Credit Constraints and Bias in Hedonic Amenity Valuations *

John Heilbron †

November 19, 2021

Click ***here*** *for latest draft.*

Abstract

I demonstrate that credit constraints bias price-based (cf. rent-based) hedonic valuations of local public amenities. Mis-measurement of the private value of local public amenities distorts welfare analysis and could cause under-investment in amenities. I introduce a method to measure this bias by applying the hedonic framework to a problem of mortgage choice in the presence of credit constraints. I use Fannie Mae pricing grids and private mortgage insurance requirements to construct borrower-level mortgage choice menus and estimate the extent of bias among fixed-rate agency loan borrowers. I find evidence that traditional hedonic techniques understate the value of amenities by ~ 50%.

Keywords: Household Finance, Housing Demand

JEL Codes: D12, H81, R21

*I would like to thank Benedict Guttman-Kenney, John Heaton, Agustin Hurtado, Elisabeth Kempf, Joshua Morris-Levenson, Stefan Nagel, Scott Nelson, Pascal Noel, Kelly Posenau, Amir Sufi, Constantine Yannellis, Anthony Zhang, Eric Zwick and my peers in Amir Sufi's working group.

†University of Chicago, Booth School of Business

1 Introduction

Where households live is instrumental to their well-being (Kling et al., 2007; Clampet-Lundquist and Massey, 2008; Ludwig et al., 2012) and the future prospects of their children (Chetty et al., 2015; Chyn, 2018). This is due, in part, to the fact that different neighborhoods offer different amenities, goods like effective schools or safe streets to which households gain access by sheer proximity. Government bodies are often responsible for overseeing provision of these local amenities, many of which are public goods by nature. And as a normative matter, complementing concerns of fairness and externalities, the optimal provision of amenities should be responsive to household preferences and the private benefits that households derive from them.

In this paper, I argue that the standard hedonic techniques used to estimate the private value of amenities are downward biased. Standard techniques compare the prices of otherwise similar homes, one with access to more of the amenity; the difference in prices is interpreted as the household's willingness-to-pay (WTP) for the additional amenity. For credit constrained households, however, payment requires a sacrifice of non-housing consumption that is especially burdensome because it is concentrated at the time of home purchase. What they are willing-to-pay understates what they would be willing-to-pay if the payments could be better financed.

I further argue that household mortgage choice is informative about the size of the bias in amenity valuations. A household's WTP for consumption at the time of home purchase is reflected in the increase in future payment obligations it is willing to accept to increase the size of its mortgage balance. Given a borrower's choice of mortgage, this can be read off the slope of a menu of mortgage contracts quoted in terms of mortgage balances and payment obligations. If a household is unconstrained, it can always smooth consumption along some other margin and will be unwilling to pay more than \$1 in present value of future payment obligations for an increase in a loan balance of \$1. A constrained household, by contrast, may be willing to pay more.

In my empirical analysis, I exploit features of the U.S. agency loan market, the GSE loan-level price adjustment grids and private mortgage insurance requirements, to construct borrower-level mortgage menus. Using these menus and borrowers' chosen contracts, I find that, on average, households will pay \$1.65 in present value of future payment obligations for an increase in a loan balance of \$1. I devise a method for correcting hedonic estimates that accounts for joint heterogeneity in WTP for credit and amenities. Incorporating my borrower-level estimates of WTP for credit, I correct estimates of mean marginal WTP for school quality and find that standard estimates are downward biased by ~50%.

The standard hedonic approach to measuring the private value of amenities is due to Rosen (1974), who regresses rents on housing characteristics, including measures of local amenities. By describing an equilibrium model of housing supply and demand in characteristic space, he gives an interpretation to the coefficients obtained in such a regression, namely, households' mean marginal WTP for amenities. Effectively, the difference in rental rate between two otherwise-identical homes, one of which has access to, say, better schools, describes household WTP for the additional school quality.

Rosen (1974) proposed a regression in housing rents, but hedonic estimates are often obtained in terms of house prices. This is because most households are home-owners and because house price data is more readily available than rental data. By analogy, such estimates purport to recover the capitalized rather than flow value of amenities. But home purchase, unlike rental, requires financing, introducing the possibility of bias. For constrained households, the costs of accessing better amenities cannot be smoothed but instead are concentrated in the period of home purchase. Therefore, what households *will* pay out of *constrained* consumption at home purchase for more amenities understates what they *would* pay as a flow of *unconstrained* consumption over time.

The distinction is important from a policy perspective, the view which hedonic estimates very often inform. A government considering the costs and benefits of a proposed intervention may not be subject to financing constraints of the kind faced by households when choosing a neighborhood. Instead, it may raise the funds for amenity improvements from a flow of taxes. As a consequence, the benefits of improving amenities should be valued relative to the costs of unconstrained consumption. Using "traditional" estimates without correcting for bias would overlook welfare-improving investment opportunities.

To better apprehend the mis-measurement of policy-relevant amenity valuations, I begin by re-interpreting the estimates captured in hedonic regression. What is termed 'willingness-to-pay' in hedonic regressions is, substantively, a marginal rate of substitution (MRS) between housing and non-housing consumption. Somewhat subtly, the housing rents or prices in the regression represent *non-housing* consumption being sacrificed in favor of various amenities. The challenge is that hedonic regressions in prices capture a MRS between housing and *constrained* non-housing consumption, whereas *unconstrained* non-housing consumption is the alternative of policy relevance.

To correct the bias in estimates, I apply the logic of the hedonic regression to a novel domain. Instead of a regression of house prices on home attributes, I consider a regression of mortgage obligations on mortgage balances, which I term a "financial" hedonic regression.[1] The coefficient now captures the MRS between present and future non-housing consumption. To a first degree of approximation, households are credit constrained at the time of home purchase and not in the later course of home-ownership. Therefore, the "financial" hedonic regression delivers the

[1]More generally, this regression could be implemented by regressing the price of a financial product on state- and time- indexed payoffs. To the extent that the menu of contracts features variation along all relevant states and household beliefs are known, it is possible to recover SDFs, which are just a collection of MRS's between consumption in different states and times. This paper sidesteps considerations of household beliefs and state-indexed consumption by focusing only on the MRS between present and future consumption.

MRS between constrained and un-constrained non-housing consumption. As discussed above, the "traditional" hedonic regression in prices recovers the MRS between housing and constrained non-housing consumption. Combining the two, we obtain the MRS of policy interest.

I implement my proposed bias correction by exploiting institutional features of the U.S. mortgage market. To obtain the credit guarantees required to securitize a mortgage in the agency market, originators must conform to standards set by the GSEs. In particular, they must pay loan-level fees to the agencies and, for loans originated above an 80 loan-to-value ratio, borrowers must obtain private mortgage insurance (PMI). The size of the fees and the PMI premiums both depend on the leverage of the underlying loan. Jointly, these requirements establish a market-wide menu of mortgage contracts in which borrowers can lever up but must pay higher effective interest rates to do so.

With the GSE requirements establishing an effective menu of contracts, I can conduct two exercises. First, in the spirit of "financial" hedonic regression, I estimate the slope of the mortgage price schedule to find households' WTP for credit, or MRS between present (constrained) and future (unconstrained) consumption, at the time of home purchase. A variety of measurement concerns arise in this setting, including the value of the default and prepayment options embedded in mortgages. I address these by conducting a sensitivity analysis in my estimates.

Second, I turn to the exercise of correcting estimates of WTP for amenities. Because there is cross sectional heterogeneity in WTP for both amenities and credit, the bias correction features a covariance term that cannot be estimated with simply the mean WTP for credit. I use the well-defined nature of the mortgage menu and publicly-available information on GSE pricing grids and PMI rate cards to calculate WTP for credit at the individual level. I devise and implement a strategy that uses this individual-level variation to correct estimates of mean WTP for school quality that accounts for this joint heterogeneity.

Ultimately, I find that the WTP for \$1 of credit at the time of home purchase has mean ~\$1.65. This is the most conservative estimate of the various sensitivity analyses that I run. The mean suggests that bias in hedonic price estimates is on the order of 65%. I use individual level estimates of WTP for credit to correct a "traditional" hedonic regression. When I correct for bias induced by credit constraints, I find that coefficients on district school quality increase by 50%.

The remainder of the paper proceeds as follows. In Section (2), I discuss the paper's contribution to each of several strands of literature. In Section (3), I formalize the main intuitions of the paper. In Section (4), I provide an overview of the GSEs in the mortgage market and the market-wide menu of mortgage contracts they establish through their loan-level fee structure and PMI requirements. In Section (5), I describe my data sources and sample construction. Section (6) presents results measuring the extent of household credit constraints at the time of home purchase from data on mortgages. Section (7) uses information on household credit constraints to correct estimates of household willingness-to-pay for amenities. Section (8) concludes.

2 Related Literature

This paper revisits a longstanding literature, dating back to the framework of Rosen (1974) and Roback (1982), that uses land rents to value local public amenities like school or air quality. The plausibility of these estimates has been improved by empirical papers using quasi-random variation to reduce selection concerns (Black, 1999; Chay and Greenstone, 2005; Bayer et al., 2007). Recent papers have raised the possibility of bias in these estimates due to equilibrium effects (Kuminoff and Pope, 2014), limited information and distorted beliefs (Gao et al., 2021), and interest costs in lending (Ouazad and Rancière, 2019). The hedonic approach continues to be used to inform public policy debates and continues to be evaluated in house prices rather than rents (Currie et al., 2015; Kulka, 2019; Diamond and Mcquade, 2019). I contribute to this literature in two ways: (i) I argue that borrowing constraints and the resulting shadow-cost of credit may introduce measurement bias and (ii) I propose a technique for measuring this bias by introducing the logic of the hedonic regression to mortgage-choice rather than housing-choice.

In recent years, a large literature has used discrete-choice structural models to investigate household financial product choice. Various papers have studied student debt (Ebrahimian, 2020), mortgage choice (Benetton, 2021; Robles-Garcia, 2018), auto debt (Grunewald et al., 2020), credit card debt (Nelson, 2018), and ETF choice (Egan et al., 2020). Several of these papers raise the possibility that borrowers may be credit constrained, either a channel of interest or an alternative explanation. None of these papers, however, characterizes product attributes as their time- and state-dependent costs or payoffs. This approach is generally clarifying because the marginal rate of substitution between consumption in various states and times is the household stochastic discount factor (up to information about household beliefs). More specifically, it provides an approach for measuring the effects of credit constraints directly in cases where it is possible to observe a menu of potential financing arrangements.

I use the GSE LLPA grid and PMI rate card grids as a basis for the menu of financing arrangements available to borrowers. In the paper, a borrower's choice of cell is revealing about their willingness to substitute non-housing consumption inter-temporally. Conceiving of these grids as a market-wide menu contributes to a literature on their effects and research uses. Fuster et al. (2013) uses these grids to measure lender profitability of loans. Hurst et al. (2016) notes that because the guarantee fees do not account for information on local housing market conditions, the grid effects a large inter-regional insurance program. Bartlett et al. (2021) notes that within LLPA grid cells, lenders are not differentially exposed to borrower credit risk, and uses the grid to study discrimination in mortgage lending.

Finally, the determinants of housing leverage has been an area of active research since the Great Financial Crisis. During the early 2000s, rising leverage was driven by the expansion in availability of sub-prime mortgage credit (Mian and Sufi, 2009) as well as improved housing collateral values against which households could borrow (Mian and Sufi, 2011). At a micro-economic level, DeFusco and Paciorek (2017) measures the effect of interest rates on first mortgage balances and Bailey et al. (2019) finds a limited role played by household beliefs about housing prices. Defusco et al. (2020) considers the role played by government policy in limiting mortgage leverage. This paper contributes to the literature by characterizing the mortgage leverage choice when a household (i) has multiple margins for borrowing and (ii) faces rising marginal costs of mortgage credit.

3 Framework

3.1 The Household Problem

I consider an infinitely-lived HH purchasing rather than renting a unit of housing. The household allocates wealth in the form of initial savings, a_0, and wages, $\{w_t\}_t$, between non-housing consumption each period, $\{c_t\}_t$, and an amenity, s, priced according to an equilibrium hedonic schedule of prices, $P(s)$. The household has a non-mortgage margin for borrowing at the risk-free rate, but faces a credit constraint, which I normalize to zero. The household may also finance the purchase of the home with a mortgage loan. The household chooses initial balance B^o, which it receives in exchange for periodic payment obligations, $B^o r(B^o)$. Here, $r(B^o)$ describes the menu of available mortgage contracts and $r'(B^o) > 0$.

Formally, I write:

$$
\begin{aligned}
\max_{\left\{\{c_t\}_{t\geq 0},\{a_t\}_{t\geq 1},s,B^o\right\}} \quad & u(c_0,s_0)+\sum_{t=1}^{\infty}\beta^t u(c_t,s) \\
\text{s.t.} \quad & a_0+w_0=c_0+\left[P(s)-B^o\right]+\frac{a_1}{1+r} && (\lambda_0) \\
& a_t+w_t=c_t+B^o r^m(B^o)+\frac{a_{t+1}}{1+r} \qquad \forall t>0 && (\lambda_t) \\
& a_t \geq 0 \qquad \forall t>0 && (\mu_t) \\
& r^{m\prime}(B^o)>0
\end{aligned}
\tag{1}
$$

I describe a solution to the HH problem in Section (C.1) using the Kuhn-Tucker conditions. I assume a solution in which the credit constraint is non-binding after the initial savings decision, $\mu_t = 0 \ \forall \ t > 1$. Rearranging the first-order conditions, I obtain two necessary conditions for the behavior of the optimizing HH.

I also define what I term a "credit wedge", $\kappa_t^{t+1} \equiv \frac{MRS_t^{t+1}}{1+r_t}$, as the ratio between the optimizing household's subjective marginal rate of substitution and the market price ratio (or interest rate) between time t and $t+1$.[2] This "credit wedge" may be represented as $\kappa^{t+1} = \frac{\lambda_t/\lambda_{t+1}}{1+r_t} = 1 + \mu_{t+1}/\lambda_{t+1}$, using the Lagrange multipliers and first order conditions of the household's problem. The "credit wedge" may therefore be interpreted as present-valued willingness to pay out of future consumption for additional consumption today. It may also be interpreted as the shadow price of credit. Note that the requirement that $\mu_{t+1} \geq 0$ imposes that $\kappa^{t+1} \geq 1$.

With the "credit wedge" defined, I write the first-order conditions as follows:

$$P'(s)\Big|_{s^*} = \frac{u_s/u_c^1}{r}\frac{1}{\kappa^1} \tag{2}$$

$$\left[\frac{B^o r_m(B^o)}{r}\right]'\Bigg|_{B^{o*}} = \kappa^1 \tag{3}$$

Equation (2) interprets the information content in the optimizing household's choice of amenity level from the slope of the hedonic price schedule. This is depicted in Figure (1); households optimize by setting their indifference curves tangent to the offer curve in amenity space but their amenity curves are no longer only determined by wealth and preferences.

When the HH credit constraint is non-binding even in the initial savings decision ($\mu_1 = 0$ and $u_c^1 = u_c^0$), then the hedonic price schedule captures the capitalized and consumption-valued service flow from the marginal unit of the amenity, $\frac{u_s/u_c^0}{r}$. This is the motivation for "traditional" hedonic regressions.

However, when the HH credit constraint is initially binding, the slope of the price schedule is biased relative to the HH's WTP for the marginal amenity out of future, *un-constrained* consumption. The magnitude of the bias, somewhat intuitively, relates to the HH shadow price of credit, μ_1/λ_1, which describes how tightly the credit constraint binds. This suggests that "traditional" hedonic regressions may be biased. The influence of household credit constraints on their indifference curves is depicted in Figure (2).

Equation (3) interprets information content in the optimizing household's choice of initial mortgage balance. Note that the information is observable in $\frac{B^o r_m(B^o)}{r}$, which describes the menu of mortgage contracts available in the space of PDV of future payment obligations. Thus, there is a second price schedule, on the financing side, with a slope that captures information about the optimizing household. The mapping between menus in rates and menus in future mortgage obligations is depicted in Figure (3).

The slope of the mortgage contract menu captures the borrower's shadow value of credit, κ^1. This equivalence is intuitive. An optimizing household with multiple margins for credit must

[2] I write only $\kappa^{t+1} \equiv \kappa_t^{t+1}$ for clarity.

equate the price of additional credit along each margin. Because mortgage rates are increasing in initial balances, each marginal dollar of mortgage borrowing costs more. The optimizing household will increase mortgage borrowing until its cost exactly equals the shadow-cost of non-mortgage borrowing. Optimization along the margin of mortgage borrowing is depicted in Figure (4).

The slope of the mortgage contract menu captures the bias in the hedonic price schedule relative to the value of amenities out of unconstrained consumption. Here, the intuition is that the mortgage menu captures the trade-off between present and future consumption, which are assumed to be constrained and un-constrained, respectively. The distortion in the amenity valuation comes precisely because of the difference in value of constrained and unconstrained consumption.

3.2 Selection and Moral Hazard in the GSE Problem

In principle, it is possible that rising interest rates in the mortgage offer curve does not reflect any information about the willingness-to-pay of HHs for credit. Suppose, for instance, that as borrowers lever up, they suffer more from moral hazard and repay a lower fraction of their mortgage obligations. The rising interest rates then simply reflect the increased credit risk associated with lending.

There are two potentially mitigating factors here. First, if households do not anticipate the extent to which they will default on loan obligations, their choice of contract may still be informative about their credit constraints. Second, if there are heterogeneous households whom the GSEs cannot or do not distinguish (note that the LLPAs use fairly little information), then whatever the equilibrium pricing schedule, for good types who will not default, the choice of mortgage is still informative about credit constraints.

In this section, I further reconcile the presence of credit risk with the informative-ness of the offer curve slope by sketching a formal (but very rudimentary) model of GSE offer curve choice. Rather than considering moral hazard, I consider a setting of asymmetric information in which the GSE faces a distribution of borrower types who represent varying credit risks. Recent work suggests that selection rather than moral hazard plays a predominant role in driving the correlation between leverage choice and moral hazard (Gupta and Hansman, 2021).

Consider a benevolent lender, the GSE, who aims to maximize some social welfare function subject to a zero-profit condition. The lender faces a distribution of borrowers, f, who vary in their WTP for credit, κ_i, and the fraction of promised obligations they will actually repay, θ_i. The lender states an offer curve, $PDV(B^o)$, of mortgage payment obligations given an initial balance of borrowing and borrowers choose the contract that suits them best. We write:

$$
\begin{aligned}
\max_{PDV(B^o)} \; & U(\{B_i^o\}) \\
\text{s.t.} \quad & \int_i \theta_i PDV(B_i^o) f(\theta_i) d\theta_i = \int_i B_i^o f(\theta_i) d\theta_i \qquad (0\Pi) \\
& \theta_i PDV'(B_i^o) = \kappa_i \qquad (\mathrm{IC})
\end{aligned} \tag{4}
$$

Assuming that the social welfare function places an infinite penalty on redistribution, the lender is now restricted to offer curves that break even, in expectation, loan-by-loan. The zero-profit conditions becomes a more restrictive condition:

$$PDV(B^o)\theta(B^o) = B^o \tag{5}$$

I also assume that all borrowers have the same WTP for credit, $\kappa_i = \overline{\kappa} \; \forall i$. Using the (IC) constraint, it is now possible to solve for the type of borrower at a given point on the offer curve. Plugging this into the loan-wise zero-profit condition, we obtain:

$$\overline{\kappa} \frac{PDV(B^o)}{PDV'(B^o)} = B^o \tag{6}$$

This differential equation is straightforward to solve and the solution to the GSE's problem is then:

$$PDV(B^o) = (B^o)^{\overline{\kappa}} \tag{7}$$

For values of $\overline{\kappa} > 1$, this is convex, which captures a feature of the observed offer curve in the data. Note that as $\overline{\kappa} \to 1$, the curve becomes more and more linear. Although the borrower types drives the sorting behavior, it is the WTP for credit that drives the degree of convexity.

3.3 Mis-measurement and Under-investment in Local Amenities

In this section, I formalize the notion that a welfare-maximizing government without borrowing constraints underinvests in amenities if it infers household willingness-to-pay from the house price envelope without correcting for the extent of household credit constraints. The intuition is that the government can borrow and finance the investment out of future tax income, which falls on the households not when they purchase the home, but in periods when they are unconstrained. The households are therefore willing to pay more for additional amenities out of future taxes than they appear to be at the time they purchase a home.

In the model, households choose housing in municipalities of varying amenity levels. A price envelope forms in equilibrium that makes households indifferent between the municipalities, so that the housing market can clear. The (federal) government has access to an investment project that increases the amenity level of all municipalities. The cost of investment is convex in the quantity of additional amenities. The government may borrow to finance the investment and recover the costs from tax revenue at a later date. The government cannot observe household preferences directly, but they can measure the equilibrium prices in the housing market to infer information about preferences.

The equilibrium price envelope in the housing market depends on whether households are constrained or un-constrained at the time of home purchase. If households are un-constrained, then the average slope of the price envelope is a sufficient statistic for the optimal level of government investment. If households are constrained at the time of home purchase but the government treats them as though they are un-constrained, then it will under-invest in local amenities. In the case that households are constrained, the sufficient statistic for determining the optimal level of government investment is the slope of the price schedule corrected for the extent of credit constraints, the corrected willingness-to-pay put forward in this paper.

3.3.1 Model Set-up

The model contains a unit mass of municipalities, j, a unit mass of households, i. Each municipality has enough room to house a single household and is endowed with a uniformly distributed level of the amenity, $s_j \sim U[0,1]$. The households are identical and have preferences over consumption and the amenity, an endowment of income becoming available in each of two periods, $\{y_0, y_1\}$, and a savings technology with borrowing limit $-\phi$.

The (federal) government in the model has access to an investment project to improve amenity quality by some margin, σ, at cost $I(\sigma)$. It can borrow and save frictionlessly and can impose a uniform flat tax on all households to fund the investment. Finally, the government cannot observe household utility directly but can observe the equilibrium price envelope in the housing market to learn about preferences. This is analogous to the way in which economists estimate hedonic regressions or discrete choice models to measure household subjective willingness-to-pay for amenities.

The model has two periods, $\{0,1\}$, and the first period features two sub-periods, $\{0a, 0b\}$. At time $0a$, equilibrium is established in the housing market. Households choose municipalities and the equilibrium price schedule forms such that the market clears and households are no better off moving to a different municipality. Households then make their savings decisions and consume their time 0 consumption. At time $0b$, the government observes the price schedule, and chooses its investment and tax policy to maximize the welfare of households. Households do not anticipate this government intervention, nor do they re-optimize in response to it. At time 1, households use their income endowment and savings to pay their tax bill and consume the rest. Amenities also realize improvements due to government investment and households consume these improved amenities.

3.3.2 Housing Market Equilibrium

For the market to clear, municipalities must each host a single household. If two municipalities offer households different utility under any amenity price schedule, no household will prefer to live in the inferior municipality. This will create excess demand elsewhere in the market, implying that the amenity price schedule is not an equilibrium. We can therefore use the household problem to characterize the equilibrium pricing schedule, which must be set so that the optimizing household is indifferent between amenity choices.

Formally, an equilibrium price schedule, $\widetilde{P}(s)$, is implicitly defined by the following problem:

$$\begin{aligned} \overline{U} = \max_{\{c_0, c_1, a_1\}} \; & u(c_0) + \beta c_1 + v(s) \\ \text{s.t.} \qquad & y_0 = c_0 + \widetilde{P}(s) + \frac{a_1}{1+r} \qquad (\lambda_0) \\ & y_1 + a_1 = c_1 \qquad (\lambda_1) \\ & a_1 \geq -\phi \qquad (\mu_1) \end{aligned} \tag{8}$$

I solve for properties of the equilibrium amenity price schedule in Appendix (C.3). The slope of the equilibrium price schedule has the following property:

$$\widetilde{P}'(s) = \frac{v'(s)}{u'(c_0^*)} = \frac{v'(s)}{1 + \frac{\mu_1}{\lambda_1}} \tag{9}$$

Note that the equilibrium price schedule has the characteristics of the price schedule in the literature on amenities as amended by the household problem introduced in Section (3.1). Namely, it reflects the marginal willingness-to-pay for the amenity out of present non-housing consumption, which is constrained. Alternatively it is a downward biased measure of the willingness-to-pay for the amenity out of future, unconstrained non-housing consumption. (N.B. The marginal utility of future consumption before subjective time discount is 1.)

3.3.3 Optimal Government Investment

The government has access to a project that will improve amenity quality by some continuous margin, σ, at the expense of some convex investment costs, $I(\sigma)$ with $I', I'' > 0$. The government funds the investment through a flat tax that falls equally on all households and has frictionless access to financing. The government then chooses how much to invest in improving the amenity in order to maximize welfare subject to the constraint that it must raise the revenue required for the investment from taxes:

Formally, the government's problem is:

$$\max_{\{\sigma,t\}} \int_i U_i(\sigma,t)di$$

$$\text{s.t.} \qquad \int_i \frac{t}{1+r} di = I(\sigma) \qquad \text{(BC)}$$

$$\text{where:} \qquad U_i(\sigma,t) \equiv u(c_0^{i*}) + \beta(c_1^{i*} - t) + v(s^{i*} + \sigma)$$

$$\{c_0^{i*}, c_1^{i*}, a^{i*}, s^{i*}\} \equiv \arg \max_{\{c_0^i, c_1^i, a_1^i, s^i\}} u(c_0^i) + \beta c_1^i + v(s^i) \tag{10}$$

$$\begin{aligned} \text{s.t.} \qquad & y_0 = c_0^i + \widetilde{P}(s^i) + \frac{a_1^i}{1+r} && (\lambda_0) \\ & y_1 + a_1^i = c_1^i && (\lambda_1) \\ & a_1^i \geq -\phi && (\mu_1) \end{aligned}$$

$$s^{i*} = s^i = s_j \quad \text{for} \quad i = j$$

In Appendix (C.4), I obtain the following first-order condition for the government that pins down the optimal level of amenity improvement:

$$\sigma^* = I'^{-1}\Big(\mathbb{E}^i\big[\widetilde{P}'(s^{i*})(1 + \mu_1^i/\lambda_1^i)\big]\Big) \tag{11}$$

I also define the level of amenity improvements, σ^g, undertaken by a government with a potentially misspecified model. The government measures the slope of the price envelope but assumes that households are unconstrained at the time of home purchase. This level of amenities is given by:

$$\sigma^g \equiv I'^{-1}\Big(\mathbb{E}^i\big[\widetilde{P}'(s^{i*})\big]\Big) \tag{12}$$

In Appendix (C.4), I show that if households are truly unconstrained, and the government's assumption is correct, that the level of government investment is equal to the optimal level of investment, $\sigma^g = \sigma^*$. If the households are constrained, however, the government invests less than is optimal, $\sigma^g < \sigma^*$. The intuition is simply that the slope of the price envelope is made less steep when households are constrained at home purchase. Households pay for the government funded amenities out of future consumption rather than present consumption and are therefore willing to pay more for the amenities.

4 Institutional Setting

In this section, I describe the institutional setting of the agency loan market. In Section (4.1), I describe the loan-level pricing adjustments and private mortgage insurance requirements required by the GSEs. In Section (4.2), I assume that GSE fees are passed through to borrowers and describe how to construct the menu of contracts available to borrowers at the time of origination using borrower FICO scores; mortgage balance, loan-to-value and interest rate; LLPA grids; and insurance rate cards. In Section (4.3), I provide evidence that these fees are passed through from lenders to borrowers at the time of origination.

In describing the institutional setting, I document how requirements imposed by the GSEs at the level of the market generate mortgage menus at the level of the borrower. These menus correspond to the interest rate menu of the household problem, $r^m(B^o)$. In the empirical setting, though, the effective rate on the loan balance will be a combination of the rate due to the mortgage, r^m, and the rate due to the mortgage insurance, r^{mi}. Additionally, the rate menu in the empirical setting is a step-wise rather than continuous function. In line with the assumption of the household problem that $r^{m\prime} > 0$, the rate menu is an increasing function.

The description of the institutional setting here also facilitates the empirical analysis conducted in Section (7). Ultimately, I use information on the borrower's chosen mortgage relative to non-chosen alternatives to extract information about the extent of borrower credit constraints. I am able to construct borrower-level mortgage menus precisely because of the standardized and transparent way in which fees and PMI requirements are assigned and passed through.

4.1 Agency Loan LLPAs and Required PMI Premiums

In the secondary mortgage market, originators sell loans into collateral pools that are divided into tranches and sold to investors as mortgage-backed securities (MBS). The vast majority of these pools are "agency" pools, which require that loans enjoy a credit guarantee from the GSEs, Fannie Mae and Freddie Mac. In the event of borrower default, the guarantee ensures that the GSEs will step in to cover required principle and interest payments. This arrangement, backed implicitly by the fiscal power of the federal government, insulates the ultimate investors from credit risk and thereby supports the functioning of the mortgage market.

To be eligible for the GSE credit guarantee, a mortgage loan must be "conforming", that is, it must meet standards set forth by the GSEs. These standards include the so-called "jumbo" limit on the size of the mortgage, restrictions on the borrower credit score at origination, and limits on borrower debt-to-income ratios. The GSEs also require that loans with loan-to-value ratios above 80 must be covered by mortgage insurance and they specify the level of required coverage.

Originators must also pay a variety of fees mandated by the GSEs in order to sell mortgages into agency pools. They pay two fees to the GSEs: a one-time up-front insurance premium known as a loan-level price adjustment (LLPA) and an ongoing, monthly "g-fee" that is a small fraction of the loan balance. The GSEs permit originators to trade off between the two by "buying down" or "buying up" the "g-fee" at specified multiples. The GSEs also require originators to pay a 25bp minimum servicing fee to the party servicing the loan. (Fuster et al., 2013)

The LLPAs are determined by a wide variety of loan characteristics, the most prominent of which are the loan-to-value ratio on the loan and the FICO score of the borrower at origination. For vanilla, 30-year, fixed rate mortgages, these completely characterize the LLPA requirement and the mapping is published and periodically updated in what are known as the GSE "pricing grids". I hand-collect the contents of the Fannie Mae "pricing grids" from 2009 to present and plot an example in the left panel of Figure (5). Loans to less credit-worthy borrowers tend to have higher LLPAs. LLPAs are also increasing in the loan-to-value ratio of the loan, though only to the 80 LTV threshold at which PMI requirements kick in.

For loans above the 80 LTV threshold, the GSEs specify the amount of PMI coverage required for the loan to meet "conforming" standards. Insurers publish "rate cards" describing the premiums (as a percentage of loan balance) of monthly, borrower-paid mortgage insurance at various levels of coverage. These "rate cards" resemble the GSE "pricing grids" in that mortgage loan-to-value ratio and borrower credit score completely characterize the premium, given the required level of coverage. The binning of loan-to-value ratios and credit scores is, in fact, identical to that used to determine LLPAs. For one mortgage insurer, Essent, I hand-collect premiums from rate cards corresponding to the minimum insurance coverage required by the GSEs and plot an example in the right panel of Figure (5). Premiums are higher for less credit-worthy borrowers; there are no premiums required below the 80 LTV threshold but premiums are increasing in the loan-to-value ratio above this.

4.2 The Borrower-Level Mortgage Menu

Because of the central role of the agency market in the US context, the LLPA and PMI requirements define near market-wide menus for mortgage credit at the borrower level. In particular, information in the rate cards can be used to approximate the cost to a household of levering up on a mortgage in terms of the additional monthly expenses to cover both the LLPA and, possibly, additional PMI payments. Below, I detail how information in a borrower's realized mortgage loan, the GSE "pricing grid", and insurer "rate cards" can be used to approximate the menu of mortgage contracts available to the borrower at the time of home purchase.

Consider a borrower, i, with credit score, S_i, who is purchasing a house at price, P_i, and choosing the amount of initial loan balance, B_i^o, to take out to finance the transaction. Given the choice of loan balance, the loan will have a loan-to-value ratio, $L_i = \frac{B_i^o}{P_i}$. Given the loan balance and loan-to-value ratio, an interest rate will be assigned to the loan, r_i^m, and the borrower may also be required to pay monthly PMI premiums, r_i^{mi}. Given the initial balance, interest rate, and PMI obligations, the borrower will have monthly payment obligations, $(r_i^m + r_i^{mi})B_i^o$.

The GSE pricing grids and PMI rate cards return LLPAs and monthly premium rates, respectively, as a function of mortgage loan-to-value ratios and borrower credit scores. Both are comprised of a sequence of loan-to-value ratio limits, $\{\overline{L}_j\}_j$, and maximum credit scores, $\{\overline{S}_k\}_k$, that divide borrowers into bins. The borrower-level limit is the smallest one greater than the corresponding loan characteristic, $\overline{L}_i = \min_j\{\overline{L}_j \mid \overline{L}_j \geq B_i^o/P_i\}$ and $\overline{S}_i = \min_k\{\overline{S}_k \mid \overline{S}_k \geq S_i\}$.

Estimating the borrower's minimum PMI rates required to enter the agency market is fairly straightforward. GSE minimum PMI coverage requirements, $\underline{f}_t(\overline{L}_j)$, and insurer rate-card matrices, $r_t^{mi}(\overline{L}_j, \overline{S}_k, \underline{f}_t)$, define a matrix of minimum required insurance rates. The borrower's required PMI rate is then:

$$r_{it}^{mi}(B_i^o \mid P_i, S_i) = r_t^{mi}\left(\overline{L}_i, \overline{S}_i, \underline{f}_t(\overline{L}_i)\right) \tag{13}$$

Together, the base g-fee requirement, g_t^b, base LLPA, $LLPA_t^b$, excess LLPA matrix, $LLPA_t(\overline{L}_j, \overline{S}_k)$, and buy-up multiples, ϕ_t, imply a 0-UIP g-fee matrix. This matrix, g_t, specifies the g-fee required for a loan to enter the agency market without up-front payments according to the loan LTV and borrower credit score. This can be written as $g_t(B_i^o/P_i, S_i) = g_t^b + \phi_t * \left[LLPA_t^o + LLPA_t(\overline{L}_i, \overline{S}_i)\right]$. Assuming the lender passes along the 0-UIP g-fee to the borrower 1-for-1, and assuming the lender does not otherwise apply any risk-pricing to the loan (the g-fee, after all, secures insurance for the loan), we obtain the menu of mortgage interest rates available to borrowers for different balances at origination:

$$r_{it}^m(B_i^o \mid P_i, S_i) = r_i^{om} + g_t^b + \phi_t\left[LLPA_t^b + LLPA_t\left(\overline{L}_i, \overline{S}_i\right)\right] \tag{14}$$

In the data, we can observe the borrower's chosen mortgage contract, $\{B_{it}^{o*}, r_{it}^{m*}\}$. If the borrower were to borrow some alternate amount, B_i^o, against the same house, we can estimate the consequent mortgage rate. (Note that many GSE pricing parameters fall out in differences.)

$$r_{it}^m(B_i^o \mid P_i, S_i) = r_{it}^{m*} + \Delta_{B^{o*}}^{B^o} r_{it}^m = r_{it}^{m*} + \phi_t\left[LLPA_t(\overline{L}_i, \overline{S}_i) - LLPA_t(\overline{L}_i^*, \overline{S}_i)\right] \tag{15}$$

4.3 Empirical Evidence on LLPA Pass-through

To estimate household's MRS between present and future consumption, it would be best to observe the actual menu of contracts available to households when entering a bank. The approach of this paper, by comparison, is to construct a household-level choice set of financial contracts on the basis of parameters governing the market-rate securitization of mortgage contracts. Though feasible, this approach is rudimentary; after all, banks are not required pay the 0-UIP g-fee or to pass along g-fees to borrowers one-to-one.

Despite these limitations, I document that banks pass along the g-fees assessed by the GSEs according to the step-wise function in the LLPA matrix. Using a sample of loans sold to Fannie Mae (the "merged sample", described below), I examine interest rates in the cross section of LTV ratios. I capture this variation by regressing loan interest rates on LTV bins, including a large number of controls, including fixed effects for MSA, month of origination, DTI, FICO, race, sex, age, income percentile, and purchase price percentile:

$$r_i \sim \alpha + \sum_{\ell=50}^{97} \beta_\ell 1\{L_i = \ell\} + \gamma' X_i + \varepsilon_i \tag{16}$$

To emphasize the breaks where the LLPAs change, I also estimate a piecewise-polynomial defined function on the same data. The domain interval cut-points are defined by the GSE pricing grid, $\overline{L} \in \mathcal{L}_{GSE} = \{60, 70, 75, 80, 85, 90, 95, 97\}$. The estimated function is quartic and the level, cubic, and quartic terms are allowed to vary between interval. The regression specification is:

$$r_i \sim \alpha + \sum_{j=1}^{4} \beta_{j,50}(L_i - 50)^j + \sum_{\overline{L} \in \mathcal{L}_{GSE}} \sum_{j \in \{0,3,4\}} \beta_{j,\overline{L}} 1\{L_i \leq \overline{L}\}(L_i - \overline{L})^j + \gamma' X_i + \varepsilon_i \tag{17}$$

The results of this analysis are plotted in Figure (7). The step-wise change in rates at the cut-points is salient, particularly when including the high-dimensional fixed-effects as controls. It is also noteworthy that above the 80 LTV threshold, the interest rates charged on loans declines, consistent with pass-through of the declining LLPAs. Borrower monthly obligations in these contracts are higher on net, however, because of PMI obligations, which are changing at the 80 LTV threshold.

I also document considerable bunching in mortgage loans at the LTV thresholds at which the GSEs impose higher g-fees on originators or require additional PMI of borrowers in Figure (8). The bunching is consistent with the loan becomes discretely more expensive as borrowers lever up above these thresholds. The bunching suggests a pricing interpretation rather than an alternative interpretation that borrowers use heuristics and reference numbers, e.g. multiples of 5, when choosing their leverage in a home. In particular, bunching does take place at LTV 97 and not at LTV 65. Although 97 is not one of these candidate reference numbers, it is an LTV above which private mortgage insurers charge more for coverage. Conversely, 65 is a reference number but sees no increase in the GSE fees and exhibits no bunching. (Note that although the density is fairly slim, there is observable bunching below this, at 60).

The bunching appears to be particularly pronounced for loans that are sold to the GSEs ex-post. Comparing Figure (8a), which contains the CRISM Sample of all conventional, conforming loans, and Figure (8b), which contains only loans from HMDA that can be matched to the FNMA loan performance data, it is clear that there is more bunching for loans sold ex-post to the GSEs. To the extent that the originator knows ex-ante whether or not the conforming loan will in fact be sold to the GSEs, this makes sense.

5 Data

5.1 Sources

I use the following data sources in various analyses:

- *Credit Risk Insights Servicing McDash (CRISM), 2005-2020* consists of credit bureau data fields from Equifax merged to servicing records from McDash covering loan origination and performance. Since 2005, CRISM covers roughly 60% of mortgage originations (Adelino et al., 2013).

- *Fannie Mae (FNMA) Single-Family Loan Performance Data, 2005-2020* describes origination and performance characteristics of loans securitized through Fannie Mae.

- *Home Mortgage Disclosure Act (HMDA) Loan Application Register (LAR), 2018-2019* describes origination characteristics of mortgages subject to HMDA reporting requirements. The data has near universal coverage of the US mortgage market.

- *Fannie Mae (FNMA) Loan Level Pricing Adjustment (LLPA) Grids, 2009-2020* publicize the fees that originators must pay in order to obtain FNMA insurance (and securitize the loan through the GSEs). I hand collected grids dated to late 2012 and beyond from the Wayback Machine (https://archive.org/web/)[3].

- *Essent Private Mortgage Insurance (PMI) Rate Cards, 2011-2020* disclose the fees that borrowers must pay in order to obtain PMI through the insurer Essent. I hand collected these rate cards back to 2011 using links to historical rate cards available on Essent's web-page.

- *The Stanford Education Data Archive (SEDA), 2009-2018* is a data product constructed by the Educational Opportunity Project at Stanford University. It includes measures of average standardized test performance at the school level for the school years 2008-2009 to 2017-2018. State-level standardized test performance is converted so that measures may be compared nationally.

[3] Thank you to Andreas Fuster for providing me with grids from 2009-2012.

- *Census Shapefiles, 2010* are a census product describing geographic boundaries of administrative units including school districts and census tracts.

5.2 Sample Construction

I construct the following samples for analysis:

- The *FNMA Sample* is the primary sample in the analysis. It consists of all 30-year, fixed rate, conventional, conforming, first-lien mortgages originated between 2005 and 2020 for the purchase of single family, owner-occupied, 1-4 unit dwellings that were subsequently sold to Fannie Mae for securitization. I describe summary statistics of the FNMA Sample in Table (1). The characteristics of these mortgages are standard.

- I construct the *HMDA-FNMA-SEDA Merge* by (i) merging HMDA records to information on school district quality in SEDA (ii) merging FNMA records to information on FNMA LLPA grids and Essent rate cards and (iii) merging HMDA and FNMA records. My approach is similar to the approach in Bartlett et al. (2021) and uses the same data sources as the more recent Buchak and Jørring (2021). Below, I describe the sample contents and construction more closely.

  i. I subset HMDA records to retain only vanilla (non-negatively amortizing, non-balloon, non-interest only), conventional, conforming, fixed-rate, 30 year, first mortgages originated in 2018 or 2019 for purchase of single-family, 1-4 unit, owner-occupied housing. I use the 2010 Census Shapefiles to assign the Census Tract of a transaction in HMDA to a school district. I use SEDA to obtain district characteristics, including school performance.

  ii. I subset FNMA records in the same manner. From the FNMA LLPA grids, I record the LLPA rate for FNMA insured loans according to the date, loan LTV, and borrower FICO at origination. From the Essent rate cards, I record monthly premiums for borrower PMI according to the date, borrower LTV and FICO at origination. Using the loan origination month, LTV, and borrower FICO, I merge these to the subset of FNMA records.

  iii. To merge records, I require that loans share an identical state, MSA, 3-digit zip code, origination year, originator if available, and debt-to-income bin. I then conduct a fuzzy merge using the original balance, loan-to-value ratio, and rate at origination. The quality of the fuzzy merge along merge variables is depicted in Figure (12). The merge rate and representativeness of the merged sample is depicted in Table (2).

- The *CRISM Sample* consists of a 7% random sample of loans in CRISM that are vanilla (non-IO and non-Balloon), conventional, conforming, 30-year, fixed-rate, first mortgages originated between 2005 and 2020 against 1-4 unit, single-family, owner-occupied housing. I discard a handful of loans that are not onboarded promptly to CRISM or that have unreasonably high loan-to-value ratios. I describe summary statistics for the CRISM Sample in Tables (5) and (6).

6 Results I: The Average Shadow Price of Credit

6.1 The Mortgage Price Schedule

I construct a representation of the mortgage price schedule from data on mortgage originations in the *FNMA Sample*. Having linked these data to *Essent Rate Cards* using origination date, borrower FICO, and loan LTV, I can observe the following in the data: annual mortgage interest rate, r_i^m, minimum required annual PMI rate, r_i^{pmi}, home price at origination, P_i^o, and initial balance, B_i^o. I also choose a time-varying discount factor, r_t. I use the 30-year treasury yield at loan origination as a measure of a risk-free rate at the time horizon of the mortgage contract tenor. This choice is merely illustrative and I consider alternatives in the sensitivity analysis below.

For mortgage i originated at time t, I construct the present discounted value of mortgage obligations, PDV_i:

$$PDV_i = \sum_{\tau=1}^{360} \frac{m_i + pmi_{i,t+\tau}}{(1 + {r_t}/{12})^{\tau}} \tag{18}$$

These obligations are a function of monthly mortgage payments, m_i, and monthly mortgage insurance payments, $pmi_{i,t+\tau}$. Monthly mortgage payments are computed according to the fixed-rate mortgage amortization formula. PMI is no longer required when borrowers' current LTV reaches 80, so PMI payments are assumed to terminate when the outstanding principle reaches 80% of the original home price:

$$m_i = {r_i^m}/{12} * B_i^o * \frac{1}{1 - (1 + {r_i^m}/{12})^{-360}} \tag{19}$$

$$pmi_{i,t+\tau} = {r_{it}^{pmi}}/{12} * B_i^o * \mathbb{1}\left\{B_{i,t+\tau}/P_i^o > 0.80\right\} \tag{20}$$

Adjustments to the mortgage and PMI rates due to the LLPA and rate card grids are quoted relative to the loan-to-value ratio. This means that the menu of payment obligations is increasing in the mortgage balance only after fixing the price of the home. To ensure that I am depicting household locational choices within a given menu and not comparing borrowers along different menus, I bin the data according to home price percentiles. I plot a bin-scatter of mortgage payment obligations, PDV_i, against initial balances, B_i^o, within home price bin.

I present the results of this exercise for the 10th, 30th, 50th, 70th, and 90th quantile of the home price distribution in Figure (9). The bin-scatter has two salient features, a "backbone" and a series of "tails". The "backbone" consists of the lower portions of each mortgage price schedule, which are approximately collinear across home price bins. The presence of the backbone is consistent with a constant price of mortgage credit, so that marginal dollars are no more expensive. The "tails" consist of the upper portions of each mortgage price schedule, which are convex and slope more steeply upward and away from the "backbone". The tails are consistent with the increasing mortgage and PMI rates as the loan-to-value on the mortgage increases.

Section (3.1) suggests the interpretation that the shadow price of credit can be read off of the slope of these mortgage price schedules - the steeper the slope, and the more mass on steeper portions of the schedule, the higher is households' shadow price of credit. Loosely, two factors determine the steepness of the curves: the slope of the "backbone" and the convexity of the "tails". In a frictionless world, the present discounted value of mortgage obligations must always be exactly equal to the mortgage balance for any contract to be freely entered. The slope of the "backbone" would be unity and the convexity of the "tails" would be nil and the figure would trace a 45 degree line.

Mathematically, the slope of the "backbone" is influenced to a great extent by the choice of the discount factor. Economically, this discount factor should capture the relevant time horizon of the contract and account for the option value of the pre-payment option embedded in the mortgage contract. The convexity of the "tails" is influenced by the payments the borrower anticipates making relative to their mortgage obligations. Economically, to the extent that borrowers anticipate defaulting on mortgage obligations in certain bad states of the world, this may reduce the convexity of the curves.

6.2 Sensitivity Analysis of Average WTP for Credit

My approach to measuring the average willingness-to-pay for credit at the time of home purchase is to measure the average slope of the mortgage price schedule, weighted by the frequency of mortgages in the data appearing along it. My tack is to take adjacent points within home price quantile on the binscatter and find the slope of the line between the two. I then assign this slope to all mortgages in that price quantile with balances between those two points. Having assigned an approximate slope to each mortgage, I take a simple average and interpret it as willingness to pay for credit.

This procedure would accurately capture willingness-to-pay in the simple model I've written down to motivate the analysis. Capturing plausible estimates of willingness-to-pay for credit outside the model requires addressing additional factors in mortgage pricing. Two of the most salient are the default and prepayment options. While I do not build a sophisticated pricing model to account for these factors, I aim to address them by conducting a sensitivity analysis.

My sensitivity analysis proceeds as follows:

1. I choose a recovery rate on mortgage payment obligations, θ, and a discount rate, r_t

   i. I construct the present discounted value of mortgage obligations as:

   $$PDV_i = \theta \sum_{\tau=1}^{360} \frac{m_i + pmi_{i,t+\tau}}{\left(1 + r_t/12\right)^{\tau}} \tag{21}$$

   Where m_i and $pmi_{i,t+\tau}$ are defined as before.

   ii. For some p and q, I bin the data by p-tiles in home price, P_i^o, and q-tiles in initial balance, B_i^o

      I. For each p-tile, I compute local slopes between adjacent points in a bin-scatter.

      II. I compute the average slope as the observation-weighted average local slope.

   iii. I search for values of p and q where estimates of the average slope stabilize. (In particular, I use 100-tiles in P_i^o and 30-tiles in B_i^o.)

   iv. My estimate for a given (θ, r_t) pair is the estimate from this stabilized cell.

2. I report estimates for each (θ, r_t) pair.

I choose the following values for the recovery rate, θ, and discount rate, r_t. I use the 10-year treasury yield, $y_t^{(10)}$, to capture the common realized term of the contract given events of repayment and default. I use the 30-year treasury yield, $y_t^{(30)}$, to capture the term written on the contract. And I use an ad-hoc discount rate, r^*, which I compute at the 30-year treasury yield plus a spread. I choose the spread so that the "backbone" of the mortgage price schedule has a slope of ~1. Thus, I conservatively assume that prepayment risk in mortgages is perfectly priced.

I consider a recovery rate of 1 and 0.99. The recovery rate of 0.99 reflects the fact that 99% of outstanding balances are recovered on average, i.e. a foreclosure rate of 2% and a loss given default of 40% (An and Cordell, 2021). The recovery rate of 1 is counter-factual but underscores the fact that borrower impressions of their loan is what matters. If borrowers anticipate that they will repay 100% of their balances at the time of mortgage origination, then it may be more accurate to be somewhat less conservative.

I report results of this analysis in Table (3). For each combination of parameter choices, I report means and, in parentheses, standard deviations of the local slopes. The choice of discount factor tends to move the estimates by a greater extent than the choice of the recovery rate. The most conservative estimate, with discount rates at $r_t^* = y_t^{(30)} + 1.5\%$ and a recovery rate of $\theta = 0.99$ results in an average slope of 1.65. This suggests that, to a first degree of approximation, hedonic estimates may be downward biased by as much as 65%.

7 Results II: Correcting Hedonic Estimates

7.1 Distribution of Shadow Prices

I begin by constructing hypothetical mortgage contract menus facing borrowers at the time of origination. To do so, I am guided by the presentation of institutional details in Section (4.2) and evidence in Section (4.3). I assume that, given the price of the home, each borrower faces a menu of contracts comprised of initial balances and interest rates. The initial balances are determined by the loan-to-value thresholds used by the GSEs and the changes in the interest rate are pinned down by the LLPA adjustments and PMI rates. The level of the interest rates is pinned down by the contract chosen by the borrower that I can observe in the data.

This notion of a borrower-level menu is depicted in Figure (6). The borrower's choice of contract, observable in the data, is depicted in red. Two alternative contracts are constructed and depicted in blue. These represent what a borrower might plausibly have faced if they levered up or down.

I use these menus to estimate borrowers' shadow price of credit, i.e. their willingness-to-pay for an additional dollar of funds today out of funds tomorrow. The intuition of this exercise is to convert the menu of rates into a menu of payment obligations, as depicted in Figure (3), and find the slope at the point where the borrower locates. In practice, I consider a menu of finite options. I formalize the borrower's choice among these finite options to derive bounds on the borrower WTP for credit in Appendix (C.2). From Equation (35), we have the following, where j^* is the chosen mortgage contract and $j^* - 1$ and $j^* + 1$ are adjacent non-chosen alternatives.

$$\frac{r^m_{j^*-1}}{r}\left[1+\frac{\%\Delta^{j^*}_{j^*-1} r^m}{\%\Delta^{j^*}_{j^*-1} B^o}\right] \le \kappa_i \le \frac{r^m_{j^*}}{r}\left[1+\frac{\%\Delta^{j^*+1}_{j^*} r^m}{\%\Delta^{j^*+1}_{j^*} B^o}\right] \tag{22}$$

Some assumptions are required to implement this analysis, which I consider here.

- I can only obtain historical mortgage rate cards from Essent, so I must assume that these rates are fairly representative. Table (5) reports that Essent provides PMI for about 5% of loans in the mortgage market suggesting that they are a large but certainly not dominant player. In inspect rate cards from other insurers in more recent years, i.e. 2019-, and find that both insurance rates and the matrix structure of the cards are comparable.

- Similarly, I can only observe interest rates for borrower-paid mortgage insurance, so it must be the case that this is a common way of structuring the contract. Table (1) confirms this. Note that nearly half of borrowers have PMI and nearly all borrowers who have PMI have "borrower-paid" PMI, the standard form of insurance.

- Converting the LLPA fee reported in the FNMA grids to interest rate increases also requires assumptions. Technically, the grids contain up-front premiums assessed to the originator. Originators commonly convert these up-front obligations to a stream of payment obligations and then assess borrowers a higher rate to cover the additional expenses. The conversion multiple by which lenders convert the up-front premium into a g-fee are difficult to collect. Inspecting one reveals that the buy-up multiple fluctuates around 5-10. I use a factor of 10 to convert the LLPA fees to estimated interest rate rises faced by borrowers. I aim to err on the conservative side, suggesting that there is a fairly small rise in the interest rate faced by borrowers as they lever up. A smaller interest rate rise will correspond with lower estimated willingness to pay for credit.

- Again, I use the 30-year Treasury yield for the risk-free rate. This is motivated by (i) considering the borrower's alternative margin to be priced by a savings rate and (ii) the notion that the service quality is invariant to the state of the market and therefore that the relevant discount rate for pricing the value of these services is a risk-free rate. However, the value of mortgage obligations may be affected by taxes or default or prepayment propensity, and discounting at the 30-year Treasury rate does not account for these.

I implement the analysis suggested above to obtain lower and upper bounds on the shadow values of additional borrowing in this simplified framework. This distribution of lower bounds is plotted in Figure (10). The x-axis should be interpreted as willingness to pay in dollars to relax the borrowing constraint by a dollar. Households appear to be considerably credit constrained and there appears to be a high degree of heterogeneity, with a variance in the estimates of about 0.5.

The borrowers populating the CDF loosely increase in LTV along the x axis from left to right. This corresponds with the convexity of the mortgage-offer curve. The high upper bounds on the willingness-to-pay for credit, whose values fall between 6 and 8, come from the borrowers with LTVs of 97. Note that the increase in the rate for PMI at 97 is considerable and that these borrowers tend to have high interest rates on their mortgage already, both of which contribute to the high estimated upper bound.

7.2 Corrected Hedonic Estimates

I consider a cross-section of households, i, each purchasing a home at some date, $t(i)$, and located in a census tract, $k(i)$, school district, $d(i)$, and county, $c(i)$. I run a standard hedonic regression of home prices, $\ln P_{id}$, on measures of school quality at the district level, s_d, and other amenity measures at the census tract level including median income, rate of owner-occupancy, and minority percentage, $X_{k(i)}$. Finally, I include fixed effects for the county, $\alpha_{c(i)}$, and month of home purchase, $\alpha_{t(i)}$. The regression specification is below:

$$\ln P_{id} \sim \tilde{\alpha}_{c(i)} + \tilde{\alpha}_{t(i)} + \tilde{\beta} s_d + \tilde{\gamma}' X_{k(i)} + \varepsilon_{id} \tag{23}$$

The interpretation of the hedonic coefficient, as amended by the analysis in this paper, is the cross-sectional mean marginal willingness-to-pay for the amenity, downward biased by the credit constraint wedge:

$$\tilde{\beta} = \mathbb{E}^i\left[\frac{u_s^i/u_c^i}{r}\frac{1}{\kappa_i}\right] \tag{24}$$

The object of policy relevance, provided that government borrowing and taxation can be used to smooth the incidence of costs, is simply the marginal willingness-to-pay for the amenity, $\mathbb{E}^i\left[\frac{u_s^i/u_c^i}{rf}\right]$.

Because of the possibility of cross-sectional heterogeneity, correcting the bias in the estimate is not as simple as finding the average credit wedge. Using a covariance decomposition, the policy-relevant estimate can be shown to be:

$$E^i\left[\frac{u_s^i/u_c^i}{r}\right] = \left(E^i\left[\frac{1}{\kappa_i}\right]\right)^{-1}\left(E^i\left[\frac{u_s^i/u_c^i}{r}\frac{1}{\kappa_i}\right] + Cov^i\left(\frac{u_s^i/u_c^i}{r}, \frac{1}{\kappa_i}\right)\right) \tag{25}$$

Beyond estimating the credit wedge, it is necessary to know something about the population covariance between the marginal willingness-to-pay for amenities and credit in equilibrium.

My tack for correcting hedonic estimates is to exploit the fact that, having constructed menus facing borrowers, it is possible to estimate willingness-to-pay for credit at the borrower level. Effectively, I bin borrowers by their willingness-to-pay for credit, estimate hedonic coefficients and take a weighted average. In Section (D.2), I sketch the sense in which a single fixed-effects regression captures the object of interest.

Returning to the same cross section of households, I introduce a borrower's credit wedge, κ_i, as well as their credit wedge bin, $j(i)$. I normalize the school quality measures by the borrower credit wedge, $\frac{s_d}{\kappa_i}$, and I include fixed effects for the borrower's credit wedge bin, $\alpha_{j(i)}$.

$$\ln P_{id} \sim \alpha_{j(i)} + \alpha_{c(i)} + \alpha_{t(i)} + \beta\frac{s_d}{\kappa_i} + \gamma' X_{k(i)} + \varepsilon_{id} \tag{26}$$

The interpretation of this corrected regression coefficient, is the cross-sectional mean marginal willingness-to-pay for the amenity. This is the value of policy interest.

$$\beta = \mathbb{E}^i\left[\frac{u_s^i/u_c^i}{r}\right] \tag{27}$$

Table (4) presents the results of this regression exercise. In the first two columns, I conduct the exercise with limited controls and in the latter two I introduce more tract-level controls. In columns two and four I run the constraint correction specification. I cluster at the county level. In both cases, the estimates of mean marginal willingness-to-pay for amenities increases by ~ 50%.

As a caveat, traditional hedonic specifications are known not to be well identified, for one, because of unobserved covariates. Limitations on my ability to merge data prevent me from instrumenting in the manner most common in school quality applications, border discontinuities at within-district catchment area borders (due to Black (1999)). Instrumenting in this way tends to lower the estimate of willingness-to-pay for school quality. The exercise presented here is intended to suggest the effect of correcting for credit constraints on estimates of amenity value.

8 Conclusion

In this paper, I argue that frictions in household credit markets may systematically bias estimates of local public amenities derived from traditional hedonic regressions. Credit constrained households bid less for marginal amenities than they would in the absence of such constraints, flattening the hedonic price schedule of amenities. I also argue that the bias in hedonic regressions, which is related to the household's shadow price of credit, may be estimated from households' choice of mortgage product, because of the increasing costs of a marginal dollar of borrowing. I estimate the "credit wedge", or household willingness-to-pay for a marginal dollar of borrowing, in the US mortgage market. And I propose and (on a preliminary basis) implement a strategy for correcting bias in hedonic estimates that accounts for the cross-section heterogeneity.

References

Manuel Adelino, Kristopher Gerardi, and Paul S. Willen. Why don't Lenders renegotiate more home mortgages? Redefaults, self-cures and securitization. *Journal of Monetary Economics*, 60 (7):835–853, oct 2013. ISSN 03043932. doi: 10.1016/j.jmoneco.2013.08.002.

Xudong An and Larry Cordell. Mortgage loss severities: What keeps them so high? *Real Estate Economics*, 49(3):809–842, 2021. ISSN 15406229. doi: 10.1111/1540-6229.12334.

Michael Bailey, Eduardo Dávila, Theresa Kuchler, and Johannes Stroebel. House Price Beliefs And Mortgage Leverage Choice. *Review of Economic Studies*, 86(6):2403–2452, nov 2019. ISSN 0034-6527. doi: 10.1093/restud/rdy068. URL https://academic.oup.com/restud/article/86/6/2403/5194341.

Robert Bartlett, Adair Morse, Richard Stanton, and Nancy Wallace. Consumer-lending discrimination in the FinTech Era. *Journal of Financial Economics*, (xxxx), 2021. ISSN 0304405X. doi: 10.1016/j.jfineco.2021.05.047. URL https://doi.org/10.1016/j.jfineco.2021.05.047.

Patrick Bayer, Fernando Ferreira, and Robert Mcmillan. A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy*, 115(4), 2007. URL http://www.journals.uchicago.edu/t-and-c.

Matteo Benetton. Leverage Regulation and Market Structure: A Structural Model of the UK Mortgage Market. *The Journal of Finance*, page jofi.13072, aug 2021. ISSN 1540-6261. doi: 10.1111/JOFI.13072. URL https://onlinelibrary.wiley.com/doi/full/10.1111/jofi.13072.

Sandra E Black. Do Better Schools Matter? Parental Valuation of Elementary Education. *Quarterly Journal of Economics*, 114(2):577–599, 1999. URL https://academic.oup.com/qje/article-abstract/114/2/577/1844232.

Greg Buchak and Adam Jørring. Do Mortgage Lenders Compete Locally? Implications for Credit Access. 2021.

Kenneth Y Chay and Michael Greenstone. Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy*, 113(2):376–424, 2005. doi: 10.1086/427462.

Raj Chetty, Nathaniel Hendren, and Lawrence F Katz. The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment *. 2015.

Eric Chyn. Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children †. *American Economic Review*, 108(10):3028–3056, 2018. doi: 10.1257/aer.20161352. URL https://doi.org/10.1257/aer.20161352.

Susan Clampet-Lundquist and Douglas S Massey. Neighborhood Effects on Economic Self-Sufficiency: A Reconsideration of the Moving to Opportunity Experiment. *American Journal of Sociology*, 114(1):107–143, 2008.

Janet Currie, Lucas Davis, Michael Greenstone, and Reed Walker. Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings †. *American*

Economic Review, 105(2):678–709, 2015. doi: 10.1257/aer.20121656. URL http://dx.doi.org/10.1257/aer.20121656.

Anthony A. DeFusco and Andrew Paciorek. The interest rate elasticity of mortgage demand: Evidence from bunching at the conforming loan limit. *American Economic Journal: Economic Policy*, 9(1):210–240, feb 2017. ISSN 1945774X. doi: 10.1257/pol.20140108. URL https://doi.org/10.1257/pol.20140108.

Anthony A. Defusco, Stephanie Johnson, and John Mondragon. Regulating Household Leverage. *Review of Economic Studies*, 87(2):914–958, mar 2020. ISSN 0034-6527. doi: 10.1093/RESTUD/RDZ040. URL https://academic.oup.com/restud/advance-article/doi/10.1093/restud/rdz040/5542954https://academic.oup.com/restud/article/87/2/914/5542954.

Rebecca Diamond and Tim Mcquade. Who Wants Affordable Housing in Their Backyard? An Equilibrium Analysis of Low-Income Property Development. *Journal of Political Economy*, 127(3):1063–1117, 2019. URL http://www.journals.uchicago.edu/t-and-c.

Mehran Ebrahimian. Student Loans and Social Mobility. *SSRN Electronic Journal*, 2020. doi: 10.2139/ssrn.3680159.

Mark L Egan, Alexander MacKay, and Hanbin Yang. Recovering Investor Expectations from Demand for Index Funds. 2020. URL http://www.nber.org/papers/w26608.

Andreas Fuster, Laurie Goodman, David Lucca, Laurel Madar, Linsey Molloy, and Paul Willen. The Rising Gap between Primary and Secondary Mortgage Rates. *FRBNY Economic Policy Review*, (December):17–39, 2013. URL http://www.newyorkfed.org/research/epr/2013/1113fust.pdf.

Xuwen Gao, Ran Song, and Christopher Timmins. The Role of Information in the Rosen-Roback Framework. 2021. URL http://www.nber.org/papers/w28943.

Andreas Grunewald, Jonathan Lanning, David Low, Tobias Salz, Guy Aridor, Jasper Clarkberg, Thi Mai Anh Nguyen, and Yining Zhu. Auto Dealer Loan Intermediation: Consumer Behavior and Competitive Effects †. 2020. URL https://www.newyorkfed.org/medialibrary/interactives/householdcredit/d.

Arpit Gupta and Christopher Hansman. Selection, Leverage, and Default in the Mortgage Market. *The Review of Financial Studies*, 00(2012):1–51, 2021. ISSN 0893-9454. doi: 10.1093/rfs/hhab052.

Erik Hurst, Benjamin J. Keys, Amit Seru, and Joseph Vavra. Regional Redistribution Through the US Mortgage Market. *American Economic Review*, 106(10):2982–3028, 2016. ISSN 00028282. doi: 10.1257/aer.20151052.

Jeffrey Kling, Jeffrey B Liebman, and Lawrence F Katz. EXPERIMENTAL ANALYSIS OF NEIGHBORHOOD EFFECTS. *Econometrica*, 75(1):83–119, 2007.

Amrita Kulka. Sorting into Neighborhoods : The Role of Minimum Lot Sizes. 2019. URL https://drive.google.com/open?id=1Wr8T687wz-jVVMEVWoZCQXB53pxCM5hK.

Nicolai V. Kuminoff and Jaren C. Pope. Do "capitalization effects" for public goods reveal the public's willingness to pay? *International Economic Review*, 55(4):1227–1250, 2014. ISSN 14682354. doi: 10.1111/iere.12088.

Jens Ludwig, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, and Lisa Sanbonmatsu. NEIGHBORHOOD EFFECTS ON THE LONG-TERM WELL-BEING OF LOW-INCOME ADULTS. *Science*, 337(6101):1505, sep 2012. doi: 10.1126/SCIENCE.1224648. URL /pmc/articles/PMC3491569//pmc/articles/PMC3491569/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC3491569/.

Atif Mian and Amir Sufi. The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis *. *Quarterly Journal of Economics*, 124(4):1449–1496, nov 2009. ISSN 0033-5533. doi: 10.1162/qjec.2009.124.4.1449. URL https://academic.oup.com/qje/article-lookup/doi/10.1162/qjec.2009.124.4.1449.

Atif Mian and Amir Sufi. House Prices, Home Equity–Based Borrowing, and the US Household Leverage Crisis. *American Economic Review*, 101(5):2132–56, 2011. doi: 10.1257/aer.101.5.2132. URL http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.5.2132.

Scott T Nelson. Private Information and Price Regulation in the US Credit Card Market. 2018.

Amine Ouazad and Romain Rancière. Market Frictions, Arbitrage, and the Capitalization of Amenities. 2019. URL http://www.nber.org/papers/w25701.

Jennifer Roback. Wages, Rents, and the Quality of Life. *Journal of Political Economy*, 90(6):1257–1278, 1982.

Claudia Robles-Garcia. Competition and incentives in mortgage markets: the role of brokers. 2018. URL https://www.fca.org.uk/.

Sherwin Rosen. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 82(1):34–55, 1974.

A Tables

Table 1: Summary Statistics for the Fannie Mae Sample[†]

	N	$\bar{x}$	s_x	min	p25	p50	p75	max
Loan Characteristics:								
Loan Amount ($k)	52.6k	235	126	10	140	209	307	794
Property Value ($k)	52.6k	0.29	0.18	0.02	0.17	0.25	0.37	2.61
Interest Rate (%)	52.6k	4.41	0.87	2	3.88	4.25	4.88	8.62
LTV (%)	52.6k	83.6	13.1	5	80	85	95	97
DTI (%)	52.4k	35.1	9.55	2	28	36	42	64
FICO Score	52.5k	751	46.1	503	721	761	789	838
PMI Characteristics:								
Has PMI	52.6k	0.5	0.5	0	0	1	1	1
Has Borrower PMI	52.6k	0.47	0.5	0	0	0	1	1

The Fannie Mae Sample consists of loans from the Fannie Mae Single Family Loan Performance dataset with the following properties: originated 2005-2020, purchase loans, first liens, single-family, owner-occupied, 1-4 unit dwelling, conventional, conforming, fixed-rate, and 30 year terms. For computational ease, the summary statistics above are computed on a 1pp random sub-sample of the Fannie Mae Sample.

Table 2: Summary Statistics for HMDA-FNMA-SEDA Merge

	FNMA Sample† N=10066	HMDA Sample† N=27478	Merge† N=3523
Loan Characteristics:			
Loan Amount ($k)	245 (126)	261 (130)	225 (107)
Property Value ($k)	. (.)	322 (181)	328 (3538)
Interest Rate (%)	4.54 (0.53)	4.50 (0.59)	4.50 (0.51)
Total Costs ($k)	. (.)	4.21 (4.01)	3.48 (2.04)
Rate Spread (%)	. (.)	0.49 (0.58)	0.50 (0.49)
LTV (%)	86.2 (12.7)	84.5 (14.1)	86.8 (12.3)
DTI (%)	36.6 (9.05)	35.7 (9.44)	34.7 (9.09)
Has PMI: Yes	62.1%	.%	63.5%
PMI Coverage	26.2 (5.79)	. (.)	26.4 (5.58)
Purchaser Type:			
Balance Sheet	0.00%	49.3%	19.4%
FNMA	100%	50.7%	69.4%
Borrower Characteristics:			
Annual Income ($k)	. (.)	160 (2661)	108 (1059)
FICO Score	748 (45.1)	. (.)	750 (43.8)
Age (y)	. (.)	41.7 (13.7)	39.9 (13.5)
Ethnicity: Hispanic or Latino	.%	13.4%	11.3%
Race:			
Asian	.%	7.37%	3.96%
Black or African American	.%	5.10%	4.43%
White	.%	84.7%	89.6%
Sex:			
Female	.%	25.2%	25.1%
Joint	.%	41.1%	40.9%
Male	.%	33.7%	34.0%
First Home: Yes	55.6%	.%	56.7%
Tract Characteristics:			
Population (k)	. (.)	5.41 (3.63)	5.65 (3.31)
Minority Population (%)	. (.)	27.3 (23.9)	25.8 (22.0)
Median Income ($k)	. (.)	82.9 (41.4)	82.5 (28.3)
1-4 Unit Homeownership Rate (%)	. (.)	74.3 (14.8)	73.3 (14.4)
Median Age of Housing Units (y)	. (.)	31.6 (19.7)	37.2 (17.3)
District Characteristics:			
Enrollment (k)	. (.)	23.5 (52.0)	13.9 (20.6)
Minority Population (%)	. (.)	40.9 (27.0)	37.8 (25.9)
Median Income ($k)	. (.)	62.3 (19.2)	58.8 (15.4)
Test Performance (z)	. (.)	0.07 (0.31)	0.04 (0.31)
Segregation (W/B)	. (.)	0.13 (0.15)	0.12 (0.12)

I merge data from the Home Mortgage Disclosure Act (HMDA) Loan Application Record, the Fannie Mae (FNMA) Single Family Loan Performance data, and the Stanford Education Data Archive (SEDA). The SEDA data are merged to HMDA prior to the merge summarized above. The samples from HMDA and FNMA include loans with the following characteristics: originated 2018-2019, purchase loans, first liens, single-family, owner-occupied, 1-4 unit dwelling, conventional, conforming, vanilla (no balloon payments, no interest only payments, no negatively amortizing loans, no other non-amortizing features, no reverse mortgages), fixed-rate, and 30 year terms. For computational ease, the summary statistics above are computed on a 1pp random sub-sample of the FNMA, HMDA, and merged data.

Table 3: Sensitivity Analysis of Credit Wedge, $\kappa = 1 + \frac{\mu}{\lambda}$

Discount rate, r_t	Recovery rate, θ	
	1	0.99
$y_t^{(10)}$	2.017 (4.73)	1.996 (4.68)
$y_t^{(30)}$	1.996 (3.05)	1.976 (3.02)
r_t^*	1.665 (2.36)	1.648 (2.34)

Credit wedges are computed as the local slopes in binscatters of mortgage payment obligations against origination balances. The computational procedure is described in the text. For these results, the data are binned by 100-tiles in initial home price, P_i^o, and within this, 30-tiles in initial mortgage balance, B_i^o. Per equation (21), the value of mortgage obligations is computed as $\theta \sum_{\tau=1}^{360} \frac{m_i + pmi_{i,t+\tau}}{(1+r_t/12)^\tau}$, where θ is the fraction of payments borrowers expect to make and r_t is the choice of discount factor on obligations.

Table 4: 'Corrected' Hedonic Regressions

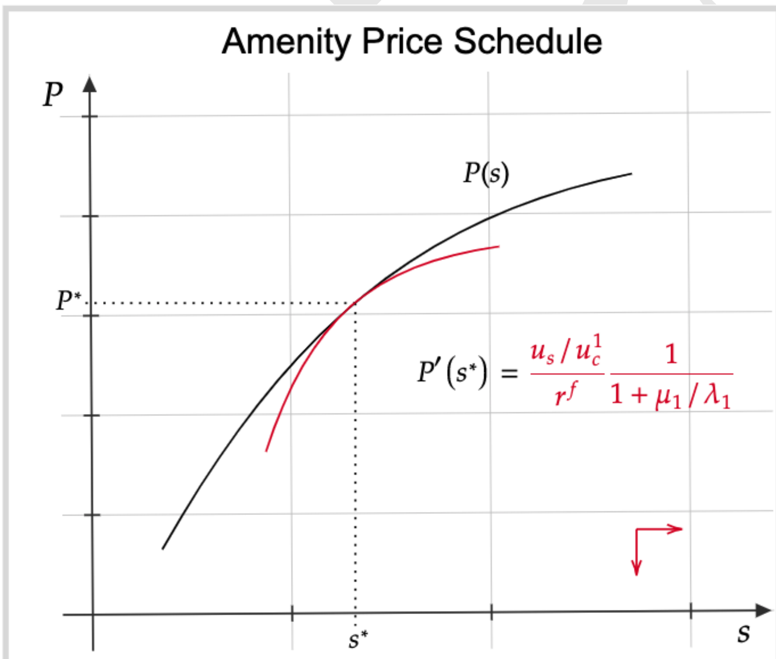
Dependent Variable:	ln Home Price$_{id}$			
	(1)	(2)	(3)	(4)
Variables				
Test Score$_d$	0.442***		0.156***	
	(0.017)		(0.013)	
Adj. Test Score$_{id}$		0.666***		0.230***
		(0.028)		(0.020)
Fixed-effects				
County FE	✓	✓	✓	✓
Month FE	✓	✓	✓	✓
Shadow Price FE		✓		✓
% Minority FE			✓	✓
q50 Income FE			✓	✓
% Ownership FE			✓	✓
Fit statistics				
Observations	382,407	382,407	382,407	382,407
R^2	0.461	0.474	0.554	0.566
Within R^2	0.065	0.049	0.009	0.006

Clustered (County) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

B Figures

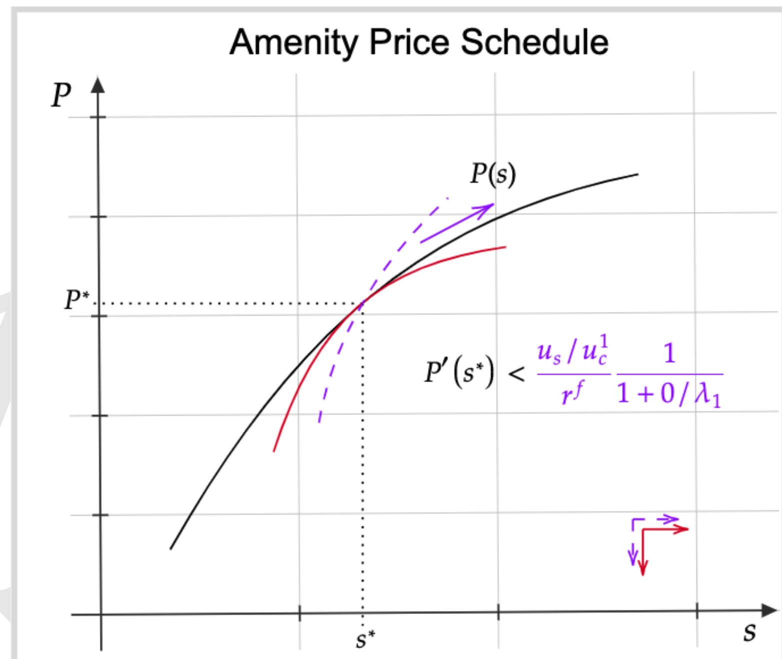
B.1 The household problem

Figure 1: Constrained household amenity optimization



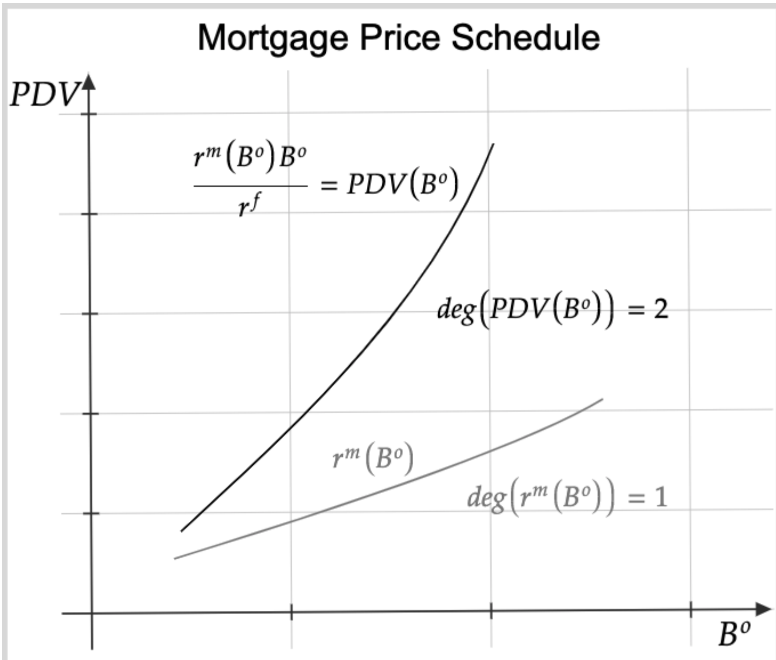
The housing market (black) consists of a menu of units, each characterized by an amenity level, s, and a price, P. The household prefers lower prices and higher levels of the amenity. It optimizes by setting its indifference curve (red) tangent to the amenity price schedule. Because households may be credit constrained at the time of home purchase, the slope of the market menu captures a downward biased version of their willingness-to-pay for the amenity.

Figure 2: Constrained household amenity re-optimization



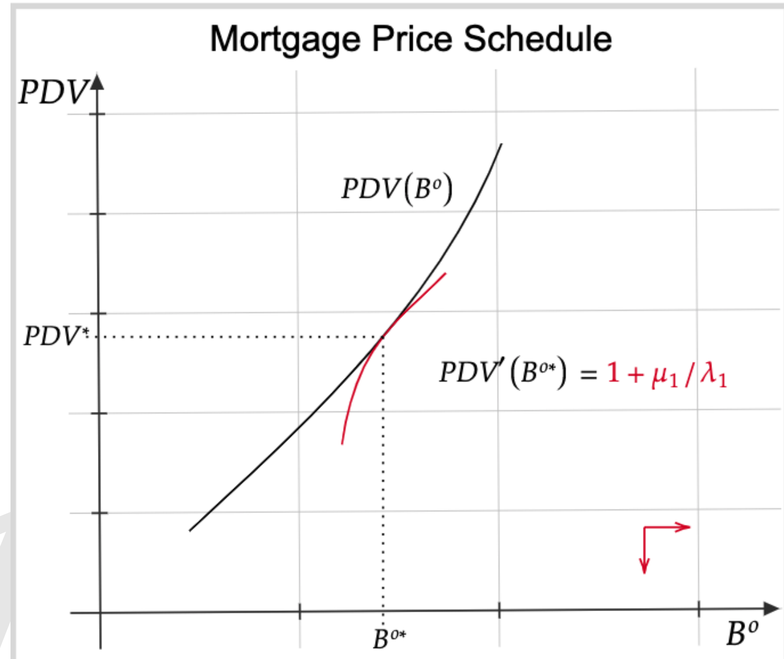
To recover the household's willingness-to-pay for marginal units of the amenity, it helps to consider a thought experiment in which the household solves its constrained optimization problem (red). Then, without changing its allocations, its credit constraint is lifted. The household's new indifference curves (dashed purple) are now steeper than the amenity price schedule. More precisely, the household's indifference curve grows steeper by a factor of $1 + \mu_1 / \lambda_1$.

Figure 3: Convexity of the mortgage price schedule



Borrowers are most often quoted the price of a loan in terms of an interest rate, thus facing a menu of loans varying in interest rate by initial balance (gray). This interest rate menu is equivalent to and more easily interpreted as a menu of loan balances and the present discounted value of payment obligations (black). Note that present discounted value scales with mortgage payment, which is the product of loan balance and interest rate. So, if interest rates are first-order in loan balances, payment obligations are second order, and therefore convex.

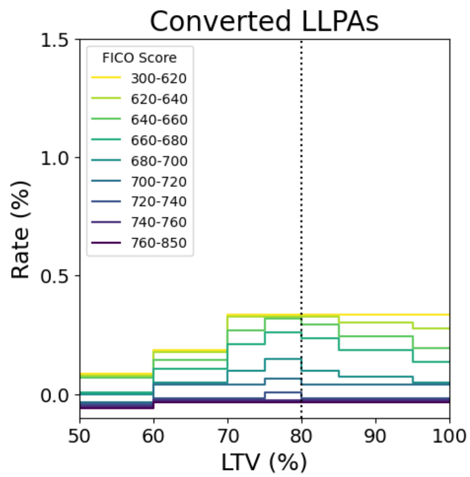
Figure 4: Constrained household mortgage optimization



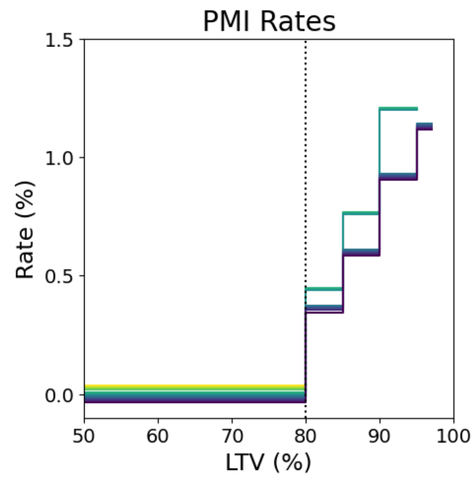
The mortgage market (black) consists of a convex menu of contracts, each defined by an initial balance, B^o, and the present discounted value of future payment obligations, PDV. The household prefers larger initial balances and lower payment obligations. It optimizes by setting their indifference curves (red) tangent to the market menu. The slope of the mortgage menu captures the household's shadow price of credit, $1 + \mu_1/\lambda_1$. This is precisely the downward bias of the amenity price schedule relative to household willingness-to-pay for amenities.

B.2 Institutional Setting

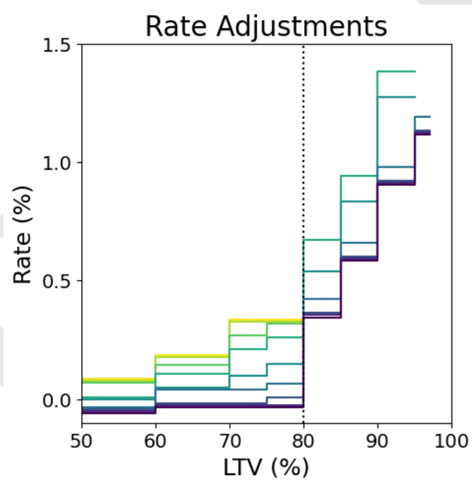
Figure 5: Example Pricing Regime (Jan.–Mar. 2011)



(a) Converted Fannie Mae LLPAs



(b) Essent PMI Rates



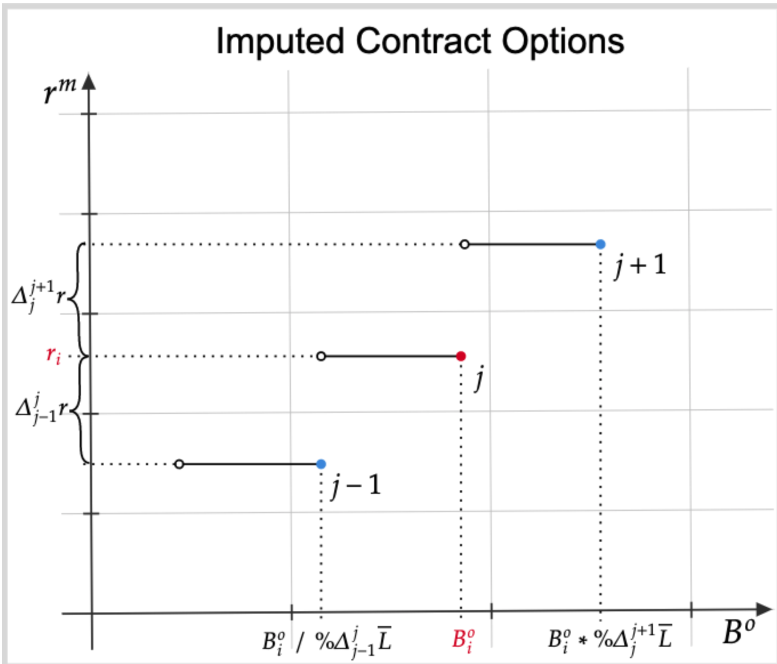
(c) Combined LLPAs and PMI Rates

This figure depicts a sample pricing regime in force from (as early as) January through March 2011.

- Figure (5a) contains LLPAs from the Fannie Mae pricing grid. These vary by mortgage LTV and borrower FICO at origination. I convert from LLPA to interest rate adjustment using a conservative factor of 10.
- Figure (5b) contains the annual PMI rates quoted on Essent rate cards for the minimum Fannie Mae PMI requirement. These also vary by LTV and borrower FICO.
- Figure (5c) contains the sum of the converted LLPAs and PMI rates. This corresponds to the menu of rate adjustment options faced by the borrower at origination of a mortgage.

The LLPAs are increasing to the 80 LTV threshold and thereafter constant or decreasing; the PMI rates are sharply increasing above 80 LTV. Because the rise in PMI more than offsets the fall in LLPAs, the combination is monotonically increasing.

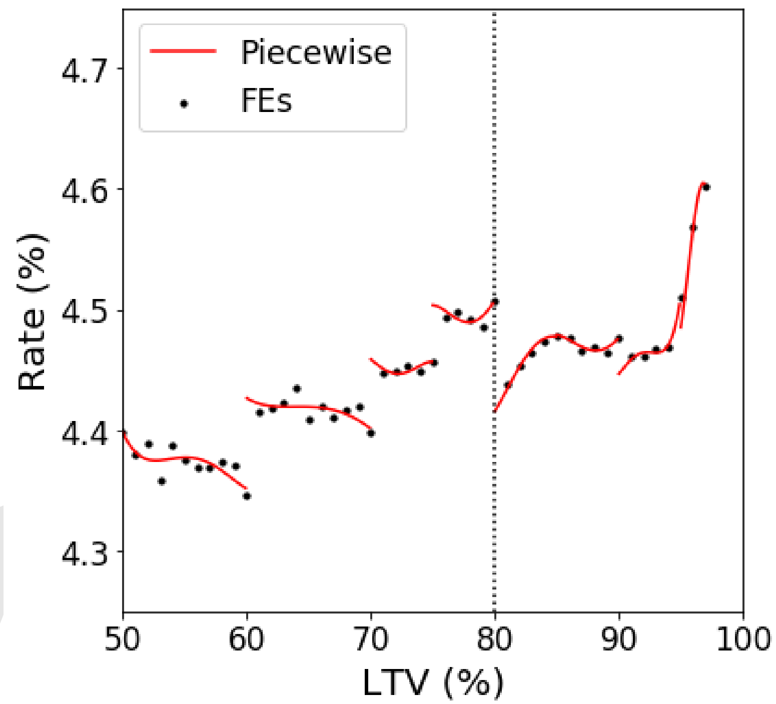
Figure 6: Imputation of borrower-level mortgage menus



I consider available alternative mortgage contracts for borrower i (red) using the borrower's chosen mortgage rate, r_i, and initial balance, B_i^o, the Fannie Mae LLPA grids, and Essent rate card grids. Considering the borrower's mortgage as contract j, I construct more and less levered alternatives, $j+1$ and $j-1$ respectively:

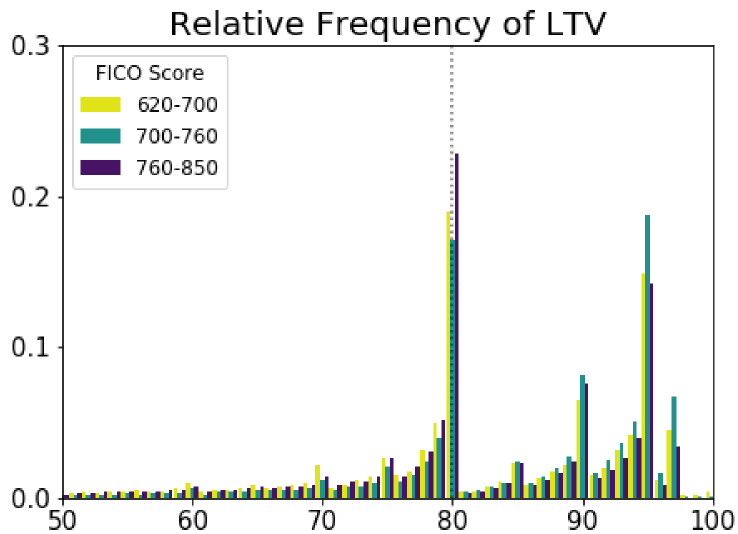
- The alternative available balance is the observed balance, B_i^o, scaled by the change in the maximum allowable leverage for the loan to qualify under a different grid cell, $\%\Delta_j^{j\pm1}\overline{L}$.
- The alternative contract rate is the observed rate, r_i, plus a component that captures the costs of covering LLPA fees or additional PMI required to remain conforming, $\Delta_j^{j\pm1}r$.

Figure 7: Pass-through of Loan-Level Pricing Adjustments

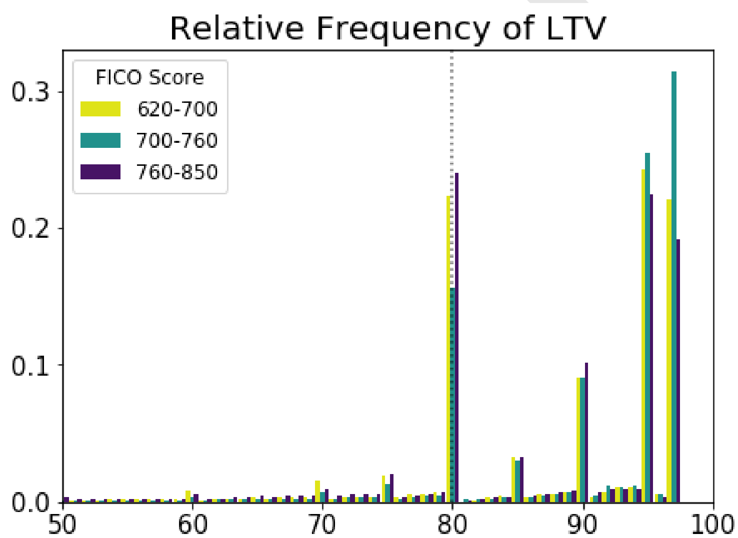


The figure above uses the FNMA-HMDA-SEDA Merged Sample to plot interest rates in the cross section of LTV at origination. The black dots are estimates from the regression with LTV bin fixed-effects; the red lines are estimates from a regression with a piecewise-polynomial defined function. I control flexibly for an array of confounders, including zip code, origination month, and borrower FICO. The rates offered by lenders reflects the step-wise adjustments to LLPAs established by the GSEs.

Figure 8: Bunching at LLPA and PMI thresholds



(a) CRISM Sample

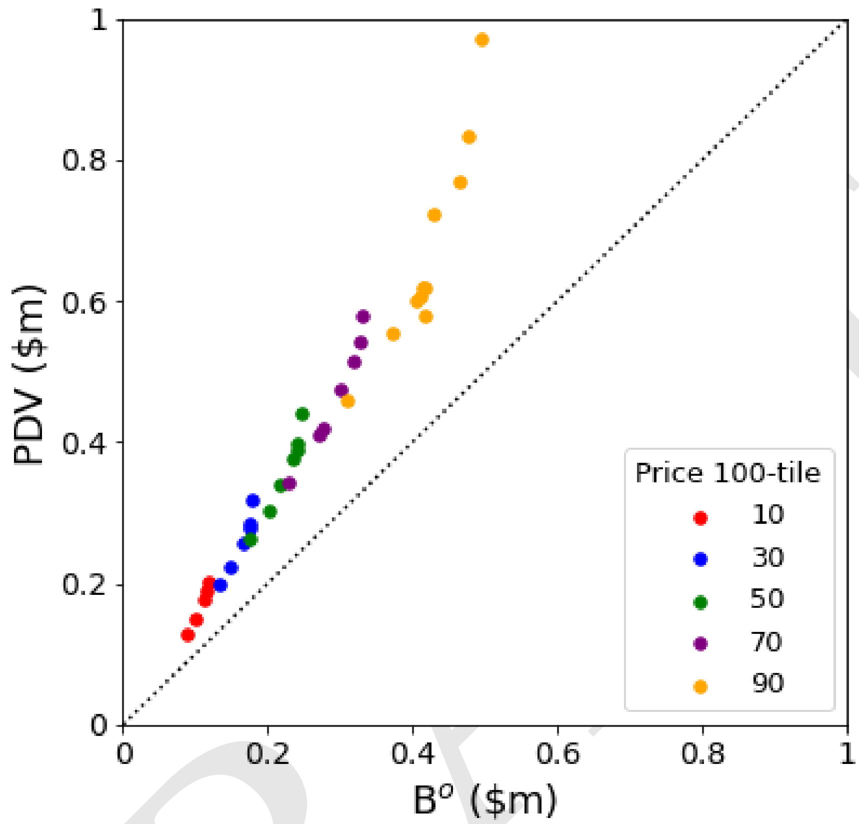


(b) Merged Sample

Histogram of loan-to-value ratios for loans in the CRISM Sample (8a) and FNMA-HMDA-SEDA Merged Sample (8b). The left plots show bunching over the entire distribution, the right plots show that it is present at all FICO levels. Bunching occurs at loan-to-value amounts where g-fees increase, insurance becomes required, and where insurance premiums increase. Notably, there is no bunching at 65 loan-to-value, a possible reference value where there is no increase in g-fees. Moreover, there is bunching at 97, unlikely to be a reference value that does entail an increase in mortgage insurance premiums or decrease in availability.

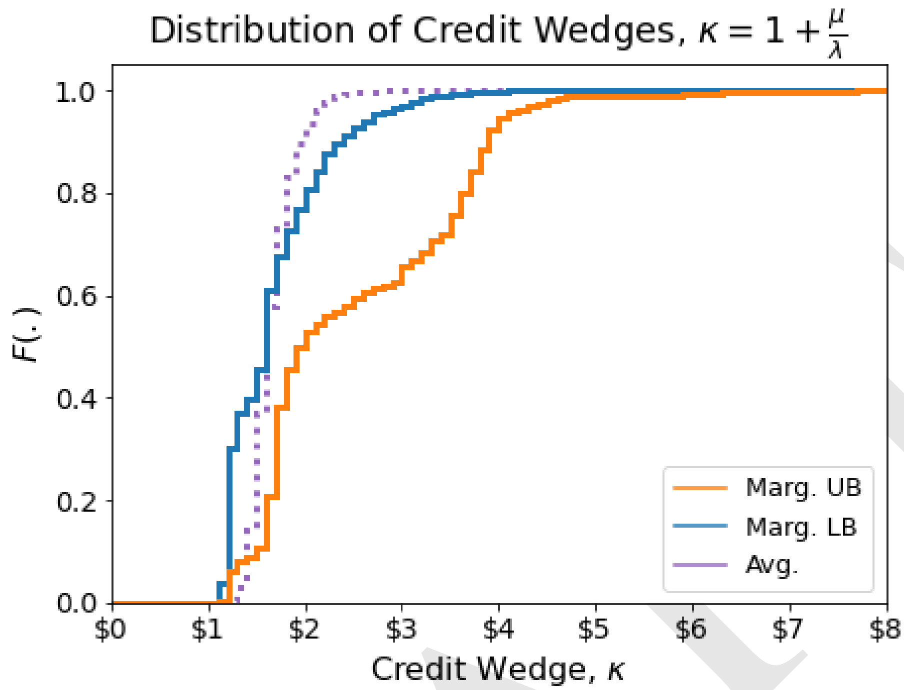
B.3 Results

Figure 9: Empirical Mortgage Price Schedule



This binscatter shows the value of mortgage obligations relative to the initial mortgage balance at different percentiles of the home purchase price distribution. Within each price percentile, the binscatter traces a convex menu, consistent with the interest rates that increase in loan-to-value. The lower portions of these convex menus are approximately collinear across price quantiles. The slope of these menus, where the borrower chooses to locate, corresponds to borrowers' willingness-to-pay for credit.

Figure 10: Distribution of Credit Wedges



This plot depicts the estimated distribution of credit wedges, κ, for borrowers in the HMDA-FNMA-SEDA merged sample. The blue and orange lines depict the CDFs of upper and lower bounds on the credit wedge, per Equation (22). Note that this formulation captures the borrower's WTP for the *marginal* dollar of credit. As a benchmark, the dotted purple line depicts the distribution of credit wedges assuming that the borrower's WTP for the *marginal* and *average* dollar of credit are the same. In particular, I compute a ratio of the borrower's effective interest rate to the 30-year treasury yield, $(r_i^m + r_i^{pmi})/y_t^{(30)}$. The WTP for the *average* dollar of credit understates the extent of credit constraints in roughly half the population.

C Derivations

C.1 Solution to the HH Problem

To solve the HH problem, I begin by forming the Lagrangian:

$$
\begin{aligned}
\mathcal{L}\big(\{c_t\}_t, \{a_t\}_t, s, B^o\big) \quad = \quad & u(c_0, s_0) + \sum_{t=1}^{\infty} \beta^t u(c_t, s) \\
& + \lambda_0 \Big[a_0 + w_0 - c_0 - [P(s) - B^o] - \frac{a_1}{1+r} \Big] \\
& + \sum_{t=1}^{\infty} \lambda_t \Big[a_t + w_t - c_t - B^o r^m(B^o) - \frac{a_{t+1}}{1+r} \Big] \\
& + \sum_{t=1}^{\infty} \mu_t a_t
\end{aligned}
\tag{28}
$$

By the Kuhn-Tucker theorem, the solution to the HH problem is $\big\{B^{o*}, s^*, \{c_t^*\}_t, \{a_t^*\}_t, \{\lambda_t\}_t, \{\mu_t\}_t\big\}$, where the household choice variables satisfy the following first-order conditions:

$$
\begin{aligned}
&\text{FOC}[c_t] \qquad & \beta^t u_c^t = \lambda_t \qquad & \forall t \geq 0 \\
&\text{FOC}[a_t] \qquad & \lambda_t + \mu_t = \frac{\lambda_{t-1}}{1+r} \qquad & \forall t > 0 \\
&\text{FOC}[B^o] \qquad & \lambda_0 = \sum_{t=1}^{\infty} \lambda_t [B^o r^m(B^o)]' \Big|_{B^{o*}} & \\
&\text{FOC}[s] \qquad & \lambda_0 P'(s^*) = \sum_{t=1}^{\infty} \beta^t u_s &
\end{aligned}
\tag{29}
$$

And the Lagrange multipliers satisfy:

$$
\begin{aligned}
\lambda_t > 0 \quad & \forall t \geq 0 \\
a_t > 0 \quad \text{or} \quad \mu_t > 0 \quad & \forall t > 0
\end{aligned}
\tag{30}
$$

I further assume that (i) the period utility function, u, is separable in the consumption good and amenity (ii) $\beta(1+r) = 1$ and (iii) the credit constraints are non-binding after the initial savings decision, $\mu_t = 0 \ \forall \ t > 1$.

I obtain the necessary conditions for household optimization as follows:

$$
P'(s^*) \overset{\text{FOC}[s]}{=} \frac{1}{r}\frac{u_s}{\lambda_0} \underset{\beta(1+r)=1}{\overset{\text{FOC}[a_1]}{=}} \frac{1}{r}\frac{u_s}{(1+r)(\lambda_1+\mu_1)} \underset{\beta(1+r)=1}{\overset{\text{FOC}[c_0]}{=}} \frac{1}{r}\frac{u_s}{u_c^1}\frac{1}{1+\frac{\mu_1}{\lambda_1}}
\tag{31}
$$

$$1+\frac{\mu_1}{\lambda_1}\overset{\text{FOC}[a_1]}{=}\frac{1}{1+r}\frac{\lambda_0}{\lambda_1}\overset{\text{FOC}[B^o]}{=}\frac{1}{1+r}\sum_{t=1}^{\infty}\frac{\lambda_t}{\lambda_1}\left[B^o r^m(B^o)\right]'\Bigg|_{B^{o*}}\overset{\text{FOC}[a_t]}{=}\left[\frac{B^o r^m(B^o)}{r}\right]'\Bigg|_{B^{o*}} \tag{32}$$

C.2 Estimating HH WTP for Credit

The consumer solves the following problem:

$$\begin{aligned}
&U_i = \max_j U_{ij} \\
&\quad \text{Where:} \quad U_{ij} = \max_{\{c_{ijt}, a_{ijt}\}_t} \sum_{t=0}^{T} \beta^t u(c_{ijt}) \\
&\qquad\qquad \text{s.t.} \quad c_{i0} = y_{i0} + a_{i0} - P_i + B^o_{ij} - \frac{a_{i1}}{1+r} && && (\lambda_0) \\
&\qquad\qquad\qquad c_{it} = y_{it} + a_{it} - B^o_{ij} r^m_{ij} - \frac{a_{i,t+1}}{1+r} && \forall t > 0 && (\lambda_t) \\
&\qquad\qquad\qquad a_{it} \geq 0 && \forall t > 0 && (\mu_t) \\
&\qquad\qquad = \max_j u(c^*_{ij0}) + \sum_{t=1}^{\infty} \beta^t u(c^*_{ijt}) \\
&\qquad\qquad \text{Where:} \qquad \beta^t u'(c^*_{ijt}) = \lambda_t && \forall t \geq 0 \\
&\qquad\qquad\qquad\qquad \lambda_t + \mu_t = \frac{\lambda_{t-1}}{1+r} && \forall t > 0 \\
&\qquad\qquad \text{And assume:} \qquad \mu_t = 0 && \forall t > 1
\end{aligned} \tag{33}$$

Comparing two financing methods, the consumer considers:

$$\begin{aligned}
U_{ij'} - U_{ij} \quad &= \quad u(c^*_{ij'0}) - u(c^*_{ij0}) + \sum_{t=1}^{\infty} \beta^t \left[u(c^*_{ij't}) - u(c^*_{ijt})\right] \\
&\overset{\text{Env. Thm.}}{=} \quad u'(c^*_{ij0})\left[c^*_{ij'0} - c^*_{ij0}\right] + \sum_{t=1}^{\infty} \beta^t u'(c^*_{ijt})\left[c^*_{ij't} - c^*_{ijt}\right] \\
&\overset{\text{FOC}[c]}{=} \quad \lambda_0 \Delta B^o - \sum_{t=1}^{\infty} \lambda_t \left[\Delta B^o \hat{r}^m_j + B^o_j \Delta \hat{r}\right] \\
&\overset{\text{FOC}[a]}{=} \quad (1+r)(\lambda_1 + \mu_1)\Delta B - \lambda_1 \left[\Delta B^o \hat{r}^m_j + B^o_j \Delta \hat{r}^m\right] \sum_{t=1}^{\infty} \frac{1}{(1+r)^{t-1}} \\
&= \quad (1+r)(\lambda_1 + \mu_1)\Delta B - \lambda_1 \left[\Delta B^o \hat{r}^m_j + B^o_j \Delta \hat{r}^m\right] \frac{1+r}{r} \\
&\propto \quad \Delta B^o - \frac{\Delta B^o \hat{r}^m_j + B^o_j \Delta \hat{r}^m}{r} + \frac{\mu_1}{\lambda_1} \Delta B^o
\end{aligned} \tag{34}$$

The difference in utility is proportional to an expression that is intuitive. The first term, ΔB^o, is the additional funds supplied by the contract j' relative to j. The second term, $\frac{\Delta B^o \hat{r}^m_j + B^o_j \Delta \hat{r}^m}{r}$, is the objectively discounted value of the additional mortgage payment obligations demanded by the contract j' relative to j. The borrower also benefits from the third term, $\frac{\mu_1}{\lambda_1}\Delta B^o$, which makes the additional funds from j' more attractive to the extent that the borrower is borrowing constrained.

For contracts j and j', we have:

$$j' > j \quad \Longleftrightarrow \quad \kappa^1 > \frac{\hat{r}^m_j}{r}\left(1 + \frac{\%\Delta_j^{j'}\hat{r}^m}{\%\Delta_j^{j'}B^o}\right) \tag{35}$$

Where everything on the right side of the > can be observed in the data. We can infer something about the extent of borrower's credit constraints from information about their choice set, which is made available in the GSE and private mortgage insurance pricing grids.

C.3 Housing Market Equilibrium

To solve for the implicitly defined equilibrium prices, I solve the HH problem over all variables but the choice of amenity level. In particular, I begin by forming the Lagrangian:

$$\begin{aligned} \mathcal{L}(c_0, c_1, a_1) \quad = \quad & u(c_0) + \beta c_1 + v(s) \\ & + \lambda_0\Big[y_0 - c_0 - \widetilde{P}(s) - \frac{a_1}{1+r}\Big] \\ & + \lambda_1\Big[y_1 + a_1 - c_1\Big] \\ & + \mu_1\Big[a_1 + \phi\Big] \end{aligned} \tag{36}$$

By the Kuhn-Tucker theorem, the solution to the HH problem is $\{c_0^*, c_1^*, a_1^*, \lambda_0, \lambda_1, \mu_1\}$, where the household choice variables satisfy the following first-order conditions:

$$\begin{aligned} \text{FOC}[c_0] \quad & u'(c_0^*) = \lambda_0 \\ \text{FOC}[c_1] \quad & \beta = \lambda_1 \\ \text{FOC}[a_1] \quad & \lambda_1 + \mu_1 = \frac{\lambda_0}{1+r} \end{aligned} \tag{37}$$

And the Lagrange multipliers satisfy:

$$\begin{aligned} \lambda_t > 0 \quad & t \in \{0,1\} \\ a_1 > 0 \quad \text{or} \quad \mu_1 > 0 \quad & \forall t > 0 \end{aligned} \tag{38}$$

I further assume that $\beta(1+r) = 1$.

I solve for the equilibrium price schedule in two regimes. In case 1, I assume that households are all unconstrained at the time of home purchase. In case 2, I assume that all households are at their borrowing limits at the time of home purchase.

- **Case 1** I use the household's first order conditions, budget constraints, and the assumption that the household's shadow price of credit is zero, $\mu_1 = 0$, to obtain expressions for optimal consumption:

$$c_0^* \overset{\text{FOC}[c_0]}{=} u'^{-1}(\lambda_0) \underset{\mu_1=0}{\overset{\text{FOC}[a_1]}{=}} u'^{-1}\big((1+r)\lambda_1\big) \overset{\text{FOC}[c_1]}{=} u'^{-1}\big((1+r)\beta\big) \overset{\beta(1+r)=1}{=} u'^{-1}(1) \tag{39}$$

$$c_1^* \overset{(\lambda_1)}{=} y_1 + a_1^* \overset{(\lambda_0)}{=} y_1 + (1+r)\big(y_0 - c_0^* - P(s)\big) \overset{w_0 \equiv y_0 + \frac{y_1}{1+r}}{=} (1+r)\big(w_0 - P(s) - c_0^*\big) \tag{40}$$

I obtain the implied equilibrium price schedule by plugging optimal consumption levels into the household's objective and rearranging:

$$\widetilde{P}(s) = w^o - u'^{-1}(1) - \Big[\overline{U} - u\big(u'^{-1}(1)\big) - v(s)\Big] \tag{41}$$

Taking derivatives, we obtain the relation between the slope of the price schedule and the household willingness-to-pay for the amenity:

$$\widetilde{P}'(s) = \frac{v'(s)}{1} \overset{u'(c_0^*)=1}{=} \frac{v'(s)}{u'(c_0^*)} \tag{42}$$

- **Case 2** I use the requirement that $a_1^* = 0$ and the household period budget constraints to obtain expressions for optimal consumption:

$$c_0^* = y_0 - \widetilde{P}(s) + \frac{\phi}{1+r} \tag{43}$$

$$c_1^* = y_1 - \phi \tag{44}$$

Again, I obtain the implied equilibrium price schedule by plugging optimal consumption into the objective function and rearranging:

$$\widetilde{P}(s) = y_0 + \frac{\phi}{1+r} - u^{-1}\Big(\overline{U} - \beta(y_1 - \phi) - v(s)\Big) \tag{45}$$

Now, the slope of the equilibrium price schedule reflects:

$$\widetilde{P}'(s) = \frac{v'(s)}{u'(c_0^*)} = \frac{v'(s)}{1 + \frac{\mu_1}{\lambda_1}} = \frac{v'(s)}{\kappa^1} \tag{46}$$

Note that although these were solved as two cases, the characterization of the equilibrium price schedule slope under Case 2 nests the result in Case 1 as a special case. In particular, an unconstrained household has $\mu_1 = 0$ and the denominator simplifies to 1.

C.4 The Government's Problem

To solve the government's problem, I begin by rewriting its objective function:

$$\begin{aligned}
\int_i U_i(\sigma,t)di &\overset{U_i\text{ def.}}{=} \int_i \left[u(c_0^{i*}) + \beta(c_1^{i*} - t) + v(s^{i*} + \sigma)\right]di \\
&\overset{\text{Taylor}}{=} \int_i \left[u(c_0^{i*}) + \beta(c_1^{i*}) + v(s^{i*})\right]di + \int_i \left[\sigma v'(s^{i*}) - \beta t\right]di \\
&\overset{\widetilde{P}(s)\text{ def.}}{=} \int_i \overline{U} di + \int_i \left[\sigma v'(s^{i*}) - \beta t\right]di \\
&\overset{\int_i di=1;\, \mathbb{E}\text{ def.}}{=} \overline{U} + \sigma \mathbb{E}^i\left[v'(s^{i*})\right] - \beta t \\
&\overset{\widetilde{P}'(s)=\frac{v'(s)}{\kappa^1}}{=} \overline{U} + \sigma \mathbb{E}^i\left[\widetilde{P}'(s^{i*})\kappa_i^1\right] - \beta t \\
&\overset{\text{(BC)}, \beta(1+r)=1}{=} \overline{U} + \sigma \mathbb{E}^i\left[\widetilde{P}'(s^{i*})\kappa_i^1\right] - I(\sigma)
\end{aligned} \tag{47}$$

The first-order condition is now:

$$\text{FOC}[\sigma] \qquad \mathbb{E}^i\left[\widetilde{P}'(s^{i*})\kappa_i^1\right] = I'(\sigma^*) \tag{48}$$

The optimal amount of amenity improvement is:

$$\sigma^* = I'^{-1}\left(\mathbb{E}^i\left[\widetilde{P}'(s^{i*})\kappa_i^1\right]\right) \tag{49}$$

Suppose the government chooses the level of investment and amenity improvement according to traditional hedonic estimates. Then it chooses the following level of amenity improvement:

$$\sigma^g = I'^{-1}\left(\mathbb{E}^i\left[\widetilde{P}'(s^{i*})\right]\right) \tag{50}$$

In the case that households are unconstrained, the government investing according to traditional hedonic estimates chooses the optimal level of investment and amenity improvement:

$$\sigma^g \overset{\sigma^g\ def.}{=} I'^{-1}\left(\mathbb{E}^i\left[\widetilde{P}'(s^{i*})\right]\right) \overset{\mu_1=0}{=} I'^{-1}\left(\mathbb{E}^i\left[\widetilde{P}'(s^{i*})\kappa_i^1\right]\right) \overset{\sigma^*\ def.}{=} \sigma^* \tag{51}$$

In the case that households are constrained, the government investing according to traditional hedonic estimates chooses amenity improvement below the optimal level:

$$\sigma^g \overset{\sigma^g\, def.}{=} I'^{-1}\Big(\mathbb{E}^i\Big[\widetilde{P}'(s^{i*})\Big]\Big) \overset{\mu_1>0;\, I''>0}{<} I'^{-1}\Big(\mathbb{E}^i\Big[\widetilde{P}'(s^{i*})\kappa_i^1\Big]\Big) \overset{\sigma^*\, def.}{=} \sigma^* \tag{52}$$

D “Financial” Hedonic Regression

D.1 Framework

Consider a financial product, f, characterized by a price, p^f, and state- and time- indexed payoffs, d_{st}^f, for dates $t \in \{1,\ldots,T\}$ and states $s \in \{1,\ldots,S\}$ with realization probabilities, $\{\pi_{st}\}_{s\times t}$. And consider a menu of available products, F. For simplicity, assume the space of available products is continuous. For a given household, i, with stochastic discount factor, M_t^i, the surplus from the choice of product f is:

$$\Pi^{if} = \sum_{t=1}^{T}\sum_{s=1}^{S}\pi_{st}E[M_t^i \mid s]d_{st}^f - p^f \tag{53}$$

The optimizing household will choose financial product, f^*, so that the marginal surplus from adjusting any of the payoffs is zero:

$$\frac{\partial}{\partial d_{st}^f}\Pi^{if}\Bigg|_{f^*} = \frac{\partial}{\partial d_{st}^f}\Big[\sum_{t=1}^{T}\sum_{s=1}^{S}\pi_{st}E[M_t^i \mid s]d_{st}^f - p^f\Big]\Bigg|_{f^*} = \pi_{st}E[M_t^i \mid s] - \frac{\partial p^f}{\partial d_{st}^f}\Bigg|_{f^*} = 0 \tag{54}$$

The slope of the menu of financial products where the borrower chooses to locate reveals the borrower’s state-price. This motivates a “financial” hedonic regression of the form:

$$p_i^f \sim \alpha + \beta' d_i^f + \varepsilon_i \tag{55}$$

The regression coefficients then recover the cross sectional average household state price.

$$\beta_{st} = E^i\Big[\pi_{st}E[M_t^i \mid s]\Big] \tag{56}$$

In the simple HH problem introduced above, the financial product is the mortgage. The price is the time-1 discounted value of mortgage obligations, $\frac{B^o r_m(B^o)}{r_f}(1+r^f)$, and the payoff is the time-0 mortgage balance, B^o. The household’s discount factor from time-0 forward to time-1 is deterministic and takes the form $1/M_1^i = \lambda_0^i/\lambda_1^i = (1+r^f)[1+\mu_1^i/\lambda_1^i]$.

Discounting mortgage obligations instead to time-0 eliminates the $(1+r^f)$ term in both the price and the regression coefficient. This suggests regression specification:

$$\frac{B_i^o r_m(B_i^o)}{r_f} \sim \alpha + \beta B_i^o + \varepsilon_i \tag{57}$$

Where the regression coefficient recovers:

$$\beta = E^i\Big[1 + \mu_1^i / \lambda_1^i\Big] \tag{58}$$

D.2 Correcting "Traditional" Hedonic Estimates

The analytical framework in this paper suggests (i) that traditional hedonic estimates actually measure household willingness-to-pay for amenities with a bias term related to household willingness-to-pay for credit and (ii) household willingness-to-pay for credit at home purchase can be measured from information on household mortgage choice. In this section, I describe how to use the information on household mortgage choice to correct hedonic estimates to recover the estimate of policy relevance, the unbiased household willingness-to-pay for amenities.

Cross-sectional heterogeneity in households poses a challenge to correcting hedonic estimates. It is not enough to know the mean willingness-to-pay for credit in order to correct the bias. A variance decomposition shows the relationship between the biased and unbiased estimates of willingness-to-pay for amenities:

$$E^i\Big[\frac{u_s^i/u_c^i}{rf}\Big] = \left(E^i\Big[\frac{1}{1 + \mu_1^i/\lambda_1^i}\Big]\right)^{-1}\left(E^i\Big[\frac{u_s^i/u_c^i}{rf}\frac{1}{1 + \mu_1^i/\lambda_1^i}\Big] + Cov^i\Big(\frac{u_s^i/u_c^i}{rf}, \frac{1}{1 + \mu_1^i/\lambda_1^i}\Big)\right) \tag{59}$$

Directly computing the covariance term, $Cov^i\Big(\frac{u_s^i/u_c^i}{rf}, \frac{1}{1+\mu_1^i/\lambda_1^i}\Big)$, poses a challenge in particular, because amenity price schedules are often not well-defined, making it difficult to compute borrower-level estimates of willingness-to-pay for amenities. Hedonic regressions, for instance, settle for population averages.

My approach is to exploit the fact that the menu of options for household mortgage choice is more transparent and that willingness-to-pay for credit can be estimated at the household level, per Section (C.2). I also use the intuition that the logic of hedonic regression, that it captures the slope of the price schedule at the place where borrowers locate on it, applies equally well to subsets of the population as it does to the population as a whole.

I consider a cross-section of households, i, located in various school districts, d. I can measure their willingness to pay for credit at the time of home purchase, $\kappa_i = 1 + \mu_1^i/\lambda_1^i$, from their mortgage menu and mortgage choice, and I bin them according to the size of this wedge, $j(i)$. I regress the price of the household's home, P_{id}, on a measure of the school quality in the district, s_d, normalized by the household's willingness to pay for credit at the time of home purchase, κ_i. I also include fixed effects for the credit wedge bin, $\alpha_{j(i)}$:

$$P_{id} \sim \alpha_{j(i)} + \beta \frac{s_d}{\kappa_i} + \varepsilon_{id} \tag{60}$$

Intuitively, with the presence of the fixed-effects, the identifying variation comes from within credit wedge bins. The within-bin variation is roughly captured by a within-bin regression slope coefficient and these are weighted by the proportion of the sample size in each bin. This suggests the following representation of the regression slope coefficient:

$$\beta = \sum_j Pr(\kappa_i = \overline{\kappa}_j) \frac{Cov(P_{id}, \frac{s_d}{\kappa_i} \mid \kappa_i = \overline{\kappa}_j)}{Var(\frac{s_d}{\kappa_i} \mid \kappa_i = \overline{\kappa}_j)} \tag{61}$$

Using the conditioning information and properties of variances and covariances, I simplify:

$$= \sum_j \overline{\kappa}_j Pr(\kappa_i = \overline{\kappa}_j) \frac{Cov(P_{id}, s_d \mid \kappa_i = \overline{\kappa}_j)}{Var(s_d \mid \kappa_i = \overline{\kappa}_j)} \tag{62}$$

I observe that the $\frac{Cov}{Var}$ term is the slope coefficient from an 'uncorrected' traditional hedonic regression on a subset of the population, i.e. with a credit wedge of $\overline{\kappa}_j$. I apply the interpretation of traditional hedonic regressions, amended by the analysis in the paper, and applied to a population subset, to obtain:

$$= \sum_j \overline{\kappa}_j Pr(\kappa_i = \overline{\kappa}_j) E^i\left[\frac{u_s^i/u_c^i}{r^f} \frac{1}{\kappa_i} \middle| \kappa_i = \overline{\kappa}_j \right] \tag{63}$$

Again, I use the conditioning information and properties of expectations to simplify algebraically:

$$= \sum_j Pr(\kappa_i = \overline{\kappa}_j) E^i\left[\frac{u_s^i/u_c^i}{r^f} \middle| \kappa_i = \overline{\kappa}_j \right] \tag{64}$$

And, finally, I use the definition of conditional and unconditional expectations to write:

$$= E^i\left[\frac{u_s^i/u_c^i}{r^f} \right] \tag{65}$$

This is the unbiased marginal willingness-to-pay for amenities, the policy-relevant estimand, and the object typically, though I argue wrongly, supposed to be estimated by traditional hedonic regressions.

E Appendix Tables & Figures

Table 5: Summary Statistics for CRISM Sample

	Purchase N=5328	Rate/Term Refi N=1799	Cash-out Refi N=1862
Loan Characteristics:			
Loan Amount ($k)	226 (123)	260 (139)	219 (112)
Property Value ($k)	292 (180)	413 (336)	362 (236)
Interest Rate (%)	4.97 (1.18)	4.28 (0.97)	5.25 (1.20)
LTV (%)	81.0 (15.1)	70.0 (19.0)	65.0 (16.0)
DTI (%)	35.3 (13.3)	35.5 (16.1)	37.0 (15.0)
FICO Score	741 (52.9)	752 (52.1)	729 (56.6)
PMI Characteristics:			
Has PMI:			
No	60.8%	84.3%	93.4%
Yes	39.2%	15.7%	6.61%
Documentation Type:			
Full Doc	75.4%	76.6%	72.7%
Low Doc	16.3%	21.4%	17.1%
No Doc	8.29%	0.34%	6.91%
MI Company:			
Arch	4.33%	3.56%	0.96%
Essent	5.71%	7.12%	0.96%
GE	14.4%	10.4%	7.69%
MGIC	18.9%	13.3%	12.5%
Radian	17.3%	9.39%	21.2%
UGIC	12.6%	6.47%	11.5%
Other	26.8%	49.8%	45.2%

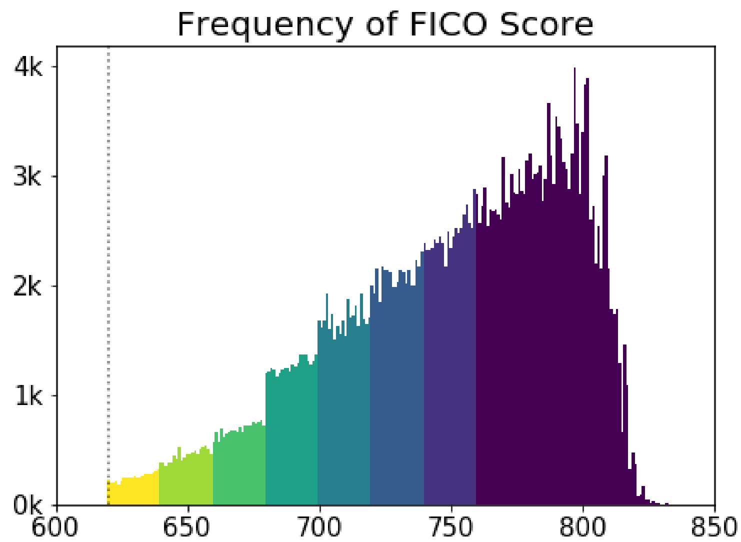
The CRISM Sample consists of loans from BlackKnight mortgage data with the following properties: originated 2005-2020, first liens, single-family, owner-occupied, 1-4 unit dwelling, conventional, conforming, vanilla (no balloon or interest-only payments), fixed-rate, and 30 year terms. For computational ease, only a 7pp random sample of the data are used in analyses. For computational ease, of this, only a 1pp random sub-sample is used to compute the summary statistics above.

Table 6: Detailed Summary Statistics for CRISM Sample

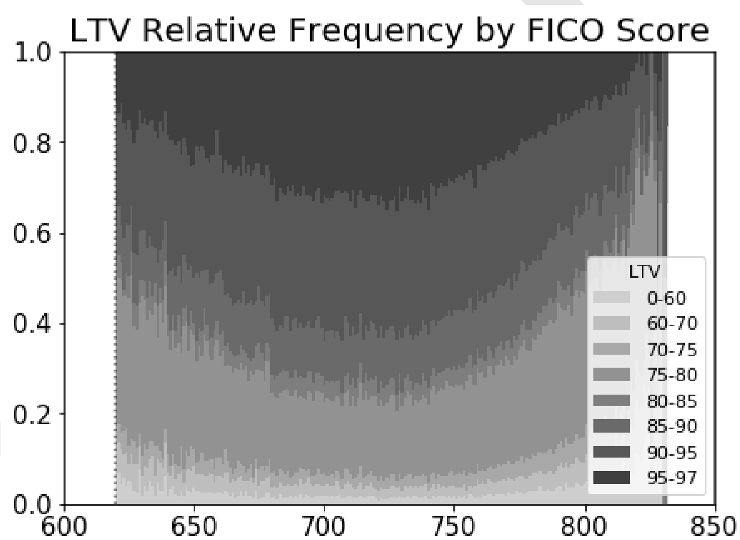
	N	$\bar{x}$	s_x	min	p25	p50	p75	max
Loan Characteristics:								
Loan Amount ($k)	9k	230	120	9	140	200	300	810
Property Value ($k)	9k	330	240	16	180	280	420	7700
Interest Rate (%)	9k	4.9	1.2	1	4	4.8	5.9	9.9
LTV (%)	9k	76	18	3	67	79	89	120
DTI (%)	7.2k	36	14	1	26	35	43	99
FICO Score	8.5k	740	54	0	700	750	780	840

The CRISM Sample consists of loans from BlackKnight mortgage data with the following properties: originated 2005-2020, first liens, single-family, owner-occupied, 1-4 unit dwelling, conventional, conforming, vanilla (no balloon or interest-only payments), fixed-rate, and 30 year terms. For computational ease, only a 7pp random sample of the data are used in analyses. For computational ease, of this, only a 1pp random sub-sample is used to compute the summary statistics above.

Figure 11: Descriptive Statistics by FICO Score



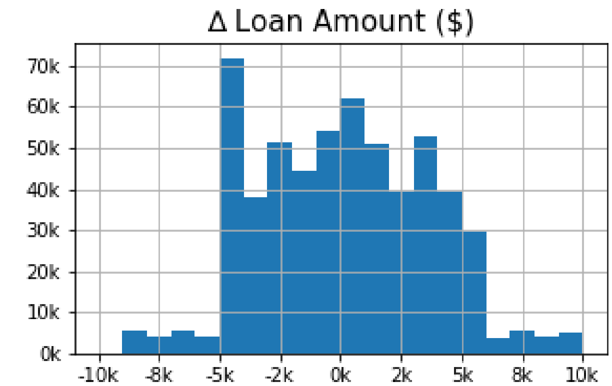
(a)



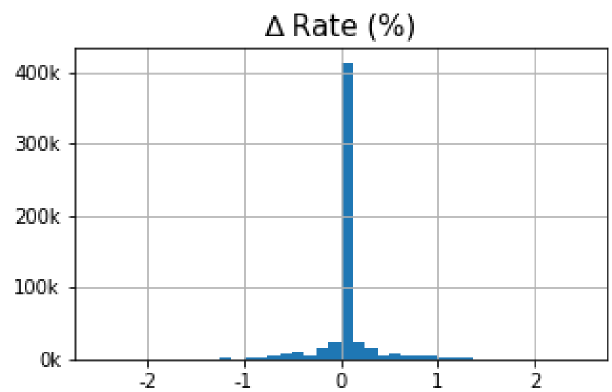
(b)

Figure (11a) depicts the frequency of borrowers with different FICO scores at origination in the HMDA Sample. Different colors depict the bins in the FNMA LLPA pricing grids. There are discontinuities at important credit score thresholds, potentially because of pricing benefits of being at a higher credit score, which increases the quantity of loans demanded. Figure (11b) depicts the relative frequency of different LTV loans within borrowers of a given credit score. Borrowers who receive loans make substantial down-payments when they have low credit scores, this tapers among borrowers of middling credit scores, and then increases at higher credit scores.

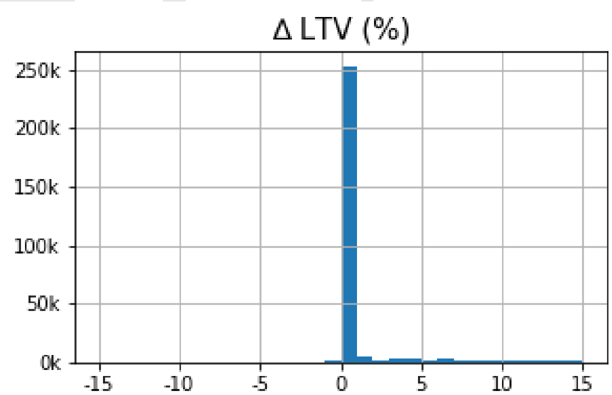
Figure 12: HMDA-FNMA-SEDA Merge Quality



(a)



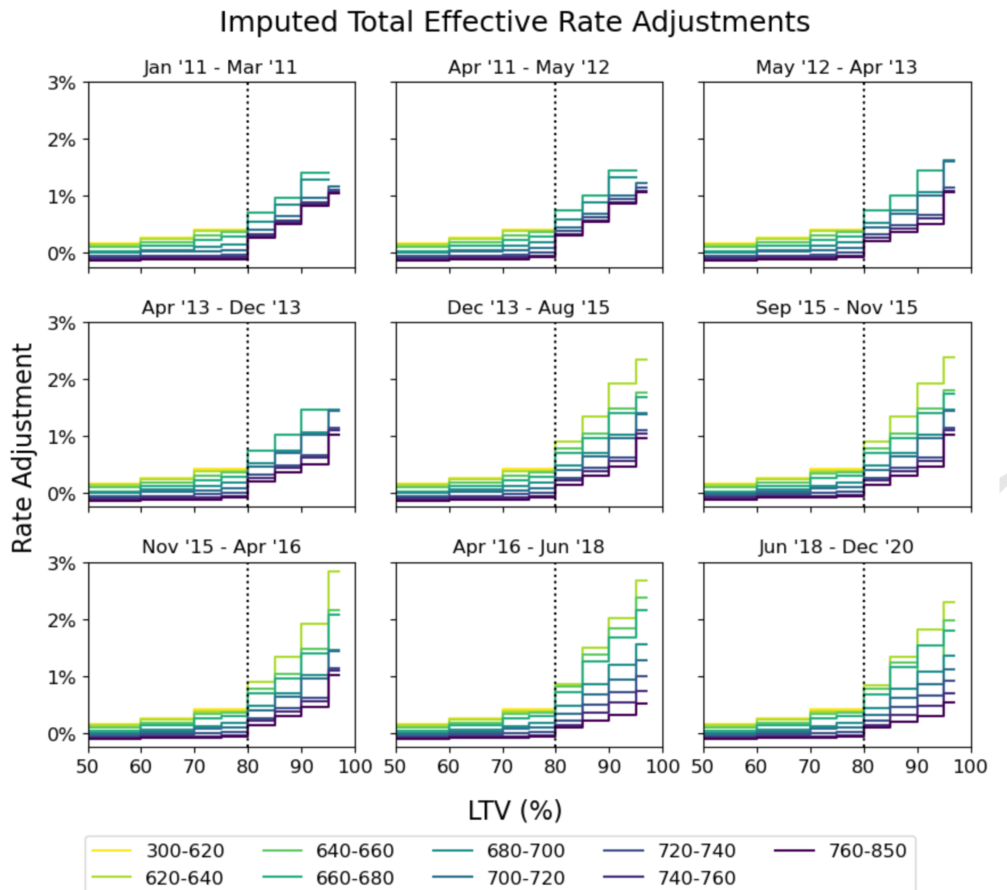
(b)



(c)

The plot above depicts the accuracy of the HMDA-FNMA-SEDA merge along the variables used for the fuzzy merge. The difference plotted denotes the Fannie Mae value minus the HMDA value. Figure (12a) plots the difference in loan amounts, Figure (12b) in interest rates, and Figure (12c) in loan-to-value ratios. The noise in the loan amount merge is expected because loan amounts in HMDA are redacted to the nearest $10k. Overall, merged loans appear fairly accurate.

Figure 13: Joint pricing regimes 2011-2020



The total effective interest rate adjustment, computed as the total rate hike coming from both LLPA pass-through and PMI costs. These panels depict nine different regimes covered by the CRISM Sample.