design of borrower relief policies, accurate pricing of mortgage bonds, debates over the root drivers of consumer default, and more. While the magnitude

[1] See, for example, Geanakoplos and Pedersen (2012); Mian and Sufi (2015); Korinek and Simsek (2016).

[2] In credit markets, this effect is referred to as moral hazard by, e.g., Adams, Einav and Levin (2009).

[3] The general empirical challenge is discussed in Chiappori and Salanié (2013).

of our estimated effect is sizable overall, we find evidence of substantial non-linearities. Crossing into negative equity sharply increases a borrower's probability of default, but there are limited marginal effects for borrowers who are deeply underwater. This is consistent with double-trigger theories of mortgage default and recent studies of default behavior among underwater borrowers (Ganong and Noel, 2018).

The paper begins with a simple conceptual framework to clarify the sources of adverse selection and moral hazard in mortgage markets. Moral hazard is a feature of standard theoretical models (e.g. Campbell and Cocco, 2015) because lenders face limited access to effective recourse and/or underwater households may be unable to avoid default by selling or refinancing. 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, while others do not default even when substantially underwater. 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) are more willing to accept the inflated interest rates that come with high-leverage loans.[4]

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 borrower's *realized* leverage—regardless of the initial choice.[5] As a result, it is possible to separately identify the two if there is exogenous variation in borrowers' leverage that is distinct from their initial mortgage choice.

In our context, the divergence of mortgage indexes generates 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*. Because monthly payments do not adjust, fluctuations in interest rates generate meaningful variation in the balances borrowers will ultimately owe without impacting the amount they must pay in the short term.[6] As a result, otherwise identical borrowers who receive different paths of interest rates because of the loan's (i) origination month and (ii) particular mortgage index (e.g. Libor vs. Treasury) will face different realized leverage. The fixed and unusually small minimum payments for Option ARMs facilitate our analysis by allowing us to set aside the role of monthly payments in the default decision.

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

[5]A natural way to understand these implications is through a set of ideal experiments. To identify adverse selection, we would ideally reassign borrowers who have endogenously chosen different leverage to owe identical balances. A remaining correlation between the initial leverage choice and default would indicate selection. For moral hazard, the ideal experiment would randomly reassign borrowers who have chosen identical leverage to owe *different* balances. A correlation between these randomly assigned balances and default would then reflect a causal effect.

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

In practice, our empirical approach isolates the relevant variation in realized leverage—driven by the *interaction* between a borrowers mortgage index and origination month—using a leave-out-mean (jackknife) IV estimator. Specifically, we instrument for a borrower i's realized leverage with the average for all $j \neq i$ with the same index type and origination month.[7] We then include origination month fixed effects, which account for any aggregate time trends, and index type fixed effects, which account for any fixed differences between borrowers, to focus specifically on the interaction. We are additionally able to include a rich set of borrower-level controls and fixed effects that account for the information set of the lender and various other potential confounds. For example, in our preferred specifications, we additionally include originator$\times$origination month fixed effects to address time varying selection of borrowers across originators and unobserved changed in underwriting standards by different institutions. Throughout, the key assumption is that—conditional on these controls—our instrument captures exogenous variation in realized leverage that is distinct from the original leverage choice. If this is the case, our IV gives the causal effect directly, and the residual correlation between initial leverage and default allows us to capture for adverse selection.

As noted above, we find that adverse selection is responsible for 60 percent of the baseline correlation between leverage and default, while moral hazard is responsible for the remaining 40 percent. The latter effect suggests a meaningful causal relationship between leverage and default. Our estimates imply that a 10 point reduction in a borrower's loan-to-value ratio two years after origination reduces the average probability of default for compliers in our setting by over 4 percentage points. However, this average masks meaningful heterogeneity across the distribution of realized leverage. We estimate much stronger causal effects of leverage for borrowers very close to the threshold of negative equity, but weak or null effects for significantly underwater borrowers (consistent with Ganong and Noel, 2018) or for borrowers with positive equity. Overall, this heterogeneity supports the view in which negative equity is a necessary but not sufficient condition for mortgage default. In robustness exercises we consider various potential concerns including time varying geographic confounds, early attrition from the sample, and the difficulty of accounting for selection in the presence of non-linear causal effects.

As a final step, we develop a stylized structural model to highlight the consequences of our findings—particular adverse selection—for macroprudential regulation. We estimate a parametric control function version of our IV approach to provide inputs into the model, and analyze the introduction of a loan-to-value cap (a widely used policy tool intended to reduce mortgage defaults by restricting leverage). While the consideration of counterfactual policies in the presence of adverse selection has traditionally been chal-

[7] This is analogous to instrumenting with a full set of index type $\times$ origination month dummies.

lenging,[8] we are able to implement our analysis thanks to the robust equilibrium concept recently proposed by Azevedo and Gottlieb (2017).

We find that a loan-to-value cap 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 unexpected welfare losses due to *knock on effects.* While the mechanical impact of the the regulation only 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.

This paper makes three contributions. We are the first, to our knowledge, to empirically isolate the presence of adverse selection on leverage in mortgage markets as distinct from moral hazard. This contributes to the growing empirical literature on asymmetric information and credit.[9] The key innovation in our setting comes in isolating a form of ex-post variation in leverage 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 of 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.[10] However, these approaches require strong assumptions as to why the relevant variation in ex-ante contracts does not also generate selection on unobservables. Furthermore, we do so in the largest and arguably most important consumer debt market in the US.[11]

Secondly, our paper contributes to the literature on mortgage default by providing evidence for a strong impact of leverage. We do so by developing an instrument for borrowers' balances that does not simultaneously impact short term payments. A large literature in this area has used non-experimental variation in borrower equity, most commonly driven by house price variation. Vandell (1995) provides an overview of early research on the role of home equity in the choice to exercise the default option. More recent work,

[8] Even the appropriate characterization of equilibrium with adverse selection is controversial, and equilibria may fail to exist under conventional concepts, e.g. Rothschild-Stiglitz.

[9] 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).

[10] Other considerations of selection in mortgage markets include 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).

[11] 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.

including Bajari, Chu and Park (2008), Foote, Gerardi and Willen (2008), Elul et al. (2010), Bhutta, Shan and Dokko (2010), and Gerardi et al. (2017), has stressed the joint importance of triggers such as liquidity and job loss alongside home equity in mortgage default. Palmer (2015) decomposes variation in defaults across cohorts of subprime mortgages into a portion due to home equity and a portion due to relaxed lending standards. Our borrower-level variation in leverage avoids the potential for measurement error and geographically-based endogeneity concerns inherent to the use of home price variation. A small number of recent well-identified papers have found evidence for a limited or non-existant causal role of borrower leverage on default. For example, Ganong and Noel (2018) uses experimental evidence from the HAMP program based on a sample of loan applicants for loan modifications. Our findings are consistent with this work when considering borrowers in negative equity: we also find little evidence that changes in leverage impact default. However, a key contribution of our work is to incorporate borrower responses around the negative equity threshold, where we find stronger effects.

Our third and final contribution comes in connecting 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 (2019), and Behn, Haselmann and Vig (2017). We provide the first empirically grounded evidence highlighting the complications that adverse selection poses for regulations of this form. By highlighting the importance of considering selection, we also contribute to work on the role of mortgage leverage in the crisis and the macroeconomy more generally (e.g. Corbae and Quintin, 2015; Geanakoplos, 2010).

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. Section 4 presents the empirical strategy. Section 5 presents our main empirical results. Section 6 shows the results of simulations, and Section 7 concludes. Appendix A contains additional robustness tests and supplementary analyses, while Appendix B outlines a model formalizing the intuition in Section 2.

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 or likelihood of default—that provides a source of adverse selection.

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 size of a borrower's loan liability and default. That is, all else equal, those who face higher leverage 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 or she poses to the lender.[12]

By moral hazard, we simply mean the causal impact of leverage on borrower default. This 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 because 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 perhaps more inclusive than usual, our definition does not rule out a causal effect caused by a mechanical relationship between the loan balance and default.[13]

2.2 Sources of Information Asymmetries in Mortgage Markets

There are two natural sources of moral hazard in a mortgage context. The first—a feature of many models of default—is the limited liability that is implicit in a mortgage contract. Lenders cannot effectively contract against borrowers defaulting when their mortgages are underwater. For a borrower who is willing to default

[12] 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.

[13] For example, if there is some exogenous probability of complete insolvency, a borrower with positive equity may refinance or sell the home, while a borrower with negative equity may be forced to default. In this context, a shift in leverage that forces an individual across this threshold would increase the probability of default.

strategically, an increase in leverage raises the probability of owing much more than the home is worth, which in turn increases the probability of default. Of course, legal restrictions specifying the degree of limited liability 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.[14] However, even in states with laws that are favorable to lenders, deficiency judgments are relatively rare in practice (Pence, 2006). The second is more mechanical. Because a borrower with negative equity may lose the ability to refinance or sell his or her home to avoid default, there may be a causal relationship between leverage and default even in the absence of a strategic motive.

What is the source of heterogeneity across borrowers 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 a borrower strategically defaults immediately if the value of the home drops below the value of the mortgage. Unless borrowers have individual (and private) insights into the future path of real estate prices, this model leaves little room for adverse selection. Borrowers default according to a uniform rule.

However, a large literature suggests that borrowers do not follow a frictionless option model (see Vandell, 1995, for a review). There is significant heterogeneity in exercise of 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 substantially more on their mortgage than the home is worth. 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, reflecting, for instance, cross-sectional differences in the value of future access to credit or accessing local schools. Alternatively, even if the full population has a strong aversion to defaulting, variation in liquidity shocks may generate heterogeneity in realized default behavior. Adverse selection exists if this heterogeneity is difficult to observe or contract upon and correlates with a borrower's demand for leverage when selecting a mortgage. In other words, adverse selection exists if risky borrowers are willing to pay higher interest rates for high leverage.

This notion can be shown—in a stylized manner—with a simple default rule. Consider a borrower i who owes a loan balance B_i for a house with value H_i.[15] The borrower must choose between repaying the loan and receiving the value of the home or defaulting, which entails some cost C_i. This parameter should be though of as a reduced form representation of the net difference in unobserved costs associated with default

[14] 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 be relieved only in bankruptcy.

[15] H_i might alternatively be thought of as the option value of continuing to service the mortgage.

vs. repayment. For example, an individual concerned with moving costs or his or her future credit score might have a very high C_i, while an individual facing a liquidity shock might have a very low effective C_i (because the unobserved costs associated with making the monthly payment are high).[16] This gives rise to the following rule: borrowers default if

$$\underbrace{B_i - H_i}_{\text{Cost of Repaying}} > \underbrace{C_i}_{\text{Cost of Defaulting}}.$$

Regardless of the interpretation of C_i, the key feature for our purposes is simply that it may be different for different individuals i. Some borrowers may have very low C_i and default anytime $B_i > H_i$, whereas others may have relatively high C_i.[17] For a given distribution of housing values ex-ante—and a fixed balance—borrowers with lower expected C_i will be more likely to default. At least some portion of these heterogeneous costs C_i are likely to be unobservable to the lender.

Adverse selection arises when this unobserved heterogeneity in default risk is correlated with a borrower's demand for leverage when selecting a mortgage, that is, if riskier borrowers—e.g. those with unobservably low realized values of C_i on average—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.[18] 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.[19] The equilibrium implications of this sort of selection are familiar from the literature on collateralized lending following Bester (1985) and more general work in the tradition of Rothschild and Stiglitz (1976). In competitive contexts, leverage acts as a screening device. Because safe borrowers value leverage less, they choose sub-optimally low leverage (compared to the first best) in order to differentiate themselves from risky types and get a lower interest rate.

Of course, 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

[16] In an extreme case, a completely illiquid and credit constrained borrower who cannot refinance or sell their home might be modeled as having $C_i = -\infty$.

[17] In principle C_i may even be negative, causing default with $B_i < H_i$. For example, if the borrower has taken a "silent second" mortgage that the lender is unaware of.

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

[19] In other words, borrowers with low costs of default place a higher value on the 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.

example, a borrower who has limited access to cash for a down payment when selecting a mortgage may also find it disproportionately costly to raise the funds to repay the loan. For the purposes of our empirical exercises the existence of a correlation is what matters—the precise mechanism generating the correlation is not crucial.

3 Background and Data

In this section, we provide background on Option ARMs, highlighting the unique features 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 data we use in our analysis and show summary statistics.

3.1 Background on ARMs

Traditional adjustable rate mortgage contracts, typically refereed to as hybrid ARMs, feature fixed interest rates and fixed payments for a set initial period—usually 5 or 7 years. After the initial period interest rates begin to adjust according to market conditions, and usually change annually or semi-annually. These interest rates are calculated as the sum of a fixed component (the *margin*) and a variable component (the *index*). Monthly payments are designed to be fully amortizing, that is, calculated to exactly pay off the loan over the full term at current interest rates. As a result, payments change to keep pace with interest rates and may unexpectedly increase if interest rates rise.

According to lenders, the potential danger of these unexpected payment increases motivated the creation of the Option ARM.[20] Banks wanted a product that incorporated floating interest rates while protecting borrowers from sharp payment increases and mortgage holders from 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 small initial monthly 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. After 5 years, the minimum payment adjusts to the fully amortizing amount.[21]

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

[21] In theory, 7.5 percent is a cap, and the minimum payment might adjust by less if a 7.5 percent increase were to exceed the fully amortizing payment. In practice, the cap is nearly always binding. The schedule may also 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 Hybrid ARMs adjust annually or semi-annually, Option ARMs update much more frequently—usually monthly. New interest rates are calculated as the sum of a fixed margin component and an index term that proxies for the cost of funds to the lender.

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

(IV) *Proposed Payment Options:* The name, Option ARM, refers to a menu of payment options offered on borrowers' 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 payment every month.

For the purposes of our identification strategy, (I) and (II) are key. Because payments are fixed for the first 5 years, borrowers' balances change as a function of realized interest rates.[22] As a result, realized leverage depends on the particular mortgage index a loan is tied to, and the specific origination month. In the next section we discuss these features and our identification strategy in greater depth.

In the years leading up to the crisis, Option ARMs became a significant fraction of the market, representing approximately 9 percent of originations in 2006.[23] As the crisis hit, borrowers with Option ARMs defaulted at high rates. In the sample studied here, 50 percent of borrowers were seriously delinquent (60 days past due) on their mortgages at some point within the first 5 years. 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., 2018). 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 Option ARM style products, both suggesting

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

[23] See the 2008 Mortgage Market Statistical Annual. 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.

that they approximate the optimal mortgage contract when borrowers have stochastic income (Piskorski and Tchistyi, 2010)[24] and noting potential benefits from a macroprudential perspective (Campbell, Cocco and Clara, 2017).

3.2 Data

The data on Option ARMs used in this paper are taken from a loan-level panel of privately securitized mortgages provided by Moody's Analytics representing over 90 percent of non-agency residential mortgage-backed securities (the same data was formerly provided by Blackbox Logic). 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. In our primary analysis we limit our sample to the approximately 600,000 Option ARMs originated between 2004 and 2007 with initial combined loan-to-value ratios between 50 and 100.

3.3 Summary Statistics: Balance Across Indices

Table I shows summary statistics, split across the different mortgage indexes that appear in our sample. Treasury and LIBOR are by far the dominant indexes, representing just over 80 percent of the sample and just over 15 percent of the sample, respectively. A small number of loans are also reported as tied to COFI (the 11th District Cost of Funds Index) or as having fixed interest rates in the initial period, with each category representing less than two percent of loans. While we include all in our baseline analysis, we also show robustness limiting our sample to LIBOR and Treasury indexed loans.

The characteristics of loans at origination are reasonably balanced across indexes. Origination leverage, measured here as the combined-loan-to-value (Origination CLTV) to account for all liens, is close to 0.8 for all categories, averaging 0.80 for Treasury and 0.82 for LIBOR. Similarly, Origination LTV, which does not consider other liens, is 0.77 on average for both LIBOR and Treasury. This is slightly larger than the average for conforming loans purchased by Fannie Mae or Freddie Mac but below the average LTV for subprime adjustable-rate mortgages.[25] The value of the home at origination is also similar for LIBOR, Treasury and COFI at just below $500,000 (although larger for the fixed rate category). Note that the median loan is much closer to $400,000 across all categories, as the distribution is skewed due to a small number of very expensive properties.

[24]See also Guren, Krishnamurthy and McQuade (2018), which confirms the findings in Piskorski and Tchistyi (2010), but also suggests that the Option ARM may perform particularly poorly in crises.

[25]The average original LTVs for Fannie Mae and Freddie Mac in 2007 were 72 and 71, respectively, according to Frame, Lehnert and Prescott (2008).

In general, borrowers have high FICO credit scores, with an average of 706 for Treasury and 717 for LIBOR. These FICO scores, combined with average home values, suggest that Option ARMs attracted a set of relatively high credit quality borrowers.[26] This is consistent with Amromin et al. (2018), which finds that borrowers with complex mortgages tend to be sophisticated, with high incomes and credit scores relative to the subprime population.

One peculiarity distinguishing Option ARMs from conforming loans is the rarity of income verification. Given borrowers with high credit scores, low required monthly payments, protections against payment increases, and generally favorable expectations about housing prices, lenders appear to have been unconcerned ex-ante about borrowers' ability to meet monthly obligations. This led to a prevalence of low or no documentation loans: 82 percent for Treasury and 85 percent for LIBOR. For these loans, borrowers provided little or no formal evidence of sufficient income to meet monthly payments, often simply stating income with no verification. [27]

A slightly higher fraction of LIBOR indexed loans were used for home purchases versus refinances when compared to Treasury indexed loans (40 percent versus 32 percent), but the fraction used for single family homes is nearly identical. Similarly, the fraction of borrowers noted as investors is under 20 percent for both, although this is potentially an understatement due to unreported investors (Piskorski, Seru and Witkin, 2015).

The four most common states for both indices are California, Florida, Arizona, and Nevada, reflecting the broader mortgage market. While Treasury loans are slightly more concentrated in California, the overall geographic patterns are similar across states. There is some difference across indexes in the timing of origination, with a higher proportion of LIBOR loans originated in 2004 and 2007, and a higher proportion of Treasury loans originated in 2004 and 2005.

The basic empirical fact our analysis seeks to explain is the positive correlation between leverage and default. Figure I displays this correlation in our sample, plotting the relationship between origination combined loan-to-value and default within the first five years of the loan. We see a strong and roughly linear upward slope, with defaults well over 50 percent for the most highly levered borrowers. In the next section we describe our strategy for decomposing this correlation to uncover the relative roles of adverse selection versus the causal relationship.

[26]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).

[27]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 in this context. Finally, we show how the model translates to a series of estimating equations and describe the IV 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 loan age t, a borrower i defaults if the difference between the loan balance (B_{it}) and the home price (H_{it}) exceeds some borrower (and time) specific cost of default: $B_{it} - H_{it} > C_{it}$.[28]

To follow the literature, we divide through by the home value to reframe this equation in terms of LTV ratios:

$$\underbrace{LTV_{it}}_{B_{it}/H_{it}} > \underbrace{\tilde{C}_{it}}_{(1+C_{it})/H_{it}} .$$

In other words, borrower i defaults at age t if the LTV ratio exceeds some borrower (and time) specific threshold $\tilde{C}_{it}$. Rearranging, and decomposing the individual specific threshold into observable ($\mathbf{x_i}$) and unobservable (ε_{it}) components, we may rewrite the default rule as:[29]

$$D_{i,t+1} = \mathbb{1}\{\alpha LTV_{it} + \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_{it} > 0\} \tag{1}$$

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

While the borrower's contract choice is the result of a complex maximization problem, we abstract from this and specify a linear demand model for leverage. Letting L_i represent the original leverage chosen by borrower i and $\mathbf{x_i}$ be the full set of observables that lenders are able to price on:

$$L_i = \mathbf{x}_i'\boldsymbol{\psi} + v_i. \tag{2}$$

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

[28]The time index indicates that C_{it} need not be fixed for an individual, and may depend on, for example, time specific liquidity shocks.

[29]We decompose $-C_{it} = \mathbf{x}_i'\frac{\boldsymbol{\beta}}{\alpha} + \frac{1}{\alpha}\varepsilon_{it}$, where x_i 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. Normalizing the variance of ε_{it}, α is then the reciprocal of the standard deviation of the default costs in the population (in LTV units).

borrower's current LTV on default. α can also be interpreted as (the square-root of) 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 current leverage constant (large ε_{it}).

4.2 Challenges to 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 LTV_{it}. In the absence of other differences, borrowers with identical L_i will tend to have identical LTV_{it}. For borrowers consistently making minimum payments, there are only two factors that might cause those with identical L_i to have different LTV_{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, LTV_{it} is usually measured with error. Similarly, variation across time in interest rates is likely correlated with macro conditions, while cross-sectional variation potentially reflects endogenous contract choices. Isolating exogenous variation in LTV_{it} is non-trivial but necessary to accurately estimate the causal effect of leverage on default (α) as distinct from the initial contract choice (e.g. by comparing borrowers who have the same L_i but, for exogenous reasons, have different LTV_{it}). Similarly, exogenous variation in LTV_{it} is necessary to recover the correlation between initial leverage choice and default (γ) is distinct from the causal effect (e.g. by comparing borrowers who have endogenously chosen different L_i but the have the same realized LTV_{it} for exogenous reasons). Put differently, exogenous variation allows us to directly estimate the causal effect of leverage, and hence attribute the residual correlation to adverse selection.

4.3 A Source Of Variation: Diverging Mortgage Indexes

To isolate plausibly exogenous variation in LTV_{it}, our identification strategy exploits the divergence of financial indexes used to determine interest rate adjustments for Option ARMs. The difference between the two most common indexes, LIBOR and Treasury, is illustrated in Figure II, which shows the levels of (Panel A) and spread between (Panel B) 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 sharply, eventually peaking at over 3 percentage points in late 2008 following the Lehman Brothers bankruptcy filing and the AIG bailout.

But how do differences in interest rates cleanly translate into differences in current leverage (LTV_{it})? This is where the unique features of the Option ARM come into play. Because minimum payments are set in the initial period for Option ARMs, changes in interest rates have *no contemporaneous impact* on monthly obligations. As the mortgage must account for changes in the interest rate somehow, any additional interest accrual is incorporated directly into the balance.[30]

This means that—for Option ARMs—the divergence between mortgage indexes caused borrowers with otherwise identical loans to have sizable differences in leverage ex-post. Appendix Figure A.II provides a stylized example of this pattern. Consider two identical $100,000 loans at origination, one of which faces a high realization of interest rates, while the other faces a low realization. Two years into the loan, the two borrowers will still have the same monthly payment, but the first borrower may owe thousands of dollars more than the second.

Of course, a key question is the source of variation in indexes across borrowers. Prior to the crisis, borrowers had little reason to prefer one index to another. For example, although there tended to be a spread between LIBOR and Treasury rates—the spread between 1-year CMT and 1-year LIBOR was generally below 50 basis points—the two indexes moved quite closely together, and any fixed difference could be accounted for in the margin. Furthermore, Bucks and Pence (2008) suggests that borrowers tended to be uninformed about their contract terms. When asked what index their loan depended on, only 25 percent of borrowers responded with plausibly correct indexes. 30 percent of borrowers simply answered that they did not know.

If borrowers were unaware of the distinction between indexes, why did some end up with a Treasury index and others with LIBOR? Much—although not all—of the variation comes as a result of the lender, with most originators specializing in a particular index. To highlight this, Appendix Table A.I shows the fraction of loans with Treasury indexes for the top 10 originators in our sample. Note that most tend to specialize in a particular index, and two of the top 10 make exclusively Treasury loans. As noted in Gupta (2019), differences across lenders are often a function not of the borrowers they lend to but their intentions on the secondary market. While the index is not exclusively determined by the originator (allowing us to include originator $\times$ origination month fixed effects), there is very limited variation within a month at a particular bank-branch (or originator-zipcode). In other words, the specific bank location where the mortgage was

[30] This stands in contrast to more typical "hybrid" adjustable rate mortgages, for which changes in the interest rate also change the monthly payment, which adjusts to ensure that the payments are fully amortizing.

originated likely determined the index.

While sorting between borrowers and originators—or other forms of selection into indexes on the basis of unobserved borrower characteristics—is certainly possible, random assignment to indexes is not necessary for our analysis. A critical feature of our approach is the fact that the impact of different indexes on leverage 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. Appendix Figure A.III demonstrates this difference-in-difference variation in borrowers' balances. The plot shows the loan balance over time for four sample $100,000 loans: one LIBOR-indexed and one Treasury-indexed loan originated in January 2005, and one of each originated in January 2007.[31] Each of the four shows a distinct balance trajectory. Our empirical strategy focuses specifically on the variation in LTV_{it} particular to the interaction between a borrower's index and origination month. This allows us to control for any fixed differences between those with different indexes, as well as any aggregate time specific effects.

4.4 Jackknife Estimator to Capture Index×Origination Month Variation

To isolate the variation in leverage driven by the interaction of index and origination month, we implement a jackknife estimator that instruments for a borrowers' LTV_{it} with the leave-one-out average of LTV_{jt} for all other borrowers with the same index type and origination month. In particular, for borrower i with mortgage index $I(i)$ originated in month $m(i)$ we define the instrument:

$$Z_{it} = \frac{1}{n_{I(i)\times m(i)} - 1}\left[\left(\sum_{j=1}^{n_{I(i)\times m(i)}} LTV_{jt}\right) - LTV_{it}\right], \tag{3}$$

where $n_{I(i)\times m(i)}$ denotes the total number of loans originated in month $m(i)$ with index $I(i)$ active at age t.

Z_{it} itself captures *all* variation that operates at the index or origination month level, including any aggregate time-series variation in leverage, and any fixed differences between indexes. Crucially, our strategy below isolates *only* the variation driven by the interaction $I(i) \times m(i)$ by further including both index and origination month fixed effects. This strategy mimics the use of a full set of $I(i) \times m(i)$ fixed effects while providing superior small sample estimation properties, as suggested in Kolesár (2013). Furthermore, this approach collapses a potentially high-dimensional set of instruments into a single Z_{it}. This allows us to both estimate a first stage regression with interpretable coefficients, and to provide evidence towards the exogeneity of Z_{it} by testing for correlation with observables.

The differences in realized leverage implied by these instruments is non-trivial. Appendix Figure A.IV

[31]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.

highlights variation in Z_{it} 24 months after origination. For simplicity, we focus only on LIBOR and Treasury indexes. Panel A shows a histogram of a origination-month level proxy for the difference in leverage at 24 months generated by the index choice. Specifically, the maximum difference in the value of Z_{it} across indexes for each origination month.[32] The typical variation in leverage is substantial: averaging around 2–3 LTV points. However, the instrument has a substantial range, reaching close to 10 LTV points for certain origination months. The top end of this range is realized, as Panel B shows, for loans originated later in our sample.

4.5 Estimating Equations For Linear Specification

Our goal is to estimate our leverage choice equation (Equation 2) and default equation (Equation 1) while instrumenting for current LTV. We do so in two ways. First, we consider a reduced form approach that collapses both into a single equation that can be estimated via standard 2SLS approaches.[33] This allows us to implement a thorough set of fixed effects and other controls in a computationally feasible manner, and to perform a wide array of robustness tests. Second, we estimate a joint model that allows us to estimate these equations directly using a control function approach.

Our basic reduced form approach considers the outcome of default between loan age t and $t+1$ for the cross-section of borrowers active in our data at loan age t. In particular, we consider a linear probability model for default of the form

$$D_{it+1} = \alpha LTV_{it} + \gamma L_i + \mathbf{x}_i' \boldsymbol{\beta} + \zeta_{j(i)} + \lambda_{I(i)} + \theta_{o(i)} \times \mu_{m(i)} + e_{it}. \tag{4}$$

To isolate variation in leverage driven by the divergence of financial indexes we instrument for LTV_{it} using the leave-one-out mean Z_{it} :

$$LTV_{it} = \delta Z_{it} + \eta L_i + \mathbf{x}_i' \boldsymbol{\pi} + \zeta_{j(i)} + \lambda_{I(i)} + \theta_{o(i)} \times \mu_{m(i)} + u_{it}. \tag{5}$$

These equations, which represent our most saturated specifications, include three sets of fixed effects: a zipcode $j(i)$ fixed effect $\zeta_{j(i)}$, an index I(i) fixed effect $\lambda_{I(i)}$, and an originator×origination month fixed effect $\theta_i \times \mu_{m(i)}$. $\mathbf{x}_i$ denotes a rich set of borrower and loan controls that represent information available to the bank at origination. Because we estimate the above cross-sectionally at different loan ages, we do not

[32] Given the leave out strategy, this similar to the absolute value of the difference across indexes in the index-specific average value of Z_{it}.

[33] To derive Equation 4 from Equations 1 and 2 we write $\varepsilon_{it} = \gamma v_i + e_{it}$, 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 $\varepsilon_{it} = \gamma(L_i - \mathbf{x}_i' \boldsymbol{\psi}) + e_{it}$. Substituting for ε_{it} in Equation (1) gives Equation 4. In a slight abuse of notation, we take $\mathbf{x}_i$ in Equations 1 and 2 to include the fixed effects spelled out in full in Equation 4.

include loan age effects (or time effects separate from origination month).

Our variables of interest are LTV_{it} and L_i, where α captures the causal effect of leverage on default and $\gamma > 0$ indicates the presence of adverse selection. We define L_i as the borrower's total leverage at origination, measured as the combined loan-to-value. We define LTV_{it} as the current loan-to-value on the Option ARM at loan age t.[34] Our standard default measure is a borrower falling 60 days past due on monthly payments, although we also consider more severe measures of default in robustness tests. Our standard results consider default between 24 and 36 months, although we also consider a variety of other cross-sections in robustness tests. Throughout, we cluster standard errors at the originator level. Our key assumptions to identify the causal effect are that Z_{it} is correlated with realized leverage ($Corr(Z_{it}, LTV_{it} \neq 0)$) and uncorrelated with the error in the default equation ($Corr(Z_{it}, e_{it}) = 0$).

In our estimation we gradually build up to these saturated specifications. At a minimum, we include index, origination month, and zipcode fixed effects. This simple set of controls explicitly adjusts for any fixed differences across indexes, as well as any aggregate time-varying effects, and therefore focuses on the variation in the instrument due to the interaction between index and origination month. We then add borrower and loan controls, which proxy for the set of observables available to the bank at origination. These include a flexible set of dummies for home value and FICO score, and indicators for loan occupancy, property type, and documentation status. Given the originate-to-distribute model at the time, soft information was a relatively minimal component of the origination process. As a result, these variables represent a reasonable proxy for the information used to price loans by originators. Effectively controlling for these characteristics is necessary to distinguish adverse selection from sorting on observables (which might occur, for example, because lenders offer *observably* different borrowers different prices for the same leverage).

Finally, we add controls for originator $\times$ originator month to account for any time-varying originator level confounds. This rules out, for example, threats due to time variation in the risk characteristics of the borrowers at different originators. This means that our saturated specifications focus on variation across loans with different indexes with a given month at a given originator.

4.6 Joint Model

While the reduced form model presented above has computational advantages, it also has a few downsides. In particular, it imposes an unattractive linearity assumption on the causal relationship between leverage and default—we would typically expect the marginal effect to depend on the level of leverage—and does

[34] We consider total vs. loan specific leverage at origination as lenders may account for second liens in pricing. While we would ideally do the same for current leverage, we are unable to observe the rate of amortization for second liens, and hence focus only on the Option ARM first lien. Of course, this raises the possibility that our observed γ might capture a causal effect due to second liens, rather than a true selection effect. To rule this out, we perform robustness checks on the subsample with no second lien. We use Zillow's zip code level home price index, available at http://www.zillow.com/research/data/ to impute current home values.

not provide parameters that are directly relevant for our simulations in Section 6. To address both these issues, we complement our analysis above by estimating a joint model of Equations 1 and 2 while including 5 to instrument for LTV_{it}. This gives us three equations (letting $\mathbf{x}_i$ include all fixed effects here, for economy of notation):

$$\begin{aligned}
D_{it+1} &= \mathbb{1}\{\alpha LTV_{it} + \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_{it} > 0\} \\
L_i &= \mathbf{x}_i'\boldsymbol{\psi} + v_i \\
LTV_{it} &= \delta Z_{it} + \eta L_i + \mathbf{x}_i'\boldsymbol{\pi} + u_{it}
\end{aligned}$$

Our estimation procedure takes a control function approach following Blundell and Powell (2004), incorporating an additional linear equation. To do so, we impose a simple parametric structure on the errors:

$$\begin{pmatrix} \varepsilon_{it} \\ v_i \\ u_{it} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 1 & & \\ \rho_{\varepsilon v}\sigma_v & \sigma_v^2 & \\ \rho_{\varepsilon u}\sigma_u & \rho_{vu}\sigma_v\sigma_u & \sigma_u^2 \end{bmatrix} \right),$$

which enables direct implementation via maximum likelihood. Again, we estimate cross-sectionally at a given t and hence make no assumption about the evolution of errors over time. The key intuition here is that $\mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_{it}$ captures the default threshold relative to LTV_{it} for individual i, with $\mathbf{x}_i'\boldsymbol{\beta}$ representing the observed component, and ε_{it} the unobserved component. Estimating the parameters of the model therefore allows us to recover the distribution of the unobserved portion of the default threshold in the population (scaled relative to LTV):

$$\tilde{C}_{it}|\mathbf{x}_i \sim N\Big(-\mathbf{x}_i'\boldsymbol{\beta}, \frac{1}{\alpha^2} \Big).$$

This distribution determines the strength of the moral hazard effect in the population.[35] Perhaps more importantly, we directly recover $\rho_{\varepsilon v}$, the correlation between ε_{ijt} and v_i. This correlation captures the strength of the adverse selection effect—the degree to which leverage choice ex-ante is correlated with the unobserved portion of default costs. $\rho_{\varepsilon v}$ and the parameters of the distribution of $\tilde{C}_{it}$ are the key elements of the simulations discussed in Section 6.

[35] For example, a smaller α indicates a more dispersed distribution of thresholds, and hence a smaller marginal effect at any given point in the distribution.

5 Results

In this section, we describe the central empirical results of the paper. We begin with a series of tests for correlations between our instrument and observables to support the assumption of instrument exogeneity. We then turn to our reduced form linear model. We find a strong first stage, indicating that our instrument has strong impacts on realized mortgage balances. We disentangle moral hazard and adverse selection, and show that our results hold across a wide range of robustness checks. Finally, we discuss the estimation of the joint model of leverage demand and default choice which provides parameters that directly inform the simulations presented in Section 6.

5.1 Testing for Correlation With Observables

The key identifying assumption in our analysis is that the index×origination month specific variation captured by our leave-one-out mean instrument does not correlate with default except through its influence on leverage. More formally, that $Corr(Z_{it}, e_{it}) = 0$. While we cannot explicitly verify this assumption, we are able to test whether the particular variation we isolate in our IV strategy correlates with observables. To do so, we regress key borrower and loan level observables on our instrument, conditioning on our minimal set of fixed effects: index type, origination month, and zipcode.

The results of these regressions, presented in Table II, show a clear picture. Firstly, our instrument is strongly and statistically significantly positively correlated with both current leverage and default (measured here at 24 months and between 24 and 36 months, respectively). This suggests a valid first stage and a significant reduced form effect of our instrument on default. However, we see no evidence of a correlation between our instrument and any observable characteristics of the borrower or loan. In particular, we see no significant relationship when considering a borrower's FICO score, origination leverage, original home value, documentation status or investor status, nor when consider whether the loan was for a home purchase (vs. refinance), or for a single family home. We present the same information graphically in Figure III, where we additionally normalize both our instruments and continuous observable variables by their standard deviations to provide more easily interpretable coefficients.[36] On the whole, these results provide evidence supporting the plausibility of our instrument exogeneity assumption.

5.2 Main Results: Reduced Form Model

We now turn to our primary specifications, which attempt to isolate the roles of adverse selection and moral hazard in the relationship between leverage and default following Equation 4. In our main estimates, shown

[36] Binary variables are not normalized.

Table III, we use linear probability models and our leave-one-out mean instrumental variable approach (our joint model, presented later in this section, considers probit-style control function estimates). We focus on the one year default rate for borrowers 24 months after origination and cluster standard errors at the originator level throughout.[37]

The main tables are structured to show three types of specifications, which we label *Baseline*, *OLS*, and *IV*. The baseline specifications show the raw relationship between origination leverage L_i and default without considering the role of current leverage—an analogue of Figure I. Our OLS specifications additionally include LTV_{it}, and so match Equation 4 (with no instrument). Finally, our IV approach instruments for LTV_{it} using Z_{it}. Additionally, for each of these we show results with increasingly rich controls. We begin with our simplest set, which controls only for index type, zipcode, and origination month. We then add borrower and loan controls. Our most saturated specification additionally includes originator $\times$ origination month fixed effects.

Our baseline regressions confirm that the positive correlation between original leverage L_i and default—shown in Figure I—persists when conditioning on controls available to the lender at the time of origination. Column (1), which includes our minimal set of controls, suggests that a 10 point increase in the leverage ratio at origination is associated with a just over 9 percentage point increase in the one year default rate. This rises to over 10.5 percentage points when including our richest set of controls (in Column (3)). This specification is analogous to the positive correlation outlined in Chiappori and Salanié (2000), and indicates the presence of asymmetric information. Furthermore, the fact that the correlation rises when controls are included suggests that lenders may steer *observably* high risk borrowers away from high leverage mortgage, either through pricing or other means.

The relationship between original leverage and default declines when accounting for current leverage via OLS. These specifications, which include LTV_{it} without incorporating an instrument, are shown in Columns (4)-(6). These results suggest that a 10 point increase in the leverage ratio at origination is associated with a significant 6.7-8.2 percentage point increase the one year default rate, depending on the set of controls. We also find evidence of a significant relationship between *current* leverage and default in the OLS, indicating that a 10 point increase in current LTV is associated with a 2.8-3.0 percentage point increase in default.

A simple way to quantify the relative importance of adverse selection is to compare the correlation between origination leverage and default when conditioning on *current* leverage to the unconditional baseline. Taking our OLS estimates at face value, the coefficient on origination leverage in Column (6) suggests that

[37] 24 months represents a middle point between allowing sufficient time for leverage to diverge, and not allowing so much time such that a significant fraction of borrowers have exited the sample. Appendix Table A.II considers specifications aimed at capturing cross-sections at other loan ages.

adverse selection is responsible for roughly 77 percent of the baseline correlation between leverage and default (0.818/1.066). This suggests that the effect of the loan balance on default—moral hazard—is responsible for the remaining 23 percent.

Finally, in our IV specifications (7)–(9) we decompose the relationship between leverage and default while instrumenting for current LTV using our leave-one-out mean. The first stage, reported in Panel B of this table, suggests instrument relevance. Higher leverage among borrowers with the same origination month and index type predicts higher current leverage—even conditional on origination month, index, geography, and other controls—because these borrowers are exposed to comparable interest rate trajectories which flow through to their balances. Note that the coefficients are less than one, consistent with the leave-one-out mean representing a noisy proxy for true leverage. We find F-statistics for the first stage between 25–28, well above standard rules of thumb (Stock and Yogo, 2005).

Our IV results indicate the presence of moral hazard, as we find a significant and positive causal relationship between leverage and default. Our linear IV specification suggests that a 10 point increase in current LTV increases the probability of default by 4.7 percentage points. Note that this effect is roughly 62 percent higher than the effect implied by our OLS estimates. While there are a large number of potential omitted variables that might generate this bias, it is also useful to recall the presence of measurement error in the construction of LTV ratios, which creates a generic source of attenuation. Of course, our estimate should be interpreted as a LATE, and additionally represents the average for a population with a relatively high baseline level of leverage, so extrapolation of the precise magnitudes should be taken with care. The parameters of our joint model provide a more reasonable benchmark for other contexts. Still, this provides strong evidence for the existence of a causal effect.

Despite this meaningful causal effect, our results indicate that adverse selection is the dominant force behind the baseline correlation between leverage and default. This effect is captured by the coefficient on origination leverage (L_i). We estimate that a 10-point higher initial leverage ratio is associated with a 4.7-7.0 percentage point increase in the one year default probability. Our richest specification suggests that adverse selection is responsible for 60 percent of the overall correlation between leverage and default, we attribute the remaining 40 percent to moral hazard.[38]

5.3 Non-Linearity in the Causal Effect

Our model above imposes a strong linearity assumption on the causal relationship between leverage and default. While this is useful for determining the *presence* of a causal effect, a more nuanced view is necessary

[38]Once again, we arrive at this breakdown by comparing the correlation between origination leverage and default when conditioning on *current* leverage to the unconditional baseline (i.e. 0.655/1.066).

for several reasons. A first concern is our interpretation of the coefficient on origination leverage as evidence for adverse selection. If there are substantial non-linearities, it is possible that what we term selection simply reflects a portion of the causal relationship that is not captured by our specification. To address this, Columns (6)-(9) of Appendix Table A.III repeat our analysis while controlling for a flexible polynomial in current leverage to account for any non-linearities. We continue to find strong evidence for adverse selection with this more general specification. The magnitude of the coefficients on origination leverage in both our OLS and IV specifications are consistent with those in Table III, although slightly less statistically significant in the case of the IV.

Perhaps more importantly, precisely understanding how the causal effect differs across the distribution of leverage is a key question from both an academic and policy perspective. For an intervention that forgives borrowers' balances to be successful, a strong marginal effect at the relevant level of leverage is more important than the size of the causal relationship on average. Furthermore, the shape of the causal relationship between leverage and default is informative in distinguishing two alternative models of borrower default. In the first, an unconstrained strategic model, leverage matters for default because borrowers simply weigh the financial benefits of repaying against the financial costs when deciding whether to default. In the second, high leverage matters because it exacerbates a liquidity constraint. Underwater borrowers may be unable to avoid default by refinancing or selling the home.

Figure IV shows a non-parametric version of the unconditional relationship between leverage at 24 months and default in our data. The plotted line is upward sloping and appears approximately linear. However, because this figure does not account for selection or other confounds, it is unlikely to reflect the underlying causal relationship.

To provide insight into the potentially non-linear causal relationship, we consider an expanded version of our IV specification. Specifically, we estimate Equation 4, but replace the linear term LTV_{it} with a set of dummies for each quartile of the distribution of current leverage. We instrument for these with dummies for quartiles of the distribution of the leave-one-out mean Z_{it}.

Results from our IV specifications, summarized in Panel A of Figure V, show evidence of a distinct non-linearity. Specifically, we observe a sharp jump upward around the negative equity threshold. The figure plots the coefficients on each quartile of LTV_{it}, with the first quartile normalized to 0. The estimated probability of default for the second quartile of leverage, representing current LTVs between 79 and 95, is effectively identical to that of the first quartile (below 79). This jumps upward by roughly 30 percent for the third quartile, which represents current LTVs between 95 and 115. There is then virtually no change between the third and fourth quartiles, suggesting there is little marginal effect of leverage beyond the third quartile.

This pattern is consistent with so-called double trigger models of mortgage default. The fact that the default probability jumps sharply around the threshold, but remains well below 100 percent, suggests that negative equity appears to be a necessary but not sufficient condition for default. This indicates that negative equity leads to default only in combination with some other circumstance (e.g. a liquidity shock). The relatively flat relationship between the third and fourth quartile suggests a limited strategic motive for default, at least for borrowers that are deeply underwater. Of course, quartiles are a somewhat rough way of capturing potentially complex nonlinearities. To display richer patterns in the data, Figure A.V shows the non-parametric reduced form relationship between our leave-out-mean instrument and default. The reduced form indicates a similar pattern. In particular, we see an increase in the default probability leading up to the negative equity threshold, a kink around the threshold, and a relatively flat relationship beyond that point.

Our specifications also include dummies for each quartile of Origination LTV, to allow for a non-linear relationships on the selection dimension. These results are summarized in Panel B of Figure V, and show a close-to-linear relationship between initial leverage and default. We present regression estimates corresponding to both figures in Appendix Table A.IV, which also displays baseline and OLS versions, as well as specifications with various levels of controls. Interestingly, our OLS results do not indicate nearly as sharp of an upward jump around the negative equity threshold, highlighting the importance of our IV approach.

5.4 Differential Attrition

A further potential concern is the possibility of non-random attrition. The sample in Table III is of loans that are active at 24 months. Our results may be biased if loans differentially exit our sample prior to 24 months (either by defaulting or prepaying) in a way that correlates with our instrument.

To address this possibility, we re-run our analysis on the full sample of loans, whether or not they exit the sample early. We consider cumulative defaults as an outcome, that is, default at any point up to time $t+1$. Of course, we are unable observe a borrower's loan balance after they exit the sample, and therefore do not see current leverage directly. However, because we observe all of the terms of the loan, we are able to mechanically infer the balance a borrower making minimum payments *would* have had at any point, and use this to construct a measure of current leverage. While this may not perfectly match actual leverage—a borrower's true balance may differ because they make an above minimum payment or miss a payment—our use of this inferred leverage can be thought of in the spirit of an intent-to-treat effect. Our approach considers the impact of the leverage the borrower *would have had* sticking to standard monthly payments.

Our specifications in Appendix Table A.II use this inferred measure of leverage for all borrowers. We once again instrument with a leave-one-out mean—although this time we calculate the instrument using other borrowers' inferred leverage.[39] In Columns (1)-(3) we repeat our main analysis. The basic patterns are similar: we find a positive and significant effect of current leverage in both our OLS and IV approaches, but find that adverse selection explains the majority of the correlation between leverage and default. However, note that our estimates suggest the role of selection is larger, and the causal effect smaller, than in our main specifications. This difference may reflect attenuation due to measurement error in inferred leverage.

This full sample approach also allows us to look at a consistent sample at different cross sections over the life of the loan. In Columns (4)-(6) we consider default between 36 and 48 months, and in Columns (7)-(9) we consider default between 48 and 60 months. We see consistent results in these later windows.

5.5 Robustness

We also consider a range of further robustness exercises for our linear specification, discussed below in turn.

The Role of Junior Liens

A potential concern with our primary specifications is that our measure of origination leverage—the origination combined LTV—incorporates junior liens, while our measure of current LTV does not. This raises the possibility that the estimated effect of initial leverage is driven by the causal effect of the second lien, rather than adverse selection. To rule this out, Columns (1)-(3) of Table A.V repeat our analysis on the subsample of loans without junior liens. The baseline correlation between leverage and default in this sample (Column (1)) is nearly identical to that in our full sample, and the magnitudes of our OLS and IV estimates are comparable for both initial and current leverage (although the standard errors on current leverage are slightly larger, reducing our significance level to 10 percent).

Restricting to LIBOR and Treasury Indexes

While our primary specifications incorporate four distinct indexes, two—LIBOR and Treasury—are dominant. In Columns (4)-(6) of Table A.V, we show that our results are robust to limiting our sample to LIBOR and Treasury indexed loans. Baseline and OLS specifications are virtually identical to the full sample results shown in Table III. Our IV estimates are statistically indistinguishable from those in the full sample, but indicate a slightly smaller role for adverse selection and a slightly larger role for moral hazard (on the order of 50 percent of the baseline correlation each).

[39] Doing so accounts, e.g. for endogenous loan terms such as the interest rate. Note also that this approach simultaneously addresses another potential concern, which is that differential payment patterns among borrowers with a particular index and origination month could generate variation in the instrument that is not driven by the path of interest rates.

Alternative Measures of Default

Columns (7)-(9) of Table A.V show that our results are effectively unchanged when considering a more extreme definition of default: 90 days past due.

No Geographic Controls

The most minimal set of controls in Table III includes reasonably fine grained geographic controls in the form zipcode level fixed effects. However, the inclusion of these controls does not materially alter our main estimates. Columns (1)-(3) of Appendix Table A.III show that our results are robust to excluding these fixed effects and conditioning only on index type and origination month.

Time-Varying Geographic Controls

Even in the presence of zipcode fixed effects, a potential concern is the presence of time-varying geographic confounds. For example, if an index is concentrated in a particular region, a local economic shock may generate variation in our instrument at the index$\times$origination month level (because of a change in local home prices) while simultaneously impacting default. In Columns (4)-(6) of Appendix Table A.III we account for the presence of time varying geographic effects by replacing our geographic fixed effects with a full set of MSA$\times$origination month fixed effects. Our results are similar to those presented in Table III.

Subsample of Current Loans

Our analysis in Table III conditions on loans that are active at 24 months. This includes loans that have been delinquent or in default prior to 24 months, so long as they have not gone into foreclosure. In Appendix Table A.VI, we condition on loans that are current at 24 months, and have not been delinquent previously. While the level of the baseline correlation is smaller, our estimates are otherwise in line with our main results.

Heterogeneity by Loan and Borrower Characteristics

In Appendix Table A.VII we explore heterogeneity in our results by three relevant loan-level characteristics: the documentation status of the borrower, whether the loan is for a home purchase or refinance, and the recourse status of the loan's state. We find stronger evidence of adverse selection for loans with no or low documentation versus those with full documentation, but fairly consistent results across our other categories.

5.6 Joint Model

The final step of our empirical analysis is to estimate a joint model of leverage demand alongside the default choice. Doing so has two primary advantages. First, it complements our investigation of non-linearities above by specifying a more realistic probit-style default equation. Second, it allows us to recover parameters that more directly relate to the model developed in the appendix and that can be used to inform the simulations developed in the next section.These parameters can also be used to simulate default probabilities in other counterfactual scenarios, for example, in a world with strong housing price growth.

Because of the increased computational complexity of this estimation, we slightly reduce the richness of included controls, e.g. substituting zipcode fixed effects with MSA fixed effects, and replacing Originator×Month fixed effects with Originator×Year fixed effects. Otherwise, we present our results as in Table III: baseline estimates that exclude current leverage, OLS specifications, and finally our control function approach (the analogue of our IV specifications).

The qualitative interpretation of these results, shown in Table IV, is identical to that in our linear specification. We both find strong evidence for adverse selection and a strong underlying causal effect. However, rather than considering marginal effects in this context, we focus on the estimates themselves which provide a series of directly interpretable parameters.

The presence of adverse selection is summarized by $\rho_{\varepsilon v}$, the correlation between the errors in the leverage choice and default equations. Across all specifications, we find a positive and significant value for $\rho_{\varepsilon v}$, with an estimated correlation of 0.12 in our most saturated IV specification. This suggests that the borrowers who are most likely to default—for ex-ante unobservable reasons—choose the highest leverage.

Similarly, we are able to estimate the parameters of the distribution of unobserved default thresholds in the population: $\tilde{C}_{it}|\mathbf{x}_i \sim N\left(-\mathbf{x}_i'\boldsymbol{\beta}, \frac{1}{\alpha^2}\right)$. The key parameter is the standard deviation, measured in units of LTV points. Here, a larger value indicates a more dispersed distribution of thresholds and therefore a relatively small marginal effect at any given point. Our estimates of this parameter are analogous to those in the linear specification: we find larger estimates (lower marginal effects) in our OLS specifications, and smaller estimates (higher marginal effects) in our IV specifications. Still, across the board, we find a positive parameter that is statistically different from 0. In our most saturated IV specification, we estimate this standard deviation to be 0.43, or 43 LTV points. Given this, we are able to recover the distribution of thresholds for any set of observable characteristics under our normality assumption.

6 Simulations and Policy Analysis

In this section, we highlight the implications of our estimates for policy in a simulation framework. As an illustration we analyze the implementation of an LTV cap, which we view as a reasonably representative example of the types of macroprudential regulations implemented in various countries since the recession. In conjunction with concerns about asset price booms and busts, these policies are commonly motivated by the causal relationship, under the assumption that limiting consumer leverage directly reduces the scope for ex-post default. In our analysis, we quantify the value of LTV caps in lowering default rates and identify the default externalities policymakers need to account for when designing 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 because borrowers face higher interest rates and take smaller mortgages in equilibrium. To address the challenges of evaluating counterfactual policies in competitive markets with adverse selection, we use the equilibrium concept proposed by Azevedo and Gottlieb (2017).

6.1 A Model to Evaluate Ex-Ante Regulations

Consumer Preferences

We consider a two period model. In the first, consumers purchase a home of fixed size 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.

In the first period, 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 B:

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

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 0.12. In a richly specified model, initial mortgage choice might also be a function of heterogeneity in each 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.

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 $G(C)$):

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

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, using 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 behavioral borrowers—uniformly distributed across contracts—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 down or up for profitable or unprofitable contracts. Convergence is achieved when the absolute value of 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.

Calibration

We calibrate three features of the simulation to the estimates from Table IV. We base these on Column (9), our IV specification with our richest set of fixed effects. First we define the standard deviation of C_i based upon our estimates [40] Next, 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 $\rho_{\varepsilon v}$. Finally, we set the mean of C_i by taking $-\mathbf{x}_i'\boldsymbol{\beta}$ using the our estimated β. We use average values of observed covariates (and the omitted base group for all categorical variables). All other parameters are set based on the data when possible and explicitly described in the bottom panel of Table V. For the purposes of the simulation, we assume that borrowers have exponential utility, with CARA coefficient a.

[40] While our estimates are in LTV ratios and our simulations are specified in levels, we simply scale by the expected realization of home values, which is the same for all borrowers in our exercise.

6.2 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.

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 VI. This figure shows results with an exaggerated degree of adverse selection, to better present the patterns across the three regimes, while Table V presents numbers based on simulations calibrated to the empirical results.

The dark gray bars illustrate the allocation of original LTV under Regime I and exhibit a basic pattern of adverse selection. While all 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. In other words, adverse selection leads to a partially separating equilibrium.

The expected impact of the regulation from the perspective of a naive regulator is illustrated with the light gray bars. 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.[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, borrowers who choose an LTV of 90 under Regime II are expected to default at the same rate as borrowers choosing an LTV of 90 under Regime I.

Columns (1) and (2) of Table V compare Regimes I and II. There is indeed a reduction in loan size, from \$387 thousand to \$361 thousand and a corresponding expected reduction in average interest rates from 14.8 to 11.4 percent. Because Regime II does not allow lenders to change interest rates, this reduction is entirely the result of borrowers choosing smaller loans with lower rates. Furthermore, by failing to account for the inherent riskiness of the borrowers who are shifted from LTVs above 90 to LTVs below 90, the naive regulator expects the reduction in defaults to be significantly larger than it actually is, even without a supply response. The naive view suggests that an LTV cap of 90 would cut the fraction of defaults by more than 22 percent, from over 18 percent of borrowers to roughly 14 percent. Appropriately accounting for the risk of the borrowers initially allocated above 90 reveals the true reduction to be closer to 17.5 percent, with more than 15 percent of borrowers continuing to default.

Allowing a Supply Response

In addition to overstating the reduction in defaults generated by the regulation, the naive view understates the changes in interest rates and loan 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 and 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 differentiate themselves, but instead forces them to do so over a smaller range of loans. The leftward shift of the white bars in Figure VI 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 V, the knock-on effects cause *average* interest rates for *all* borrowers to rise by roughly 50 basis points (from 11.38 to 11.83). Furthermore, borrowers reduce their loan size by over \$100 on average. Despite getting smaller loans, higher

[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.

interest rates lead the average borrower to have a final balance that is roughly \$1500 larger under Regime III than under Regime II.

Optimal regulation involves balancing reductions in defaults with the welfare loss that results from borrowers facing higher interest rates and taking smaller loans. In the simulations provided here, a naive regulator overstates the reduction in defaults by close to 5 percent and underestimates the additional costs born by borrowers (not accounting for the fact that they take smaller loans) by \$1,500 per borrower. While our model is highly stylized, the fundamental intuition is clear: when adverse selection is present, policy makers must weigh the benefits of preventing defaults against the knock-on effects that impact all borrowers when risky types are forced to choose new contracts.

7 Conclusion

In this paper, we empirically separate moral hazard from adverse selection in the mortgage market. To do so, we focus on a natural experiment generated by two features of Option ARMs: fixed payments and variable interest rates. Because monthly payments do not change when interest rates rise or fall, fluctuations in market rates directly impact borrowers' balances. 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 the variation in interest rates that comes as the result of the financial index used to determine rate adjustments. Because of the unexpected divergence between index rates during the crisis, borrowers experienced substantially different balances as a function of the loan's index and origination month.

We isolate this variation in borrower's balances using a leave-one-out jackknife IV estimator. This allows us 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. Adverse selection is responsible for 60 percent of the baseline correlation between leverage and default, while moral hazard is responsible for the remaining 40 percent.

To analyze the implications of adverse selection for ex-ante macroprodential policy, we construct and simulate a simple model of borrowers leverage choice and default using our estimated parameters. We show that, in the presence of adverse selection, policies such as loan-to-value caps may have knock-on-effects that impact *all* borrowers in equilibrium. Such regulations will cause all borrowers to face higher interest rates and choose smaller loans. In general, regulators should be cognizant of the potential for distortions in mortgage markets that comes as a result of adverse selection.

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. 2018. "Complex Mortgages" *Review of Finance*, 22(6): 1975–2007.

Ausubel, Lawrence M. 1999. "Adverse Selection in the Credit Card Market" Working Paper.

Azevedo, Eduardo M, and Daniel Gottlieb. 2017. "Perfect Competition in Markets with Adverse Selection" *Econometrica*, 85(1): 67–105.

Bajari, Patrick, Chenghuan Sean Chu, and Minjung Park. 2008. "An Empirical Model of Subprime Mortgage Default From 2000 to 2007" NBER Working Paper No. 14625.

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" *Working Paper*.

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.

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

DeFusco, Anthony, Stephanie Johnson, and John Mondragon. 2019. "Regulating Household Leverage" *The Review of Economic Studies, forthcoming*.

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

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*.

Ganong, Peter, and Pascal Noel. 2018. "Liquidity vs. Wealth in Household Debt Obligations: Evidence from Housing Policy in the Great Recession" NBER Working Paper No. 24964.

Geanakoplos, John. 2010. "The Leverage Cycle" *NBER Macroeconomics Annual*, 24: 1–66.

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. 2017. "Can't Pay or Won't Pay? Unemployment, Negative Equity, and Strategic Default" 3, Oxford University Press.

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. 2019. "Foreclosure Contagion and the Neighborhood Spillover Effects of Mortgage Defaults" *Journal of Finance*, 74: 2249–2301.

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*, 31(9): 3532–3567.

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.

Kolesár, Michal. 2013. "Estimation in an Instrumental Variables Model with Treatment Effect Heterogeneity" *Working Paper*.

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.

Palmer, Christopher. 2015. "Why did so many subprime borrowers default during the crisis: Loose credit or plummeting prices?"

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" *Journal of Finance*, 70(6): 2635–2678.

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.

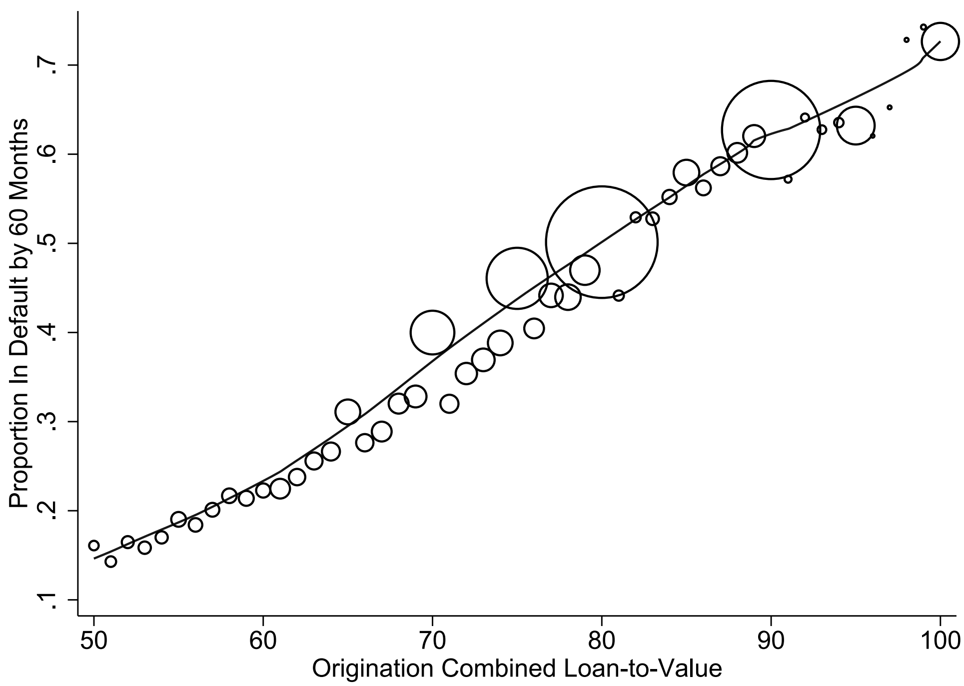
Rothschild, Michael, and Joseph Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information" *The Quarterly Journal of Economics*, 90(4): 629–649.

Stock, James, and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression" In *Identification and Inference for Econometric Models*. , ed. Donald W.K. Andrews, Chapter 10, 80–108. New York:Cambridge University Press.

Vandell, Kerry D. 1995. "How Ruthless is Mortgage Default? A Review and Synthesis of the Evidence" *Journal of Housing Research*, 6(2): 245.

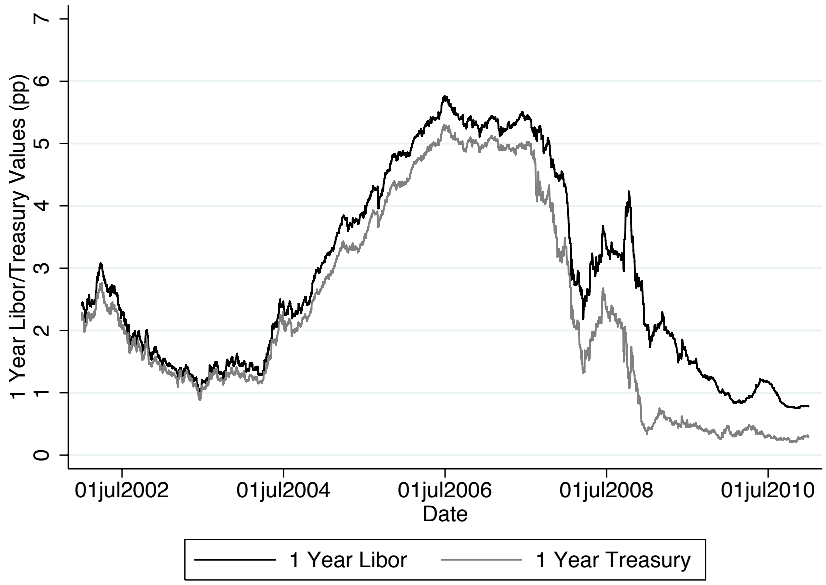
Figures

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

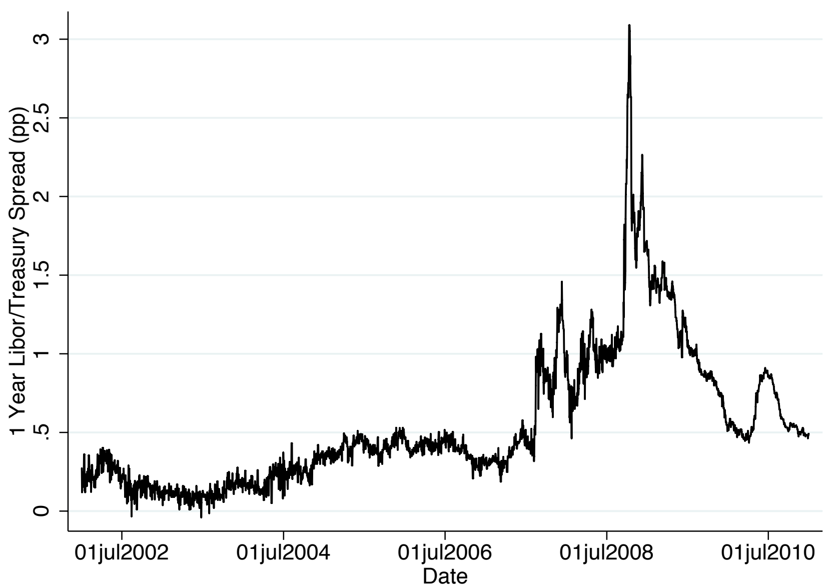


Hollow dots show the average proportion of loans defaulting within 60 months for each 1-point bin of origination combined loan-to-value. Size of dots is proportional to number of borrowers within each bin. Default is defined as 60 or more days past due. The solid line shows a local linear smoothing of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

FIGURE II
PANEL A: DIVERGENCE IN LIBOR AND TREASURY RATES

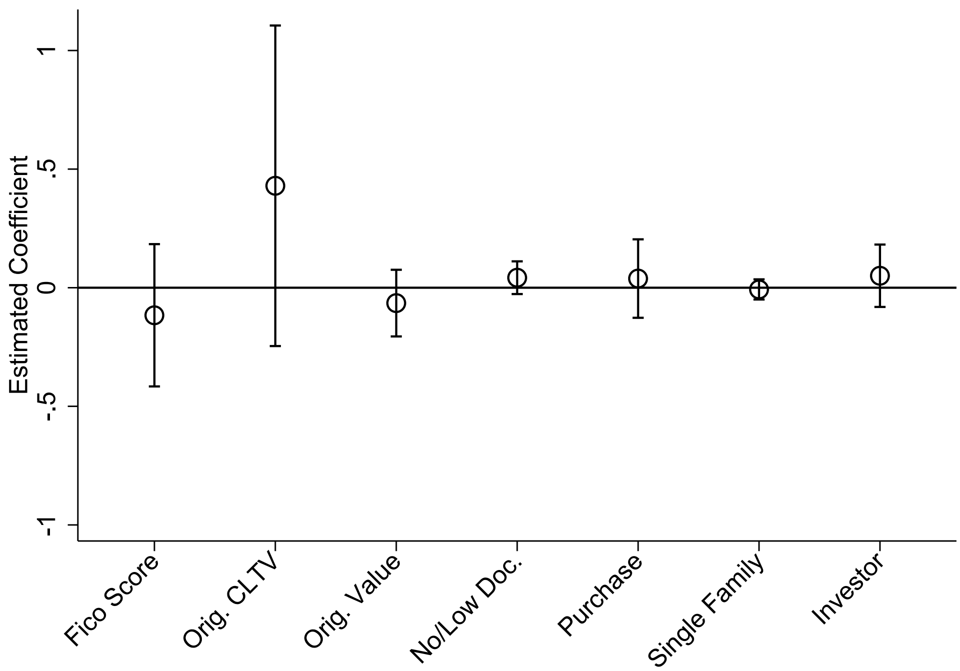


PANEL B: SPREAD BETWEEN LIBOR AND TREASURY INCREASED DRAMATICALLY DURING CRISIS



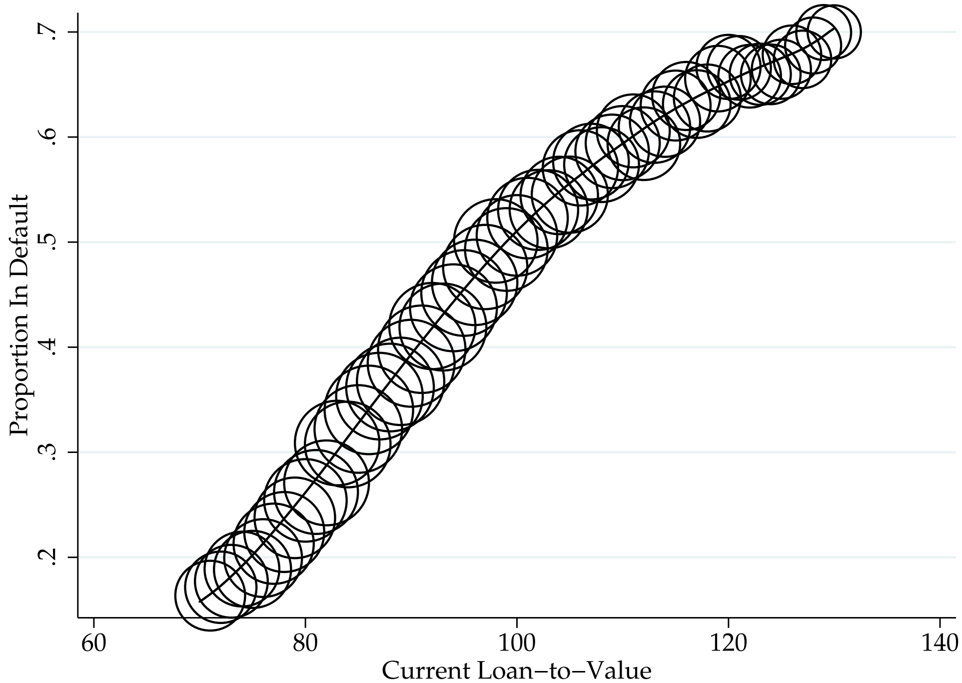
The top figure shows 1-year LIBOR and 1-year Constant Maturity Treasury (CMT) rates between 2002 and 2010. The bottom figure plots the spread between the two rates.

FIGURE III
INSTRUMENT UNCORRELATED WITH OBSERVABLES AT ORIGINATION



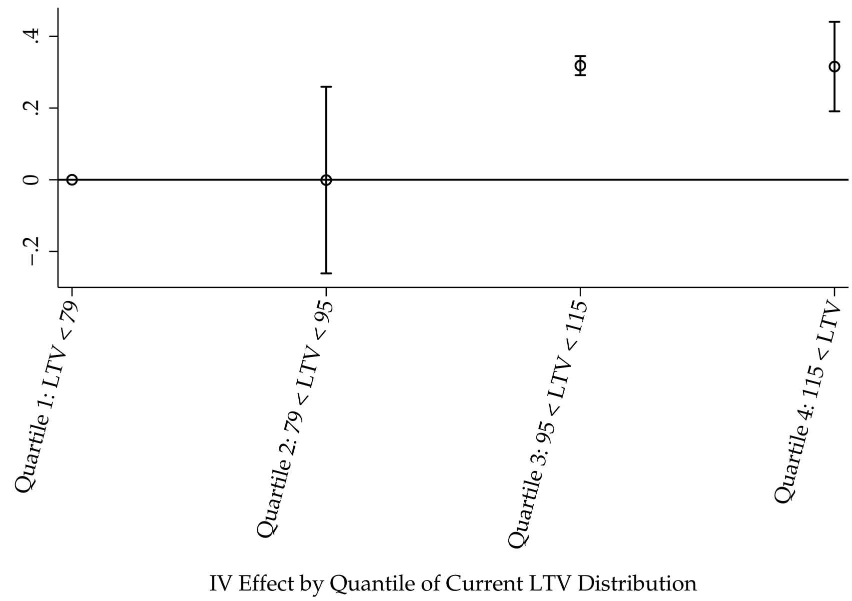
This figure shows coefficients from a regression of our leave-out-mean instrument Z_{it} on a range of borrower-level and loan-level covariates. We control throughout for our minimal set of controls: origination month, index type, and zipcode. We cluster standard errors at the originator level. The first three variables—FICO score, origination LTV, and origination home value—are also normalized by dividing by their standard deviations to provide interpretable coefficients. The last four variables are dummy variables indicating the presence of no or low documentation in loan origination, whether the mortgage is purchase instead of a refinancing, whether the property is single family, and whether the borrower was a reported investor. Across all categories, we find no statistically significant relationship between our instrument and borrower characteristics. Regression results are also displayed in Table II.

FIGURE IV
CORRELATION BETWEEN CURRENT LTV AND MORTGAGE DEFAULT

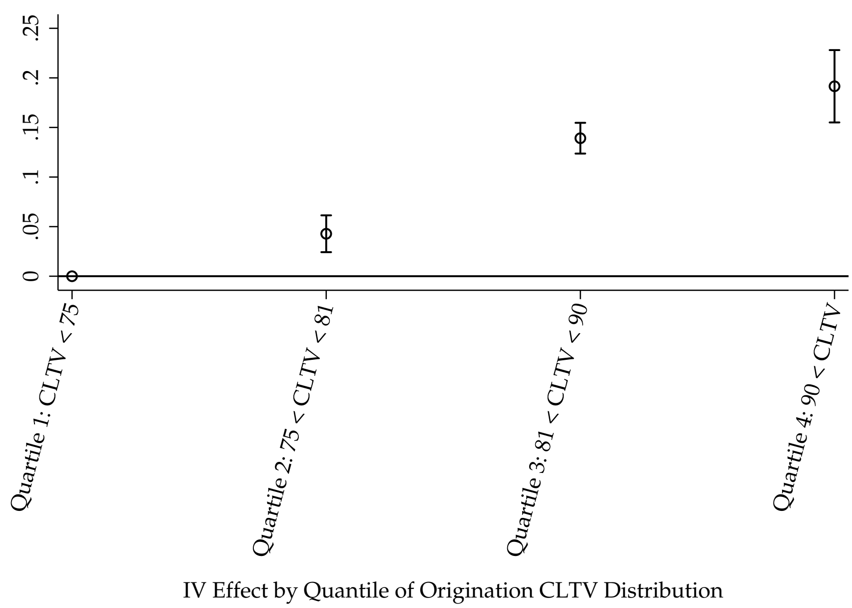


Hollow dots show the average proportion of loans defaulting between 24 and 36 months for each 1-point bin of current loan-to-value at 24 months. Size of dots is proportional to number of borrowers within each bin. Default is defined as 60 or more days past due. The solid line shows a local linear smoothing of the raw data.

FIGURE V
PANEL A: NON-LINEARITIES IN CAUSAL RELATIONSHIP BETWEEN LEVERAGE AND DEFAULT

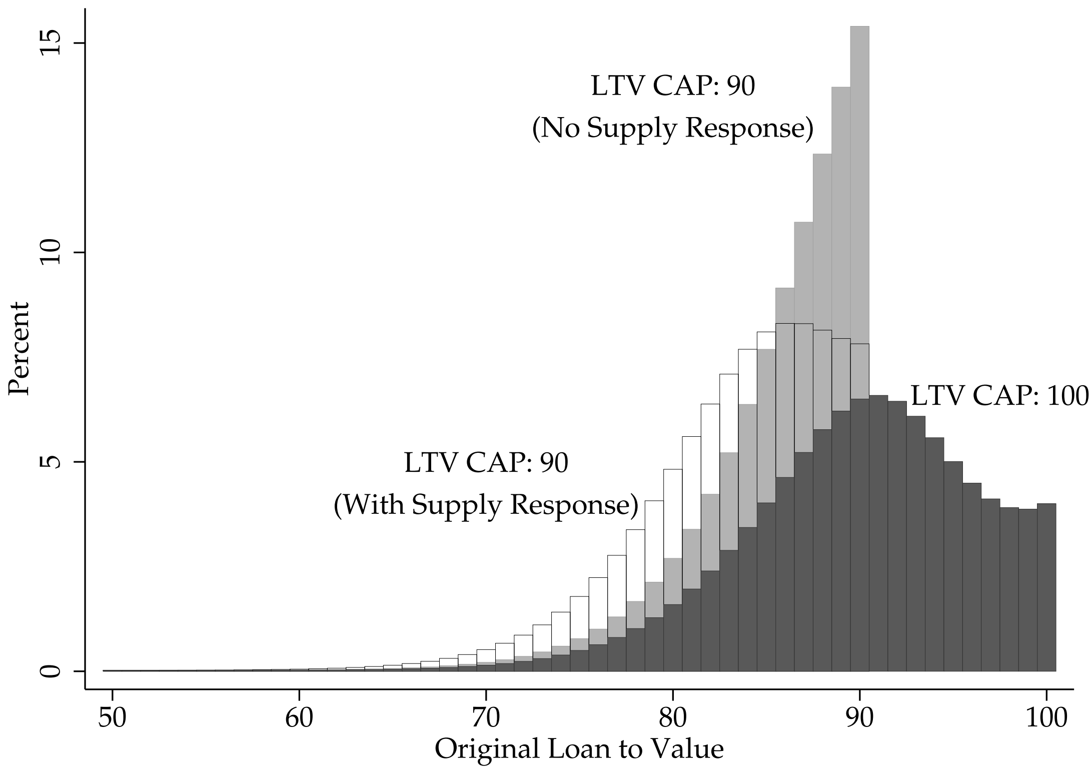


PANEL B: NON-LINEARITIES IN CORRELATION BETWEEN ORIGINATION LEVERAGE AND DEFAULT



Both panels plot coefficients from the regressions displayed in Column (9) of Appendix Table A.IV. The top panel shows the coefficients on each quartile of Current LTV, as instrumented by quartiles of our leave-out mean Z_{it}. The bottom panel shows coefficients on each quarter of origination leverage, measured as the origination combined LTV. For both the first quartile is normalized to 0. The sample includes all active loans at 24 months. Default is defined as falling 60 days past due between 24 and 36 months. Regressions control for index type, zipcode fixed effects, originator×origination month fixed effects, and our full set of credit and loan controls. These controls include the original home value (we include both a linear term and dummies for each $25k bin), dummies for each 20 point bin of FICO credit score at origination, and categorical dummies for occupancy, property type, loan purpose, and documentation level. Standard errors are clustered at the originator level.

FIGURE VI
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 the prices of 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 V shows appropriately calibrated results.

Tables

TABLE I
SUMMARY STATISTICS

	Treasury		LIBOR		COFI		Fixed	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Origination CLTV	0.80	0.098	0.82	0.12	0.80	0.097	0.81	0.094
Origination LTV	0.77	0.082	0.77	0.076	0.76	0.086	0.76	0.068
Current LTV (24 Months)	0.99	0.28	1.05	0.31	0.88	0.26	1.04	0.25
Origination Value	4.95	2.89	4.81	2.89	4.93	3.21	6.13	3.05
Fico Score	706.0	45.3	717.1	43.9	708.9	45.9	696.4	43.1
No/Low Documentation	0.82		0.85		0.79		0.92	
Loan for Purchase	0.32		0.40		0.40		0.26	
Single Family Home	0.66		0.65		0.66		0.78	
Investor	0.16		0.19		0.19		0.053	
State:								
- California	0.48		0.43		0.45		0.67	
- Florida	0.14		0.12		0.10		0.073	
- Arizona	0.042		0.048		0.040		0.022	
- Nevada	0.038		0.040		0.047		0.019	
Origination Year:								
- 2004	0.076		0.14		0.25		0.026	
- 2005	0.34		0.17		0.36		0.19	
- 2006	0.47		0.28		0.27		0.53	
- 2007	0.11		0.41		0.12		0.25	
Observations	492563		92648		11806		9975	

Summary statistics by index for Option ARMs in our sample with origination combined loan-to-value ratios between 50 and 100.

TABLE II
INSTRUMENT UNCORRELATED WITH OBSERVABLES AT ORIGINATION

Dep. Var.	Current LTV	Default	Orig. Fico	Orig. CLTV	Orig. Value	No/Low Doc.	Purchase	Single Family	Investor
Current LTV Leave-Out Mean	0.845^{***}	0.556^{***}	-27.667	0.224	-0.988	0.225	0.206	-0.038	0.269
(Index $\times$ Orig. Month)	(0.188)	(0.183)	(36.435)	(0.180)	(1.090)	(0.187)	(0.450)	(0.116)	(0.357)
Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Index Type FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zipcode FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.99	0.44	706.5	0.81	4.88	0.84	0.33	0.64	0.17
N	491,215	491,215	433,809	491,215	491,215	491,215	491,215	491,215	491,215

This table shows coefficients from a regression of our leave-out-mean instrument Z_{it} on a range of borrower-level and loan-level covariates and outcomes. We include all loans in our sample that are active at 24 months. We control throughout for our minimal set of controls: origination month, index type, and zipcode. The first two columns represent basic first stage and reduced form versions of our IV specification, respectively. In the first, current leverage, measured as current LTV on the Option ARM, is included as a dependent variable. In the second, default, defined as falling 60 days past due between 24 and 36 months, is included as a dependent variable. FICO score, origination LTV, and origination home value are all continuous, while the last four variables are dummy variables indicating the presence of no or low documentation in loan origination, whether the mortgage is for a home purchase instead of a refinancing, whether the property is single family, and whether the borrower was a reported investor. Mean of Dep. Var. refers to the mean of the dependent variable. We cluster standard errors at the originator level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE III
IMPACT OF ORIGINATION AND CURRENT LEVERAGE ON DEFAULT: LINEAR MODEL

Panel A: Estimates of Origination and Current Leverage on Default (60 DPD)

	Baseline			OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Origination Leverage	0.917***	1.062***	1.066***	0.666***	0.827***	0.818***	0.472**	0.701***	0.655***
	(0.064)	(0.067)	(0.073)	(0.084)	(0.081)	(0.075)	(0.183)	(0.198)	(0.127)
Current Leverage				0.301***	0.278***	0.285***	0.533**	0.427*	0.473***
				(0.048)	(0.038)	(0.028)	(0.225)	(0.247)	(0.171)
Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Index Type FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zipcode FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Credit/Loan Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Originator x Orig. Month FEs	No	No	Yes	No	No	Yes	No	No	Yes
Mean of Dep. Var.	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44
N	491,215	491,215	491,215	491,215	491,215	491,215	491,215	491,215	491,215

Panel B: First Stage Estimates of Current Leverage on Leave-Out Mean

	(7)	(8)	(9)
Current Leverage Leave-Out Mean (Index $\times$ Orig. Month)	0.658***	0.648***	0.457***
	(0.129)	(0.122)	(0.091)
Kleibergen-Paap F-Statistic	26.1	28.1	25.0

Panel A shows our main results, regressions of default between 24 and 36 months on origination and current leverage. We include a cross-section of all Option ARMs in our sample active at 24 months with an origination combined loan-to-value between 50 and 100. Default is defined as falling 60 days past due between 24 and 36 months. Origination leverage is defined as the combined loan-to-value ratio at origination. Current leverage is defined as the current loan-to-value ratio on the Option ARM at 24 months. All specifications control for origination month, index type, and zipcode fixed effects. Credit/loan controls refer to the original home value (we include both a linear term and dummies for each $25k bin), dummies for each 20 point bin of FICO credit score at origination, and categorical dummies for occupancy, property type, loan purpose, and documentation level. Originator$\times$orig. month fixed effects refer to fixed effects for each combination of originator and origination month. In columns labeled IV we instrument for current leverage with Z_{it}, the leave-out-mean of current leverage at 24 months for all loans with the same origination month and index type. Panel B shows coefficients and F-statistics from first stage regressions. Mean of Dep. Var. refers to the mean of the dependent variable. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE IV
IMPACT OF ORIGINAL AND CURRENT LEVERAGE ON DEFAULT: JOINT MODEL

	Panel A: Estimated Parameters from Leverage Choice and Default Equations								
	Baseline			No Instrument			Control Function		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ρ: Correlation of Errors	0.262***	0.274***	0.276***	0.189***	0.213***	0.213***	0.116**	0.162***	0.118**
in Default and Leverage Choice	(0.011)	(0.009)	(0.006)	(0.013)	(0.010)	(0.005)	(0.054)	(0.056)	(0.055)
Current Leverage				0.970***	0.978***	1.000***	1.854***	1.719**	2.327***
				(0.115)	(0.095)	(0.060)	(0.656)	(0.732)	(0.586)
Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Index Type FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Credit/Loan Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Originator x Orig. Year FEs	No	No	Yes	No	No	Yes	No	No	Yes
S.D. of Default Error				1.03	1.02	1.00	0.54	0.58	0.43
N	488,999	488,999	488,999	488,999	488,999	488,999	488,999	488,999	488,999
	Panel B: Excluded Instrument for Current Leverage								
Current Leverage Leave-Out							0.706***	0.679***	0.472***
Mean (Index × Orig. Month)							(0.124)	(0.121)	(0.123)

Panel A shows results from our joint model of leverage demand and the default choice. We include a cross-section of all Option ARMs in our sample active at 24 months with an origination combined loan-to-value between 50 and 100. Default is defined as falling 60 days past due between 24 and 36 months. Origination leverage is defined as the combined loan-to-value ratio at origination. Current leverage is defined as the current loan-to-value ratio on the Option ARM at 24 months. All specifications control for origination month, index type, and MSA fixed effects. Credit/loan controls refer to the original home value (we include both a linear term and dummies for each $25k bin), dummies for each 20 point bin of FICO credit score at origination, and categorical dummies for occupancy, property type, loan purpose, and documentation level. Originator×orig. year fixed effects refer to fixed effect for each combination of originator and origination year. In columns labeled Baseline, we estimate the leverage choice and default equation without include current leverage in the default equation. Columns labeled No Instrument include current leverage in the default equation. In columns labeled Control Function we specify an additional linear equation in which current leverage is a function of Z_{it}, the leave-out-mean of current leverage at 24 months for all loans with the same origination month and index type. ρ displays the estimated correlation between the errors in the leverage and default equations, capturing adverse selection. S.D. of the Default Error shows the reciprocal of the coefficient on current leverage. Panel B shows coefficients on the excluded instrument. Mean of Dep. Var. refers to the mean of the dependent variable. Standard errors are clustered at the originator level. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

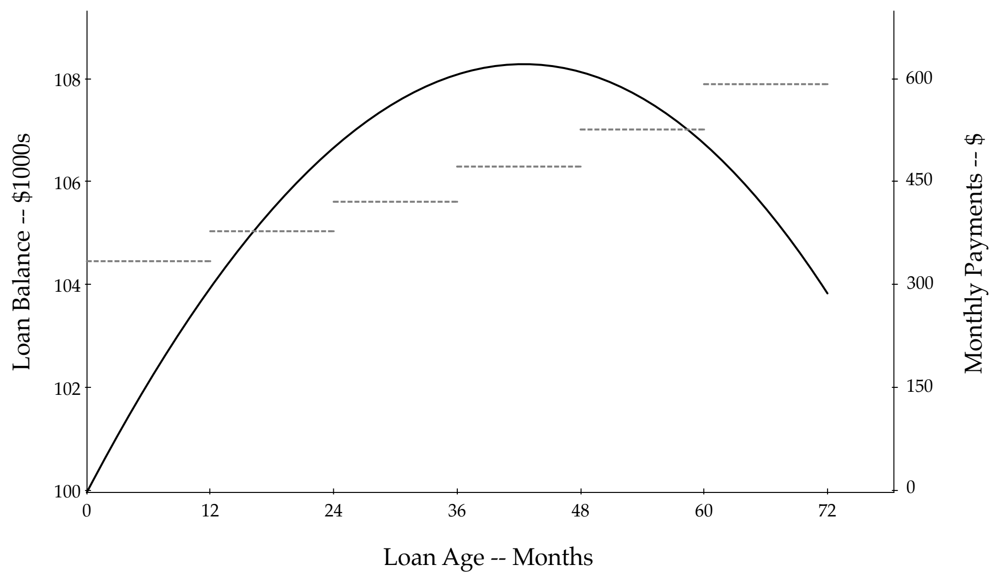
Table V
Simulation Results: The Impact of a Reduction in the LTV Cap From 100 to 90

	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 Size (Thousands)	\$387.5	\$361.4	\$361.3
Average Interest Rate	14.82%	11.38%	11.83%
Average Balance (Thousands)	\$445.0	\$402.5	\$404.0
Defaults	18.31%	15.09%	15.19%
Naive Defaults	-	14.24%	-
		Parameters	
	Initial Price: $H_0 = \$495.0$	Final Value: $H \sim N(\mu_1 = 594, \sigma_1 = 297)$	Proportion Behavioral: 1%
	CARA Coefficient: a=0.005	Viscosity: $\gamma = 1.075$	$\beta = \frac{1}{1+r_f} = 0.975$
	Prop. Recovered in Default: $\delta = 0.80$	$C_i \sim N(0.38\mu_1, 0.43\mu_1)$	Borrowers: $N = 1000$

Simulations from structural model described in Section 6. CARA utility assumed, 1000 simulated borrowers, with 1% behavioral. Viscosity set to match estimated $\rho = 0.118$. Parameters if C_i distribution taken from estimation in Table IV. The mean based upon the linear projection using average value of continuous covariates and the omitted category of all categorical covariates. Initial home prices set based upon average for the largest index category (Treasury). 20 percent net expected net growth over the term of the loan assumed. 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.

A Internet Appendix: Tables and Figures

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



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

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

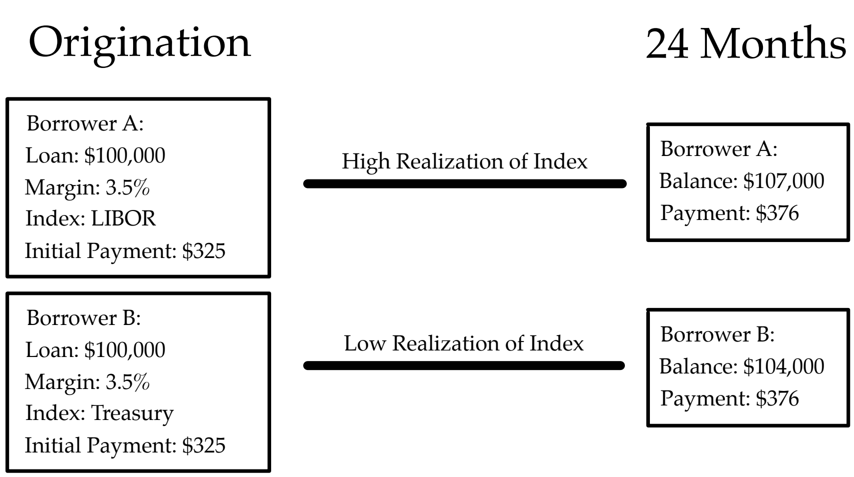
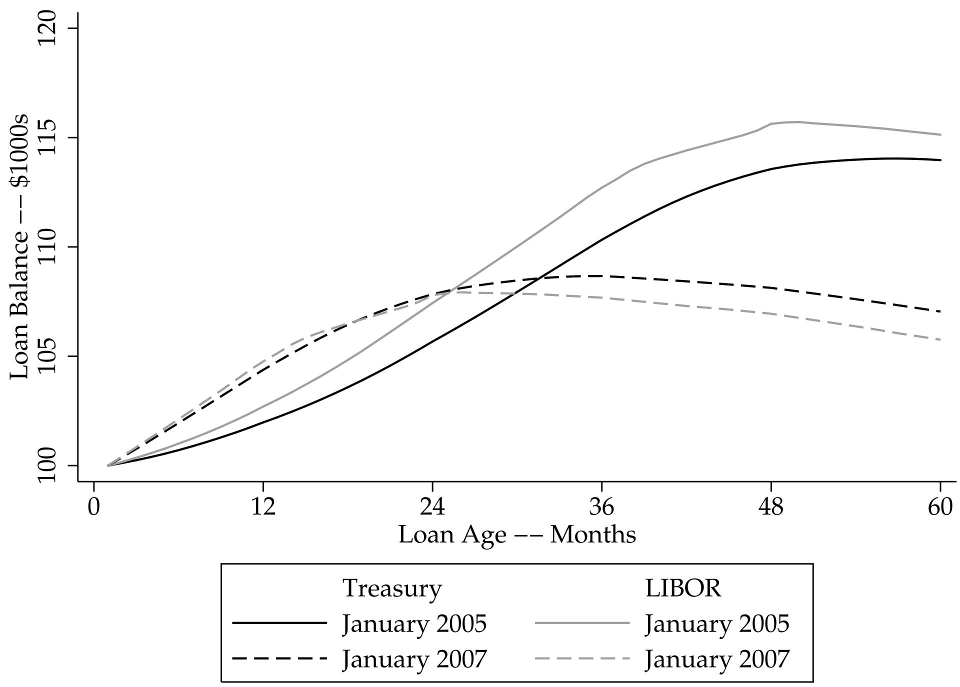


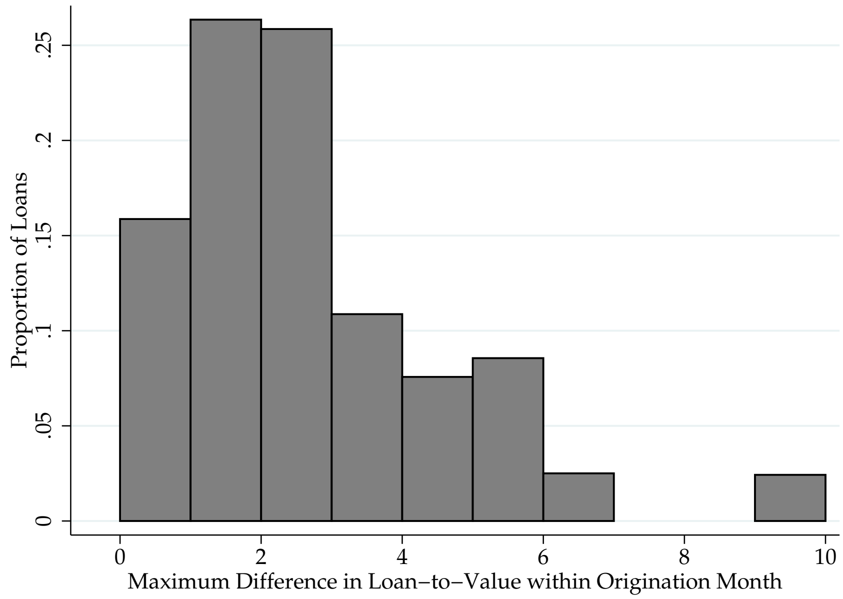
FIGURE A.III
VARIATION IN BALANCE TRAJECTORIES FOR OPTION ARMS



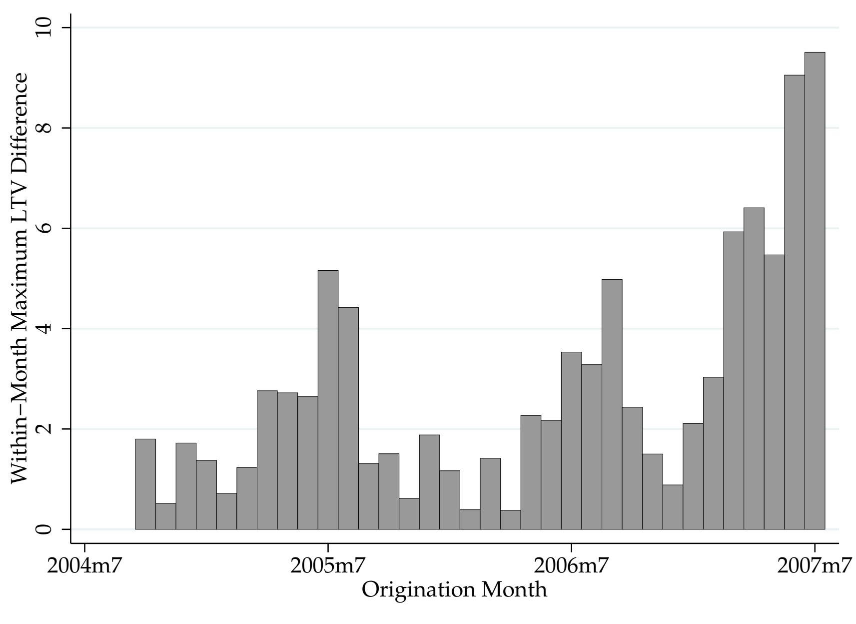
Simulated balance trajectories for $100,000 LIBOR- and Treasury-indexed loans originated in January 2005 or January 2007. Trajectories assume margin of 3.5 percent and initial payment based on 1.75 percent teaser. Treasury refers to 1-year MTA, LIBOR refers to 3-month duration.

FIGURE A.IV

PANEL A: HISTOGRAM OF WITHIN-ORIGINATION MONTH DIFFERENCES IN LTV ACROSS LIBOR AND TREASURY INDEXES

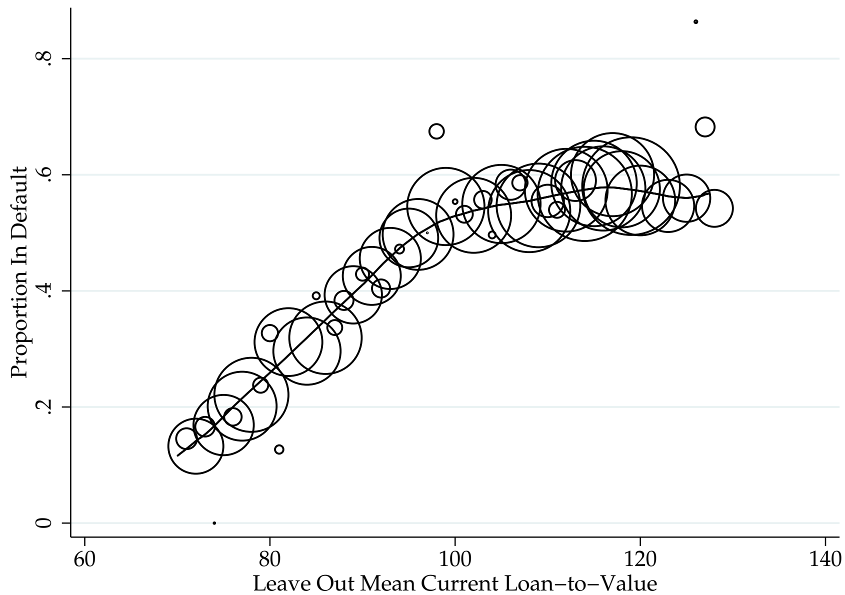


PANEL B: TIME-SERIES VARIATION IN WITHIN-MONTH DIFFERENCES IN LTV ACROSS LIBOR AND TREASURY INDEXES



For each borrower we compute Z_{it}, the leave-out-mean of current leverage at 24 months for all loans with the same origination month and index type. We then consider the maximum difference int this instrument for borrowers linked to LIBOR vs. borrowers linked to Treasury in each origination month. Given the leave out strategy, this is similar to taking the average difference between the two groups. Panel A shows a histogram of the difference. Panel B shows the difference across origination month. Magnitudes can be interpreted in LTV points.

FIGURE A.V
CORRELATION BETWEEN LEAVE-OUT-MEAN INSTRUMENT AND MORTGAGE DEFAULT



Hollow dots show the average proportion of loans defaulting between 24 and 36 months for each 1-point bin of the leave-out-mean instrument at 24 months. 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.

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

Originator	Percent of Loans Indexed to Treasury
Countrywide	78.5%
American Home Mortgage	99.1%
Greenpoint	66.2%
IndyMac	92.1%
EMC Mortgage	32.1%
Bear Stearns	33.1%
Residential Funding	93.4%
Downey	100%
Bank United	100%
Mortgage IT Inc	61.1%

Percentage of Treasury-indexed loans for the top 10 originators in the sample.

TABLE A.II
ACCOUNTING FOR ATTRITION: IMPACT OF ORIGINAL AND CURRENT LEVERAGE ON CUMULATIVE DEFAULT

	Panel A: Estimates of Original and Current Leverage on Default (60 DPD)								
	Default by 36 Months			Default by 48 Months			Default by 60 Months		
	Baseline	OLS	IV	Baseline	OLS	IV	Baseline	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Origination Leverage	1.041***	0.748***	0.810***	1.130***	0.917***	0.934***	1.125***	0.913***	0.989***
	(0.109)	(0.110)	(0.130)	(0.078)	(0.077)	(0.064)	(0.059)	(0.056)	(0.055)
Current Leverage		0.342***	0.270***		0.207***	0.191***		0.192***	0.123***
		(0.032)	(0.098)		(0.017)	(0.050)		(0.011)	(0.046)
Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Index Type FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zipcode FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Credit/Loan Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Originator x Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.35	0.35	0.35	0.45	0.45	0.45	0.50	0.50	0.50
N	606992	606992	606992	606992	606992	606992	606992	606992	606992
	Panel B: First Stage Estimates of Current Leverage on Leave-Out Mean								
Current Leverage Leave-Out Mean (Index × Orig. Month)			0.536***			0.507***			0.628***
			(0.112)			(0.132)			(0.065)
Kleibergen-Paap F-Statistic			22.9			14.7			93.9

This table repeats the analysis in Table III using the full sample of loans and considering the outcome of cumulative default. We define cumulative default as falling 60 days past due at any point up to 36, 48, or 60 months, depending on the specification. We include a cross-section of all Option ARMs with an origination combined loan-to-value between 50 and 100. Origination leverage is defined as the combined loan-to-value ratio at origination. Current leverage is defined as the imputed current loan-to-value ratio at 24, 36, or 48 months (In columns (1)-(3), (4)-(6), and (7)-(9), respectively) given the borrowers loan terms if they had made exactly the minimum payment every month. This variable may be defined at any point for any borrower, regardless of whether or not he or she has exited the sample. All specifications control for origination month, index type, and zipcode fixed effects. Credit/loan controls refer to the original home value (we include both a linear term and dummies for each $25k bin), dummies for each 20 point bin of FICO credit score at origination, and categorical dummies for occupancy, property type, loan purpose, and documentation level. Originator×orig. month fixed effects refer to fixed effects for each combination of originator and origination month. In columns labeled IV we instrument for current leverage with Z_{it}, the leave-out-mean of imputed current leverage at the relevant point (24, 36 or 48 months) for all loans with the same origination month and index type. Panel B shows coefficients and F-statistics from first stage regressions. Mean of Dep. Var. refers to the mean of the dependent variable. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE A.III
IMPACT OF ORIGINAL AND CURRENT LEVERAGE ON DEFAULT: FURTHER ROBUSTNESS

Panel A: Estimates of Original and Current Leverage on Default (60 DPD)

	Minimal Controls			Geography × Time FEs			Flexible Current Leverage		
	Baseline	OLS	IV	Baseline	OLS	IV	Baseline	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Origination Leverage	0.904***	0.477***	0.471**	1.082***	0.794***	0.646***	1.066***	0.756***	0.519*
	(0.074)	(0.058)	(0.211)	(0.076)	(0.057)	(0.182)	(0.073)	(0.069)	(0.291)
Current Leverage		0.495***	0.503**		0.330***	0.499**		−1.345***	61.254**
		(0.023)	(0.245)		(0.012)	(0.235)		(0.271)	(26.955)
Current Leverage^2								2.199***	−77.448**
								(0.378)	(34.477)
Current Leverage^3								−1.093***	40.304**
								(0.179)	(18.021)
Current Leverage^4								0.172***	−7.300**
								(0.025)	(3.259)
Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Index Type FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zipcode FEs	No	No	No	No	No	No	Yes	Yes	Yes
MSA x Orig. Month FEs	No	No	No	Yes	Yes	Yes	No	No	No
Credit/Loan Controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Originator x Orig. Month FEs	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44
N	491,215	491,215	491,215	488,595	488,595	488,595	491,215	491,215	491,215

Panel B: First Stage Estimates of Current Leverage on Leave-Out Mean

	(3)	(6)
Current Leverage Leave-Out Mean (Index × Orig. Month)	0.673***	0.362***
	(0.117)	(0.083)
Kleibergen-Paap F-Statistic	32.9	19.2

This table shows robustness for the results in Table III. The columns labeled Minimal Controls include only origination month and index type fixed effects. The columns labeled Geography×Time FEs replace zipcode fixed effects with MSA×origination month fixed effects. The columns labeled Flexible Current Leverage include a fourth order polynomial in current leverage. We include a cross-section of all Option ARMs in our sample active at 24 months with an origination combined loan-to-value between 50 and 100. Default is defined as falling 60 days past due between 24 and 36 months. Origination leverage is defined as the combined loan-to-value ratio at origination. Current leverage is defined as the current loan-to-value ratio on the Option ARM at 24 months. Credit/Loan controls refer to the original home value (we include both a linear term and dummies for each $25k bin), dummies for each 20 point bin of FICO credit score at origination, and categorical dummies for occupancy, property type, loan purpose, and documentation level. Originator×orig. month fixed effects refer to fixed effect for each combination of originator and origination month. MSA×orig. month fixed effects refer to fixed effects for each combination of MSA and origination month. In columns labeled IV we instrument for current leverage with Z_{it}, the leave-out-mean of current leverage at 24 months for all loans with the same origination month and index type. In Column (9) we include a fourth order polynomial of the instrument in the first stage. Panel B shows coefficients and F-statistics from first stage regressions. Mean of Dep. Var. refers to the mean of the dependent variable. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

Table A.IV
Effects Across Distribution of Leverage
Separating Out Quartiles of Origination and Current Leverage

	Baseline			OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Quartile 2: 75 $<$ Origination Leverage $\leq$ 81	0.089***	0.115***	0.102***	0.036**	0.067***	0.054***	0.029**	0.059***	0.043***
	(0.017)	(0.017)	(0.010)	(0.017)	(0.017)	(0.008)	(0.012)	(0.010)	(0.009)
Quartile 3: 81 $<$ Origination Leverage $\leq$ 90	0.195***	0.216***	0.215***	0.131***	0.158***	0.156***	0.120***	0.146***	0.139***
	(0.013)	(0.011)	(0.010)	(0.014)	(0.011)	(0.009)	(0.010)	(0.010)	(0.008)
Quartile 4: 90 $<$ Origination Leverage	0.224***	0.293***	0.294***	0.144***	0.218***	0.218***	0.129***	0.200***	0.192***
	(0.021)	(0.021)	(0.023)	(0.025)	(0.024)	(0.025)	(0.025)	(0.024)	(0.019)
Quartile 2: 79 $<$ Current Leverage $\leq$ 95				0.107***	0.100***	0.100***	0.059	0.016	-0.001
				(0.005)	(0.005)	(0.005)	(0.168)	(0.140)	(0.133)
Quartile 3: 95 $<$Current Leverage $\leq$115				0.214***	0.203***	0.204***	0.296***	0.298***	0.318***
				(0.017)	(0.016)	(0.014)	(0.021)	(0.019)	(0.014)
Quartile 4: 115 $<$ Current Leverage				0.294***	0.278***	0.281***	0.312***	0.295***	0.316***
				(0.021)	(0.019)	(0.014)	(0.072)	(0.056)	(0.064)
Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Index Type FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ZIP FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Credit/Loan Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Originator x Orig. Month FEs	No	No	Yes	No	No	Yes	No	No	Yes
Mean of Dep. Var.	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44
N	491,215	491,215	491,215	491,215	491,215	491,215	491,215	491,215	491,215

Regressions of default between 24 and 36 months on origination and current leverage. We include a cross-section of all Option ARMs in our sample active at 24 months with an origination combined loan-to-value between 50 and 100. Default is defined as falling 60 days past due between 24 and 36 months. Origination leverage is included as quartiles of the combined loan-to-value ratio at origination. Current leverage is included as quartiles of the current loan-to-value ratio on the Option ARM at 24 months. All specifications control for origination month, index type, and zipcode fixed effects. Credit/Loan controls refer to the original home value (we include both a linear term and dummies for each $25k bin), dummies for each 20 point bin of FICO credit score at origination, and categorical dummies for occupancy, property type, loan purpose, and documentation level. Originator$\times$Orig. Month fixed effects refer to fixed effects for each combination of originator and origination month. In columns labeled IV we instrument for current leverage with quartiles of Z_{it}, the leave-out-mean of current leverage at 24 months for all loans with the same origination month and index type. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE A.V
IMPACT OF ORIGINAL AND CURRENT LEVERAGE ON DEFAULT: FURTHER ROBUSTNESS

	Panel A: Estimates of Original and Current Leverage on Default (60/90 DPD)								
	No Junior Lien			LIBOR and Treasury Only			90 Days Past Due		
	Baseline	OLS	IV	Baseline	OLS	IV	Baseline	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Origination Leverage	1.103***	0.674***	0.535**	1.069***	0.818***	0.505***	1.073***	0.822***	0.647***
	(0.096)	(0.116)	(0.268)	(0.073)	(0.078)	(0.156)	(0.079)	(0.082)	(0.138)
Current Leverage		0.331***	0.439*		0.286***	0.644***		0.289***	0.489***
		(0.039)	(0.241)		(0.031)	(0.190)		(0.029)	(0.183)
Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Index Type FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zipcode FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Credit/Loan Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Originator x Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.40	0.40	0.40	0.44	0.44	0.44	0.39	0.39	0.39
N	346,919	346,919	346,919	473,179	473,179	473,179	491,215	491,215	491,215

	Panel B: First Stage Estimates of Current Leverage on Leave-Out Mean								
Current Leverage Leave-Out Mean (Index × Orig. Month)			0.357***			0.556***			0.457***
			(0.034)			(0.123)			(0.091)
Kleibergen-Paap F-Statistic			108.6			20.5			25.0

This table shows robustness for the results in Table III. The columns labeled No Junior Lien subset to loans for which we have no observable junior liens, and so for which the origination combined loan-to-value and origination loan-to-value on the option ARM are the same. The columns labeled LIBOR and Treasury Only subset to loans indexed to either LIBOR or Treasury. The columns labeled 90 Days Past Due define default as falling 90 days past due between 24 and 36 months. Outside of the restrictions specified above, we include a cross-section of all Option ARMs in our sample active at 24 months with an origination combined loan-to-value between 50 and 100. In the first six columns, default is defined as falling 60 days past due between 24 and 36 months. Origination leverage is defined as the combined loan-to-value ratio at origination. Current leverage is defined as the current loan-to-value ratio on the Option ARM at 24 months. All specifications control for origination month, index type, and zipcode fixed effects. Credit/loan controls refer to the original home value (we include both a linear term and dummies for each $25k bin), dummies for each 20 point bin of FICO credit score at origination, and categorical dummies for occupancy, property type, loan purpose, and documentation level. Originator×orig. month fixed effects refer to fixed effects for each combination of originator and origination month. In columns labeled IV we instrument for current leverage with Z_{it}, the leave-out-mean of current leverage at 24 months for all loans with the same origination month and index type. Panel B shows coefficients and F-statistics from first stage regressions. Mean of Dep. Var. refers to the mean of the dependent variable. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE A.VI
IMPACT OF ORIGINAL AND CURRENT LEVERAGE ON DEFAULT: LOANS CURRENT AT 24 MONTHS

	Panel A: Estimates of Original and Current Leverage on Default (60 DPD)								
	Baseline			OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Origination Leverage	0.664*** (0.039)	0.783*** (0.042)	0.781*** (0.045)	0.401*** (0.048)	0.530*** (0.046)	0.517*** (0.041)	0.274** (0.115)	0.424*** (0.121)	0.355*** (0.106)
Current Leverage				0.318*** (0.038)	0.299*** (0.034)	0.306*** (0.025)	0.470*** (0.152)	0.424*** (0.155)	0.494*** (0.128)
Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Index Type FEs	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Zipcode FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Credit/Loan Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Originator x Orig. Month FEs	No	No	Yes	No	No	Yes	No	No	Yes
Mean of Dep. Var.	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28
N	381,658	381,658	381,658	381,658	381,658	381,658	381,658	381,658	381,658
	Panel B: First Stage Estimates of Current Leverage on Leave-Out Mean								
Current Leverage Leave-Out Mean							0.779*** (0.132)	0.767*** (0.128)	0.525*** (0.103)
Kleibergen-Paap F-Statistic							35.0	35.8	26.2

This table repeats the analysis in Table III but includes only loans that are current at 24 months and have never previously been delinquent. Panel A shows regressions of default between 24 and 36 months on origination and current leverage. Default is defined as falling 60 days past due between 24 and 36 months. Origination leverage is defined as the combined loan-to-value ratio at origination. Current leverage is defined as the current loan-to-value ratio on the Option ARM at 24 months. All specifications control for origination month, index type, and zipcode fixed effects. Credit/Loan controls refer to the original home value (we include both a linear term and dummies for each $25k bin), dummies for each 20 point bin of FICO credit score at origination, and categorical dummies for occupancy, property type, loan purpose, and documentation level. Originator$\times$orig. month fixed effects refer to fixed effects for each combination of originator and origination month. In columns labeled IV we instrument for current leverage with Z_{it}, the leave-out-mean of current leverage at 24 months for all loans with the same origination month and index type. Panel B shows coefficients and F-statistics from first stage regressions. Mean of Dep. Var. refers to the mean of the dependent variable. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE A.VII
HETEROGENEITY BY INITIAL LOAN CHARACTERISTICS

	Panel A: Loan Documentation					
	No/Low Documentation			Full Documentation		
	Baseline	OLS	IV	Baseline	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Origination Leverage	1.120***	0.892***	0.806***	0.724***	0.393***	0.138
	(0.061)	(0.066)	(0.124)	(0.113)	(0.089)	(0.374)
Current Leverage		0.268***	0.370**		0.347***	0.615
		(0.030)	(0.162)		(0.019)	(0.447)
Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index Type FEs	Yes	Yes	Yes	Yes	Yes	Yes
Zipcode FEs	Yes	Yes	Yes	Yes	Yes	Yes
Credit/Loan Controls	Yes	Yes	Yes	Yes	Yes	Yes
Originator x Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.47	0.47	0.47	0.30	0.30	0.30
N	411,091	411,091	411,091	77,760	77,760	77,760

	Panel B: Loan Purpose					
	Purchase			All Others		
	Baseline	OLS	IV	Baseline	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Origination Leverage	0.965***	0.767***	0.609***	1.109***	0.850***	0.727***
	(0.056)	(0.050)	(0.122)	(0.068)	(0.074)	(0.100)
Current Leverage		0.331***	0.595***		0.265***	0.392***
		(0.029)	(0.194)		(0.028)	(0.129)
Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index Type FEs	Yes	Yes	Yes	Yes	Yes	Yes
Zipcode FEs	Yes	Yes	Yes	Yes	Yes	Yes
Credit/Loan Controls	Yes	Yes	Yes	Yes	Yes	Yes
Originator x Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.42	0.42	0.42	0.45	0.45	0.45
N	162,804	162,804	162,804	326,298	326,298	326,298

	Panel C: State Recourse Laws					
	Some Recourse			Non-Recourse		
	Baseline	OLS	IV	Baseline	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Origination Leverage	0.960***	0.645***	0.618***	1.133***	0.937***	0.884***
	(0.078)	(0.066)	(0.117)	(0.071)	(0.079)	(0.142)
Current Leverage		0.361***	0.392**		0.228***	0.289
		(0.029)	(0.156)		(0.027)	(0.186)
Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Index Type FEs	Yes	Yes	Yes	Yes	Yes	Yes
Zipcode FEs	Yes	Yes	Yes	Yes	Yes	Yes
Credit/Loan Controls	Yes	Yes	Yes	Yes	Yes	Yes
Originator x Orig. Month FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.42	0.42	0.42	0.45	0.45	0.45
N	211,188	211,188	211,188	279,902	279,902	279,902

This table shows additional heterogeneity analysis of Table III estimates. We show regressions of default between 24 and 36 months on origination and current leverage, split across three categories. Panel A shows results for loans with no/low documentation versus full documentation. Panel B shows results for loans for home purchases versus all other purposes. Panel C shows loans in some recourse vs. non-recourse states, based on the classification in as coded in Rao and Walsh (2009). We include a cross-section of all Option ARMs in our sample active at 24 months with an origination combined loan-to-value between 50 and 100. Default is defined as falling 60 days past due between 24 and 36 months. Origination leverage is defined as the combined loan-to-value ratio at origination. Current leverage is defined as the current loan-to-value ratio on the Option ARM at 24 months. All specifications control for origination month, index type, and zipcode fixed effects. Credit/loan controls refer to the original home value (we include both a linear term and dummies for each $25k bin), dummies for each 20 point bin of FICO credit score at origination, and categorical dummies for occupancy, property type, loan purpose, and documentation level. Originator×orig. month fixed effects refer to fixed effects for each combination of originator and origination month. In columns labeled IV we instrument for current leverage with Z_{it}, the leave-out-mean of current leverage at 24 months for all loans with the same origination month and index type. Panel B shows coefficients and F-statistics from first stage regressions. Mean of Dep. Var. refers to the mean of the dependent variable. * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

B Supplementary 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 a privately known cost associated with defaulting, C_i, which captures the difference in dollar terms between defaulting and not defaulting.

Default Choice

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

$$H_1 - B < -C_i.$$

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

[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 $\frac{L}{H_0}$.

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

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

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

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

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

Additionally, borrower utility is decreasing in the balance:

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

This 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]

At actuarially 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

[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.

[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.

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 borrowers with 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.