Mortgage Default: A Heterogeneous-Agent Model

Philip Lewis Kalikman*‡
Joelle Scally†‡

2023 June 7

Abstract

We introduce a loan-level model of mortgage default with heterogeneity in borrower characteristics and mortgage terms, including idiosyncratic penalties for default. Borrowers' penalties determine how closely their behavior hews to the predictions of the double-trigger or strategic models. The state space varies loan-to-loan based on all of the loan's, borrower's, property's, and neighborhood's idiosyncratic characteristics. We test the model on a high-performance computing cluster against real data drawn from linked databases with billions of observations of hundreds of simultaneous attributes. The model predicts defaults out-of-sample, fits cross-sectional characteristics of the distribution of mortgage performance, and classifies likelihood of default with high accuracy and better than all known benchmarks.

*University of Cambridge. Corresponding author. philip@kalikman.com (preferred), pk563@cam.ac.uk.
I am grateful to my advisors John Geanakoplos and Andrew Metrick, and my unofficial advisor Andrew Haughwout. I thank Archishman Chakraborty, Tess Scharlemann, and Walter Torous for thorough discussion. For helpful comments, I thank Andrew Ellul, Ravi Jagadeesan, Wilbert van der Klaauw, Donghoon Lee, Lira Mota Mertens, Nagpurnanand Prabhala, Gary Pivo, Albert Saiz, Matt Sekerke, Tray Spilke, Gokhan Torna, and seminar participants at the American Real Estate and Urban Economics Association, Cambridge University, the Federal Reserve Bank of New York, the Financial Management Association, Institut Minés-Telecom, Massachusetts Institute of Technology, the Office of Financial Research, Oxford University, Stony Brook University, Texas A&M University, the University of Arizona, the University of Hawaii, Yale University, the Yale School of Management, the Yale SOM Financial Crisis Conference, and Yeshiva University. I thank Frank Innocenti, Kevin Kelliher, Dmitry Krivitsky, Ernest Miller, and Na'im Tyson for assistance with the FRBNY computing cluster. I thank the FRBNY for generously providing the data, computing resources, and espresso necessary to complete this project.

†Federal Reserve Bank of New York.

1 Introduction

Mortgages are the largest and most consequential financial instruments with which most people will interact in their lifetimes. From 2008 to 2012, millions of borrowers stopped making payments on their mortgages, sparking the Global Financial Crisis. These people borrowed with diverse mortgage terms, faced dissimilar financial situations, and lived in regions all over the country with distinct local house price dynamics. Can a single model explain why *each* of these borrowers decided to default?

We introduce a structural model that explains defaults better than any previous benchmark of which we are aware. We test our model out-of-sample using data from multiple periods, including before, during, and after the financial crisis. We find that it not only better predicts the aggregate level of default than benchmarks, but also explains the cross-section of defaults: why different borrowers with different mortgage terms, different financial characteristics, and living in different neighborhoods default at different rates, and what those rates are.

The key innovation in our model is that each borrower's utility function embeds her own idiosyncratic penalty for default. In prior research, the default penalty has either been neglected or, when included, functioned as a fudge factor introduced to make aggregate default rates fit the data. Our approach differs in that we take the default penalty to be a faithful representation of each borrower's motivations. Prior research justifies this approach. Bhutta, Dokko, and Shan (2017) establish that borrowers deviate significantly from pure financial rationality. Guiso, Sapienza, and Zingales (2013) establishes the same, and shows further that borrowers exhibit heterogeneity in their willingness to endure financial hardship for the sake of avoiding default. Quantifying borrowers' willingness is the key to explaining their behavior.

Explicitly modeling idiosyncratic borrower default penalties is not only more faithful to borrowers' motivations. It also provides a potential resolution to a longstanding debate in the literature over borrowers' motivation for default. Do borrowers default because exogenous shocks reduce their ability to afford payments—the liquidity-constrained, or "double-trigger" model of default? Or do they default because they rationally calculate that continuing to pay a mortgage in deep negative equity is throwing good money after bad—the "strategic" model of default?[1]

[1]While the terms "double-trigger" default and "strategic" default, sometimes referred to as "can't-pay" default and "won't-pay" or "ruthless" default, are standard in the literature, they have acquired moral baggage that we wish to avoid. When we write that a borrower or her behavior is "strategic," or "more strategic," we mean only that she or it more closely resembles the prediction of a rational economic model driven by

Both models have problems fitting empirical data on defaults. The strategic model predicts far too many defaults: even in the depths of the crisis, when over one in four borrowers was underwater, the mortgage delinquency rate never exceeded 12%, and the foreclosure rate never exceeded 2.5%. But the double-trigger model does not explain why borrowers default well in advance of when they actually run out of money. Gerardi, Herkenhoff, et al. (2013) found that over 75% of 90%+ LTV borrowers who defaulted could afford their payments when they defaulted. And their measurement of payment affordability, as they acknowledge, also does not account for borrowers' ability to consume by taking on additional unsecured debt.[2]

Our model generalizes both the strategic and double-trigger models, embedded both as special cases and in so doing dissolving the debate between them. We argue that the most faithful characterization of borrower behavior is that it falls along a spectrum: every borrower acts according to a blend of double-trigger and strategic motives, but they differ in how strongly the different motives weigh on their decision. Liquidity constraints reduce the amount a borrower can consume in the present if she continues making payments, while low expectations of future house price appreciation reduce expected future consumption. Both effects reducing the borrower's willingness to make payments. But, crucially, borrowers vary in *how* willing they are to endure hardship to make payments; equivalently, borrowers vary in how strategic their behavior appears to the researcher to be. One borrower will default whenever it is financially optimal; another will endure a significant shortfall in utility before succumbing to the incentive to default.[3] In our paradigm, borrowers are all strategic about their mortgages, but their strategy and their level of liquidity interact. Neither the double-trigger model nor the strategic model adequately captures these interactions. We quantify them precisely, within a computational structural model.

In what we believe is a first in the literature, we introduce an idiosyncratic non-pecuniary

pure-financial motives. The "double-trigger" model—which refers to two circumstances "triggering" default, namely, negative equity and illiquidity—is sometimes also referred to as the liquidity-constrained model; some authors distinguish the terms to focus on positive-equity defaults.

[2] Accordingly, the double-trigger model also fails to explain why borrowers facing a temporary liquidity shock and a temporary home price shock do not simply take on unsecured debt and wait for either shock to abate.

[3] We have remarked that we wish to avoid the moral associations that follow labels of borrowers as "strategic" defaulters. Beyond our desire to avoid a moral debate, we also consider the labeling exercise to be limiting from the standpoint of economic precision. Borrowers do not come in two uniform homogenous populations, the double-trigger population and the strategic population. All borrowers deciding to default make the same financial choice using the same financial variables, but with different values of those variables as inputs. Categorizing the borrower into coarse types reduces the rich, high-dimensional information about the borrower's financial situation and expectations yet does not gain anything in predictive power. By contrast, retaining the rich information about each borrower enables studying how each borrower would behave in other circumstances. We discuss this further in Appendix A.10.

penalty for default, which governs the strength of illiquidity and strategy interacting in determining the borrower's default decision. Borrowers with higher penalties endure further deviations from the pure-strategic financial optimum before they default. Our model quantifies the size of this deviation borrower-by-borrower. We will argue that in recognizing borrowers' varying degrees of strategy and illiquidity, our model more faithfully describes reality.

If penalties could only be revealed ex post, the ex post characterization of defaults as determined by borrower motives and default penalties would be the extent of our contribution. But we also provide a procedure for estimating the relationship between borrowers' observable characteristics and their unobservable default penalties ex ante. Specifically, we show that higher-credit-score borrowers have higher penalties, and we show how to find the relationship between score and penalty.

Our estimation procedure is motivated by a straightforward observation. Lenders offer lower interest rates to borrowers with higher credit scores—even though the lenders have full knowledge of borrowers' income and employment, the characteristics of relevant collateral, and every other relevant financial variable—because even after taking these characteristics into account, higher-score borrowers are less likely to default. This phenomenon is widely known, yet has not been explained in a quantitative framework. Our model provides a structural, quantitative underpinning for the variation in the behavior of borrowers with different credit scores: higher-score borrowers default less because of, and to the degree determined by, their greater default penalties. As far as we are aware, ours is not only the first structural model to provide a borrower-level quantitative underpinning for borrowers' distinct degrees of reluctance to default, but also the first to estimate this characteristic using observable data.

Our work contributes to a long literature exploring borrowers' motivations for default. Foote and Willen (2018) characterize much of this literature. Until recently, academics were largely split on the importance of strategy versus illiquidity. More recently, more scholars are converging on illiquidity as a crucial driver of defaults. As we discuss further in Section 2 (and further in Section A.10), we believe that our characterization of borrower behavior as falling along a spectrum between the strategic and double-trigger extremes is more faithful to reality and offers a more illuminating characterization of the richly interacting motivations that borrowers weigh. We argue not only that our model is more parsimonious, in subsuming the double-trigger and strategic models into a more general framework, but also argue in Section 5 that our model better fits the empirical data.

In addition to contributing to the literature on default motivations, our work also contributes to a rich literature on structural modeling of mortgages. This literature traces its origins to Black and Scholes (1972), Kau, Keenan, Muller, et al. (1995), and related papers by these authors and their contemporaries. Campbell and Cocco (2015) introduced the modern benchmark structural model, which featured heterogeneity in many borrower characteristics, but not credit score or a penalty for default. Several authors advanced the literature in the direction of including non-pecuniary default penalties. Schelkle (2018) modeled a default-like penalty as a utility benefit for owning a home, but which thus introduces the same disincentive to sell a house as to default on it. In our model, borrowers may choose to build wealth by selling their homes without penalty. Laufer (2018) used a hybrid model featuring a default penalty to study equity extraction, without mortgage age as a state variable. In our model, borrowers pay down principal as mortgages age, changing their equity-driven incentives to default.

With these refinements, and because it also supports complete heterogeneity in all borrower characteristics, we believe it is also the first that features realistic enough predictions of distinct borrowers' behavior that a lender or policymaker can use it to predict how different borrowers will vary their default propensity in different financial situations and house price climates. In the companion paper Kalikman and Scally (2022), we use the model to study a heterogeneous mortgage modification policy that might have more effectively mitigated the wave of defaults throughout the Global Financial Crisis.

Throughout the rest of this paper, we argue that our modeling approach better fits the data on default than a model without default penalties, without idiosyncratic default penalties, or without idiosyncratic default penalties correlated to credit score. As far as we are aware, this is the first effort in the academic literature to estimate a structural model of default separately, loan-by-loan, for each loan in a heterogeneous, nationally representative sample with variation in loan terms, borrower characteristics, and regional price dynamics, and with idiosyncratic penalties for default.

2 Literature

This essay contributes to two strands of the literature in mortgage default: one exploring borrowers' default motivations empirically, and often in a classification exercise between strategic and double-trigger motivations, and another exploring structural models of default. With regard to the literature on motivations for default, we provide an underpinning for borrower motivations that we believe subsumes the double-trigger and strategic models in a more general explanation, in which all borrowers exhibit varying degrees of strategy or illiquidity. With regard to the literature on structural models, we provide what is, to our knowledge, the first fine-grained and fully heterogeneous model of default in the literature. The model extends what Foote and Willen (2018) term "hybrid" models, models which embed aspects of the double-trigger and strategic frameworks in a single lifecycle model.

Hybrid models are an improvement over the crude extremes of the double-trigger or strategic models, and there is strong and longstanding justification in the literature for a hybrid approach. In Elul et al. (2010)'s study of mortgage default, for example, both negative equity and illiquidity are associated with default, and the authors find interactions between negative equity and illiquidity: borrowers in deep negative equity are more likely to default given a liquidity shock than borrowers with moderate negative equity. Hybrid models offer a means to quantifying the magnitudes of such effects and their interactions.

But hybrid models are still limited by homogeneity when they assume that all different borrowers different behaviors are explainable by a single combination of strategic and double-trigger motivations. We move past this limitation in two primary directions.

First, we calibrate our model to a completely heterogeneous, nationally representative sample of loans with variation both in loan terms and borrower characteristics. Laufer (2018), for example, only calibrates his model to mortgages in one county; Schelkle (2018) considers only a single interest rate and loan-to-value ratio; and other studies likewise, implicitly or explicitly, make comparable representativeness assumptions as they consider one or a handful of cases of variables of interest. Computational resources are now available that eschew the need for such representativeness assumptions. Second, we explicitly model a quantitative determinant of the relative strengths of the strategic and double-trigger motivations in *each* borrower's lifecycle optimization problem. We represent that quantitative determinant as an idiosyncratic, non-pecuniary (util-denominated) penalty for default.

The literature provides strong evidence both that non-pecuniary default penalties exist and that they are heterogeneous. Bhutta, Dokko, and Shan (2017) study subprime negative-equity borrowers and find that they need to be far more underwater before they default than predicted by theory, arguing that borrowers weigh "emotional and behavioral factors," before defaulting—thus arguing for a non-pecuniary default penalty. White (2010) argues that the financial industry deliberately and successfully cultivates non-pecuniary default penalties in borrowers through a campaign of proactive moral suasion, including by misrepresenting the severity of the consequences that borrowers would face for defaulting. Guiso, Sapienza, and Zingales (2013) provide evidence not only that non-pecuniary penalties are common to many borrowers, and considerable in magnitude—77% of borrowers they surveyed reported that they would not default on a mortgage, provided they could afford their payments, even if the value of the loan exceeded the value of the home by $100,000—but also that they are heterogeneous. Of the 23% of borrowers who *would* default at the $100,000 shortfall, 61% of those would *not* default were the shortfall only $50,000.

We explicitly model an underpinning for the variation in tolerance to being underwater as an idiosyncratic penalty for default. As far as we are aware, ours is the first model both to make penalties heterogeneous and to estimate them to observable data rather than only calibrating them as an ex post correction factor. Beyond being more faithful to a heterogeneous underlying reality, modeling heterogeneous penalties may clear up some discrepancies in the literature. Schelkle (2018) and Laufer (2018) find widely varying estimates of the *size* of the borrower's default penalty, from 1.5% of permanent income at one extreme to 29% at the other.

It is possible that the discrepancy is an artifact of too-strong representativeness assumptions: Schelkle (2018) considered only a single interest rate and loan-to-value ratio; Laufer (2018) calibrated his model to mortgages in one county. Widely varying house price dynamics and mortgage terms not captured by these representativeness assumptions could account for different estimates of borrowers' default penalties. We run our model with heterogeneous data from thousands of distinct borrowers in distinct geographies with distinct loan terms, thus avoiding what we might term "the curse of representativeness."

It is also possible that the discrepancy among other authors' estimates of default penalty magnitude is a faithful artifact of the underlying reality that borrower penalties are heterogeneous. We adopt this viewpoint as our foundation, but we are not alone in doing so. Recogonition of heterogeneity is a bedrock principle in the mortgage finance industry. Mortgage lenders set rates by forecasting borrowers' propensity to default using observable characteristics, such as employment and income history, to get at unobservable characteristics, such as future willingness to make payments. The most prominent non-pecuniary observable characteristic every borrower has—and in fact must submit to lenders, by law—is a credit score. Lenders, in turn, treat credit score as revealing default propensity, even though lenders observe and can control for all other relevant borrower financial characteristics.

In what we believe is also a first in the literature on default, we use this relationship as the basis on which to estimate penalties. We treat the credit score as partially revealing a borrower's penalty, and estimate the relationship between credit score and penalty using historical data on defaults. As far as we are aware, this is the first model to feature idiosyncratic penalties for default, that explicitly treats the penalty as an idiosyncratic and estimable borrower characteristic, and that explicitly models the quantitative channel by which credit score, which is widely acknowledged to be an important predictor of default, influences the borrower's decision.

In the remainder of this essay, we first describe the mathematical structure of the model. We then describe the data and estimation procedures we use to identify penalties. Finally, we test the model in in-and-out-of-sample tests against the benchmark models to assess its performance.

3 Model

Our model is a finite-period backwards-recursive optimal option exercise model, in the vein of Black-Scholes-Merton and Campbell and Cocco (2015). Risk-averse, fully rational borrowers with CRRA utility for real consumption, constrained by income, savings, and mortgage payments, maximize utility on each node in a recombining tree of forecast house prices, labor shocks, and mortgage statuses. In each state in each period, the borrower decides not only whether to default, but also whether to prepay the mortgage, and if so, whether to keep or sell the home.

So that the model can investigate policies that are not one-size-fits-all but rather targeted to different borrowers, we also need to use different inputs to the model for each different borrower. We require of our model that it accurately reflect how each different borrower would behave under the particular realization of house price paths and employment circumstances that befell the borrower while also representing the borrower decision in light of any changes in mortgage terms that a policymaker might wish to offer. Thus each borrower in the model imagines a *different* tree of future possibilities with its own distinct house price dynamics, interest rate path, and interactions with her own present and future income, wealth, and mortgage terms.

Additionally, each borrower knows her own default penalty, i.e. the size of her psychological reluctance to default even when it is financially optimal to stop making payments. As we explained in the literature review, significant research establishes the existence of the default penalty; in our utility function it is the mechanism which causes borrowers to forgo financial gain on the order of a hundred thousand dollars. The literature also establishes that the amount borrowers would forgo varies borrower to borrower; we design our model to be faithful to this reality. In the model, in any state of the world when the borrower defaults, she pays a penalty in utils, distinct from the financial consequences of default such as having to search and move to a rental home with volatile rent and losing the upside in future home price appreciation. Different borrowers suffer different penalties, corresponding to their varying degrees of reluctance to default even when financially incentivized to do so. But we suppose the policymaker knows only the borrower's observable characteristics: her credit score and her history of payments. The policymaker estimates the borrower's default penalty from these observable variables, and constructs a forgiveness policy accordingly.

Heterogeneous default penalties enable our model to embed the existing structural models of mortgage default as particular cases. The double-trigger, strategic, and hybrid model are reproduced by particular parametrizations of our model assuming certain loan terms, property characteristics, income dynamics, and values for penalties. Heterogeneous penalties are also the crux of accurately reflecting the effects of targeted policy design. In particular we believe modeling heterogeneous penalties helps resolve a conundrum in modification policy: because different borrowers actually vary in how likely they are to continue payments without modification, a policy that pre-emptively reduces payments needs to account ex ante for the likelihood that it modifies a borrower who would have continued making payments without the need for a modification. Prior work studying modifications has been afflicted by the "curse of averages"—the need to extrapolate the behavior of the average borrower to the average of borrower behaviors. We sidestep the curse by directly modeling every distinct borrower's distinct decision using distinct data on that borrower's distinct financial circumstances *and* embedding that borrower's distinct degree of deviation from the purely financial-maximizing behavior.

Our model identifies the unnecessarily modified borrowers as, all else equal, those with the highest penalties, as those borrowers would have been least likely to default even without modification. By the same token, our model reproduces distributional features of borrower mortgage behavior micro-founded at the individual borrower level, rather than by appeal to coarse buckets or groups that would not permit finer grains of modification.

We feed into the model data from McDash, Equifax, and CoreLogic, linked at the loan, borrower, and property level. The linked datasets enable seeing the nearly complete portrait of a borrower's financial conditions that the model requires to generate its realistic distinct state space for each borrower, with that borrower's particular mortgage terms, financial situation, property characteristics, and credit score. While these large-scale, high-resolution data have been more widely available in recent years, as far as we are aware ours is the first paper to calibrate a structural loan-level model of default to a broad collection of complete borrower financial circumstances, rather than to representative samples with heterogeneity only in limited dimensions. We also believe this work to be the first to embed idiosyncratic default penalties estimated to borrowers' credit scores, providing a structural, micro-level underpinning for the aggregate relationship between credit scores and borrowers' payment behavior.

In Kalikman and Scally (2022), we use the model to study modification policy in the Global Financial Crisis, detailing how the model's heterogeneity enables discovering Pareto-improving policy that would be opaque to coarser models. In the remaining sections of this essay, we argue that the model better fits the data on default than benchmark models, and further that it predicts defaults out-of-sample precisely enough both in aggregate and across the distribution of borrowers to justify taking seriously the policy conclusions that can be drawn from its predictions of borrower behavior. In the remainder of this section, we describe the mathematical structure of the model.

3.1 Time

Time is discrete; $t = 0$ in the reference period, $1, \ldots, T$ in the loan's maturity before or T' after modification. The reference period $t = 0$ varies in historical time based on when the borrower is eligible for a modification.

3.2 Mortgages

Mortgages mature in T remaining periods. They have remaining principal M_0 in period 0, scheduled remaining principal M_t in period t, coupon payments $\{m_t\}_{t=1\ldots,T}$, and interest rates $\{r_t^m\}_{t=0,\ldots,T-1}$. For fixed-rate loans, the constant per-period mortgage rate is $r_t^m = r^m$, and the mortgage coupon payment $m_t = m$ therefore satisfies the usual full amortization schedule

$$M_0 = \frac{m}{r^m}\left(1 - \frac{1}{(1+r^M)^T}\right) \qquad\qquad M_t = (1+r^m)M_{t-1} - m.$$

Our model differs from some others in featuring exact mortgage amortization schedules rather than an approximation via a steady-state mortgage with a geometric approximation to amortization. It supports loans of arbitrary product structure, including adjustable-rate mortgages, interest-only mortgages, and mortgages with balloon payments. HAMP modifications with principal forbearance, which featured in the Global Financial Crisis, yield mortgages with large balloon payments.

3.3 Utility

Borrowers have constant relative risk aversion (CRRA) utility over real consumption C_t^s for each feasible state s in each time t, with relative risk aversion denoted by γ and interperiod discount rates, which may vary, denoted by $\{\beta_\tau\}_{\tau=0,\ldots,T-1}$, so that the gross period-0 discount factor in period t is $B_t = \prod_{\tau=0}^{t} \beta_\tau$:

$$U^i\left(\{C_t^s(i)\}_{t\in T}^{s\in S(t)}\right) = E_0 \sum_{t=0}^{T} B_t \left(\frac{(C_t^s(i))^{1-\gamma(i)}}{1-\gamma(i)} - \mathbf{1}_{\text{i defaults in } s}\lambda(i)\right).$$

The expectation is taken in period 0 over all realizable states s. The penalty λ^i or $\lambda(i)$ directly reduces the borrower's utility in and only in the state and period when she defaults. It is not, as in comparable models, a multiplier on utility in states when the borrower owns the home, as this would introduce the same disincentive to sell the house free-and-clear as to default on the mortgage. As is standard, the amount of housing or housing services consumed is fixed regardless of whether the borrower is an owner or renter, so we suppress a separate housing term from the utility function. In sensitivity tests, we allow for a subsistence level floor to consumption and for moving into subsistence-level housing.

3.4 House Prices

House prices are uncertain, exogenous, and specific to each distinct ZIP code of each borrower in the data. They follow an approximate geometric Brownian motion on a recombining trinomial tree. Trinomial trees support a discrete space of house prices and transition probabilities that model a wide range of possible series in expected mean and variance of house price changes while also enabling paths of prices which "jump" without requiring borrowers to recalculate revised expectations. Borrowers in the baseline model have time-revising expectations which are short-term trend-following and long-term mean-reverting. We also consider specifications in which they have constant expectations or certain-but-not-constant expectations of a particular series of price changes and price volatilities. We provide further mathematical details in Appendix A.1

By contrast with many other models of default, we do *not* assume a representative house price process. Consider Figure 1. We show the overall path for house prices for the United States, as well as for each of the twenty cities composing the S&P/Case-Shiller House Price Index. We highlight the three metropolitan areas with the largest peak-to-trough decline. The variation in price paths across the twenty metro areas, including both overall rise and fall, pace of recovery, and size of peak-to-trough decline, is striking. Rather than paper over these differences for computational tractability, we consider it of paramount importance to model how different borrowers in these different regions would respond to the distinct dynamics in their own areas. Borrowers in our model experience considerably different paths of prices, form different expectations, and thus face different default decisions.

[Figure 1 about here.]

3.5 Labor Income

Labor income follows a two-state Markov process reflecting "employment" or "unemployment," where unemployment is understood to encompass any shock to liquidity—not only job loss, but also unexpected expenses from medical costs or divorce. The borrower's income in the employed state is L_t^h, a multiple of her initial income L_0^h determined by her age and multipliers estimated from BLS data. In the unemployed state, the borrower receives replacement income L_t^l as a fraction τ_{unemp} of normal income:

$$L = \begin{cases} L_t^h, & \text{borrower is "employed"} \\ L_t^l = \tau_{\text{unemp}} \cdot L_t^h, & \text{borrower is "not employed"} \end{cases}$$

The replacement fraction may be thought of as unemployment insurance income or as the borrower's income net of the shock.

Initial labor income L_0^h is determined by the borrower's stated debt-to-income ratio. The expected growth in labor income $\{L_t^h\}_{t=0,\dots,T}$ is determined by the borrower's age. As is standard, we assume a constant probability of unemployment π_u, and a constant probability of re-employment π_e. Labor income uncertainty is thus entirely determined by the Markov process: there is no uncertainty in income conditional on the realization of the employment outcome, and present labor shocks do not affect future realizations.

3.6 States and Uncertainty

States are vectors encompassing the exogenous state variables house prices P_t^s and labor income L_t^s as well as the endogenous control variables liquid assets A_t^s, remaining mortgage terms $\vec{M}_t^s$, whether the borrower still owns or has sold the home H_t^s, the years left, if any, the borrower will live rent-free after defaulting and before the lender completes foreclosure proceedings D_t^s, and the years, if any, since the borrower accepted a modification with an embedded lender option / equity share agreement Q_t^s:

$$s =< t; A_t^s, \vec{M}_t^s, H_t^s, D_t^s, Q_t^s; L_t^s, P_t^s > .$$

The model features standard assumptions on irreversibility or absorbing states; we describe these further in Appendix A.2 A typical log-scale discretization of borrower wealth yields an overall state space on the order of 300,000 nodes for one borrower with one mortgage.

3.7 Savings, Borrowing, Unsecured Debt

The borrower chooses a consumption plan $\{C_t^s\}_{t\in T}^{s\in S(t)}$, specifying the consumption level at any feasible state s in any time period t, to maximize U subject to a budget constraint on liquid assets, or cash-on-hand, A_t^s.

The borrower invests cash savings at the risk-free interest rate r_t^f. In addition to their mortgages, borrowers may borrow in unsecured debt at a (higher) borrowing rate r_t^b. Savings and borrowing rates may vary over time.

Unsecured borrowing is rare in structural models of mortgage default because it increases the computational complexity of the state space and borrower optimization, but it is an important and realistic aspect of borrower behavior. In the Equifax CRIS data, 92% percent of borrowers have observed unsecured debt when they default on their mortgages; 55% of the population have debt in excess of $15,000. We do not require the borrower to stay above an arbitrary lower bound in unsecured debt, but we also do not allow borrowers to default on unsecured debt.

The borrower thus begins state s' in the next period with liquid assets $A_{t+1}^{s'}$ equal to

$$A_{t+1}^{s'} = \begin{cases} (1 + r_t^f)a_t^s, & a_t^s \geq 0 \\ (1 + r_t^b)a_t^s, & a_t^s < 0, \end{cases}$$

where a_t^s is the end-of-period cash-on-hand as defined in the budget constraint below.

3.8 Budget Constraint

$$a_t^s = \qquad \text{End of period cash-on-hand} \tag{1}$$

$$A_t^s \qquad \text{Starting assets} \tag{2}$$

$$+ L_t^s \cdot (1 - \tau_{\text{inc}}) \qquad \text{Labor income less income tax} \tag{3}$$

$$- 1_{\text{Has Mortgage in } s} \cdot m_t \qquad \text{Mortgage coupon payment} \tag{4}$$

$$+ 1_{\text{Has Mortgage in } s} \cdot \tau_{\text{inc}} r_{t-1}^m M_{t-1} \qquad \text{Mortgage interest deduction} \tag{5}$$

$$- 1_{\text{Owns Home in } s} \cdot \left((\tau_{\text{maint}} + \tau_{\text{prop}}) P_t^s \right) \qquad \text{Maintenance and property taxes} \tag{6}$$

$$+ 1_{\text{Owns Home in } s} \cdot (\tau_{\text{inc}} \tau_{\text{prop}} P_t^s) \qquad \text{Property tax deduction} \tag{7}$$

$$- 1_{\text{Rents Home in } s} \cdot 1_{\text{Did not recently default}} \cdot (\tau_{\text{rent}} P_t^s) \qquad \text{Rent} \tag{8}$$

$$- I_t C_t^s \qquad \text{Inflated price of consumption} \tag{9}$$

$$- 1_{\text{Prepays Mortgage in } s} \cdot (1 + \tau_{\text{prepay}}) \cdot M_t \qquad \text{Principal plus prepayment fees} \tag{10}$$

$$+ 1_{\text{Sells Home in } s} \cdot (1 - \tau_{\text{move}} - \tau_{\text{sell}}) \cdot P_t^s \qquad \text{Price less moving costs, broker's fees} \tag{11}$$

$$- 1_{\text{Sells Home in } s} \cdot 1_{t<Q^t} \cdot Q^p \cdot \max(P_t^s - Q_f^t, 0) \qquad \text{(Equity share mod)} \tag{12}$$

$$- 1_{\text{Defaults in } s} \cdot \tau_{move} \cdot P_t^s \qquad \text{Moving costs} \tag{13}$$

$$+ 1_{\text{Defaults in } s} \cdot \max(\tau_{distressed} \cdot P_t^s - M_t, 0) \qquad \text{Distressed sale price less principal} \tag{14}$$

$$- 1_{\text{Defaults in } s} \cdot 1_{\text{Recourse}} \cdot (M_t - \tau_d P_t^s) \qquad \text{Principal less distressed sale price} \tag{15}$$

The agent will conclude the period with a level of liquid assets (1), that she will invest until the next period if positive, or borrow if negative, determined as follows. She begins the period with a given level of liquid assets (2), and receives labor income net of taxes (3). If she still has a mortgage, she pays the coupon (4), but receives a deduction from income tax for her mortgage interest (5). Whether she has the mortgage or not, provided she owns the home, she pays maintenance and property taxes (6), but deducts property tax from income tax (7). If instead she rents her home, she pays rent unless she recently defaulted, in which case she lives rent-free for a fixed period (typically two years) (8). These costs are all determined by her initial state. Subsequently, she makes several choices: first, a level of consumption (9), and then decisions regarding her mortgage. If she prepays the mortgage, she pays the outstanding balance to the lender, plus a fee proportional to that balance (10). If she prepays or has previously prepaid, she may sell the home. She receives the home price less proportional broker's fees and moving costs (11). (If she sells the home after receiving a modification with an equity share agreement, (12), she may have to pay the lender its share of equity; modeling such agreements enables studying policy recommendations made in the wake of the Global Financial Crisis, but does not feature in this essay.) Finally, if she defaults, she must pay the same moving costs she would after selling (13). The home is sold in a distressed sale. If the distressed sale price covers her outstanding principal, the lender receives the principal and she receives the difference (14); if not, and if the lender has recourse, the borrower pays the shortfall out of liquid assets (15).

3.9 Rent

We take rent to be proportional to price up to a ceiling and floor determined at period 0. Rent costs are typically modeled as one of two extremes: either as a fixed fraction of the initial house price, or as a fixed percentage of the contemporaneous house price. Each of these choices embeds a realistic feature and an unrealistic feature. When rent is a constant fraction of the contemporaneous house price, rent costs may balloon enormously. But such models capture that ownership serves as a hedge against volatility in the costs of renting; this hedge also serves as a disincentive to default. When rent is a fixed fraction of the initial house price, borrowers do not have to worry about unaffordable rent in extreme states. Furthermore, in reality, rent is imperfectly correlated with house prices. Many properties may be either owned or rented, but differences in supply, regulation, credit, and transaction costs break perfect substitutability between the two markets. Therefore, models with a fixed cost of renting more accurately capture the imperfect correlation between prices in the rental and ownership markets, but they also inaccurately eliminate the value of the ownership hedge against rent volatility. In the baseline, we take an approximate middle ground between these approaches by allowing rent to vary with prices, but only up to a maximum or minimum multiple of the initial rent.

4 Data And Methodology

4.1 Data Sources

Microdata were generously provided by the Federal Reserve Bank of New York and consist of Black Knight Financial Services' McDash Data, Equifax Credit Risk Insight Servicing (CRIS) Data, and CoreLogic Home Price Indices linked at the loan, property, or ZIP level. Black Knight Financial Services provides the industry-standard mortgage servicing dataset known typically as "the McDash Data" (hereinafter). These are loan origination and performance data as reported monthly by loan servicers. Coverage in McDash is estimated at approximately 80% of first liens originated in the United States, with detailed loan performance histories since 2005. McDash is considered a representative sample of first liens and includes not only prime but also subprime and Alt-A loans in the crisis era. The data cover both origination characteristics and performance characteristics. We use McDash origination data for original mortgage size, purchase house price, mortgage interest rate, mortgage product structure, mortgage purpose and occupancy—we restrict to purchases of owner-occupied homes—and borrower characteristics front-end payment-to-income ratio and credit score.

We use McDash performance data for contemporaneous interest rate, payment history, history of principal outstanding, whether the loan has received a modification, and modified terms of modified loans.

Because McDash data do not provide adequate second-lien coverage, we use the Equifax CRIS data to identify combined loan-to-value ratios. The CRIS database consists of monthly observations of credit characteristics by borrower and tradeline type. Coverage is virtually 100% of the universe of borrowers as virtually all formal lenders report borrowers' payments to Equifax and the other major credit bureaus. The database reports several credit scores each month for each borrower. The results herein are not sensitive to the choice of credit score.

Data are broken down by month, borrower, and tradeline type, and include number of accounts, total balance, total balance past due, and total credit limit or high credit. Because we observe separate tradelines for first and second mortgages, we can compute combined loan-to-value ratios. Equifax CRIS also observes credit card debt, which we use as a measure of borrowers' unsecured borrowing.

We use CoreLogic Home Price Indices, joined as of the contemporaneous date and at the zip code level to the McDash-reported property zip code, to estimate both the path of house prices for each property and the parameters of each borrower's expectations for her house price process. We take the savings rate as the 1-year treasury rate and use average credit-card APRs from the Federal Reserve Economic Data portal for borrowing rates. We use data from the 2007-9 Survey of Consumer Finances to estimate borrower's asset levels, and income data from the Bureau of Labor Statistics to estimate the expected path of labor income over the lifecycle.

4.2 Methodology

We employ standard methodology for defining mortgage-related variables of interest and estimating quantities not directly observable. We define default and modification using our total view into payment history, which enables us to disentangle temporary payment lapses from long-term defaults. We estimate borrower income from payments and the reported front-end payment-to-income ratios. We estimate liquid assets using SCF data regressed against borrower income, mortgage debt, and non-mortgage debt. We estimate combined loan-to-value ratios by scaling McDash first-lien loan-to-value by total mortgage debt reported in Equifax CRIS. And we estimate baseline borrower expectations of house prices and volatility of house prices and the related parameters of the house price trinomial tree from the history of ZIP-level house price indices. We describe the methodology in greater detail in Appendix A.

4.3 Sample Selection

Our initial sample of mortgages are those active as of January 1, 2007 and with combined loan-to-value ratio of 80% or higher. We restrict to mortgages used to purchase a home that the owner would occupy, with price at least \$20,000 and no more than \$1,000,000. We further restrict to 30-year, fixed-rate, non-interest-only mortgages. We restrict to conventional mortgages, i.e. excluding those backed by Federal Housing Authority or Department of Veterans' Affairs insurance. We require the borrower's credit score and front-end payment-to-income ratio (PTI) to be observed at origination. Finally, we restrict to borrowers who have a single first mortgage. This restriction excludes borrowers who are also financing vacation or investment properties. The financial incentives facing such borrowers are different, and they are unlikely to be the focus for government relief policies.

These sampling restrictions select for the population of mortgages used in most comparable studies, but they are a relatively safe pool versus mortgages in the crisis era. As Li and Goodman (2014) and others discuss, mortgage origination in the crisis era concentrated in riskier product structures, such as mortgages with low teaser rates that made initial payments small. Such mortgages appear affordable to borrowers until rates reset to a floating index after the teaser period. Similarly, interest-only loans feature low, affordable-feeling payments but without amortization, meaning that borrowers do not develop equity in the property and are therefore more likely to experience negative equity and thus to default.

We describe further sampling restrictions in Appendix A.3. For the most part, these and our other selection criteria are adverse to our hypothesis: selecting borrowers who choose the least risky mortgages selects for those borrowers who are also likely to have the highest personal reluctance to default, i.e., default penalties. Our selection thus likely leads us to underestimate the overall effectiveness we estimate for principal forgiveness. The one exception is selecting against ARM loans, which may cause us to underestimate the role of liquidity constraints, as interest rate resets to higher levels would exacerbate borrowers' liquidity shocks. But by the same token, selecting for fully-amortizing loans rather than including interest-only loans likely causes us to underestimate the equity effects of borrowers with partially-, non-, or negatively-amortizing loans, who would have less equity and thus an even greater strategic default incentive. Incorporating such mortgages would be a fruitful direction for further research, although origination in such exotic product structures is rare today and may well remain uncommon as long as lenders remain wary of the risks Li and Goodman (2014) characterize as stemming from those product structures. Table 1 summarizes standard statistics on this population.

[Table 1 about here.]

4.4 Estimation of Combined Loan-to-Value Ratio

The McDash data do not have property-level identifiers and have low second-lien coverage. They thus do not provide a reliable source of information regarding the *combined* loan-to-value ratio on a mortgage, the total loan-to-value ratio after accounting for junior liens. Observing the correct combined LTV (CLTV) is crucial to investigating default behavior. A borrower who owes 90% of the value of a home on her first mortgage could sell the home if she found payments unaffordable—unless she had a second mortgage on which she owed 20% of the value of the home. Using only first-lien LTV to infer positive equity would therefore confuse the interpretation of the decision by such a borrower to default. The effect of missing junior liens is just as detrimental when the borrower is already in negative equity even just accounting for the first lien: the borrower's own estimate of the probability of house price changes restoring her to positive equity may be substantially lower than what one would infer from the data without accounting for the junior liens she had.

Accordingly, following the method also used by Crews Cutts and Merrill (2008), we adjust for combined LTV on loans using the Equifax CRIS data. We have already restricted to borrowers who have only one first mortgage, and we further restrict to those who have only one or zero closed-end second (CES) mortgages for the nine months following the reference period, and one or zero Home Equity Lines of Credit (HELOCs) at the same time. This restriction increases the likelihood that any CES and HELOC loans in the borrower's credit report are on the same property as the first mortgage, though it is not logically impossible that a borrower could have a junior lien on a property other than that backing her senior lien. We further restrict to those borrowers for whom the first mortgage balance reported in Equifax CRIS differs by not more than 5% from that reported in McDash. (Because of timing mismatches, the two sources may report different outstanding principal balances on what is actually the same mortgage.) For these borrowers, we add the total outstanding balance on the CES and HELOC loans as observed in CRIS to the first mortgage balance observed in CRIS in order to compute the CLTV as a multiple of the first-lien LTV on the property. We restrict to properties for which this multiple is less than 1.8, as junior liens are typically smaller in volume than senior liens, so multiples higher than 1.8 are likely due to data inaccuracies or unusual circumstances. We then compute total principal outstanding on the property as the CRIS-derived multiple times the McDash-reported first-lien principal outstanding (first-lien LTV). We use the McDash principal rather than the Equifax CRIS principal because the rest of the loan performance characteristics, including the interest rate, loan status and therefore inferred date of default, and so on, are reported in McDash but not Equifax. Accordingly, the modeled borrower makes her default decision based on the timing of data reported in McDash.

Unlike first liens, second liens do not have origination or performance characteristics reported in McDash. As a result, while we observe the total outstanding balance on such liens via the Equifax CRIS data, we do not know the interest rate or product structure for these second liens. As is consistent with the literature, we approximate the amortization schedule of these loans by assuming they are capitalized into the first mortgage balance and amortize at the same rate and term, rather than treating them separately. Accordingly, the borrower who defaults in the data is assumed to default on all mortgages backing a property simultaneously; they are treated as one large mortgage in the model.

4.5 Estimation of House Price Expectations and Realizations

Each borrower forms her own expectation of the mean and volatility of house price surprises. Her expectations are independent of other borrowers' expectations and are uncorrelated with labor income surprises and other exogenous factors. Her expectations may be formed according to several specifications. In the simplest, they are fixed to the historical averages in her ZIP code from 1975 until 2002, and do not interact with the path of prices taken in computing scenario-specific expected default rates, lender recoveries, or other statistics of modification effectiveness. In this specification, a borrower who lives in a ZIP code with high home price appreciation through 2002, who originates a mortgage in 2007, and then experiences a drastic decline in home prices, will continue to expect high prices. We also study a specification in which borrowers revise expectations after price surprises (positive or negative).

The expected default rate (and lender recoveries) are computed conditional on the realization of prices for the given ZIP code from the reference period up until the last period of interest.

4.6 Estimation of Penalties

As we described in the Introduction, prior literature establishes that non-pecuniary penalties for default exist, are significant in magnitude, and vary significantly across individuals. But if they are unobservable, how can we feed them into our model? We take advantage of two features of the data and our model. First, we know that borrower credit scores, which are fully observable to lenders and policymakers, correlate to their reluctance to default. That correlation underlies why lenders charge higher interest rates to borrowers with higher credit scores, even though the lenders know the borrowers' income, assets, employment, and so on. Second, we know the history of borrowers' actual continuation and default decisions. Given their financial incentives, that behavior is only consistent with certain sizes of penalties.

We combine these two observations into an estimation procedure that we believe sits at the intersection of simplicity, accuracy, and usefulness to policymakers. In particular, we select a method that would have been implementable just after the crisis. Our estimation procedure has two steps. First, we assume a linear relationship between borrowers' observable credit scores and their penalties. Any parametrization of that relationship maps each borrower to her penalty via her credit score. Second, we estimate the parameters of that relationship by running the model with the implied penalties, selecting the parameters that maximize the fit of borrowers' actual default rates to the default rates implied by the model when it uses the parameter-implied penalties.

Mathematically, we begin with a pool of borrowers $\mathscr{L} = 1, \ldots, i, \ldots, N$ of loans active and current on payments at some time t. Letting $\lambda(i) = \lambda^i$ be borrower i's penalty, we assume

$$\lambda^i = \alpha + \beta \cdot (\text{Credit Score})^i,$$

where we observe $(\text{Credit Score})^i$ for each borrower.[4] Then, for any particular choice of α, β, we define two quantities: the empirical cumulative default rate in $\mathscr{L}$

$$CDR^T_{\mathscr{L}} = \frac{\#\text{ of loans in } \mathscr{L} \text{ that defaulted by T}}{\#\text{ of loans in } \mathscr{L}},$$

[4] While the equation relating credit score and penalty suggests estimation by linear regression, this is not our approach and in fact is not possible as we do not know the entries of the vector $\{\lambda^i\}$ ex ante. In a different direction, we also recognize that more variables besides credit score may correlate to penalties. We consider this and related possibilities in Appendix A.9.

and the model-implied cumulative default rate, which is indirectly a function of α and β as they determine each model input λ^i:

$$\widehat{CDR_{\mathscr{L}}^T(\alpha,\beta)} = \frac{\#\text{ of loans in } \mathscr{L} \text{ model predicts default by T given } \{\lambda^i\}_{i=1,\ldots,N}}{\#\text{ of loans in } \mathscr{L}}.$$

(In the definitions above, both cumulative default rates condition also on the observed series of house prices between t and T, suppressed in the notation.) Finally, we estimate $\hat{\alpha}$ and $\hat{\beta}$ to minimize

$$\left| CDR_{\mathscr{L}}^T - \widehat{CDR_{\mathscr{L}}^T(\alpha,\beta)} \right|.$$

4.7 Simulation

To compute default probabilities and expected lender receipts for a pool of loans $\mathscr{L}$, we first solve the model by backwards recursion for each loan in the pool. We then draw $25,000$ forward paths, i.e. realizations of house price and labor income draws across time, for *each* loan. Typically, we condition on a given starting path of house prices, so that only labor income is uncertain along the beginning of each path. Conditional on a loan being active in a given period t, the borrower's state in $t+1$ is determined by the borrower's decision regarding the mortgage and regarding assets in t: a borrower who chooses to continue paying the mortgage will have the mortgage in $t+1$ with an asset level determined by the savings chosen at t according to the budget constraint in Section 3.8. Note that in forward paths, prices need not follow edges along the price tree. In this way we are able to model "surprises," whereby the price series conditioned upon in computing a statistic such as default probability or lender receipts is not anticipated by the borrower in the reference period. We thereby compute for each loan its expected probability of default by a given period in the future as well as the expected dollars paid to the lender over the life of the loan as the sum of these figures (in the case of default probability, 1 for defaults and 0 otherwise) period-by-period divided by the number of paths. We then compute pool-level figures by summing over the individual figures computed for each loan in the pool. Loans in a pool are each subject to a given house price path determined by the relevant experimental specification, but there is no interaction between loans in the pool. That is, there is no reason that a given loan defaulting on a given path should make it more likely that another randomly selected loan on another randomly selected path will default. In reality, surprises that result in one mortgage's default likely are at least partially correlated with others that would lead to other defaults. Omitting this correlation is therefore likely to bias estimates of cost-effectiveness of a mitigation program downward.

4.8 Summary of Model Parametrization

The full set of model parameters, data sources, and baseline values are summarized in Table 2. The bulk of parameters, those listed below without a numerical value, are input into the model as different values for each borrower i, depending on data in the indicated sources according to the inference methodology described in the preceding sections. We assume fixed values primarily for transaction costs and for variables other than those that pertain to mortgages and house prices. In the baseline specification, We take the borrower's interperiod discount rate β to be constant at 0.985. This value is consistent with other authors' selected values, as well as with risk-free rates of approximately 1.5%, approximately those that prevailed throughout the period of interest. We set the coefficient of relative risk aversion to 3.5 again for consistency with other authors' parameter selection.

[Table 2 about here.]

5 Results

We test our procedure by estimating parameters on one random sample of loans but then selecting a new, entirely distinct sample and verifying that the model's out-of-sample performance does not degrade. We first take a randomly selected sample of loans active as of January 2007 and with combined loan-to-value ratio of 80% or higher. We select loans with high LTVs because they are more likely to experience negative equity and therefore to default. Characteristics of this sample are presented in the middle column of Table 1, with those of the defaulting borrowers presented in the rightmost column. Certain salient features of these statistics are that the characteristics of defaulting borrowers tilt consistently in the expected directions: they have larger loans, a higher percentage of second liens, higher second-lien balances, higher LTVs, higher payment-to-income ratios, lower wealth, lower income, and lower credit scores. Origination vintage skews heavily towards more recent loans (2005 and 2006 vintage) because these loans are those less likely to have amortized below 80% LTV.

We then draw a new sample of loans with the same sampling restrictions, and run the model again using the estimated coefficients to determine each borrower's penalty. We compare the predicted default rate generated by the model to the default rate of these distinct loans in actuality. The in-sample and out-of-sample fit are displayed in Figure 2.

[Figure 2 about here.]

We find very strong out-of-sample performance: the model fits the level of defaults out-of-sample even though none of those loans were used to calibrate the model. In this graph, the model calibration targets matching the year-5 cumulative default rate. Thus the shape of each curve, as well as the terminal values at year 7, are all out-of-sample estimates that we take to validate the calibration's performance. The model not only fits well on the out-of-sample pool; it also fits the shape of defaults out-of-sample on years 0 through 4 and 6 and 7. These results give us confidence that the model fits defaults accurately enough to use for drawing realistic policy conclusions.

Importantly, the model also fits cross-sectional characteristics of defaulting borrowers.

5.1 Cross Section

Borrowers with low credit scores are more likely to default. Is this because credit score reveals information about the borrower's penalty, as we claim? Or is it only because borrowers with lower credit scores are more likely to have higher-interest-rate loans, negative shocks, and other risk factors which together fully account for their higher default propensities? That lenders continue to use credit scores despite full transparency into borrowers' financial circumstances suggests that the predictive power of credit score is not fully explained by these other risk factors.

We can also shed light on these possibilities by using the model in a simple test. If credit score does not act through penalty, then the model-predicted default rates for pools of borrowers with different credit scores but a fixed penalty should not be meaningfully worse than the model-predicted default rates for pools of borrowers with different credit scores and varying penalties: the same borrowers, same financial circumstances, and same correlated risk factors enter in each case, so if the non-penalty risk factors are sufficient to explain how credit score and default propensity correlate, then they should fully account for the variation in default behavior of borrowers with different credit scores.

In particular, this test also functions as a test of the hybrid benchmark model against our idiosyncratic penalty model. Which produces a better fit to the distribution of defaults across borrowers with different credit scores? We perform thi stest by computing the cumulative default rates for borrowers across five quintiles of credit score across three regimes: the empirical value, as predicted by the hybrid model with the single fixed default penalty which best fits the aggregate level of default, or as predicted by our model. The results for the top four quintiles are shown in Figure 3. Empirical cumulative default rate curves are shown in solid lines. The predictions of the model with idiosyncratic penalties are displayed in the tight dotted lines, while the alternate hypothesis in which all borrowers have the same penalty but credit score alone explains default rates generates the predicted default curves in the long dashed lines.

[Figure 3 about here.]

The fixed-penalty model does generate some heterogeneity in borrower behavior across quintiles of credit score, indicating that credit score does, to some extent, predict defaults by correlation with other risk factors. But if that correlation fully explained the relationship between credit score and default rate, then we would see the dashed lines overlap or approximate the solid lines. Instead, the dashed lines exhibit an inferior fit; the model with idiosyncratic penalties more closely tracks the empirical variation in default rate across credit scores. This exercise suggests both that credit score's power to predict defaults is not fully explained by correlation to other risk characteristics. It also shows that the idiosyncratic-penalty model provides a better fit to the data than the hybrid benchmark.

5.2 Further Comparison

In a final test of the relative performance of our model versus benchmarks, we compare each model's ability to classify defaults ex ante and out-of-sample. We run our model against the aforementioned hybrid model with fixed penalty set at the penalty which best fits the aggregate default rate, against the pure-strategic model, and against the double-trigger model. Because we use historical default rates in calibrating the hybrid and idiosyncratic-penalty mdoels, we also allow each of the double-trigger and strategic models to use a classification threshold which ensures that the model generates the correct total percent of defaults. That is, even the double-trigger and strategic models as we have implemented them issue default *probabilities* for borrowers, in that the borrower does not know which path of income or prices will obtain. We thus set the threshold for classifying whether a given probability of default indeed predicts a default as that threshold which makes the model best predict the overall default percent.

With each model's classification threshold set, we calculate the confusion matrix of each model. We summarize these findings in Table 3 The strategic model performs worse, with a true-positive rate of approximately 26% and a true-negative rate of approximately 92.75%. The double-trigger model performs essentially slightly better, with a true-positive rate of approximately 29% and a true-negative rate of approximately 92.5%. The hybrid model outperforms each, with a true-positive rate of approximately 36% and a true-negative rate of approximately 93.75%. But the idiosyncratic model has the best performance, with a true-positive rate of approximately 41% and a true-negative rate of approximately 94%. (As defaults are rare, all models exhibit much higher true-negative rates than true-positive.)

[Table 3 about here.]

6 Conclusion

Computational power has expanded significantly in the last decade. Simplifying assumptions that are made for computational tractability are rarely still necessary. At the same time, heterogeneous and high-resolution microdata are becoming more widely available. The convergence of these trends provides an incredible opportunity for economics to provide more accurate and more illuminating heterogeneous models.

We have offered one such model for the mortgage default decision. The model trades computational complexity for greater precision in fitting borrower behavior, a wider variety of behavior that it can explain, and what we feel is a truer characterization of the motivations underlying observed behavior. Specifically, borrowers in our model are heterogeneous in all their financial characteristics, in their mortgage characteristics, in the house price dynamics they face, and in their non-pecuniary penalties for default. We fit the model to correspondingly heterogeneous microdata, showing how this model more faithfully captures the distribution of borrowers' default motivations than benchmarks that rely on representativeness assumptions.

We have offered a structural model which we hope will become one in a new class of computational structural models in the literature. Academics promoting structural models in the past may have over-promised and under-delivered, in that their vision was too far ahead of the lagging availability of computational power and large microdata. We feel that these resources may now be at a point to enable structural modeling to flourish.

We are eager to investigate whether machine learning methods, including deep learning and recursive neural networks, can offer further improvements to our understanding of mortgage default. We feel that a parameterized structural model, with a one-to-one mapping from parameters to identified features in the data, benefits from explainability and the ease of performing counterfactuals. Such a model is not limited by the single-history problem facing empirical data generated from the world. But we look forward to learning whether emerging techniques in machine learning may nonetheless push the frontier of mortgage finance understanding yet further.

References

ATTOM Data Solutions (Jan. 15, 2019). *U.S. Foreclosure Activity Drops to 13-Year Low in 2018*. ATTOM Data Solutions. URL: https://www.attomdata.com/news/most-recent/2018-year-end-foreclosure-market-report/ (visited on 12/09/2019).

Bajari, Patrick, Chenghuan Sean Chu, and Minjung Park (Dec. 2008). *An Empirical Model of Subprime Mortgage Default From 2000 to 2007*. National Bureau of Economic Research: 14625. URL: https://www.nber.org/papers/w14625 (visited on 07/07/2022). preprint.

Bhutta, Neil, Jane Dokko, and Hui Shan (Dec. 2017). "Consumer Ruthlessness and Mortgage Default during the 2007 to 2009 Housing Bust: Consumer Ruthlessness and Mortgage Default". In: *The Journal of Finance* 72.6, pp. 2433–2466. ISSN: 00221082.

Black, Fischer and Myron Scholes (May 1972). "The Valuation of Option Contracts and a Test of Market Efficiency". In: *Journal of Finance* 27.2, pp. 399–417.

Board of Governors of the Federal Reserve System (US) (Nov. 1, 1994). *Commercial Bank Interest Rate on Credit Card Plans, All Accounts*. FRED, Federal Reserve Bank of St. Louis. URL: https://fred.stlouisfed.org/series/TERMCBCCALLNS (visited on 10/25/2021).

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (Feb. 2018). "Diagnostic Expectations and Credit Cycles". In: *The Journal of Finance* 73.1, pp. 199–227. ISSN: 00221082.

Bradley, Michael G., Amy Crews Cutts, and Wei Liu (June 2015). "Strategic Mortgage Default: The Effect of Neighborhood Factors:" in: *Real Estate Economics* 43.2, pp. 271–299. ISSN: 10808620.

Bricker, Jesse et al. (2011). "Surveying the Aftermath of the Storm: Changes in Family Finances from 2007 to 2009". In: p. 38.

Campbell, John Y. and JoãO F. Cocco (Aug. 2015). "A Model of Mortgage Default". In: *The Journal of Finance* 70.4, pp. 1495–1554. ISSN: 00221082.

Corradin, Stefano (June 1, 2014). "Household Leverage". In: *Journal of Money, Credit and Banking* 46.4, pp. 567–613. ISSN: 1538-4616.

Crews Cutts, Amy and William A Merrill (Mar. 2008). *Interventions in Mortgage Default: Policies and Practices to Prevent Home Loss and Lower Costs*. Working Paper. McLean, VA: Freddie Mac, p. 78.

Demyanyk, Yuliya and Otto Van Hemert (2008). *Understanding the Subprime Mortgage Crisis*. Proceedings. Federal Reserve Bank of Chicago.

Elul, Ronel et al. (2010). "What "Triggers" Mortgage Default?" In: *The American Economic Review* 100.2, pp. 490–494. ISSN: 0002-8282. JSTOR: 27805045.

Fannie Mae (Oct. 2, 2019). *Loan-Level Price Adjustment (LLPA) Matrix*.

Federal Reserve Board (2019). *Delinquency Rate on Single-Family Residential Mortgages, Booked in Domestic Offices, All Commercial Banks*. FRED, Federal Reserve Bank of St. Louis. URL: https://fred.stlouisfed.org/series/DRSFRMACBS (visited on 12/09/2019).

Ferreira, Fernando and Joseph Gyourko (June 2015). *A New Look at the U.S. Foreclosure Crisis: Panel Data Evidence of Prime and Subprime Borrowers from 1997 to 2012*. w21261. Cambridge, MA: National Bureau of Economic Research.

Foote, Christopher L, Kristopher Gerardi, and Paul S Willen (May 30, 2008a). *Subprime Facts: What (We Think) We Know about the Subprime Crisis and What We Don't*. Working Paper. Federal Reserve Bank of Boston, p. 57.

— (Sept. 2008b). "Negative Equity and Foreclosure: Theory and Evidence". In: *Journal of Urban Economics* 64.2, pp. 234–245. ISSN: 00941190.

Foote, Christopher L. and Paul S. Willen (Nov. 2018). "Mortgage-Default Research and the Recent Foreclosure Crisis". In: *Annual Review of Financial Economics* 10.1, pp. 59–100. ISSN: 1941-1367, 1941-1375.

Foster, Chester and Robert Van Order (1984). "An Option-Based Model of Mortgage Default". In: *Housing Finance Review* 3.4, pp. 351–372.

Fuster, Andreas and Paul S. Willen (Nov. 2017). "Payment Size, Negative Equity, and Mortgage Default". In: *American Economic Journal: Economic Policy* 9.4, pp. 167–191. ISSN: 1945-7731, 1945-774X.

Ganong, Peter and Pascal Noel (Oct. 2020). "Liquidity versus Wealth in Household Debt Obligations: Evidence from Housing Policy in the Great Recession". In: *American Economic Review* 110.10, pp. 3100–3138. ISSN: 0002-8282.

— (Jan. 22, 2017). *The Effect of Debt on Default and Consumption: Evidence from Housing Policy in the Great Recession*. Working Paper, p. 78.

Geanakoplos, John (June 1, 2016). "The Credit Surface and Monetary Policy". In: *Progress and Confusion: The State of Macroeconomic Policy*. The MIT Press. ISBN: 978-0-262-33345-0.

— (Apr. 2010). "The Leverage Cycle". In: *NBER Macroeconomics Annual 2009, Volume 24*. Vol. 24. University of Chicago Press, pp. 1–65.

Gerardi, Kristopher, Adam Hale Shapiro, and Paul Willen (Dec. 3, 2007). *Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures*. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.

Gerardi, Kristopher S., Kyle Herkenhoff, et al. (2013). "Unemployment, Negative Equity, and Strategic Default". In: *SSRN Electronic Journal*. ISSN: 1556-5068.

Gerardi, Kristopher S., Andreas Lehnert, et al. (2011). "Making Sense of the Subprime Crisis". In: *Lessons from the Financial Crisis*. Wiley-Blackwell, pp. 109–117. ISBN: 978-1-118-26658-8.

Ghent, Andra C. and Marianna Kudlyak (Sept. 2011). "Recourse and Residential Mortgage Default: Evidence from US States". In: *Review of Financial Studies* 24.9, pp. 3139–3186. ISSN: 0893-9454, 1465-7368.

Goodman, Laurie, Walt Scott, and Jun Zhu (July 2018). *How Beneficial Are Streamlined Modifications?* Research Report. Urban Institute, p. 46.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (Aug. 1, 2013). "The Determinants of Attitudes toward Strategic Default on Mortgages". In: *The Journal of Finance* 68.4, pp. 1473–1515. ISSN: 1540-6261.

Gyourko, Joseph and Joseph Tracy (Mar. 1, 2014). "Reconciling Theory and Empirics on the Role of Unemployment in Mortgage Default". In: *Journal of Urban Economics* 80, pp. 87–96. ISSN: 0094-1190.

Jagtiani, Julapa and William W. Lang (Dec. 9, 2010). *Strategic Default on First and Second Lien Mortgages During the Financial Crisis*. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.

Kalikman, Philip Lewis and Joelle Scally (Aug. 22, 2022). *Targeted Principal Forgiveness Is Effective: Mortgage Modifications and Financial Crisis*. preprint.

Kau, James, Donald Keenan, and Kim Taewon (1994). "Default Probabilities for Mortgages". In: *Journal of Urban Economics* 35.3, pp. 278–296. ISSN: 0094-1190.

Kau, James B and Donald C Keenan (1995). "An Overview of the Option-Theoretic Pricing of Mortgages". In: *Journal of Housing Research*, p. 29.

Kau, James B. (1992). "A Generalized Valuation Model for Fixed-Rate Residential Mortgages". In: *Journal of Money, Credit and Banking* 24.3, pp. 279–99.

Kau, James B., Donald C. Keenan, Walter J. Muller, et al. (July 1, 1995). "The Valuation at Origination of Fixed-Rate Mortgages with Default and Prepayment". In: *The Journal of Real Estate Finance and Economics* 11.1, pp. 5–36. ISSN: 1573-045X.

Laufer, Steven (Apr. 2018). "Equity Extraction and Mortgage Default". In: *Review of Economic Dynamics* 28, pp. 1–33. ISSN: 10942025.

Li, Wei and Laurie Goodman (Nov. 2014). *Measuring Mortgage Credit Availability Using Ex-Ante Probability of Default*. Urban Institute, p. 35.

McArdle, Mark (July 22, 2013). *Understanding HAMP Re-Default Rates*. U.S. Department of the Treasury.

Merton, Robert C. (1973). "Theory of Rational Option Pricing". In: *The Bell Journal of Economics and Management Science* 4.1, pp. 141–183. ISSN: 0005-8556. JSTOR: 3003143.

Mian, Atif and Amir Sufi (Nov. 1, 2009). "The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis". In: *The Quarterly Journal of Economics* 124.4, pp. 1449–1496. ISSN: 0033-5533.

Palmer, Christopher (Sept. 24, 2015). *Why Did So Many Subprime Borrowers Default During the Crisis: Loose Credit or Plummeting Prices?* SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.

Pearl, Judea and Dana Mackenzie (May 15, 2018). *The Book of Why: The New Science of Cause and Effect*. Basic Books. 470 pp. ISBN: 978-0-465-09761-6. Google Books: 9HOdDQAAQBAJ.

Scharlemann, Therese C. and Stephen H. Shore (Oct. 1, 2016). "The Effect of Negative Equity on Mortgage Default: Evidence From HAMP's Principal Reduction Alternative". In: *The Review of Financial Studies* 29.10, pp. 2850–2883. ISSN: 0893-9454.

Schelkle, Thomas (Sept. 2018). "Mortgage Default during the U.S. Mortgage Crisis". In: *Journal of Money, Credit and Banking* 50.6, pp. 1101–1137. ISSN: 00222879.

Shiller, Robert (Oct. 2007). *Understanding Recent Trends in House Prices and Home Ownership*. w13553. Cambridge, MA: National Bureau of Economic Research.

Vandell, Kerry D. (1995). "How Ruthless Is Mortgage Default? A Review and Synthesis of the Evidence". In: *Journal of Housing Research* 6.2, pp. 245–264. ISSN: 1052-7001. JSTOR: 24832828.

White, Brent T. (2010). "Underwater and Not Walking Away: Shame, Fear and the Social Management of the Housing Crisis". In: *SSRN Electronic Journal*. ISSN: 1556-5068.

Figures and Tables

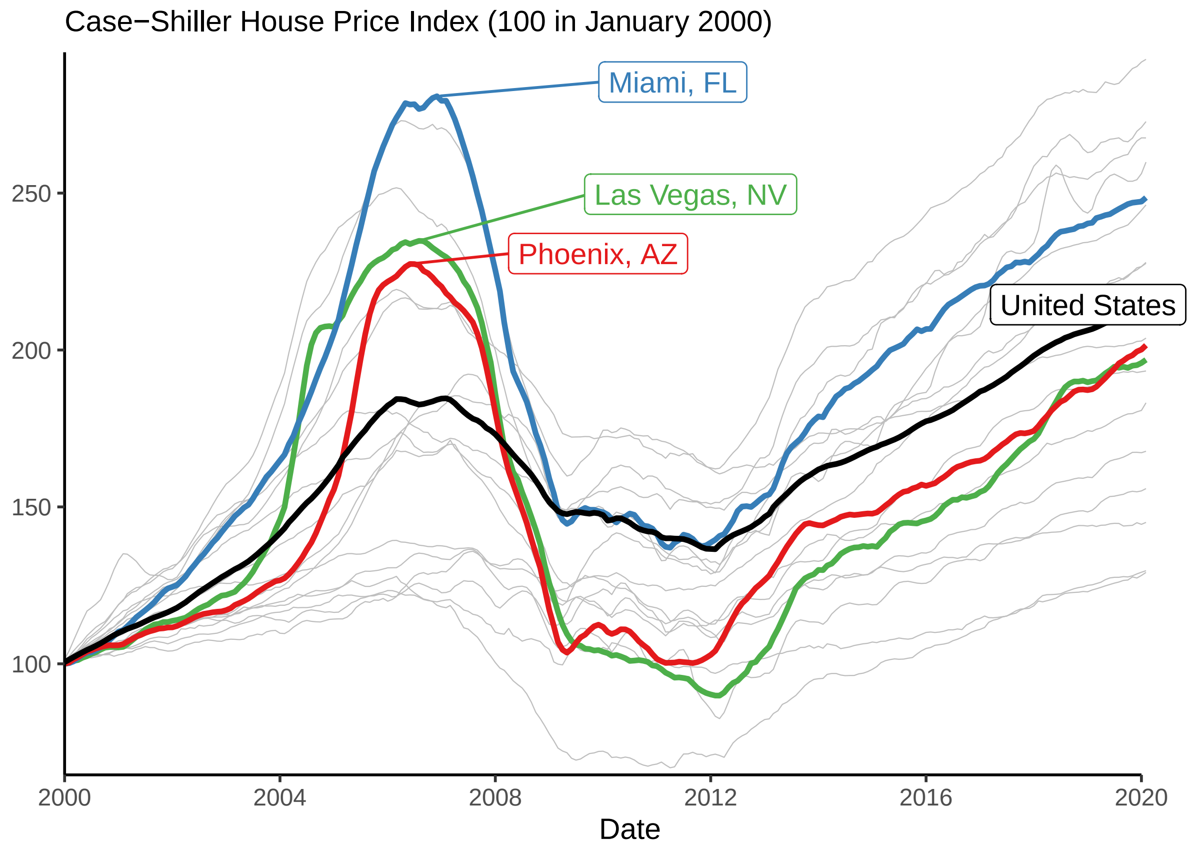

Figure 1: House Prices, United States and 20 Metro Areas, 2000 to 2020

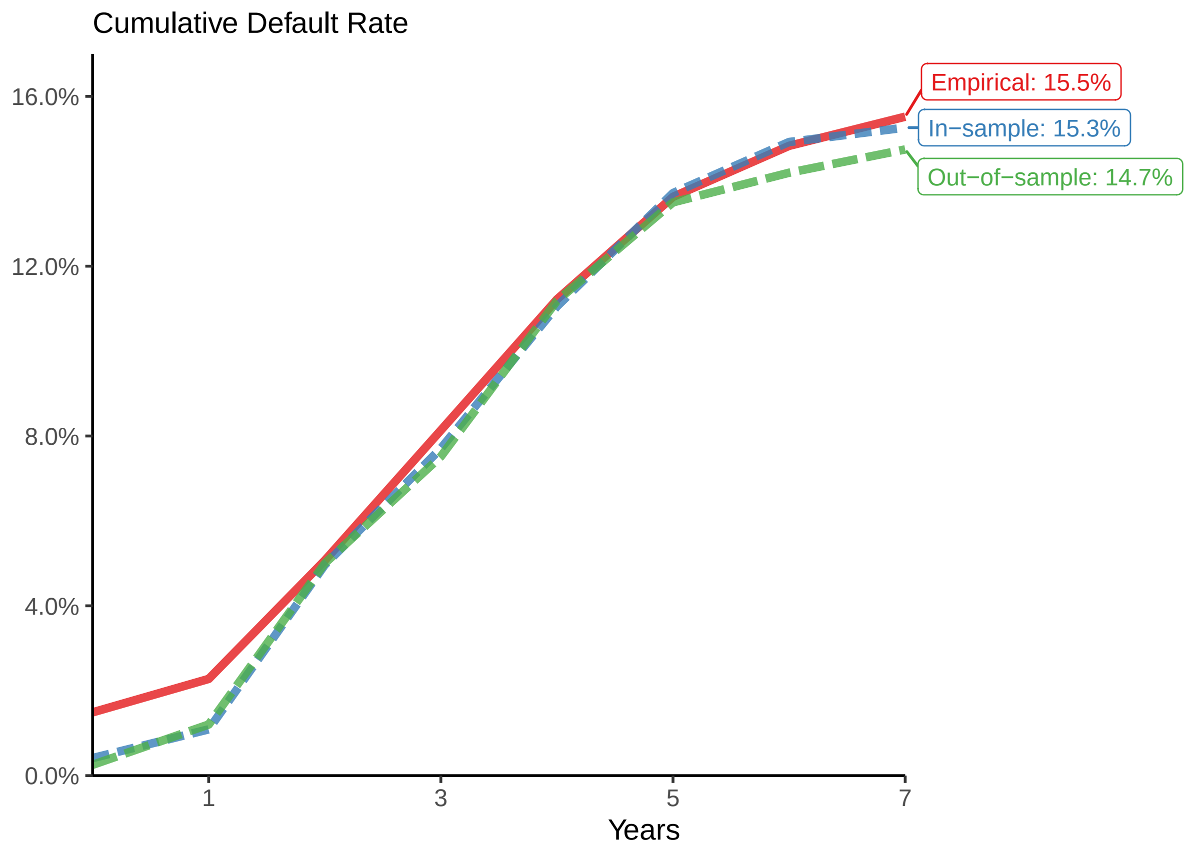

Figure 2: Cumulative Default Rates, Empirical versus Model In-sample and Model Out-of-sample

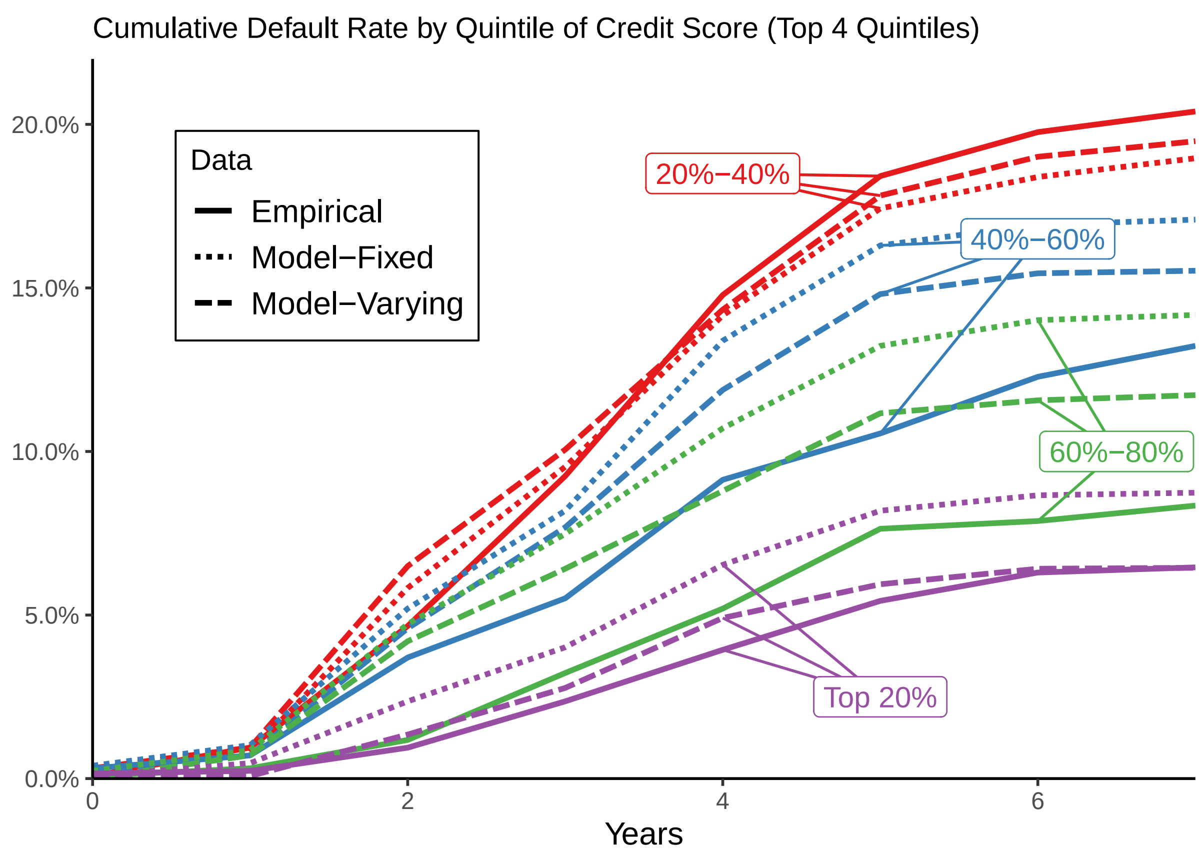

Figure 3: Cumulative Default Rates, Fixed vs. Varying Penalties, by Quintile of Credit Score

Attribute Date Observed	All Borrowers January 2007	Borrowers who Default January 2007
Original Purchase Price	\$221,000 (\$125,000)	\$226,000 (\$124,000)
Original Loan Size	\$186,000 (\$98,000)	\$196,000 (\$98,000)
Original First-lien LTV	86.2% (8.8%)	88.8% (8.9%)
Original First-lien PTI	33.16% (10.39%)	35.7% (10.0%)
Original First-lien Annual Payment	\$13,700 (\$7,200)	\$14,700 (\$7,300)
Original First-lien Interest Rate	6.21% (0.57%)	6.36% (0.64%)
Original Credit Score	717 (58)	686 (61)
Current Credit Score	714 (73)	662 (91)
Current First-lien LTV	82.8% (9.5%)	85.8% (9.8%)
Current Second Lien Balance (when positive)	\$37,000 (\$28,000)	\$42,000 (\$31,000)
Current Combined LTV	90.6% (8.3%)	92.4% (8.6%)
Current Total Annual Payment	\$15,100 (\$8,000)	\$15,900 (\$8,200)
Current Income	\$49,500 (\$65,800)	\$47,700 (\$62,800)
Current Net Liquid Assets	-\$18,000 (\$31,000)	-\$18,400 (\$30,000)
House Price Appreciation Log-Mean	0.056 (0.011)	0.057 (0.011)
House Price Appreciation Log-Volatility	0.058 (0.018)	0.059 (0.017)
Months until Default (Defaulters)		43.7 (26.8)
Distribution by Status as of 2021		
Active	76.3%	
Default	19.3%	100%
Paid	4.4%	
Distribution by Origination Vintage		
Before 2001	<1%	<1%
2001	<1%	<1%
2002	2%	2%
2003	5%	4%
2004	10%	8%
2005	36%	35%
2006	46%	51%

Table 1: Summary Statistics of 30-year Fixed-Rate Loans with 80%+ CLTV in January 2007

Group Parameter	Model Symbol	Value or Data Source
Prices and Rates:		
House Prices	$\{P_t(i)\}_{t=0,\dots,T}$	McDash + CoreLogic HPI
House Price Log-Drifts	$\{\mu_t(i)\}_{t=0,\dots,T-1}$	Zip-level CoreLogic HPI
House Price Log-Volatilities	$\{\sigma_t(i)\}_{t=0,\dots,T-1}$	Zip-level CoreLogic HPI
Risk-free Savings Rates	$\{R_t^f\}_{t=0,\dots,T-1}$	FRED 1-yr T-bill Rates
Unsecured Borrowing Rates	$\{R_t^b\}_{t=0,\dots,T-1}$	FRED Credit Card APRs
Inflation Index	$\{I_t\}_{t=0,\dots,T-1}$	1.01^t
Mortgage:		
Term	$T(i)$	McDash
Combined Initial Mortgage Principal	$M_0(i)$	Equifax CRIS + McDash
Mortgage Interest Rate Series	$\{r_t^m(i)\}_{t=0,\dots,T-1}$	McDash
Mortgage Coupon Schedule	$\{m_t(i)\}_{t=1,\dots,T}$	McDash
Mortgage Principal Schedule	$\{M_t(i)\}_{t=0,\dots,T}$	McDash
Lender Option Price Floor	$Q_t^f(i)$	Experiment-specific
Lender Option Time Limit	$Q_t(i)$	Experiment-specific
Lender Option Sharing Percent	$Q_t^p(i)$	Experiment-specific
Borrower:		
Idiosyncratic Default Penalty	$\lambda(i)$	Calibrated by Model
Income	$L_t^h(i)$	McDash & BLS
Starting Liquid Assets	$A_0(i)$	McDash/CRIS/SCF Estimate
Probability of Unemployment	π_u	0.07
Probability of Re-employment	π_e	0.35
CRRA Risk-aversion parameter	γ	3.5
Discount factor	β	0.985
Transaction Costs, Taxes, Fees, etc.:		
Income tax rate	τ_{inc}	0.20
Property tax rate	τ_{prop}	0.015
Property maintenance/insurance/HOA	τ_{maint}	0.01
Transaction cost for selling	τ_{sell}	0.05
Transaction cost for prepaying	τ_{prepay}	0.015
Transaction cost of moving	τ_{move}	0.005
Rent-to-price ratio	τ_{rent}	0.04
Distressed sale ratio	$\tau_{\text{distressed}}$	0.65
Unemployment income replacement ratio	τ_{unemp}	0.60

Table 2: Model Parameters, Data Sources, and Baseline Values

Model	Strategic	Double-Trigger	Hybrid	Heterogeneous
TN Rate (%)	26	29	36	41
TP Rate (%)	92.75	92.5	93.75	94

Table 3: True Positive and True Negative Rates by Model

A Model and Methodology Details

A.1 House Price Trinomial Tree

There are $2t+1$ possible exogenous values of the house price index in period t. Denoting the time by the superscript and the discrete realizable values by the subscript,

$$P^t = \{P^t_{-t}, \ldots, P^t_{-1}, P^t_0, P^t_1, \ldots, P^t_t\}.$$

The levels of the prices in the grid are determined by the parameters initial price $P^0(i)$, log-drifts $\{\mu_t(i)\}_{t=0,\ldots,T}$, log-volatilities $\{\sigma_t(i)\}_{t=0,\ldots,T}$, and grid parameters $\{\hat{\sigma}_t(i)\}_{t=0,\ldots,T}$. Each loan (i) has its own instances of these parameters, though we suppress the notation hereafter.

These parameters determine both the level of prices at nodes in the grid and the transition probabilities between those nodes. The j^{th} level of the house price index P^t_j is

$$P^t_j = P_0 \cdot \exp\left(\sum_{\tau=0}^{t-1} \mu_\tau + \hat{\sigma}_t \cdot j\right). \tag{16}$$

The transition probabilities between prices in subsequent periods, $Pr(P^{t+1}_{j+1}|P^t_j)$, $Pr(P^{t+1}_j|P^t_j)$, and $Pr(P^{t+1}_{j-1}|P^t_h)$, are determined by the moment conditions

$$E_{j'=j-1,j,j+1}\left[\log(P^{t+1}_{j'}) - \log(P^t_j)\right] = \mu_t \tag{17}$$
$$Var_{j'=j-1,j,j+1}\left[\log(P^{t+1}_{j'}) - \log(P^t_j)\right] = \sigma^2_t \tag{18}$$

as well as the usual constraints on probability distributions:

$$0 \leq Pr(P^{t+1}_{j'}|P^t_j) \leq 1$$
$$Pr(P^{t+1}_{j+1}|P^t_j) + Pr(P^{t+1}_j|P^t_j) + Pr(P^{t+1}_{j-1}|P^t_j) = 1.$$

The grid parameters $\{\hat{\sigma}_t\}_{t=0,\ldots,T}$ are chosen as the minimum values feasible to satisfy Equations (16), (17), and (18).

House prices during the Global Financial Crisis followed a path that many would regard as inconsistent with the expectations of borrowers before 2008. Our model thus supports time-varying expectations with "surprises"; borrowers are rational but have imperfect foresight. That is, a borrower with a mortgage at period 0 may have a forecast of log-drifts and log-volatilities as of period 0, $\{\mu_0^0, \mu_1^0, \mu_2^0, \dots\}$ and $\{\sigma_0^0, \sigma_1^0, \sigma_2^0, \dots\}$, that imply an expected house price path $E_0[P_t], E_0[P_{t+1}], \dots,$ which does not lie within the grid of prices $\{P_{-1}^1, P_0^1, P_1^1\}$ that the borrower forecasts. We account for this possibility by computing additional "out-of-grid" price levels determined by the actual price history the borrower would experience and simulating Monte Carlo histories across price paths with jumps or surprises.

If the borrower sells the home, she receives the sale price P less transaction costs $\tau_{\text{sell}} \cdot P$ and $\tau_{\text{move}} \cdot P$ for searching for and moving to a new rental home. If she defaults, she pays the moving cost only but lives rent-free for a predetermined period reflecting the time it takes the lender to foreclose. Either way, she thereafter rents a comparable property. We take rent to be proportional to price ($\tau_{\text{rent}} P_t^s$), up to a ceiling and floor so that rental costs do not get unrealistically high or low.

A.2 Irreversibility

The model features standard assumptions on irreversibility or absorbing states. First, a borrower who leaves her house, whether through sale or default, moves into a house of the same "size" or "quality." Borrowers do not adjust house size or quality. Accordingly, the term for housing in the utility function is suppressed. Second, a borrower who terminates a mortgage, whether due to prepayment, default, or completion, does not procure a new mortgage. In particular, while there is prepayment, there is no refinance and no equity extraction or other unscheduled adjustment of equity other than complete prepayment. Mortgages either amortize according to the amortization schedule determined at origination or modification, or they are fully prepaid. Finally, a borrower who rents does so until the terminal period. Borrowers do not ever purchase new homes, even if they could afford to do so entirely out of cash-on-hand.

As is widely acknowledged in the literature, the irreversibility of these decisions is not strictly accurate. But the most damning limitation of making these assumptions arises only for predictions in general equilibrium, which is not the primary concern of this study. Even in that case, the inaccuracy is mitigated to some extent by the true frictions that befall defaulters in mortgage markets. In addition to reducing defaulters' credit scores, default events generally stay on a credit report for seven years; both the lowered score and the record of the default limit access to future credit that would be necessary for most such borrowers to purchase another home.

Notwithstanding the aforementioned absorbing states, the model allows for prepayment, including prepayment without sale. Borrowers may therefore capture the value of future home price appreciation either by reducing liquid assets in one fell swoop, or by continuing to pay the mortgage coupon—though for most borrowers in our sample, prepaying without selling the home would require them to take out expensive unsecured debt, the subject of the following section.

A.3 Further Sampling Restrictions

We restrict to properties in ZIP codes in which the CoreLogic home price index is observed throughout the model period. The CoreLogic home price data cover approximately 7,500 ZIP codes. ZIP codes where there are few transactions do not have price indices computed. As transaction volume correlates with population, restricting to ZIP codes with defined home price indices likely does not break the generality of the results.

We restrict to mortgages that McDash indicates do not admit lender recourse. The literature is divided on the effects of recourse on a borrower's decision to default: on the one hand, Ghent and Kudlyak (2011) find that it reduces a borrower's default propensity; on the other, Guiso, Sapienza, and Zingales (2013) find that borrowers do not have a clear understanding of whether lenders have recourse to pursue their assets, let alone reliable estimates of lenders' likelihood of doing so or of succeeding. It is not clear how recourse would affect our estimates, as while borrowers who had expectations of lender recourse would be less likely to default ex ante, they would also be less likely to re-default after modification.

Equifax matches tradeline credit data to mortgage loans in McDash using a proprietary probabilistic matching algorithm. We restrict our analysis to loans with high match confidence; we further ensure correct matching by selecting borrowers whose first mortgage balance in the model reference period as reported in McDash is within 5% of the balance reported in Equifax. The analysis herein is not sensitive to including loans that are more or less confidently matched. We exclude loans that transferred servicers during the period of interest.

A.4 Definition of Default

Mortgage servicers do not adhere to a standard definition of the period in which a borrower defaults, and thus neither the Equifax CRIS nor the McDash data indicate a date of default. The major challenge to identifying defaults is that borrowers often fall behind on payments and then make up missed payments, or "cure," and then may stay current until prepayment or completion of the loan—or may fall behind again and default. Additionally, servicers vary in when they initiate foreclosure proceedings, and foreclosure can take varying amounts of time. Fortunately, because this is a historical study, the typical challenge of identifying a true default is mitigated. We define defaults in the data according to the *final* status of the mortgage. Because we are using mortgage history data from 2021, we observe whether the mortgage prepays, defaults, or is modified at any point after a given period. Their terminal state may be one of the following: paid off or prepaid, foreclosed upon and liquidated, transferred to a servicer that did not continue report to McDash, modified, or still active in 2021.

In the model, borrowers do not miss payments and then cure. Whenever a borrower in the model misses a payment, she will miss all future payments. The borrower will remain in the home for two years while the lender completes foreclosure proceedings, but there is no opportunity for the borrower to change her mind or access additional liquidity to make up a past missed payment. This feature is more realistic than it may appear: it means the focus of our model is only terminal defaults, not delinquencies-with-cures, and we identify only terminal defaults in the data. The model restriction and the data are consistent.

Again excluding mortgages that transferred servicers, we define mortgages that were paid or prepaid by 2021, or that are still active in 2021, as not having defaulted, even if the borrower temporarily fell behind on payments. Excluding servicer transfers may bias the sample towards safer loans but is necessary in order not to mis-classify delinquent mortgages that eventually cure after a transfer as defaults. Mortgages that were foreclosed upon before 2021did default. Additionally, if a mortgage was delinquent and then received a modification, We define it as having defaulted. The reason for this choice is that in the study of defaults, the most logical assumption is that a delinquent mortgage that received a modification received the modification precisely because the modifying party (typically the mortgage servicer) considered modifying the mortgage more profitable than the alternative of initiating foreclosure proceedings. But for this to be the case, the servicer must have expected the borrower not to cure and continue making payments on the original mortgage. We therefore treat such outcomes as defaults; they are equivalent to a default on the original mortgage and an origination of a new mortgage.

For mortgages that default, defining the *date* of default presents its own challenges. Because default in the model refers to the first point after which the borrower decides it is preferable to stop making payments and ceases making any payments, we take defaults in the data also to be the first point after which the borrower fully ceases making payments. Therefore, if a borrower becomes delinquent but cures, and subsequently becomes delinquent again but does not cure, her default date is the month before the second delinquency. The cure event is ignored.

Consider the following examples for three hypothetical loans, where e.g. 30DPD refers to a loan that is 30-59 days past due.

[Table 4 about here.]

Borrower 1 did not default. Borrower 2 defaulted in period 4: her earlier delinquency is ignored because she cured, whereas as of period 4 she began a series of missed payments that culminated in a liquidation. Borrower 3 also defaulted in period 4: she decided to miss her first payment in the following period, and never to make a subsequent payment. Borrower 4 is excluded from the sample. Because we do not observe her loan performance after the period in which her mortgage was transferred to a different servicer that did not report to McDash, we do not know whether she eventually cured her mortgage.

A.5 Definition of Modification

The McDash data include partial coverage of loss mitigation data provided by servicers, consisting primarily of flags indicating months in which a mortgage receives a modification, whether that modification was conducted as part of a program such as HAMP, and whether that modification involved various features such as principal forgiveness, interest rate reduction, or term extension. These flags sometimes do not coincide with actual observed changes in mortgage variables. To account for such discrepancies, we also follow the modification-identification procedure outlined in Goodman, Scott, and Zhu (2018). The authors identify a loan as receiving a modification when it transitions from 60 or more days delinquent to current while simultaneously increasing term, changing interest rate by more than 10 basis points, or changing principal balance or principal-plus-interest payment by more than 3 percent. We take the earliest date of such a coincident change (in both mortgage status and mortgage variables) after the reference period as the first date of modification of a loan.

A.6 Estimation of Income

We use McDash-reported front-end payment-to-income ratio to calculate the borrower's annual income. Documentation of income is known to be limited, particularly before the financial crisis, so this may overstate income estimates in many cases. An investigation of an additional dataset with borrower-level income and employment data revealed that it matched too few borrowers who default during the sample period to offer a viable alternative to this methodology. This dataset has better coverage for recent years and could prove useful in extending the analysis to more recent and/or future defaults.

A.7 Estimation of Liquid Assets

We model cash on hand for the borrower as a predicted level of liquid assets net of the level of non-mortgage debt as observed in CRIS, where the prediction is formed from the 2007-2009 Survey of Consumer Finances data by regressing log liquid asset levels against log income, log mortgage debt, and log non-mortgage debt for respondents with values of each of these variables in the middle 90% of respondents. We define mortgage debt as the sum of mortgage and HELOC debt, non-mortgage debt as the sum of credit card debt, auto debt, education debt, and other debt, and assets as the sum of checking and savings accounts, mutual fund holdings, savings bonds, stocks, and other significant assets. As indicated in Table 5, log levels of income, mortgage debt, and non-mortgage debt are all statistically and economically significant predictors of log liquid assets, with an R^2 of 41%.

[Table 5 about here.]

A.8 House Price Expectations Revisions

Shiller (2007) and Bordalo, Gennaioli, and Shleifer (2018) argue that borrowers revise expectations after price surprises, overweighting recent news. To reflect these significant house price surprises and the expectations revisions that borrowers likely undertake following such drastic surprises, for a particular realized path of house prices, we recompute the entire tree each period that the borrower retains her mortgage. If the borrower does not terminate the mortgage in period 0, her decision in period 0 determines her successor state in period 1. She forms a new forecast of future prices at this state in period 1, and therefore recomputes the entire remaining truncated tree, with the contemporaneous level of the house price determined exogenously but her asset level and other state variables determined endogenously. This modeling approach is computationally taxing but enables the model to capture two important realistic features: that the significant price drops in financial crises are not foreseen, and that such drops likely induce expectations revisions in borrowers who endure them.

With expectations revisions, borrowers default more quickly even when they have the same penalties, because negative surprises cause them to revise their expectations of future price appreciation downward. Cumulative default curves therefore exhibit steeper early rises. Other results are qualitatively similar.

A.9 Estimation of Penalties

Taking for granted that penalties exist, are significant, and vary across borrowers, but are unobservable, how should we match borrowers to their penalties?

Two approaches suggest themselves: first, using borrowers' complete payment histories to bound plausible penalties, independent of their other observable characteristics; and second, using borrowers' other observable characteristics to proxy for and thus estimate their penalties. Within the second approach, the modeler has the choice of which observable characteristics to select and which to omit.

In this essay, we prefer the simplest possible flavor of the second approach. We wish our method to be implementable by policymakers, and even policymakers in 2009 could have estimated penalties to credit scores by looking at payment histories through 2009, and also by looking at defaults in prior years. And since credit score is the only boldface borrower characteristic used by lenders for estimating their default propensity but which has no direct mathematical interpretation in a model of default, it is for both reasons the most logical charateristic to use for estimating penalties.

When we do expand the number of admissible predictors of default penalty, we find credit score remains the most robust predictor of a borrower's propensity to default. We consider a reduced form logistic regression in which income, wealth, age, and credit score together predict borrower penalty. Credit score remains the strongest explanatory factor predicting a borrower's likelihood of default. We do find that introducing other factors can explain a realistic non-monotonocity in the data, where borrowers with high credit score are sometimes disproprtionately more likely to default than their estimated penalty would suggest. We speculate their relatively greater default propensity is due to their greater financial sophistication, greater access to legal resources, and greater awareness of the limited financial consequences of default.

We could also take the opposite route: entirely ignore borrowers' observable characteristics and simply estimate penalties as a purely idiosyncratic and uncorrelated error term unique to each borrower. This approach exploits the discrepancy between the continuity in financial incentives facing borrowers who decide to continue or default and the discontinuity in her behavior if she does, using months when the borrower continues to lower-bound penalties and the month when she defaults to reveal an upper bound. This approach may more accurately match each individual borrower to her own most precise penalty, at the cost of only being implementable ex post.

Among other directions for future research, the model, and thus our estimation procedure, assumes that each borrower expects her own penalty not to change in the future. It is possible that shocks such as the Global Financial Crisis, in which borrowers observe mass defaults by neighbors, re-evaluate their estimate of the reputational cost they would suffer in the eyes of their neighbors and also learn information about the foreclosure process, its impact on reputation and credit access, its procedural inconveniences which may further lead them to re-evaluate their penalties; in the opposite direction, borrowers who plan to build credit over time may anticipate that their future penalties would grow with time. If these effects are of first order, then accounting for the possibility of changing penalties should be a fruitful direction for future research.

A.10 Methodological Viewpoint Regarding Classification

As neither the strategic nor double-trigger model alone can explain the behavior of all borrowers, some authors have attempted to determine fractions of defaulting borrowers that fall into one of the two types. Bradley, Cutts, and Liu (2015), for example, use borrower income data, linked to mortgage performance data, to study the prominence of strategic default. Their research motivation is to show that neighborhood effects matter for strategic default, a contagion phenomenon widely reported on in the popular news media and in industry. A finding of neighborhood effects would also provide evidence for the theory that borrowers face a psychological cost to default: borrowers who see their peers and neighbors default reduce the reputational cost they associate with defaulting. The authors find that strategic defaults account for 15-20% of defaults, or equivalently, that 80-85% of defaults occur due to liquidity constraints.

We believe there are many shortcomings with the classification approach. Among these: it does not provide much forecasting ability, much ability to investigate counterfactuals, or therefore much power to provide policy guidance. But more damningly, we argue that the classification approach is simply ontologically inaccurate: the correct classification applies just as much to the borrower's circumstances as to the borrower herself, but circumstances change, and thereby undermine the claimed classification.

Consider a simple thought experiment: suppose a borrower with a 110% LTV mortgage continues making payments but suffers a job loss. She falls behind on payments. She seems to be a double-trigger defaulter, the type of borrower screened for by the Treasury department's Home Affordable Modification Program. So she applies for mortgage relief, and qualifies by virtue of being behind on payments while having evidence of financial hardship. The borrower receives a modification that reduces her payment just below the affordability threshold of 31% payment-to-income.

Now suppose the borrower's house value plummets over the following year, raising her LTV to 130%. We should not be surprised when this borrower defaults, even though we called her a double-trigger borrower and her payments are now affordable. In fact, this fate befell many borrowers who received HAMP modifications. HAMP rolled out in 2009, but prices continued to fall. Fully *half* the recipients of HAMP modifications in 2009—all of whom had been screened for as double-trigger defaulters—re-defaulted, and a third of all HAMP modifications re-defaulted. This outcome is less surprising if one considers that HAMP's eligibility screen was based on a label that was never likely to persist.

The intrinsic incoherence of the classification approach is also consistent with the widely varying estimates in the literature of the proportion of strategic defaulters. The review in Gerardi, Herkenhoff, et al. (2013) found the strategic percent range from 7% to up to 38%—a huge range. Ganong and Noel (2020) argue for an even lower percent of strategic defaulters, finding that illiquidity is a necessary condition for 97% of defaults. They employ a novel and compelling methodology, based on the work of Pearl and Mackenzie (2018)—whose argument that a structural model provides the sine qua non of causal and counterfactual analysis also motivated our own structural approach—that a liquidity shock is a necessary condition for default for 97% of defaulters. We are not ready to conclude that the strategic model is irrelevant. Nearly all borrowers experience a liquidity shock before they default, but as Gerardi, Herkenhoff, et al. (2013) point out, most borrowers who experience liquidity shocks still do not default—and those who do default well before their illiquidity binds. The label "double-trigger" lacks the power to distinguish those who do from those who do not, and lacks the power to determine even when such a defaulter will finally stop making payments.

Ganong and Noel (2020) point out that their findings are rationalized by a model which incorporates a high private cost of mortgage default. Ours is such a model. It explains the decision to default by providing an explicit quantitative representation of how each borrower weighs considerations that represent illiquidity and strategy, weighing her willingness to tolerate reduced consumption just for the sake of not defaulting against her expectations for future consumption if she defaults and saves money. It provides a precise forecast of each borrower's default decision conditional on paths of income and house prices. It need not classify the borrower to faithfully forecast her behavior conditional on any circumstance that may arise. Therein lies what we claim is its superiority to a classification approach:

when the facts facing a borrower change, the borrower changes her mind. So does the model.

Appendix Figures and Tables

Period	Loan 1	Loan 2	Loan 3	Loan 4
0	Current	Current	Current	Current
1	30 DPD	30 DPD	Current	30 DPD
2	60 DPD	60 DPD	Current	60 DPD
3	90 DPD	30 DPD	Current	90 DPD
4	60 DPD	Current	Current	Transferred
5	30 DPD	30 DPD	30 DPD	
6	Current	60 DPD	60 DPD	
7	Current	90 DPD	90 DPD	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	
T	Paid	Liquidated	Liquidated	

Table 4: Hypothetical Loan Performance Examples

	Log Liquid Assets
(Intercept)	-11.69^{***}
	(0.34)
Log Mortgage Debt	0.23^{***}
	(0.03)
Log Non-Mortgage Debt	-0.09^{***}
	(0.01)
Log Income	1.64^{***}
	(0.03)
R^2	0.41
Adj. R^2	0.41
Num. obs.	7994
RMSE	2.11

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 5: Regression Coefficients: SCF Estimation of Liquid Assets