Reference Price Updating in the Housing Market*

Shengwei Guo, University of Wisconsin - Madison

Preliminary Draft

June 19, 2023

Abstract

What reference price do home sellers use when deciding on listing prices? This paper revisits this question using a recently developed seller behavioral model and a novel dataset, while incorporating the role of financial constraints and the default option. Integrating insights from financial markets into the housing market context and using the appraised price of a refinance mortgage as an observable historical peak, I find strong evidence that sellers update their reference price from the initial purchase price to this more recent and higher property valuation, exhibiting greater nominal loss aversion. Quantitatively, sellers are about 2.5 times more loss-averse to this appraised price than to the purchase price. Additionally, this paper incorporates insights about the impact of the default option on the listing price, leading to a more comprehensive understanding of the forces at play in the seller's decision-making process. This paper sheds new light on explaining the price-volume correlation through mortgage refinancing and reference price updating.

JEL Classification: D01, D91, G01, G21, G51, R21, R31

Keywords: appraisal, loss aversion, reference point updating, seller behavior, mortgage refinancing, housing market

*Guo: University of Wisconsin - Madison (email: sguo88@wisc.edu). I am particularly grateful for my advisors, Lu Han, Timothy Riddiough and Dayin Zhang for the invaluable support and guidance. I also wish to thank Elliot Anenberg, Alina Arefeva, Yongheng Deng, David Genesove, Alvin Murphy, Stuart Rosenthal for helpful comments. All errors are my own.

1 Introduction

While extensive research has demonstrated that sellers in the housing market exhibit reference-dependent behavior and nominal loss aversion (Genesove and C. Mayer 2001), it remains largely unclear what reference point sellers use when making listing decisions. The convention in the housing literature is to use original purchase price as the reference point. However, studies on financial markets have shown that "historical highs" also serve as salient reference prices (Barberis and Xiong 2009, Huddart, Lang, and Yetman 2009, Baker, Pan, and Wurgler 2012). In this study, I explore the reference point updating role of the "historical high" in the housing market, utilizing the appraised price of a refinance mortgage as a verifiable and credible (to a homeowner) "historical high" for a specific property.

The advantages of using appraised prices and refinance mortgages are manifold. First, in the appraisal process, appraisers rely on nearby comparable sales of the subject property for valuation, the results of which furnish a credible estimate of its market value. Second, because appraisers conduct on-site inspections, they can account for housing quality. Third, in an economic climate characterized by rising housing prices, households are strongly incentivized to access home equity through mortgage refinancing (Campbell and Cocco 2007, Greenspan and Kennedy 2008, Mian and Sufi 2011). This trend has led to an increase in refinancing activity during the early 21st century (Figure 2), thereby justifying the appraised refinance price as a valid "historical high" and allowing for a comprehensive study of refinance mortgages during this period. Fourth, the appraised price is the only price homeowners can observe that is specific to their own properties at the time of refinancing. In contrast, in a purchase mortgage context, there are two prices, each potentially confounding the other - the contract price and the appraised price[1].

This paper is the first to investigate the role of appraised price as a "historical high" reference price and examine its influence on household listing strategies in comparison with the original purchase price. To facilitate a robust investigation of household listing behavior, this paper integrates and scrutinizes many forces proposed by the existing literature. I construct a model that includes two reference prices, a down-payment constraint (Stein 1995), and a default option (Head, Sun, and Zhou

1. Various studies, including those by Cho and Megbolugbe 1996, Fout and Yao 2016, Agarwal, Song, and Yao 2017, Eriksen et al. 2018, Conklin et al. 2020, Eriksen et al. 2020 and Calem et al. 2021, indicate that appraisal values tend to cluster around the purchase price. There are two main explanations for this phenomenon in the context of purchase mortgages. One reason is appraisal targeting, where the appraiser targets the contract price as the appraised value to increase the probability of loan approval. The other involves renegotiation between buyers and sellers when the appraised value comes in lower than the initially agreed contract price.

2023)[2], all while maintaining the model's tractability. The model is designed to capture the complexity of the decision-making process that households undergo when deciding whether to list their properties for sale, while clearly distinguishing the underlying forces that generate the observed listing price. In the model, agents derive utility from the decision to exercise a default option, choose an optimal listing price, and make a listing decision. Incorporating insights from the prospect theory, developed by Kahneman and Tversky (1979), the model posits that home sellers derive utility based on realized gains and losses relative to each reference point if a sale is successfully completed. Moreover, sellers are more sensitive to losses than to gains, rendering them loss-averse. In addition to reference-dependent utility, the model also incorporates the role of home equity positions in influencing household listing strategy through down-payment constraints and the default option. The strength of considering this detailed model lies in its ability to effectively capture the interaction between loss aversion and equity position. This approach offers a nuanced understanding of how these two factors intersect to influence listing behavior.

In the model, a seller essentially balances the marginal benefit of a higher sale price premium, which would be realized upon a successful transaction, against the marginal cost tied to the increased likelihood of a transaction failure. To guide the empirical analysis and elucidate the intricacy of the model forces driving households' listing behavior, I present an analytical solution under simplified functional form assumptions. The model yields three key testable predictions. First, households may exhibit different degrees of loss aversion depending on the reference price, regulated by their respective loss aversion parameters. Second, households constrained by down-payments set higher listing prices. Third, when potential home equity falls below the default cost, the default option becomes active, leading to excessive seller "fishing" behavior compared to a model without the default option. To comprehend the model's implications regarding the default option, consider that defaulting on a mortgage allows a household to walk away from an underwater property without incurring substantial losses to compensate the lender. This limited downside risk effectively lowers the marginal cost associated with a failed listing. Consequently, sellers are more inclined to tolerate a lower probability of sale by setting a higher listing price—a mechanism also observed in Head, Sun, and Zhou 2023[3].

The derived optimization condition of the model uncovers the underlying process that determines

2. Unlike home purchases which can be financed either through a full cash transaction (without a default option) or a mortgage purchase transaction (with a default option), cash transactions introduce an empirical confound when studying the default option using initial home purchase transactions. See Han and Hong 2020 and Reher and Valkanov 2022 for a discussion on the differences between cash and mortgage sales in the housing market. In contrast, households always have an option to default when considering refinance mortgages. This is another benefit of using refinance mortgages.

3. The distinction is that I demonstrate that a static model without search can also generate the same implication.

listing prices. Proceeding with empirical testing, the primary challenge that hinders researchers from studying appraised price as an updated reference price is data limitation. I manage to overcome this limitation by assembling a novel, comprehensive dataset that traces the trajectory of residential properties in the U.S., from initial home purchases and refinancing, to eventual selling and default decisions by households. To ensure the accuracy of the data merging process, exhaustive tasks are undertaken to evaluate the data matching quality. To isolate the effect of the two nominal loss measures, I leverage the substantial independent variations in the years households list their homes and the years of initial purchases, conditional on the years of refinancing.

The analysis is focused specifically on the non-agency[4] segment of the housing market during a period of significant housing market volatility. There are several features unique to this period and market segment that make it an ideal setting for this study. First, the notable fluctuation in housing prices during this period facilitates an analysis of loss aversion due to a higher occurrence of nominal losses during the market bust (as depicted in Figure 1). Second, refinancing activity sees a dramatic increase during the housing boom of 2002-2006, offering a rich data source for investigation (Figure 2). Lastly, as documented by Mian and Sufi 2009 and Keys et al. 2010, the subprime segment of the market experienced a high rate of mortgage default during this period. This underscores the need for understanding the implications of the default option for the housing market.

The empirical findings of this study offer several novel insights. Regarding the effect of loss aversion, I find that the "historical high" serves as a more influential reference price, influencing seller listing prices more substantially than the original purchase price. To illustrate, a 10% rise in the expected nominal loss relative to the "historical high" leads sellers to set a 4.4% higher list price, while a similar increase relative to the purchase price only prompts a 0-1% rise in the list price. These results hold across various specifications to account for sample selection, omitted variables, multicollinearity issue, as well as using an instrumental variable approach. In subsample analyses, the pronounced effect of the "historical high" persists across situations of mortgage prepayment and varying holding periods. However, I note that the effect of anchoring to the "historical high" weakens by about half when a seller prepays the mortgage or holds the property for an extended period before selling. The empirical results regarding the influence of down-payment constraints and the default option on listing price align well with both the proposed model and existing literature.

Following the model estimation, I find that sellers update the reference price to the newer, higher

4. Mortgages that are securitized through the private market.

appraised price and show considerable aversion to nominal losses relative to this “historical high”. Specifically, sellers appear about 2.5 times more loss averse to the “historical high” than to the purchase price. Interestingly, while there is a certain level of reference dependence on the “historical high”, there’s virtually none concerning the purchase price.

Lastly, to delve deeper into the importance of the “historical high” as a reference point and the default option, various model variants are constructed. The analysis starts with a simple model that incorporates only down-payment constraints. It then includes the purchase price as the only reference price, followed by the addition of the “historical high” as another reference point, and finally, the full model version. Results from these four model variants help to decompose the relative explanatory power of each reference price and the default option for the observed listing price. The results find that the “historical high” substantially enhances the model’s predictive power for the observed listing premium by 41.87%, whereas the purchase price and default option contribute 2.78% and 0.35%, respectively.

Building on the insights of Genesove and Mayer (2001), which suggested that nominal loss aversion can explain the positive correlation between price and volume observed in the data, my model posits that sellers facing nominal losses set higher listing prices in pursuit of a higher final sales price, while tolerating a lower probability of sale. This insights explains the initial sticky house prices and decreasing sales volumes during the housing market downturns. Using the same four model variants, I decompose the contributions of reference dependence and loss aversion for each reference price, as well as the default option, to the observed price-volume correlation in the housing market. The findings emphasize that the “historical high” can explain 56% of the price-volume correlation, while the default option accounts for 24% of the correlation, specifically for the sample period in this study. Conversely, the original purchase price explains 1% of this correlation.

Related Literature: This paper has the potential to contribute to three main lines of research. First, it adds to the existing literature that applies prospect theory to study reference dependence in the housing market and its effect on household behavior. The renowned seminal work by Genesove and C. Mayer 2001 documents loss aversion in house sales with the initial purchase price as an anchor point. They find that condominium sellers facing expected nominal losses set higher list prices and experience lower sale hazard using Boston market data. Since then, several other studies (e.g., Anenberg 2011 for a study on San Francisco Bay Area market; Bokhari and Geltner 2011 for a study on commercial real estate; Bracke and Tenreyro 2021 for a study on housing sales in England and Wales; Ross and Zhou 2021; Zhou, Clapp, and Lu-Andrews 2021 and Andersen et al. 2022) have investigated home selling and listing decisions from a behavioral perspective and their implications for the aggregate housing market. I extend this strand of literature in several unique ways. Firstly, it underscores the relevance of the appraised price as a salient reference point for home sellers due to its nature as an observed “historical high”. To achieve this, I introduce a new dataset. Secondly, this study brings novel perspectives to the understanding of the formation and evolution of reference prices in the housing market. Thirdly, I offer unique evidence regarding the role of refinance mortgages in the housing cycle and their contribution to the correlation between housing prices and transaction volumes. Lastly, I incorporate the insights documented by Head (2023) regarding the effect of the default option on listing prices. This offers a more comprehensive understanding of the forces at play when homeowners decide to list their properties for sale.

This paper also contributes to the extensive body of literature studying updating and adaptation of reference points. Numerous studies have examined the dynamics of reference points, either through forward-looking expectations or backward-looking past dependence. Bowman, Minehart, and Rabin 1999 and Kőszegi and Rabin 2009 integrate this structural behavior into consumption-saving models. Baucells, Weber, and Welfens 2011 explore reference price updating in experimental settings, finding that the reference price is a weighted average of the first and most recent information. Simonsohn and Loewenstein 2006 study reference point adaptation in the housing rental market, contrasting movers from wealthier and less affluent neighborhoods. DellaVigna et al. 2017 and Thakral and Tô 2021 focus on the adaptation of reference prices in labor markets, while Post et al. 2008 examine the behavior of game show contestants. Card and Dahl 2011 investigate violence following football games, but found no evidence of reference point updating in this context. Kahneman 1992 posits that future research should explore "how multiple reference points compete and combine" (p. 310). My paper expands on this rich literature by examining seller behavior in the housing market. By leveraging real-world data that involves high-stakes decisions and has actual market implications, this study provides new insights into the interaction and combination of multiple reference points in the context of property transactions.

This paper is also closely connected to the behavioral finance literature, which argues that historical peaks serve as reference prices influencing investor trading behavior (Huddart, Lang, and Yetman 2009) and bidder offer prices in merger and acquisition contexts (Baker, Pan, and Wurgler

2012). This study extends these findings to the realm of real estate, providing evidence that historical peaks also significantly influence sellers when deciding their property listing prices.

This paper is organized as follows. Section 2 provides the background on the appraisal industry and illustrates the appraisal process. Section 3 presents the theoretical foundation to guide empirical exercise by introducing a model of housing listing behavior. Section 4 outlines the data sources, describes the data construction procedures, details the empirical strategy, and provides descriptive evidence. Section 5 discusses the empirical facts in detail. Section 6 presents the estimation process and results, and provides insights on the importance of each reference price. Section 7 discusses the research implications and concludes the paper.

2 Institutional Background on Appraisals

The policy discourse regarding the appraisals, with a focus on appraiser independence and accuracy, has a long history, dating back to 1987. The Real Estate Appraisal Reform Act of 1988 established Federal appraisal standards, including the requirement that these standards adhere to the Uniform Standards of Professional Appraisal Practice. This act also authorized states to establish appraiser certifying agencies and mandated that all federally-covered real estate-related transactions be conducted by independent and certified appraisers. The act's primary purpose was to ensure an accurate and unbiased assessment of a subject property's market value.

The discussion of appraisal independence resurfaces during the housing price run-up between 2004 and 2006, which ultimately contributed to the financial system's collapse. A high-profile legal case involving New York Attorney General Andrew Cuomo suing Washington Mutual for pressuring eAppraiseIT to inflate home values captured both academic and public attention. In response to concerns about biased appraisal practices, the Federal Housing Finance Agency and the Dodd-Frank Wall Street Reform and Consumer Protection Act introduce more restrictive rules for the appraisal industry. These rules include the Home Valuation Code of Conduct (HVCC), which established a more stringent firewall between appraisers and lenders through the use of appraiser management companies[5].

In the United States, appraisals are required for nearly all mortgage applications, encompassing both purchase and refinance mortgages. For a purchase mortgage, once a home buyer and seller have

5. For a discussion of the effects of this policy intervention on appraisal practice, see Shi and Zhang 2015, Ding and Nakamura 2016, and Agarwal, Ambrose, and Yao 2017.

agreed on a price and signed the contract, the buyer may seek financing from a lender, who in turn requests an appraisal conducted by a qualified appraiser. The appraiser is provided with a copy of the contract containing information on the pre-agreed contract price and any other terms that may affect the valuation process. The appraisal helps to ensure that the buyer is not overpaying for the property and to estimate the lender's potential recovery if the borrower defaults on the mortgage.

The Loan-to-Value (LTV) ratio of the mortgage is calculated based on the lower value between the appraised value and the contract price. Moreover, if the appraised value comes in lower than the contract price, the buyer has several options. She could contribute more equity for the down payment, renegotiate the price with the seller, or even terminate the purchase agreement due to the appraisal contingency clause[6]. In contrast to purchase mortgages, the appraised value serves as the sole determinant of the mortgage loan amount and LTV calculation for refinance mortgages, which mortgages are the primary focus of this study. For both types of mortgages, appraisers are compensated with a fixed amount per appraisal with little variation.

There are three approaches to appraise a property: (1) comparable sales approach, (2) income approach, and (3) cost approach. Among these, the comparable sales (i.e., "comps") approach is the most commonly used. To illustrate the process, appraisers conduct an on-site inspection of the subject property, identifying recently transacted nearby properties that best match the subject in various physical attributes, including size, number of bedrooms/bathrooms, property age, and so forth. Among the set of all possible candidates, appraisers select three to five comps to estimate the value. Second, appraisers make price adjustments to account for the differences in physical attributes between each comparable property and the subject property. Lastly, appraisers apply weights to each price-adjusted comp to reach a final valuation of the subject property. Since appraisers rely on real housing transactions in the estimation procedure, it is reasonable to believe that the appraisal price uncovers any remaining uncertainty in a household's belief of the property value, and thus updates the reference price.

3 Modeling Framework

I build upon the model of household listing behavior by Andersen et al. 2022 and add two components to the model. First, there are two reference prices: (1) purchase price and (2) "historical high".

6. This clause grants the buyer the right to cancel or renegotiate the transaction if the appraisal comes in lower than the agreed-upon contract price.

Second, there is an option to default if households fail to sell their properties on the market. The model consists of two periods. In the first period, a household receives a "shock" θ that summarizes a household's multi-dimensional reasons or motivation to sell into a single factor, and draws an option value to default "c" capturing the threshold at which she will exercise the default option. The distribution of $\theta \sim \mathcal{N}(\theta_m, \sigma_m^2)$ and $c \sim \mathcal{N}(\theta_c, \sigma_c^2)$ are estimated to match empirical moments. θ enters the utility function if a household successfully sells her property in the market, capturing a one-time increase in lifetime utility resulting from moving out of the current "mismatched" house. "c" can be interpreted as a "default cost", implying that only when the home equity falls below this threshold is the seller willing to incur this cost and relinquish the property to the lender. In the second period, a seller optimally decides whether to exercise the default option post listing, the price at which they list their house on the market, and whether or not to list the property for sale.

Before delving into the comprehensive mathematical description of the model, I introduce some notations. Let L represent the listing price chosen by the seller, and $\hat{P}$ represent the property's contemporaneous market value predicted using a hedonic pricing model. The listing premium is the difference between the listing price and the property's market value ($\ell = L - \hat{P}$). Conditional on this listing premium, the probability of sale is captured by a concave demand function $\alpha(\ell)$[7], as proposed by Guren 2018. The final realized sale price is a function of the market price and listing premium, $P(\ell) = \hat{P} + \beta(\ell)$, where $\beta(\ell)$ has a linear functional form, $\beta(\ell) = \beta_0 + \beta_1 \ell$, which is observed in the data. Both $\beta(\ell)$ and $\alpha(\ell)$ are estimated from the data.

In line with Andersen et al. 2022, I also take into account that some sellers may not have the capacity to precisely target the final sale price, adding a layer of realism to the model. In other words, there are two types of sellers: π fraction of the sellers can obtain $P(\ell)$ without uncertainty (i.e. $P(\ell) = \hat{P} + \beta(\ell)$), whereas $(1-\pi)$ sellers can only imprecisely target the final sale price $P(\ell)$ because of the price negotiation process and form an expectation on the distribution of the final sale price with $P(\ell) = \hat{P} + \beta(\ell) + \epsilon$[8].

Potential home equity is defined as the log difference between the property market value and the current outstanding mortgage balance ($\hat{H} = \hat{P} - M$). Selling the property serves as a potential remedy

7. A period is six months in the model and empirical analysis. Therefore, $\alpha(\ell)$ measures the probability of a sale within six months since the initial listing date.

8. It is worth noting that the primary objective of Andersen et al. 2022 in introducing imprecise targeters is to better align with the empirical observation of substantial strict and diffuse bunching of sales at or above the original purchase price. However, in this study, I introduce the "historical high" as an additional reference point, which complicates the determination of which price households gravitate toward. Consequently, I externally set the parameter π to enhance the realism of the model.

for a severely delinquent homeowner in the model. If a household chooses to default after a failed listing, I assume that the property is immediately seized by the lender and listed as a foreclosure sale. If the proceeds from the foreclosure sale do not cover the default cost, the loss to the seller is capped by this default cost, meaning that the seller has limited downside risk following a failed sale. When making a default decision, the seller will compare the option value from defaulting with the outside option formulated below.

The fundamental trade-off a seller faces lies between the listing price and the probability of sale. A higher listing premium (ℓ) can result in a higher realized final sale price, but it reduces the likelihood of matching with a serious buyer. More formally, a household's home selling, listing, and default decision in the first period is represented by the following objective function:

$$\max_{s\in\{0,1\},D\in\{0,1\}} \{\, s\cdot \max_{\ell,D\in\{0,1\}} \{\alpha(\ell)\cdot[\, E[\, U(P(\ell),\cdot)\,]+\theta\,] + (1-\alpha(\ell))\cdot[\, D\cdot max[-c,\hat{H}_f - c]+(1-D)\cdot \underline{u}\,]-\phi\,\} \\ +(1-s)\cdot[\, D\cdot max[-c,\hat{H}_f - c]+(1-D)\cdot\underline{u}\,]\,\} \tag{1}$$

I assume that a seller's outside option is $\underline{u} = \hat{H}$, implying that she is indifferent between retaining the property with the current market value while paying off the remaining mortgage payments, and selling it with proceeds equal to the current equity in a frictionless world. The seller's decision-making process can be outlined as follows: First, a household optimally decides whether to default ($D = 0$ or $D = 1$) on the mortgage based on her equity position ($\underline{u} = \hat{H}$) in the property. Second, conditional on her ex post default decision, she maximizes the expected utility from the listing by balancing the marginal benefit and cost of the listing price. If the sale is successful, she receives θ. Also, she incurs a one-time utility cost ϕ associated with search friction, listing costs, and other factors whether the listing is sold or not. Lastly, conditional on the default decision and optimal listing price, she decides whether or not to list the house on the market. If the seller decides against listing the property, she is still faced with the decision of whether to default on the mortgage or not [9]. Throughout the paper, my focus is on default behavior conditional on a listing, with the assumption that the household's default behavior would remain the same even without listing.

Decompose $U(P(\ell),\cdot)$ **:**

9. In the model-based estimation, I normalize $\hat{P}$ to one so that I can measure all relevant variables in log difference terms, as in Andersen et al. 2022.

The function $U(P(\ell),\cdot)$ contains two components and follows the formula below:

$$U(P(\ell),\cdot) = u(P(\ell),\cdot) - \kappa(P(\ell)) \tag{2}$$

The first part, $u(P(\ell),\cdot)$, represents reference-dependence utility, while the second part involves a penalty function, $\kappa(P(\ell))$, which diminishes the utility if the down-payment constraint is not satisfied.

Adhering to the standard in the literature (Kőszegi and Rabin 2006 and Kőszegi and Rabin 2007), the reference-dependence utility function for each of the reference point has a piecewise linear functional form. A seller in this model faces two reference points: R_1 and R_2. Throughout the analysis, I assume $R_1 \leq R_2$. Therefore, R_1 corresponds to the initial home purchase price, and R_2 corresponds to the appraised price of a refinance mortgage. A seller derives utility from the final realized price plus additional "psychological" utility depending on whether she realizes nominal gains or losses, $G_i(\ell) = P(\ell) - R_i$:

$$u(P(\ell), R_1, R_2) = P(\ell) + \bar{\lambda}_1 \eta_1 G_1(\ell) + \bar{\lambda}_2 \eta_2 G_2(\ell) \tag{3}$$

where:

$$\bar{\lambda_i} = \begin{cases} \lambda_i & \text{if } G_i(\ell) < 0 \\ 1 & \text{if } G_i(\ell) \geq 0 \end{cases} \quad \text{for i = 1, 2}$$

The parameter η_i quantifies the degree of reference dependence. To illustrate this concept, consider a seller with only one reference price (e.g., purchase price) and two extreme values of η_1. When $\eta_1 = 0$, the reference-dependent utility reduces to $P(\ell)$, indicating no reference dependence. Conversely, when $\eta_1 = -1$ and assuming the seller realizes nominal gains, the reference-dependent utility collapses to R_1, representing complete reference dependence. The parameter $\lambda_i > 1$ implies that sellers exhibit greater sensitivity to nominal losses than to gains, i.e., they exhibit loss aversion. A higher λ_i value indicates a stronger degree of loss aversion.

I now turn to the second component of $U(P(\ell),\cdot)$, the down-payment penalty function $\kappa(P(\ell))$. Stein 1995 posits that households lacking sufficient equity from the sale of their current house may attempt to "fish the market" by listing at a higher premium. Instead of assuming that a seller rejects the sale offer when the down-payment requirement is not met, the model flexibly accommodates this behavior by incorporating a down-payment penalty function. Assuming a household is actively searching for a new, equally-sized property of comparable quality[10], a household faces a down-payment constraint if the realized home equity (i.e., $H(\ell) = \hat{H} + \beta(\ell)$) falls below $\gamma\hat{P}$, and vice versa. γ represents the down-payment requirement for each one-dollar increase in property value. In the estimation, γ is set at 20% for the post-crisis period. To circumvent the disutility arising from this down-payment penalty, households with lower home equity opt for a higher listing price. Thus, the down-payment penalty function $\kappa(P(\ell))$ takes the following form[11]:

$$\kappa(P(\ell)) = \bar{\mu}(\gamma\hat{P} - H(\ell))^{1/2} \tag{4}$$

where:

$$\bar{\mu} = \begin{cases} \mu & \text{if } H(\ell) < \gamma\hat{P} \\ 0 & \text{if } H(\ell) \geq \gamma\hat{P} \end{cases}$$

***Household Default Decision*:**

In the model, the primary force driving a household's default decision is the potential home equity level. However, there are more subtle implications of the observed pattern between default propensity and home equity. The model predicts that households who end up defaulting on their mortgages are those sellers whose listing strategies are more aggressive, setting a higher listing premium and consequently leading to a lower sale probability. The model accounts for this selection effect and utilizes the observed pattern between default likelihood and home equity as additional data moments to inform structural parameters.

A more rigorous examination of household default behavior necessitates the use of advanced modeling techniques and additional assumptions. As extensively discussed in the literature[12], another major reason for households to default on their mortgages is liquidity concerns, such as unemployment, which result in unaffordable monthly payments. Investigating liquidity-based default would require a researcher to incorporate income and consumption dynamics, and a loan amortization structure (see Corbae and Quintin 2015 as an exemplary study that investigates the foreclosure crisis using a dynamic model.). Given the current data, I choose to focus on equity-based default, which can

10. Thus, the market value for the next home is also $\hat{P}$.

11. I choose a concave function as the majority of households in my sample achieve a final realized home equity level of around 20% upon a successful sale.

12. For an in-depth study of mortgage default from both theoretical and empirical perspectives, refer to Mayer, Pence, and Sherlund 2009, Bajari, Chu, and Park 2008, Foote, Gerardi, and Willen 2008, Bhutta, Shan, and Dokko 2010, Elul et al. 2010, Ghent and Kudlyak 2011, and Campbell and Cocco 2015.

still yield sharp empirical predictions.

A household chooses to default if, and only if, the option value of defaulting exceeds the outside option value:

$$max[-c, \hat{H}_f - c] > \underline{u}$$

The relationship outlined above demonstrates that a household only has an incentive to default when the mortgage is underwater, i.e., $\underline{u} = \hat{H} < 0$[13]. Thus, a household's default decision can be simplified as follows:

$$D(\underline{u}) = \begin{cases} 1 & if\ \underline{u} < -c \text{ and utility following a failed sale is } -c \\ 0 & if\ \underline{u} \geq -c \text{ and utility following a failed sale is } \underline{u} \end{cases}$$

***Analytical Solution under Simplified Assumptions*:**

To derive critical intuition and insights from the model, I provide an analytical solution for a simplified version of the model and discuss its implications in this section. Assuming that $\alpha(\ell)$, $\beta(\ell)$, and $\kappa(P(\ell))$ all follow linear functional forms[14], a seller who can accurately target the final sale price maximizes the following objective function conditional on a listing:

$$V^*(\hat{G_1}, \hat{G_2}, \hat{H}, \theta, c) = \max_{\ell}\ (\alpha_0 - \alpha_1\ell)[P(\ell) + \bar{\lambda}_1\eta_1 G_1(\ell) + \bar{\lambda}_2\eta_2 G_2(\ell) - \bar{\mu}(\gamma\hat{P} - H(\ell)) + \theta] + (1 - \alpha_0 + \alpha_1\ell)[\hat{H} - 1_{\{\hat{H}<-c\}}(c + \hat{H})] \quad (5)$$

The first component of the seller's objective function captures the utility from a realized sale multiplied by the probability of a sale, while the second component governs the utility from the outside option and the exercise of the default option following a failed listing, multiplied by the probability of a failed sale. Taking the first-order condition with respect to ℓ, we obtain:

$$\ell^*(\hat{G_1}, \hat{G_2}, \hat{H}, \theta, c) = \frac{\alpha_0\beta_1 - \alpha_1\beta_0}{2\alpha_1\beta_1} - \frac{\overbrace{\bar{\lambda}_1\eta_1\hat{G}_1 + \bar{\lambda}_2\eta_2\hat{G}_2}^{\text{Reference dependence and Loss aversion}} + \theta + \hat{P} \overbrace{+\bar{\mu}(\hat{H} - \gamma\hat{P})}^{\text{Down-payment constraints}} \overbrace{-\hat{H} + 1_{\{\hat{H}<-c\}}(c + \hat{H})}^{\text{Default option}}}{2(1 + \bar{\lambda}_1\eta_1 + \bar{\lambda}_2\eta_2 + \bar{\mu})\beta_1} \quad (6)$$

13. This effectively implies that $\hat{H_f} = \chi \cdot \hat{P} - M < 0$

14. i.e., $\alpha(\ell) = \alpha_0 - \alpha_1\ell$; $\beta(\ell) = \beta_0 + \beta_1\ell$; and $\kappa(P(\ell)) = \bar{\mu}(\gamma\hat{P} - H(\ell))$.

Equation (6) essentially equates the marginal benefit from a higher sale price premium when successfully realizing a sale to the marginal cost associated with an increasing likelihood of a failed transaction. Several notable model implications are revealed by this optimization condition:

First, compared to the purchase price, the appraised price represents a more recent valuation of a seller's property, while also taking into account housing quality. Moreover, it is typically higher than the original purchase price, as refinancings are often conducted during a housing price boom. Analogous to studies on reference dependence in the stock market that emphasize historical highs (e.g., 52-week high) as salient reference points driving investor behavior, it is natural to expect that R_2 is a more relevant reference price, and sellers are more averse to the nominal losses derived from R_2. This leads to the first prediction:

Prediction 1. $\lambda_1 < \lambda_2$.

Next, considering the relationship between listing premium and potential home equity, equation (6) demonstrates that sellers with expected nominal home equity below 20%, which corresponds to the violation of down-payment constraints, set a higher listing price[15]. This leads to the second prediction:

Prediction 2. There is a negative relationship between listing premium and potential home equity, and the negative slope becomes steeper at 20% of the home equity level.

The next prediction concerns the impact of the default option on seller behavior. The availability of a default option effectively mitigates the marginal cost associated with a failed listing by limiting the downside risk for sellers with significantly underwater mortgages. This is because such borrowers can relinquish their properties without incurring additional costs, even if the foreclosure proceeds cannot fully cover the cost of default. The final prediction is as follows:

Prediction 3. The presence of a default option leads to excessive "price fishing" behavior among sellers with low potential home equity.

The aforementioned three predictions offer a compelling rationale for empirical examination. Prior to conducting a comprehensive empirical analysis and estimating the structural parameters, I outline the data and data merging process in the next section. Furthermore, I meticulously develop the empirical equation, connecting it to the seller's optimal choice of listing price, to ensure a transparent analysis.

15. This relationship hold when $\mu > 1$. I recover the parameter μ and verify that this assumption is satisfied.

4 Data, Empirical Strategy and Hedonic Pricing Model

4.1 Data

The housing data utilized in this study are all micro-level and primarily coming from five sources: CoreLogic Mortgage, CoreLogic Owner Transfer, CoreLogic Multiple Listing Service (MLS), CoreLogic Property Assessment, and Moody's Black Box Logic dataset. I assemble these five datasets to gather a high-quality property by date penal dataset covering the universe of residential properties in the U.S., which serves as the basis for the empirical analysis conducted in this paper.

The first set of data employed in this study is obtained from CoreLogic, a leading provider of housing market data in the United States. The CoreLogic Owner Transfer dataset encompasses a universal set of house transactions from county assessor's offices in the U.S., including both residential and commercial properties. For this study, the focus is on residential properties such as single-family houses and condominiums. This dataset also provides detailed sales transaction information, including sales date, sale price, property address, buyer and seller's information, type of sale, and deed type. The CoreLogic MLS dataset, sourced from regional Multiple Listing Service platforms, covers more than 50% of the listings in the U.S., providing granular information on the start date to the last date of a listing, contract agreement date, original listing price, changes in the listing price, property address and characteristics, and listing status (withdrawn, expired, or sold). The CoreLogic Property Assessment dataset offers detailed property characteristics at the year when a property is assessed for tax purposes, which mitigates the potential omitted variable bias for hedonic modeling to a considerable extent. Finally, the CoreLogic Mortgage dataset covers universally originated mortgages in the U.S. with information on loan contract terms, mortgage origination date, mortgage type, purpose of a mortgage, borrower name, lender and loan officer information, as well as property information. I cross-walk these four datasets by using the CoreLogic proprietary property identifier and relevant date information. The details of the merging process is discussed in the Appendix. The sample period of all CoreLogic datasets is between 1998 and 2019.

The second dataset utilized in this study is the comprehensive loan-level information dataset provided by Moody's Black Box Logic (BlackBox). The original BlackBox dataset consists of approximately 22 million non-agency mortgage loans, all securitized through private securitizers (such as investment banks) before the global financial crisis. As the non-agency mortgage-related securitization channel ceased operation following the crisis, the sample period employed for the merging process spans from 1998 to 2007. BlackBox offers detailed information on borrower and loan characteristics at origination, including contract terms and borrower credit scores. Furthermore, the database tracks loan performance since origination on a monthly basis. BlackBox supplements the CoreLogic Mortgage data by providing information on property appraisal values at origination, loan delinquency status, and the dates when loans become 90 days or more delinquent. I now proceed to an in-depth discussion of the algorithm that I have developed to merge the BlackBox and CoreLogic Mortgage datasets.

***BlackBox and CoreLogic Mortgage Merging Process*:**

The identity of a property is available in CoreLogic but not in BlackBox. Therefore, I match the CoreLogic mortgage data with BlackBox mortgage data based on a common information approach. To the best of my knowledge, no paper has merged these two datasets. I follow the matching algorithm developed by Griffin and Maturana 2016, which merges analogous property transaction and loan data from different data providers. To maintain the accuracy of statistical inference and regression analysis, I extensively diagnose my match quality and procedure in the appendix.

The BlackBox sample comprises first-lien, 15 or 30-year mortgages originated between 1998 and 2007 for purchase or refinancing. For this analysis, I focus on residential properties. Thus, I exclude loans that use commercial, multi-family, or unknown property types used to secure the mortgages. I follow Kruger and Maturana 2021 and drop loans with original LTV ratios above 103% or combined-LTV ratios below 25%. In accordance with the literature, I require all relevant variables used in my match process and regression analyses to be non-missing and provide valid information[16].

The CoreLogic sample consists of conventional mortgages associated with arm-length transactions where the underlying property is for residential use. I further exclude FHA loans, VA loans, construction loans, lease deeds, small business loans, revolving lines of credit, reverse mortgages, and duplicate mortgages. I also require all relevant variables to be non-missing[17].

Following the initial data cleaning steps, I match CoreLogic and BlackBox mortgages based on 5-digit property ZIP code, property type (single-family or condo); loan term (15 or 30 years); interest rate type (Fixed-Rate or Adjustable-Rate); loan purpose (purchase or refinancing); loan amount (in thousand dollars); the loan origination dates from two samples are required to be within a [-30, 30]

16. These variables include loan balance, origination date, loan term, original LTV and combined-LTV, loan purpose, interest rate type, property ZIP code, as well as the property appraisal value at loan origination.

17. These variables include property identifier, mortgage origination date, property ZIP code, loan amount, and loan purpose.

day interval.

The resulting CoreLogic-BlackBox matched sample includes approximately 8.7 million mortgages. However, some of these matched mortgages are duplicate matches, meaning one CoreLogic mortgage can be matched with multiple BlackBox mortgages, or vice versa. Therefore, I carefully analyze my matched sample and only maintain samples with high-quality matches. Specifically, first, I require the interest rates, whenever they are not missing from both data sources, to be matched. Second, for multiple matches, I manually check the records with matched lender's names. For adjustable-rate mortgages (ARM) specifically, I require that the ARM-specific contract characteristics are also matched[18]. Third, I keep the matched records whose absolute differences in loan origination dates and loan amounts across two data sources are smallest among the multiple matches. Lastly, if there are still remaining multiple matches, I keep the first occurrence of the matched records.

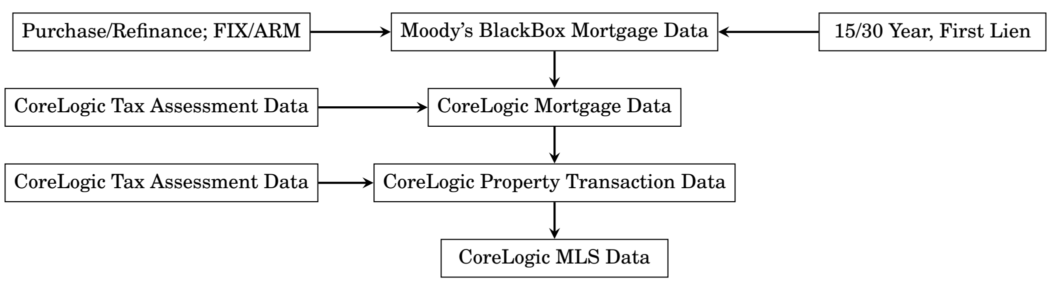
Aftering filtering the duplicate matches, the final CoreLogic-BlackBox matched sample encompasses slightly over 6 million mortgages. I am able to match 42.2% of the original BlackBox sample. Figure A1 compares the distribution of contract or borrower characteristics between matched and unmatched loans. This comparison illustrates that the distributions in both samples are quite similar, indicating that the matching procedure does not result in a selected sample. With respect to origination year distribution, a marginally higher matching rate is observed between 2003 and 2007, corresponding to the years when the non-agency market reaches its peak in terms of loan origination. Same as Griffin and Maturana 2016, I obtain a higher match ratio for mortgages in California and Florida compared to other states.

***Final Merged Data*:**

In the appendix, I provide a comprehensive description of how I crosswalk all CoreLogic datasets and outline the data construction filters. With all the filtered and merged datasets, I construct my final data in the following order. First, using CoreLogic mortgage data as the base dataset, I crosswalk the purchase mortgage subset with CoreLogic Owner Transfer data to link the mortgage information associated with the transaction data, utilizing the CoreLogic mortgage and transaction identifiers. For unmatched transactions, I consider them as cash purchases and retain them in the final data. Next, I crosswalk the mortgage subset of the resulting data from the above step with the BlackBox

18. The ARM-specific characteristic I use is the first interest rate adjustment date. However, I checked that other dimensions are also matched, such as the margin between mortgage rate and the applicable ARM index or the highest interest rate that can be charged to the borrower over the life of an ARM loan.

data to obtain the property appraised price and loan performance information, employing the CoreLogic and BlackBox mortgage identifiers. Lastly, I crosswalk the transaction subset of the resulting data from the previous two steps with CoreLogic Listing data to obtain the listing information of each transaction, using the CoreLogic transaction and listing identifiers. For withdrawn listings[19], I retain them in the final dataset. Overall, the final data comprises an unbalanced property-by-date panel dataset that tracks a property's transaction, listing, and financing history between 1998 and 2019[20]. Additionally, I obtain a mortgage's appraised price and all historical default information[21] for the subset of mortgages originated between 1998 and 2007. To ensure that the empirical analysis is consistent with the model assumption (i.e., $R_1 \leq R_2$), I focus on refinance mortgages between 2004 and 2007. To mitigate the impact of atypical houses on the hedonic model, I exclude properties whose land size, living area size, number of bedrooms, and number of bathrooms are in the top 1%. I also trim the sales price distribution at the top and bottom 1% of the price distribution. The following flow chart summarizes the entire merging process:



After estimating the hedonic value of a property at the time of listing (discussed in Section 4.3 below), I filter the sample to construct the final listing dataset for empirical analysis. First, I require the initial sale to occur before 2007 and limit the refinance mortgage to be subsequent to the time of the first sale but preceding the time of the second sale (if there is a second sale). Moreover, I only consider properties whose refinance mortgage with a matched appraised price is the most recent one prior to listing (i.e., there is no subsequent refinancing mortgage that would provide a different appraised price). I focus on the first listing following refinancing and require this listing to be between

19. Approximately 50% of the listings are sold and matched to a sale transaction, while another 50% of the listings are withdrawals. The reason for such a high withdrawal ratio is because I require a sold listing to be matched with a sale transaction. The match ratio and match process are discussed in the appendix.

20. To clarify, there are two main datasets for the empirical analysis. One is the unbalanced property-by-date panel dataset on which the hedonic model is built. The other is the listing subset of this unbalanced property-by-date panel data, which is used for hypothesis testing.

21. I define default as a mortgage being 90 days or more delinquent. The historical default information includes all months when the mortgage has defaulted.

2007 and 2015 (inclusive)[22]. Table 1 presents a selected example of a property located in Chester County, PA to illustrate the data structure. Next, I forward-fill all necessary variables for the empirical exercise for these listings. Finally, the sample is restricted to listings for which I have non-missing measures of all the independent variables, such as nominal losses, current LTV and hedonic predicted price, etc.

4.2 Empirical Design

To examine the model implications, I adhere to equation (6) and model the log of list price for property i in census tract n, purchased at month s, refinanced at month m, and listed at month t (with $s < m < t$) as follows[23]:

$$\log(Y_{insmt}) = m_1 \text{Loss\_PurchasePrice}_{inst} + m_2 \text{Loss\_AppraisedPrice}_{inmt} + \beta LTV_{int} + \delta \widehat{Log\ Price}_{int} + \alpha_0 + \epsilon_{insmt} \tag{7}$$

$\text{Loss\_PurchasePrice}_{inst}$ represents the expected nominal loss relative to the purchase price, truncated above zero. It is defined as the greater value between the difference of the log of the purchase price and the hedonic predicted price, and zero. Similarly, $\text{Loss\_AppraisedPrice}_{inmt}$ is defined as the expected nominal loss relative to the appraised price, also truncated above zero.

The BlackBox data tracks the monthly outstanding loan balance at the beginning of each period, starting from loan origination. This data feature allows for the avoidance of using an amortization formula to infer the current outstanding mortgage balance, as done by Anenberg 2011 and Ross and Zhou 2021. Therefore, I impute the continuous measure of the loan-to-value (LTV) ratio as the ratio of the outstanding mortgage balance at the beginning of the listing month to the expected predicted price[24]. Then, I define LTV_{int} as the greater value between the difference of this ratio and 0.8, and zero. $\widehat{Log\ Price}_{int}$ represents the expected log market price predicted from a hedonic model, as discussed in the following section. Additionally, I include the number of months since purchase and refinance to control for tenure effects. In an alternative specification, I incorporate the hedonic pricing error of the previous purchase as a proxy for unobserved quality. Since the empirical model follows

---

22. According to the FHFA housing price index, it is highly probable that a household is no longer subject to nominal loss if the property is listed on the market after 2015.

23. This model-based empirical specification primarily builds upon and extends the one presented in Genesove and C. Mayer 2001. The primary differences lie in the inclusion of an expected nominal loss in relation to the updated reference price, as well as the inclusion of a default option.

24. The expected predicted price is estimated using the same hedonic model, but replacing the log of sales price with the level of sales price to avoid the log transformation from the predicted log value to the value in level. This version of the hedonic model yields an R^2 of 85%

a two-stage estimation procedure where the first stage is hedonic pricing model and the second state estimates the equation (7), I bootstrap the standard errors in accordance with the literature tradition.

Taking into account the implications of prospect theory and this extended model, it is expected that a household would be more sensitive to the appraised price, which reflects the most recent information about the value of their property. In other words, the hypothesis testing aims to determine if $m_1 < m_2$, which implies $1 < \lambda_1 < \lambda_2$ in the model. The coefficient β captures the change in slope when the down-payment constraint is in effect, and thus, the second prediction of the model implies $\beta > 0$[25].

4.3 Hedonic Pricing Model

To estimate the market price of the property at time t, $\widehat{Log\ Price}_{int}$, I construct a hedonic pricing model using the sale transaction sample as the estimation sample. Since the refinancing sample subset focuses on a specific market segment, I use all transactions from 1998 to 2019 for the hedonic model to ensure that the price indices are not derived from a selected sample. Subsequently, I generate both in-sample and out-of-sample predictions of the property market value for the full property-by-date panel sample[26].

I model the log of sale price, $\log(P_{ins})$, for property i situated in census tract n and sold at time s as follows:

$$\log(P_{ins}) = \beta X_{is} + \theta_{ns} + \eta_{ins} \tag{8}$$

where X_{is} represents a vector of time-varying hedonic characteristics (sourced from the CoreLogic Tax Assessment data, which typically provides annual assessments for properties). These hedonic variables include land size in sq.ft., living area size, property age, and a collection of dummy variables for the number of bedrooms, number of bathrooms, property condition, construction type, heating type, roof type, building style type, property type, and garage type. Additionally, a set of indicator variables for air conditioning, basement presence, fireplace, and pool are included. To account for potential nonlinear effects of certain hedonic features, I create 100 bins for land size, 50 bins for

25. The empirical model omits the default option because the theoretical model assumes that the costs of default are idiosyncratic across sellers. Additionally, these costs are not easily observable in the data. As a result, the estimated $\hat{\beta}$ is subject to omitted variable bias, which is expected to bias it downward, as indicated by the optimality condition. Consequently, $\hat{\beta}$ provides a lower bound for the true β.

26. The transaction sample consists of 71 million observations, while the property-by-date final sample comprises 157 million.

living area size, and 10 bins for property age and include them as a battery of dummies variables in the regression. The bins are created by evenly dividing the distribution of the corresponding variable, with missing values forming a separate category. θ_{ns} accounts for census tract by year-month fixed effects.

A highly accurate hedonic pricing model is fundamental for the analysis, as omitted variable bias may lead to false inferences or biased coefficient estimates, as extensively discussed by Genesove and C. Mayer 2001, Anenberg 2011, and Zhou, Clapp, and Lu-Andrews 2021. I achieve an R^2 of 81% from the hedonic model, which substantially mitigates concerns and noise arising from omitted variable bias. Furthermore, I conduct a series of robustness checks to address potential sample selection with regard to focusing on the sample of properties that have been refinanced. I also adopt a repeat-sales approach, incorporating a set of property fixed effects in equation (8) to alleviate concerns stemming from time-invariant unobserved factors.

4.4 Identifying the Effect of Nominal Loss: Two Reference Prices

Unlike Genesove and C. Mayer 2001 who rely on variations in original home purchase years among households with similar current home values, the empirical strategy in this study relies on variations across two dimensions due to the presence of two reference prices. First, conditional on the year of mortgage refinancing, variations in listing years provide differences in the extent to which households are subject to nominal loss, thus helping to identify the effect of nominal loss relative to the appraised price. Second, conditional on the year of mortgage refinancing, variations in the original home purchase years offer additional variations in nominal loss relative to purchase prices, independent of appraised prices. This mitigates concerns regarding the high correlation between the two nominal loss measures. The figure in Panel A of Figure 3 presents the distribution of purchase and listing years, conditional on each of the four refinancing years in the sample, and confirms the presence of substantial variations along these two dimensions. In addition, Panel B of Figure 3 plots the joint distribution of two measures of nominal gains, each related to a reference price. Conditional on the presence of a loss to the appraised price, the figure shows that there are substantial independent variations in the extent of nominal gains/loss to the purchase price, allowing for the identification of their independent impacts on listing prices.

4.5 Descriptive Evidence: Two Reference Prices

As a preliminary step to formally examining the salience of each reference price, I begin by presenting descriptive evidence on the effect of the reference prices on listing prices. It is reasonable to expect that the listing price is centered around a salient reference price, with a spike in density mass precisely at that reference price. Conversely, there would be no clear pattern for a less relevant reference point. Figure 4 displays the density distribution of the log percentage difference between the listing price and each of the reference prices. It clearly demonstrates that sellers often set listing prices close to or equal to the appraised prices, consistent with reference dependence. In contrast, list prices tend to be higher than the initial purchase prices, indicating that the purchase price may be a weaker reference point. In the following section, I rigorously verify the relative importance of these two reference prices using regression analysis.

5 Empirical Results

In this section, I utilize the final listing dataset to thoroughly evaluate the model implications and document empirical patterns consistent with the model predictions. The first set of empirical observations focuses on the relevance and salience of each reference price in determining listing prices in line with nominal loss aversion (i.e., $\hat{G_1}$ and $\hat{G_2}$). The second set of observations examines the relationship between potential home equity, listing premium, and default probability (i.e., $\hat{H}$, ℓ, and D), with a particular focus on the down-payment constraints and default option. Third, I confirm that the demand function is concave in this particular market segment, consistent with Guren 2018, and document the pattern of the final price realization function (i.e., $\alpha(\ell)$ and $\beta(\ell)$). Finally, I gather additional data moments to investigate the relationship between potential home equity and listing propensities (i.e., s).

5.1 Fact 1: Loss Aversion under Two Reference Points

Figure 5 illustrates the average listing premium across each 1-percentage point bin for different measures of potential gains[27]. The left-hand side figure presents the relationship with the purchase price as the reference price, while the right-hand side figure displays the same relationship with the

27. Listing premium is defined as ℓ = Log listing Price $-\widehat{Log\ Price}$, and potential gain is defined as $\widehat{G_i} = \widehat{Log\ Price} -$ Log R_i, where R_i represents the reference prices.

appraised price as the reference price[28]. Intriguingly, I observe the same "hockey stick" relationship as in Andersen et al. 2022 when the reference price is the purchase price, even though with a different sample from a different country.

In accordance with the modeling framework, these downward relationships align with the assumption that a seller's preference is characterized by a reference-dependent and loss-averse utility function. Specifically, $\eta_i > 0$ predicts downward slopes, and λ_i foresees different slopes transitioning from nominal gains to losses. However, the strong negative relationships depicted by these two figures potentially indicate that no single reference price can fully explain the choice of listing price. It is highly plausible that an average seller tends to assign (possibly different) weights along both dimensions. For instance, conditional on the same potential gains of 10%, the listing premium for an average seller is approximately 0% (for $\widehat{G_2}$) and around 14% (for $\widehat{G_1}$). I investigate which reference price dominates the other using the regression framework below.

Table 2 presents the baseline regression results for Equation (7), which examines the relationship between the listing price and expected nominal losses measured by two reference prices. Column (1) demonstrates that a seller subjected to a 10 percent increase in expected nominal loss with respect to the appraised price sets a 4.44% higher listing price. Conversely, a 10 percent increase in expected nominal loss with respect to the purchase price only results in a 0.95% higher listing price. Column (2) adds the hedonic pricing error from the previous purchase as a noisy measure of unobserved quality[29]. Column (2) reveals that a seller will set a 4.35% higher listing price if there is a 10% increase in expected nominal loss relative to the appraised price. However, the effect of loss relative to the purchase price is insignificant and near zero.

Columns (3) and (4) further include two corresponding quadratic loss terms, the results of which suggest that the marginal value of nominal losses relative to the appraised price diminishes as loss increases, consistent with prospect theory. Notably, the estimated coefficients on LTV if ≥ 80% are consistently positive and significant across all specifications, confirming the role of down-payment constraints implied by the model. Considering the results in Table 2 collectively, it implies that the appraised price serves as a valid reference point, and the effect of loss aversion is more salient for this more recent and higher property valuation to a seller.

28. One should interpret the relationship cautiously, as $\widehat{G_1} > \widehat{G_2}$ always holds for the same seller, leading to a systematic relationship where the average listing premium for each percentage bin is consistently higher for $\widehat{G_1}$ compared to the same level of $\widehat{G_2}$.

29. As posited by Genesove and C. Mayer 2001, the coefficient estimates between these two specifications bracket the true effect of loss aversion. Column (1) provides an upper bound on the effect of loss aversion, while Column (2) offers a lower bound.

Heterogeneity: A natural characteristic of the appraised price is that lenders use it for underwriting mortgages and determining the maximum loan size a borrower can qualify for. Consequently, there may be concerns that mortgage balance could overstate the role of appraised price as a reference point. Table 3 addresses this concern by estimating the same specifications in two samples based on whether a seller has prepaid the mortgage and, thus, has a zero current LTV. Column (1) and (4) present the results for sellers with a non-zero current LTV, while Column (5) and (8) display the results for sellers with a zero current LTV. Table 3 reveals that the role of the appraised price as a reference point is significantly diminished once a borrower has paid off the mortgage (the estimated coefficient in the linear loss term reduces by approximately half)[30]. Nonetheless, the appraised price still has a more substantial impact on a seller's choice of listing price compared to the previous purchase price.

A potential concern is that the effect of the appraised price as a reference point could diminish as the holding period increases. Furthermore, the role of the initial purchase price could still be influential or even more dominant with a longer holding period. Consequently, it is reasonable to expect that the effects of nominal loss could also vary by the length of holding periods[31]. Table 4 formally tests this hypothesis by dividing the main listing sample based on the duration a seller lists the property for sale after the last refinance, measured in years[32]. Column (1) and (4) show the results for the sample where a seller lists their property within 3 years since the last refinancing[33], whereas Column (5) and (8) present the results for the sample where a seller lists their property more than 3 years after the last refinancing. The results clearly indicate that the effect of loss aversion relative to appraised price declines as the holding period of the property lengthens. On the other hand, the role of the initial purchase price as a reference price still has a significant and even stronger impact on the listing price as the holding period increases. Nonetheless, the magnitude of the loss aversion effect using appraised price as a reference price remains larger than that of the purchase price.

Integrating the findings from Table 2, 3, 4, and Figure 5, this first set of empirical results indicates that: (1) the "historical high" price is a valid reference price; (2) sellers exhibit greater loss

30. The negative and significant coefficient on LTV if $\geq 80\%$ could be due to sample selection bias. One potential explanation for sellers still having an LTV greater than zero is financial constraint. The motivation to sell the property could differ as sellers become more financially constrained or have a higher LTV ratio.

31. I thank Timothy Riddiough for offering this insight.

32. In contrast to previous research, I can divide the sample in this manner because the variation of expected nominal loss comes from when a seller lists a property. On the other hand, previous studies rely on variation from when a seller initially purchases the property, and all listings are observed during the housing downturn. Thus, if they were to divide the sample by holding periods, it would leave little variation to identify the effect of loss with a longer holding period.

33. The median holding period.

aversion to the "historical high" compared to the purchase price; (3) the influence of the refinancing appraised price as a reference price diminishes as a seller prepays the mortgage or as the holding period increases.

***Robustness check*:** In the Appendix A3, I further substantiate the robustness of these results by: (1) employing alternative specifications that use potential home equity instead of the LTV ratio to be consistent the model variables; (2) estimating the effects of the two expected nominal loss measures in equation (7) separately to mitigate potential multi-collinearity concern; (3) controlling for the number of refinancing events prior to listing as a proxy for borrower sophistication to address sample selection issue; (4) augmenting the hedonic modelling with property fixed effects to mitigate the concerns due to unobserved time-invariant housing quality; and (5) implementing an instrumental variables (IV) approach. The main conclusions are virtually unaffected.

5.2 Fact 2: Home Equity, Default Decision and Listing Premium

Figure 6 illustrates the observed patterns between potential home equity, household default likelihood, and listing premium[34].

Starting from the left figure, the vertical axis represents default probability, while the horizontal axis corresponds to each 1-percentage point bin of potential home equity, conditional on listing the property for sale[35]. The figure clearly demonstrates a negative relationship between default probability and equity position. Beginning at -40% of potential home equity and below, the relationship levels off, indicating that nearly all households will exercise the default option with some noise if they fail to sell their houses. This figure also supports other explanations for default beyond solely equity-based default, while also justifying the model assumption that each household has a different default cutoff drawn from a distribution.

The right panel of Figure 6 depicts the relationship between the average listing premium and each 1-percentage point bin of potential home equity. The pattern aligns with the model's implications regarding down-payment constraints and the default option. Firstly, the slope of the curve is flatter for sellers with over 20% of potential home equity, as the down-payment constraint is not binding. Once home equity falls below 20%, the down-payment penalty takes effect, resulting in a steeper

34. Potential home equity is defined as $\widehat{H} = \widehat{Log\ Price} - \text{Log}(1 + \text{Outstanding Mortgage Balance})$; Default is defined as ever 90 days + delinquency (i.e., missing more than 3 months of mortgage payments) since the listing date; listing premium and potential gains are defined as above.

35. The figure does not condition on a failed listing for two reasons. First, the model in Section 3 states that the option to default is conditional on a listing. Second, the data reveals that virtually all defaulted households stem from those who fail to sell the property on the market. Thus, the same relationship holds conditional on a failed sale.

slope. Moreover, the polynomial curve fit of order four starts to display a convex relationship between home equity and listing premium starting from around -40% of potential home equity and below, where the default option begins to have an substantial impact. In conjunction with the left figure, this convex relationship is influenced by both down-payment constraints and the default option. In other words, the default option implied by the model incrementally contributes to a higher listing premium.

***Heterogeneity*:** To further substantiate the role of the default option, I examine potential variations in default costs across subsamples in Figure 7. The two figures in Figure 7 illustrate the relationship between listing premium and potential home equity based on whether a property serves as a primary residence and the borrower's credit score range. Intuitively, it can be expected that the default cost "c" is considerably higher for owner-occupied properties compared to non-owner-occupied properties due to differences in housing utility flow. Consequently, the model predicts that households selling their properties for primary residence are less likely to exercise the default option, and thus, list at a lower price, as observed in the left figure. Additionally, Harrison, Noordewier, and Yavas 2004 contend that borrowers with FICO Scores between 620 and 660 face significantly higher default costs than borrowers with scores outside this range, a hypothesis supported by their empirical analysis. Following this argument, I define high default cost as sellers whose FICO Scores at loan origination range between 620 and 660 (inclusive), and vice versa. In line with the model's implications, households with FICO Scores between 620 and 660 set lower list prices. Intriguingly, both figures in Figure 7 reveal an evident divergence in the relationship starting from 0%, which is consistent with equity-based default.

5.3 Fact 3: Concave Demand ($\alpha(\ell)$) and Price Realization ($\beta(\ell)$)

Figure 8 displays the probability of a sale within 26 weeks[36] across each 1-percentage point bin of the listing premium. This figure confirms the findings of Guren (2018), who provided evidence of concave demand for a broader sample. The concave demand posits that reducing the listing premium increases the sale probability up to a certain point, beyond which the sale probability only improves marginally. This concave demand relationship corresponds to the $\alpha(\ell)$ function in the model. The solid line represents a generalized logistic function fit, while the dashed line represents a linear fit.

36. The raw variable represents an indicator for whether a listing is sold within 26 weeks following the original listing date. Subsequently, this indicator is collapsed for each level of the listing premium to compute the average sale probability. Guren (2018) employs a 13-week threshold; however, the relationship is robust to this cutoff and various alternative cutoffs.

The generalized logistic function is ideal because its estimates are bounded from both below and above, ensuring that the sale probability does not exceed one or fall below zero. As discussed by Andersen et al. 2022, the presence of concave demand creates an empirical confound for the "Hockey Stick" relationship in the left panel of Figure 5, as the flatter sale probability at the lower end of the listing premium results in a flatter response of the listing premium to gains, even without a significant degree of loss aversion. Thus, the estimation addresses this empirical confound by estimating λ_i while taking the demand condition as given.

Figure 9 presents the empirical counterpart of $\beta(\ell)$, which maps the listing premium to the final realized price premium, conditional on successful listings. A linear fit of this empirical pattern offers insight into the functional form of $\beta(\ell)$ in the model, with estimated coefficients $\beta_0 = -0.063$ and $\beta_1 = 0.516$.

5.4 Fact 4: Extensive Margin

To investigate the extensive margin decision of whether to sell properties or not, I construct a property-year panel data set based on the property-level listing data utilized in the previous empirical analysis. The data set is created by extending each property from 2007 to the minimum of 2015 and the sales closing year depending on whether the listing is identified as a closed sale by CoreLogic[37]. In reality, selling decisions may be influenced by various factors, such as relocation shocks, work locations, educational or school choices, environmental concerns, among others. A comprehensive consideration of selling decisions is beyond the scope of this paper; instead, I encapsulate all factors driving the moving decision into a single factor, θ. Abstracting from the moving decision, I plot the listing propensity for each 1-percentage bin of $\hat{H}$ in Figure 10. In line with the model, a higher propensity to list is associated with higher potential home equity due to the increase in the option value of listing compared to the outside option. I incorporate these additional moments in the model estimation to inform the structural parameters and utility costs associated with listings.

6 Method of Moments Estimation

6.1 Data and Model Moments

To estimate the parameters of the model, I employ the Simulated Method of Moments (SMM) approach. This approach involves minimizing the distance between the moments derived from the

37. I choose 2007 and 2015 as the panel start and end year to maintain consistency with the listing sample period.

data and those predicted by the model. Formally, the objective function can be expressed as follows:

$$\hat{b} = arg\ \min_{b}[M_d(x) - M_m(y(b))]'W[M_d(x) - M_m(y(b))] \tag{9}$$

where b denotes the vector of model parameters, $M_d(x) - M_m(y(b))$ represents the distance between the data moments and the model-implied model, and W stands for the weighting matrix. An identity matrix is employed as W first to ensure that all moments hold equal weight during the estimation process.

To clearly elucidate the process of moment matching, the following empirical moments $M_d(x)$ are gathered, corresponding to Figures 5 through 10, with the identification discussed subsequently:

$$M_d(x) = \begin{bmatrix} \ell(\hat{G}_1): & \text{Figure 5} \\ \ell(\hat{G}_2): & \text{Figure 5} \\ \ell(\hat{H}): & \text{Figure 6} \\ f_{default}(\hat{H}): & \text{Figure 6} \\ f_{list}(\hat{H}): & \text{Figure 10} \end{bmatrix}$$

The moments are collected as follows: Firstly, $\ell(\hat{G}_1)$ and $\ell(\hat{G}_2)$ are captured as the average listing premium for each 1 percentage point bin of the potential gain $\hat{G}_1$ and $\hat{G}_2$, respectively, utilizing all listings with either potential gains in the range of -50% to +50% (accounting for 101 moments for each reference point, and 202 moments in total, as illustrated in Figure 5). Subsequently, $\ell(\hat{H})$ and $f_{default}(\hat{H})$ are assembled as the average listing premium and average default probability for each 1 percentage point bin of potential home equity $\hat{H}$ for all listings carrying an outstanding mortgage. The interval for potential home equity spans from -100% to 60% (providing 161 moments each, and 322 moments in total, as depicted in Figure 6). Finally, $f_{list}(\hat{H})$ represents the listing propensity for each 1-percentage point bin of the state variable as displayed in Figure 10 (yielding 161 moments in total). Therefore, the total number of empirical moments accumulated amounts to 685.

To amass the vector of model-implied moments corresponding to the data moments, I begin with an external calibration of the down-payment constraint parameter, γ = 20%, to align with the post-crisis Loan-to-Value (LTV) requirements, and set the proportion of precise targeters, π = 0.15%, based on the estimate from Andersen et al. 2022. The concave demand function ($\alpha(\ell)$) and final price realization function ($\beta(\ell)$), as depicted in Figures 8 and 9, are estimated and externally set in relation to the sellers' objective function. In accordance with the data observations, I enforce limitations on the generalized logistic function for sale probability, establishing the lower and upper bounds at 0.2 and 0.7, respectively. Furthermore, to prevent sellers from pricing themselves out of the market due to this non-zero probability of sale, I establish a range for the listing premium over which they optimize. This range, [-65%, 250%], also aligns with observations from the data.

Beginning with an initial approximation of the vector of model parameters, I numerically derive the optimal solutions to the seller's objective function, which is a variant of equation (6), over a five-dimensional grid. This grid stands for the seller's state variables and is devised from each 1-percentage point bin of potential gains relative to appraised price and purchase price, potential home equity, and each realization of the moving shock and default cost. Optimal listing premiums are then quantitatively imputed for both precise and imprecise targeters. Finally, I employ the derived household default function subsequent to a failed listing to compute the probability of default conditional on a seller's equity position, and compare the expected utility from listing against the utility from abstaining from listing to ascertain the selling decision. I further normalize the model-implied listing probability for potential home equity level $\hat{H}$ in the range of -50% to 0% to match its observed average value in the data. The model-implied solutions are then aggregated to assemble analogous moments to those in the data.

***Identification of Model Parameters*:**

The model parameters can be separately identified in the following ways: (1) η_i mainly determines the configuration of the listing premium profile across the spectrum of potential gains/loss for each of the reference points; (2) λ_i captures the slope of the listing premium profile in the potential loss domain for each of the reference points; (3) μ and the penalty function structure shape the curvature of the listing premium profile when the down-payment requirement is binding, specifically when $\hat{H} \leq 20\%$; (4) θ_m captures the average value of the listing premium in accordance with equation (6), and σ_m accounts for the extensive margin decision beyond the state variables and also smooths the key relationships observed in the data; (5) θ_c specifically pinpoints the average potential home equity where the default option comes into play, and σ_c establishes the default probability profile in relation to potential home equity, while also having additional contribution to fitting the listing premium profile in the lower range of potential home equity; Lastly, (6) ϕ determines the mean utility loss associated with a listing and aids in shaping the profile of listing propensity. Table 5 lists all the variables and parameters in the model and reports the data moments targeted by each parameter.

6.2 Model Validation and Results

To assess the internal validity of the model and enable a more rigorous estimation of the structural parameters, Figure 11 reports the empirical patterns and model predictions together. Generally, the model broadly encapsulates the shape and level of the data patterns, with an average root mean squared error of 23.1 percent (see Table 6).

Table 7 presents the parameter estimates derived from the full model version that incorporates a default option and concave demand. The model yields reference dependence parameters and loss aversion parameters of η_1 = 0.012, η_2 = 0.344, λ_1 = 1.450, and λ_2 = 3.660. These estimates are in strong alignment with the model implications and empirical results, indicating that sellers firmly anchor to the "historical high" property valuations and exhibit heightened aversion to the consequent nominal loss. Intriguingly, the estimated reference dependence parameters demonstrate that once a newer and higher property valuation is disclosed, sellers cease to use the original purchase price as a reference point. Additionally, sellers are about 2.5 times more loss averse to this "historical high" valuation than to the original purchase price, quantitatively speaking.

Regarding the down-payment constraint parameter, the model suggests the home equity penalty parameter of μ = 1.230, which aligns with what Andersen et al. 2022 finds. In contrast, the estimates for the mean value of the selling motivation θ_m is 3.825, and the utility cost of search and listing with ϕ is 1.025, i.e., 1.025 of the property hedonic predicted price, which may seem implausible at first glance. However, considering that this study focuses on the subprime crisis period, one could expect elevated search friction in the housing market during this period, leading to a higher search and listing cost. Consequently, sellers who do list their properties on the market are strongly motivated to sell and move out of their current residences. As for the distribution of default option value, the mean θ_c = 0.124 suggests that an average seller defaults when the home equity level falls below -12.4%, and substantial variation in the default cost as indicated by σ_c can find support in the data corresponding to the left-hand side figure in Figure 6.

In Appendix A4, I find that the primary reference dependence and loss aversion parameters remain robust when using an alternative, convex functional form of the down-payment penalty function. Additionally, I discuss the potential reasons behind the sensitivity of μ and ϕ estimates to this different formulation of the penalty function.

In the following subsections, I delve into the importance of considering this observable "historical high" price. I elaborate on the role of the default option and highlight the interplay between household loss aversion behavior and equity position. This discussion underlines the need for comprehensive consideration of the equity position when it comes to listing premiums.

6.3 Interaction between Loss Aversion and Equity Position

The model adeptly integrates numerous elements to explain variations in the listing price while ensuring tractability proposed thus far in the literature. Importantly, it takes into account factors such as home equity constraints and the default option, thereby facilitating an exploration of how the impacts of nominal loss and equity position interact on household behavior, and their alignment with observed interactions in the empirical data.

Figure 12 showcases this interaction. The figure on the left reports the responsiveness of listing premia to potential gains, conditional on the level of equity constraints, by estimating the regression of listing premia on potential gains and reporting the estimated slopes. It compellingly indicates that sellers who are less financially constrained respond strongly to potential gains, a pattern observable both in the model and the data. This is particularly true when sellers are not subject to down-payment constraints. Furthermore, the figure on the right illustrates how sellers respond to home equity conditional on the level of potential gains. As sellers move from potential loss to potential gain territory, they react more strongly to their equity position. The model effectively captures this relationship, with confidence intervals that closely align with the data, barring one exception.

In summary, the new findings suggest that sellers who are subject to equity constraints or significant losses are less responsive to potential gains or home equity. A potential explanation for these new findings is that the dominant factor shaping a seller's listing behavior is the one that places the seller in the most constraining situation.

6.4 The Role of the Default Option

To probe the model's predictions regarding the role of the default option, I re-collect the model moments employing a variant of the model in which I excluding the default option. Using the same set of estimated structural parameters, I visualize the projected listing premium profile across potential home equity.

Figure 13 presents the findings. A stark divergence is apparent between a model that includes a default option and one that does not—echoing the implications derived from the model. Remarkably, this disparity intensifies as one transitions from moderate negative home equity to significantly lower equity positions, implying that the default option effect becomes more pronounced due to the increased probability of exercising this option. This illustration underlines the importance of considering the default option when studying factors that influence listing prices.

6.5 Listing Premia Decomposition

I previously discussed how sellers update their reference price when a "historical high" price is revealed that exceeds the original purchase price, and how they exhibit significant loss aversion to this more recent property valuation. In addition, I investigate the role of the default option, which encourages sellers to “fish” the market by setting a higher listing price. How important is it to consider this additional reference price and the default option? To answer this question, I introduce different model variants and breakdown the contributions of each reference point and the default option to the observed listing premiums and the correlation between price and volume in the following two sections.

I proceed the analysis by estimating the following equation:

$$\ell^{Data}(\hat{G}_1, \hat{G}_2) = \alpha_1 + \rho_1 \ell^{Model}(\hat{G}_1, \hat{G}_2) + \epsilon \tag{10}$$

I begin with a model incorporating only the financial constraints, and then progressively reintroduce each reference price and the default option back into the model. The same set of parameter estimates as in Table 7 are used, enforcing η_i and λ_i to be zero when the corresponding reference price is not considered.

Table 8 presents the estimation results for Equation (10). In the absence of any reference price, a model that only incorporates down-payment penalties can already explain 39.5% of the variation in listing premiums. The predictive capacity for listing premiums in Equation (10) improves marginally by 2.78% and 41.87%, to 40.6% and 57.6% respectively, when the original purchase price and appraised price are gradually reintroduced to the model, allowing households to exhibit loss aversion towards these prices. Lastly, incorporating the default option into the complete model slightly enhances its explanatory power by 0.35%, reaching 57.8%. Most notably, $\ell(\hat{G}_1, \hat{G}_2)$ in the model essentially moves one-to-one with the data when this "historical high" price is taken into account. The significant improvement in the model's ability to capture the variation in observed listing premiums, along with the point estimate of ρ_1, emphasize the importance of considering this "historical high" price as a salient reference point.

6.6 Price-Volume Correlation Decomposition

Housing market is characterized by a positive correlation between house price and sales volume with the change in sales volume always leads the change in house price. Loss aversion and down-payment constraints are two most significant explanations for this pattern. With the initial purchase price as the reference price, the households are sluggish in adjusting their list prices even when the housing market condition has changed dramatically. To materialize the power of reference point updating found in this paper, this section investigates the marginal contribution of the new reference price to the price-volume correlation, while also exploring the influence of the default option in explaining this correlation.

The model imputes the listing decision (s) conditional on each 1-percentage bins of the $\hat{G}_1$, $\hat{G}_2$ and $\hat{H}$ and the level of annual housing stock N_{stock} (in the extensive margin analysis). With the estimated demand function ($\alpha(\ell(\hat{G}_1,\hat{G}_2,\hat{H}))$) we can compute the counterfactual realized sales:

$$N_{sale} = s(\hat{G}_1,\hat{G}_2,\hat{H}) \times \alpha(\ell(\hat{G}_1,\hat{G}_2,\hat{H})) \times N_{stock} \tag{11}$$

I consider the following regression to estimate the observed price-volume correlation (ρ_2) in the data:

$$\Delta ln N_T = \alpha_2 + \rho_2 \Delta ln \hat{P}_T + \epsilon_T \tag{12}$$

where N_T represents the sales volumes in year T as observed from the CoreLogic dataset, and $\hat{P}_T$ denotes the corresponding hedonic price index in year T, collected from the Federal Housing Finance Agency dataset (Figure 1). The empirical correlation between prices and volumes, as measured by this methodology, is reported in the top panel of Table 9 and equals ρ_2 = 0.765.

Following this, I map the observed changes in hedonic value to changes in sales volumes (N_{sale}) as implied by the model. I then substitute N_{sale} into Equation (12) to re-estimate the model-implied ρ_2. Utilizing the same model variants as in Table 8, I find that reference dependence and loss aversion to the appraised price account for 56% of the price-volume correlation. On the other hand, reference dependence and loss aversion to the purchase price contribute to 1% of the price-volume correlation. Furthermore, the default option also holds substantial explanatory power for the price-volume correlation, contributing 24%.

This exercise clearly highlights the importance of considering the updated reference price and the default option in understanding the aggregate dynamics of housing prices and sales volume.

7 Conclusion

This paper leverages a recently developed structural model and a unique dataset to estimate and quantify the relative importance of the original purchase price and an observed "historical high" as reference points. I find compelling evidence that the appraised price from a mortgage refinance serves as a salient reference price for real players involved in the multi-layered house selling process.

Furthermore, this paper uncovers and substantiates how sellers incorporate the default option into their listing decision, highlighting the significant role of home equity. The interplay between mortgage refinancing, home equity, and the default option provides an potential channel to understand how the mortgage default crisis can spill over to the housing market, as was the case leading up to the Great Financial Crisis.

The findings of this paper offer several implications. Firstly, to elucidate the subprime mortgage crisis, classic arguments have documented the conflict of interests within the financial system, amplified by the originate-to-distribute model, leading to excessive credit provision to subprime borrowers (Mian and Sufi 2009, Keys et al. 2010, and Agarwal, Chang, and Yavas 2012). A fresh perspective on the subprime narrative emphasizes the role of middle-class borrowers (Adelino, Schoar, and Severino 2016). This paper supplements the subprime narrative from a distinct angle, focusing on households behavior during the bust. Specifically, nominal loss aversion and the default option distort sellers' listing behavior, leading to an illiquid housing market.

Secondly, by establishing the role of appraised price as a reference point, this paper emphasizes the need to address appraisal bias and establish a uniform standard for the appraisal process. While even an unbiased appraised value can distort households' listing strategy if refinancing activity comoves with the housing cycle, multiple studies have shown that appraisals are often biased, typically overvaluing properties compared to their counterfactual fair market values (Agarwal, Ben-David, and Yao 2015 and Griffin and Maturana 2016). This bias further exacerbates nominal loss, erodes household equity position, and may amplify a housing market downturn.

Lastly, this paper also underscores why the boom and bust cycles are intricately linked. During an expansion period, rising housing prices or a low-interest-rate environment incentivize households to refinance for equity extraction and mortgage rate reduction. This refinancing boom could further fuel the aggregate economy and amplify housing price volatility via the household demand channel (Mian, Sufi, and Verner 2020). When the housing market starts to cool down, these same households experience nominal loss and consequently list their properties at above-market prices. This paper provides a fresh perspective to understand the dynamics of housing market boom and bust cycles.

References

Adelino, Manuel, Antoinette Schoar, and Felipe Severino. 2016. "Loan originations and defaults in the mortgage crisis: The role of the middle class." *The Review of Financial Studies* 29 (7): 1635–1670.

Agarwal, Sumit, Brent W Ambrose, and Vincent W Yao. 2017. "The effects and limits of regulation: Appraisal bias in the mortgage market." *National University of Singapore, Pennsylvania State University, and Georgia State University Unpublished working paper.*

Agarwal, Sumit, Itzhak Ben-David, and Vincent Yao. 2015. "Collateral valuation and borrower financial constraints: Evidence from the residential real estate market." *Management Science* 61 (9): 2220–2240.

Agarwal, Sumit, Yan Chang, and Abdullah Yavas. 2012. "Adverse selection in mortgage securitization." *Journal of Financial Economics* 105 (3): 640–660.

Agarwal, Sumit, John C Driscoll, and David I Laibson. 2013. "Optimal mortgage refinancing: a closed-form solution." *Journal of Money, Credit and Banking* 45 (4): 591–622.

Agarwal, Sumit, Richard J Rosen, and Vincent Yao. 2016. "Why do borrowers make mortgage refinancing mistakes?" *Management Science* 62 (12): 3494–3509.

Agarwal, Sumit, Changcheng Song, and Vincent Yao. 2017. "Relational contracts, reputational concerns, and appraiser behavior: Evidence from the housing market." *Georgetown McDonough School of Business Research Paper* 3076944.

Allen, Eric J, Patricia M Dechow, Devin G Pope, and George Wu. 2017. "Reference-dependent preferences: Evidence from marathon runners." *Management Science* 63 (6): 1657–1672.

Ambrose, Brent W, James Conklin, N Edward Coulson, Moussa Diop, and Luis A Lopez. 2021. "Do Appraiser and Borrower Race Affect Valuation?" *Available at SSRN 3951587.*

Andersen, Steffen, Cristian Badarinza, Lu Liu, Julie Marx, and Tarun Ramadorai. 2022. "Reference dependence in the housing market." *American Economic Review* 112 (10): 3398–3440.

Anenberg, Elliot. 2011. "Loss aversion, equity constraints and seller behavior in the real estate market." *Regional Science and Urban Economics* 41 (1): 67–76.

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

Baker, Malcolm, Xin Pan, and Jeffrey Wurgler. 2012. "The effect of reference point prices on mergers and acquisitions." *Journal of Financial Economics* 106 (1): 49–71.

Barberis, Nicholas, and Wei Xiong. 2009. "What drives the disposition effect? An analysis of a long-standing preference-based explanation." *the Journal of Finance* 64 (2): 751–784.

———. 2012. "Realization utility." *Journal of Financial Economics* 104 (2): 251–271.

Baucells, Manel, Martin Weber, and Frank Welfens. 2011. "Reference-point formation and updating." *Management Science* 57 (3): 506–519.

Bennett, Paul, Richard Peach, and Stavros Peristiani. 2001. "Structural change in the mortgage market and the propensity to refinance." *Journal of Money, Credit and Banking,* 955–975.

Bhutta, Neil, and Benjamin J Keys. 2016. "Interest rates and equity extraction during the housing boom." *American Economic Review* 106 (7): 1742–1774.

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

Bokhari, Sheharyar, and David Geltner. 2011. "Loss aversion and anchoring in commercial real estate pricing: Empirical evidence and price index implications." *Real Estate Economics* 39 (4): 635–670.

Bowman, David, Deborah Minehart, and Matthew Rabin. 1999. "Loss aversion in a consumption–savings model." *Journal of Economic Behavior & Organization* 38 (2): 155–178.

Bracke, Philippe, and Silvana Tenreyro. 2021. "History dependence in the housing market." *American Economic Journal: Macroeconomics* 13 (2): 420–43.

Calem, Paul, Jeanna Kenney, Lauren Lambie-Hanson, and Leonard Nakamura. 2021. "Appraising home purchase appraisals." *Real Estate Economics* 49 (S1): 134–168.

Campbell, John Y, and Joao F Cocco. 2007. "How do house prices affect consumption? Evidence from micro data." *Journal of monetary Economics* 54 (3): 591–621.

———. 2015. "A model of mortgage default." *The Journal of Finance* 70 (4): 1495–1554.

Card, David, and Gordon B Dahl. 2011. "Family violence and football: The effect of unexpected emotional cues on violent behavior." *The quarterly journal of economics* 126 (1): 103–143.

Card, David, Andrew Johnston, Pauline Leung, Alexandre Mas, and Zhuan Pei. 2015. "The effect of unemployment benefits on the duration of unemployment insurance receipt: New evidence from a regression kink design in Missouri, 2003–2013." *American Economic Review* 105 (5): 126–130.

Card, David, David S Lee, Zhuan Pei, and Andrea Weber. 2015. "Inference on causal effects in a generalized regression kink design." *Econometrica* 83 (6): 2453–2483.

Case, Karl E, and Robert J Shiller. 1988. *The efficiency of the market for single-family homes.*

Chen, Hui, Michael Michaux, and Nikolai Roussanov. 2020. "Houses as ATMs: mortgage refinancing and macroeconomic uncertainty." *The Journal of Finance* 75 (1): 323–375.

Chetty, Raj. 2015. "Behavioral economics and public policy: A pragmatic perspective." *American Economic Review* 105 (5): 1–33.

Cho, Man, and Isaac F Megbolugbe. 1996. "An empirical analysis of property appraisal and mortgage redlining." *The Journal of Real Estate Finance and Economics* 13 (1): 45–55.

Conklin, James, N Edward Coulson, Moussa Diop, and Thao Le. 2020. "Competition and appraisal inflation." *The Journal of Real Estate Finance and Economics* 61 (1): 1–38.

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

DellaVigna, Stefano. 2009. “Psychology and economics: Evidence from the field.” *Journal of Economic literature* 47 (2): 315–372.

———. 2018. “Structural behavioral economics.” In *Handbook of Behavioral Economics: Applications and Foundations 1,* 1:613–723. Elsevier.

DellaVigna, Stefano, Attila Lindner, Balázs Reizer, and Johannes F Schmieder. 2017. “Reference-dependent job search: Evidence from Hungary.” *The Quarterly Journal of Economics* 132 (4): 1969–2018.

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

Ding, Lei, and Leonard Nakamura. 2016. “The impact of the home valuation code of conduct on appraisal and mortgage outcomes.” *Real Estate Economics* 44 (3): 658–690.

Dunn, Kenneth B, and John J McConnell. 1981. “Valuation of GNMA mortgage-backed securities.” *The Journal of Finance* 36 (3): 599–616.

“Prospect theory: An analysis of decision under risk.” 1979. *Econometrica* 47 (2): 363–391.

Elul, Ronel, Nicholas S Souleles, Souphala Chomsisengphet, Dennis Glennon, and Robert Hunt. 2010. “What “triggers” mortgage default?” *American Economic Review* 100 (2): 490–494.

Eriksen, Michael D, Hamilton B Fout, Mark Palim, and Eric Rosenblatt. 2018. “The influence of contract prices and relationships on collateral valuation.” *University of Cincinnati Lindner College of Business Research Paper.*

———. 2020. “Contract price confirmation bias: Evidence from repeat appraisals.” *The Journal of Real Estate Finance and Economics* 60 (1): 77–98.

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

Fout, Hamilton, and Vincent Yao. 2016. *Housing market effects of appraising below contract.* Technical report. Working paper, available at: http://www. fanniemae. com/resources/file . . .

Garriga, Carlos, and Aaron Hedlund. 2020. “Mortgage debt, consumption, and illiquid housing markets in the great recession.” *American Economic Review* 110 (6): 1603–34.

Genesove, David, and Christopher Mayer. 2001. “Loss aversion and seller behavior: Evidence from the housing market.” *The quarterly journal of economics* 116 (4): 1233–1260.

Genesove, David, and Christopher J Mayer. 1994. *Equity and time to sale in the real estate market.*

Ghent, Andra C, and Marianna Kudlyak. 2011. “Recourse and residential mortgage default: evidence from US states.” *The Review of Financial Studies* 24 (9): 3139–3186.

Glaeser, Edward L, and Charles G Nathanson. 2017. “An extrapolative model of house price dynamics.” *Journal of Financial Economics* 126 (1): 147–170.

Greenspan, Alan, and James Kennedy. 2008. “Sources and uses of equity extracted from homes.” *Oxford Review of Economic Policy* 24 (1): 120–144.

Griffin, John M, and Gonzalo Maturana. 2016. "Who facilitated misreporting in securitized loans?" *The Review of Financial Studies* 29 (2): 384–419.

Guren, Adam M. 2018. "House price momentum and strategic complementarity." *Journal of Political Economy* 126 (3): 1172–1218.

Han, Lu, and Seung-Hyun Hong. 2020. "Cash is king? Evidence from housing markets."

Han, Lu, and William C Strange. 2014. "Bidding wars for houses." *Real Estate Economics* 42 (1): 1–32.

———. 2016. "What is the role of the asking price for a house?" *Journal of Urban Economics* 93:115–130.

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

Head, Allen, Hongfei Sun, and Chenggang Zhou. 2023. "Indebted sellers, liquidity and mortgage standards." *European Economic Review* 151:104321.

Huddart, Steven, Mark Lang, and Michelle H Yetman. 2009. "Volume and price patterns around a stock's 52-week highs and lows: Theory and evidence." *Management Science* 55 (1): 16–31.

Hurst, Erik, and Frank Stafford. 2004. "Home is where the equity is: Mortgage refinancing and household consumption." *Journal of Money, credit and Banking,* 985–1014.

Jiang, Erica Xuewei, and Anthony Lee Zhang. 2022. "Collateral Value Uncertainty and Mortgage Credit Provision." *Available at SSRN.*

Kahneman, Daniel. 1992. "Reference points, anchors, norms, and mixed feelings." *Organizational behavior and human decision processes* 51 (2): 296–312.

Kau, James B, Donald C Keenan, Walter J Muller, and James F Epperson. 1995. "The valuation at origination of fixed-rate mortgages with default and prepayment." *The Journal of Real Estate Finance and Economics* 11:5–36.

Keys, Benjamin J, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig. 2010. "Did securitization lead to lax screening? Evidence from subprime loans." *The Quarterly journal of economics* 125 (1): 307–362.

Koijen, Ralph SJ, Otto Van Hemert, and Stijn Van Nieuwerburgh. 2009. "Mortgage timing." *Journal of Financial Economics* 93 (2): 292–324.

Kőszegi, Botond, and Matthew Rabin. 2006. "A model of reference-dependent preferences." *The Quarterly Journal of Economics* 121 (4): 1133–1165.

———. 2007. "Reference-dependent risk attitudes." *American Economic Review* 97 (4): 1047–1073.

———. 2009. "Reference-dependent consumption plans." *American Economic Review* 99 (3): 909–936.

Kruger, Samuel, and Gonzalo Maturana. 2021. "Collateral misreporting in the residential mortgage-backed security market." *Management Science* 67 (5): 2729–2750.

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

Mian, Atif, and Amir Sufi. 2009. "The consequences of mortgage credit expansion: Evidence from the US mortgage default crisis." *The Quarterly journal of economics* 124 (4): 1449–1496.

———. 2011. "House prices, home equity–based borrowing, and the us household leverage crisis." *American Economic Review* 101 (5): 2132–2156.

Mian, Atif, Amir Sufi, and Emil Verner. 2020. "How does credit supply expansion affect the real economy? the productive capacity and household demand channels." *The Journal of Finance* 75 (2): 949–994.

Post, Thierry, Martijn J Van den Assem, Guido Baltussen, and Richard H Thaler. 2008. "Deal or no deal? decision making under risk in a large-payoff game show." *American Economic Review* 98 (1): 38–71.

Reher, Michael, and Rossen I Valkanov. 2022. "The mortgage-cash premium puzzle." *Available at SSRN 3751917.*

Ross, Stephen L, and Tingyu Zhou. 2021. *Loss Aversion in Housing Sales Prices: Evidence from Focal Point Bias.* Technical report. National Bureau of Economic Research.

Schwartz, Eduardo S, and Walter N Torous. 1989. "Prepayment and the valuation of mortgage-backed securities." *The Journal of Finance* 44 (2): 375–392.

Shi, Lan, and Yan Zhang. 2015. "Appraisal inflation: Evidence from the 2009 GSE HVCC intervention." *Journal of Housing Economics* 27:71–90.

Simonsohn, Uri, and George Loewenstein. 2006. "Mistake: The effect of previously encountered prices on current housing demand." *The Economic Journal* 116 (508): 175–199.

Stanton, Richard. 1995. "Rational prepayment and the valuation of mortgage-backed securities." *The Review of financial studies* 8 (3): 677–708.

Stein, Jeremy C. 1995. "Prices and trading volume in the housing market: A model with down-payment effects." *The Quarterly Journal of Economics* 110 (2): 379–406.

Thakral, Neil, and Linh T Tô. 2021. "Daily labor supply and adaptive reference points." *American Economic Review* 111 (8): 2417–2443.

Van Nieuwerburgh, Stijn, and Pierre-Olivier Weill. 2010. "Why has house price dispersion gone up?" *The Review of Economic Studies* 77 (4): 1567–1606.

Yavas, Abdullah, and Shiawee Yang. 1995. "The strategic role of listing price in marketing real estate: theory and evidence." *Real estate economics* 23 (3): 347–368.

Zhou, Tingyu, John M Clapp, and Ran Lu-Andrews. 2021. "Is the behavior of sellers with expected gains and losses relevant to cycles in house prices?" *Journal of Housing Economics* 52:101750.

Tables

Table 1: Data Structure and Data Example

This table presents an example of a property located in Chester County, Pennsylvania, to demonstrate the data structure of the final merged dataset and to illustrate the potential role of the appraised price as a reference price.

Property County	Date	List Price	Appraised Price	Sale Price	Estimated Market Price	Transaction Type
Chester County, PA	1999-08-12			420,000	447,000	Sale
Chester County, PA	2006-07-21		650,000			Refinance Mortgage
Chester County, PA	2010-08-01	629,000			609,000	Listing (Withdrawal)

Table 2: Loss Aversion under Two Reference Prices

This table presents the regression results that examine the effect of loss aversion relative to two reference prices using the final listing dataset. The dependent variable is the log of the original listing price. Loss: Appraised Price represents the maximum value between the difference of the log appraised price from the last refinance mortgage and the hedonic predicted market price, and 0. Loss-squared: Appraised Price is the squared term of this value. Loss: Purchase Price denotes the maximum value between the difference of the log purchase price and the hedonic predicted market price, and 0. Loss-squared: Purchase Price is the corresponding squared term. LTV if ≥ 80 represents the maximum value between the difference of the current LTV and 0.8, and 0. Estimated Value refers to the base value of the property estimated from the hedonic model without including the census tract-specific monthly price index. Estimated Price Index is the census tract-specific monthly price index. Residual is the error term from the previous purchase. Months since last refinance measures the difference in months between the last refinance and the current listing, while Months since last purchase measures the difference in months between the last purchase and the current listing. Bootstrapped standard errors are displayed in parentheses.

	(1)	(2)	(3)	(4)
VARIABLES	Log (Original Listing Price)			
Loss: Appraised Price	0.444***	0.435***	0.565***	0.525***
	(0.006)	(0.006)	(0.013)	(0.014)
Loss-squared: Appraised Price			-0.124***	-0.088***
			(0.015)	(0.015)
Loss: Purchase Price	0.095***	-0.007	0.150***	0.009
	(0.008)	(0.010)	(0.015)	(0.015)
Loss-squared: Purchase Price			-0.011	0.012
			(0.010)	(0.012)
LTV if ≥ 80%	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Estimated Value	1.046***	1.053***	1.045***	1.052***
	(0.003)	(0.003)	(0.003)	(0.003)
Estimated Price Index	1.025***	1.008***	1.025***	1.008***
	(0.002)	(0.002)	(0.002)	(0.002)
Residual		0.276***		0.269***
		(0.006)		(0.005)
Months since	-0.000***	-0.000***	-0.000**	-0.000***
last purchase	(0.000)	(0.000)	(0.000)	(0.000)
Months since	-0.001***	-0.001***	-0.001***	-0.001***
last refinance	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.505***	-0.561***	-0.511***	-0.567***
	(0.037)	(0.039)	(0.039)	(0.039)
Observations	97,635	97,635	97,635	97,635
R-squared	0.860	0.867	0.861	0.867

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Loss Aversion under Two Reference Prices Heterogeneity: Loan-to-Value Ratio

This table presents the regression results that examine the heterogeneous effect of loss aversion relative to two reference prices using the final listing dataset by Loan-to-Value (LTV) Ratio. Columns (1) - (4) show the results for sellers who have a non-zero outstanding mortgage balance at the time of listing, whereas columns (5) - (8) show the results for sellers who have a zero outstanding mortgage balance (or have prepaid) at the time of listing. The dependent variable is the log of the original listing price. Loss: Appraised Price represents the maximum value between the difference of the log appraised price from the last refinance mortgage and the hedonic predicted market price, and 0. Loss-squared: Appraised Price is the squared term of this value. Loss: Purchase Price denotes the maximum value between the difference of the log purchase price and the hedonic predicted market price, and 0. Loss-squared: Purchase Price is the corresponding squared term. LTV if ≥ 80 represents the maximum value between the difference of the current LTV and 0.8, and 0. Estimated Value refers to the base value of the property estimated from the hedonic model without including the census tract-specific monthly price index. Estimated Price Index is the census tract-specific monthly price index. Residual is the error term from the previous purchase. Months since last refinance measures the difference in months between the last refinance and the current listing, while Months since last purchase measures the difference in months between the last purchase and the current listing. Bootstrapped standard errors are displayed in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Log (Original Listing Price)							
	LTV > 0				LTV == 0			
Loss: Appraised Price	0.677***	0.653***	0.766***	0.721***	0.329***	0.327***	0.368***	0.335***
	(0.008)	(0.008)	(0.013)	(0.012)	(0.010)	(0.009)	(0.023)	(0.022)
Loss-squared: Appraised Price			-0.104***	-0.076***			-0.042*	-0.009
			(0.014)	(0.015)			(0.025)	(0.022)
Loss: Purchase Price	0.141***	0.037***	0.208***	0.065***	0.121***	0.028*	0.173***	0.040*
	(0.011)	(0.011)	(0.017)	(0.018)	(0.017)	(0.015)	(0.025)	(0.021)
Loss-squared: Purchase Price			-0.038**	-0.008			-0.022	-0.005
			(0.017)	(0.016)			(0.022)	(0.015)
LTV if ≥ 80%	-0.001***	-0.001***	-0.001***	-0.001***				
	(0.000)	(0.000)	(0.000)	(0.000)				
Estimated Value	1.018***	1.026***	1.019***	1.026***	1.064***	1.072***	1.063***	1.072***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)	(0.007)	(0.006)
Estimated Price Index	0.995***	0.982***	0.996***	0.982***	1.043***	1.025***	1.043***	1.025***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.004)
Residuals		0.244***		0.239***		0.294***		0.292***
		(0.006)		(0.006)		(0.010)		(0.012)
Months since	-0.000***	-0.000***	-0.000***	-0.000***	-0.000	-0.000***	0.000	-0.000***
last purchase	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months since	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
last refinance	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.152***	-0.220***	-0.172***	-0.235***	-0.744***	-0.815***	-0.739***	-0.814***
	(0.041)	(0.040)	(0.038)	(0.034)	(0.080)	(0.068)	(0.081)	(0.080)
Observations	68,255	68,255	68,255	68,255	29,380	29,380	29,380	29,380
R-squared	0.872	0.879	0.873	0.879	0.841	0.848	0.841	0.848

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Loss Aversion under Two Reference Prices Heterogeneity: Holding Period since Last Refinancing

This table presents the regression results examining the heterogeneous effect of loss aversion relative to two reference prices using the final listing dataset by the holding period since the last refinancing. Columns (1) - (4) show the results for sellers who list their properties within 3 years since the last refinancing, while columns (5) - (8) show the results for sellers who list their properties more than 3 years after the last refinancing. The dependent variable is the log of the original listing price. Loss: Appraised Price represents the maximum value between the difference of the log appraised price from the last refinance mortgage and the hedonic predicted market price, and 0. Loss-squared: Appraised Price is the squared term of this value. Loss: Purchase Price denotes the maximum value between the difference of the log purchase price and the hedonic predicted market price, and 0. Loss-squared: Purchase Price is the corresponding squared term. LTV if ≥ 80 represents the maximum value between the difference of the current LTV and 0.8, and 0. Estimated Value refers to the base value of the property estimated from the hedonic model without including the census tract-specific monthly price index. Estimated Price Index is the census tract-specific monthly price index. Residual is the error term from the previous purchase. Months since last refinance measures the difference in months between the last refinance and the current listing, while Months since last purchase measures the difference in months between the last purchase and the current listing. Bootstrapped standard errors are displayed in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Log (Original Listing Price)							
	Years Since Refinancing ≤ 3				Years Since Refinancing > 3			
Loss: Appraised Price	0.629***	0.613***	0.775***	0.725***	0.384***	0.376***	0.447***	0.414***
	(0.011)	(0.011)	(0.020)	(0.020)	(0.008)	(0.008)	(0.014)	(0.015)
Loss-squared: Appraised Price			-0.161***	-0.121***			-0.068***	-0.039***
			(0.027)	(0.027)			(0.013)	(0.014)
Loss: Purchase Price	-0.007	-0.099***	0.050**	-0.073***	0.189***	0.080***	0.280***	0.125***
	(0.016)	(0.016)	(0.021)	(0.024)	(0.010)	(0.011)	(0.017)	(0.012)
Loss-squared: Purchase Price			-0.005	0.013			-0.051***	-0.023**
			(0.017)	(0.019)			(0.013)	(0.010)
LTV if ≥ 80%	0.000***	0.000***	0.000**	0.000***	0.001***	0.001***	0.001***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Estimated Value	0.996***	1.005***	0.997***	1.006***	1.104***	1.107***	1.103***	1.107***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)
Estimated Price Index	0.992***	0.980***	0.991***	0.980***	1.051***	1.030***	1.052***	1.030***
	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)
Residuals		0.236***		0.227***		0.294***		0.288***
		(0.007)		(0.007)		(0.008)		(0.009)
Months since last purchase	-0.000***	-0.000***	-0.000***	-0.000***	0.000	-0.000***	0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Months since last refinance	-0.002***	-0.002***	-0.002***	-0.002***	0.000***	0.001***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.131***	0.045	0.105**	0.028	-1.344***	-1.337***	-1.348***	-1.340***
	(0.045)	(0.046)	(0.041)	(0.035)	(0.050)	(0.053)	(0.063)	(0.054)
Observations	49,659	49,659	49,659	49,659	47,976	47,976	47,976	47,976
R-squared	0.869	0.876	0.870	0.876	0.860	0.867	0.861	0.867

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Variables and Parameters in the Model

	Parameter	Value/Source	Target Moments	Description
State variables:	R_1	Data		Nominal reference price: purchase price
	R_2	Data		Nominal reference price: appraised price
	M	Data		Oustanding mortgage balance
	$\hat{P}$	Data		Hedonic predicted property value
	$\hat{G}_1$	Data		Potential gains ($\hat{P} - R_1$)
	$\hat{G}_2$	Data		Potential gains ($\hat{P} - R_2$)
	$\hat{H}$	Data		Potential home equity ($\hat{P} - M$)
	θ	Estimated		Sell motivation shock
	c	Estimated		Default cost
Endogenous variables:	ℓ			Listing Premium
	D			Default decision
	s			Selling decision
External parameters:	$\alpha(\ell)$	Data		Concave demand function
	$\beta(\ell)$	Data		Realized premium
	γ	20%		Down-payment requirement
	π	15%		Fraction of precise targeters
Structural parameters:	η_i	Estimated	$\ell(\hat{G}_i)$	Reference dependence parameter (i = 1 or 2)
	λ_i	Estimated	$\ell(\hat{G}_i)$	Loss aversion parameter (i = 1 or 2)
	μ	Estimated	$\ell(\hat{H})$	Down-payment constraint parameter
	ϕ	Estimated	$f_{list}(\hat{H})$	Search and listing utility cost
	$\mathcal{N}(\theta_m, \sigma_m^2)$	Estimated	$\ell(\hat{G}_i), f_{list}(\hat{H})$	Distribution of sell motivation shock
	$\mathcal{N}(\theta_c, \sigma_c^2)$	Estimated	$\ell(\hat{H}), f_{default}(\hat{H})$	Distribution of default cost

Table 6: Goodness of Fit

This table reports root mean squared differences between the level of each moment in the model and the data, normalized by the mean value of each data moment. This prediction error is explainable in percentage terms, as reported below for each moment separately, as well as jointly for the full set of moments shown in the last column.

	$\ell(\hat{G}_1)$	$\ell(\hat{G}_2)$	$\ell(\hat{H})$	$f_{default}(\hat{H})$	$f_{list}(\hat{H})$	Overall Fit
RMSE	22.0%	25.2%	28.0%	29.9%	10.4%	23.1%

Table 7: Structural Parameter Estimates

This table reports structural parameter estimates resulting from the method of moments estimation.

	Reference Dependence η_1	Reference Dependence η_2	Loss Aversion λ_1	Loss Aversion λ_2	Financial Constraints μ	Average Motivation Shock θ_m	St. dev. of Motivation Shock σ_m	Average Default Cost θ_c	St. dev. of Default Cost σ_c	Search and Listing Cost ϕ
Parameter Estimates	0.012	0.344	1.450	3.660	1.230	3.825	1.182	0.124	0.389	1.025

Table 8: Listing Premia Decomposition

This table reports the regression results from $\ell^{Data}(\hat{G_1},\hat{G_2}) = \alpha_1 + \rho_1 \ell^{Model}(\hat{G_1},\hat{G_2}) + \epsilon$, where $\ell^{Data}(\hat{G_1},\hat{G_2})$ are average listing premia across each potential gains in the data and $\ell^{Model}(\hat{G_1},\hat{G_2})$ the corresponding moments implied by the model. Each row corresponds to a different model variant depending on the included reference points and the inclusion of the default option. The last column reports the marginal increase in the regression R-squared attributable to each model variant. Standard errors are displayed in parentheses.

Model Variants: Included Reference Points	Default Option	ρ_1	R-squared	Marginal (%)
None	x	2.605	0.395	N.A.
		(0.051)		
R_1 Only	x	2.582	0.406	2.78%
		(0.050)		
R_1 and R_2	x	1.062	0.576	41.87%
		(0.015)		
R_1 and R_2	✓	0.907	0.578	0.35%
		(0.012)		

Table 9: Price-Volume Correlation Decomposition

This table reports the regression results from $\Delta lnN_T = \alpha_2 + \rho_2 \Delta ln\hat{P}_T + \epsilon_T$, where ΔlnN_T are transaction volumes in year T from CoreLogic data and $\Delta ln\hat{P}_T$ is the price index in year T from FHFA data. The top panel reports the coefficient estimate ρ_2 from the data. Each row in the bottom panel corresponds to a different model variant depending on the included reference points and the inclusion of the default option. The last column reports the marginal increase of the model-implied price-volume correlation attributable to each model variant. Standard errors are displayed in parentheses.

Data		ρ_2	
		0.765	
		(0.344)	
Model Variants:			
Included Reference Points	Default Option	ρ_2	Marginal (%)
None	x	0.045	N.A.
R_1 Only	x	0.053	1.05%
R_1 and R_2	x	0.484	56.34%
R_1 and R_2	✓	0.670	24.31%

Figures

Figure 1: FHFA House Price Index

This figure presents the house price index based on the Federal Housing Finance Agency (FHFA) dataset.

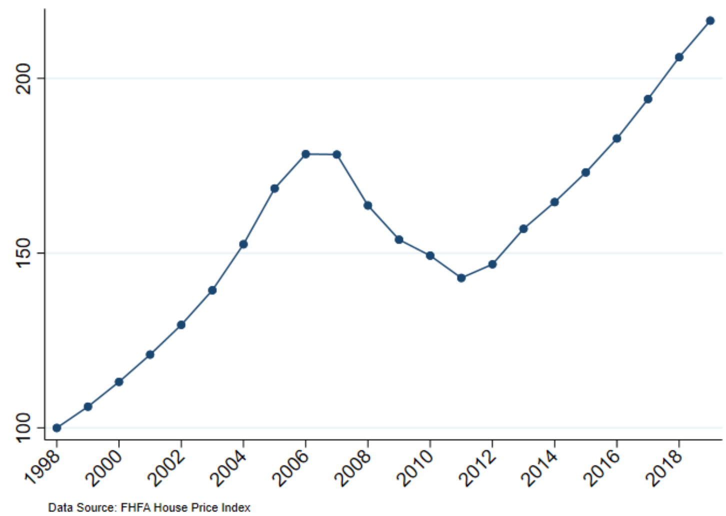


Figure 2: Annual Sales Volume and Refinance Mortgage Volume

This figure presents the annual volume of 15-year and 30-year refinance mortgages and house sales from CoreLogic Mortgage and Owner Transfer datasets. All series are indexed to the 1998 values.

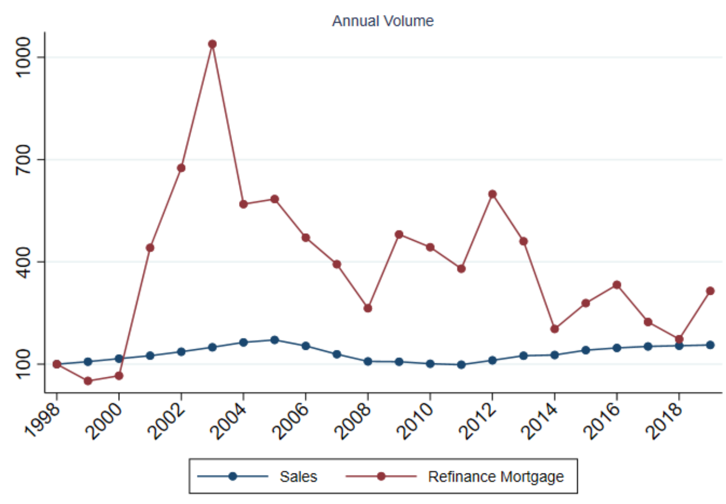


Figure 3: Identifying the Effect of Nominal Loss: Two Reference Prices

This figure illustrates the incidence of independent variations in two measures of nominal loss, each relating to a reference price. The purpose of this visualization is to demonstrate how the two measures of nominal loss can be separately identified. The figure in Panel A presents the distribution of listing years (represented by orange bars) and home purchases (represented by blue bars) conditional on the refinancing year in the final listing sample. The vertical axis displays the proportion of listings or purchases within a specific year relative to the total number of listings or home purchases. The figure in Panel B plots the joint distribution of the two measures of nominal gains at the time of listing. The color scheme refers to the relative frequency of observations in 10 percentage point bins, where darker bins indicate a higher density of observations. Bins that lack sufficient observations are shaded white. The dotted blue lines divide the joint distribution into four groups: (1) winners with respect to both reference prices ($G_1 \geq 0\%$ and $G_2 \geq 0\%$), covering 24% of the sample, (2) winners with respect to only the gain on purchase price ($G_1 \geq 0\%$ and $G_2 < 0\%$), accounting for 49% of the sample, (3) winners with respect to only the gain on appraised price ($G_1 < 0\%$ and $G_2 \geq 0\%$), accounting for 0% of the sample, (4) losers in relation to both reference prices ($G_1 < 0\%$ and $G_2 < 0\%$), comprising 27% of the sample.

Panel A: Distribution of Listing and Purchase Years

Panel B: Joint Distribution of Two Nominal Gains

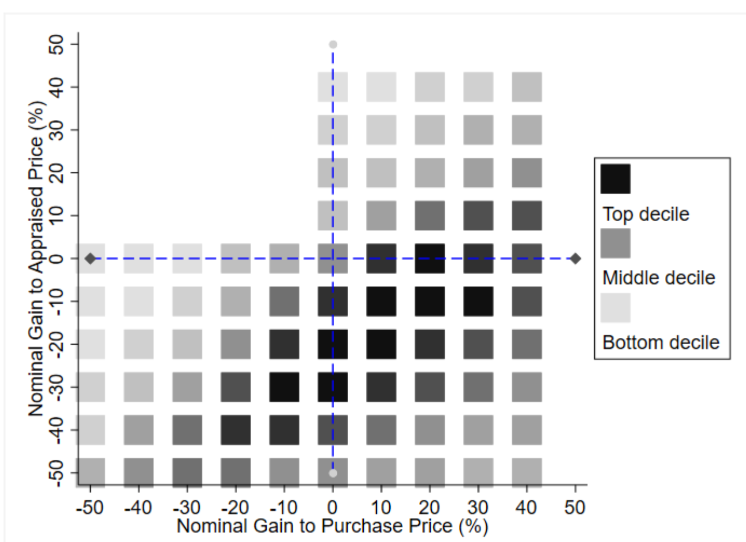
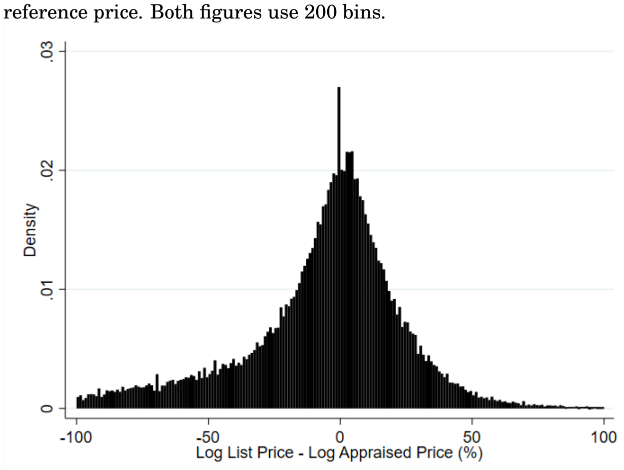


Figure 4: Listing Price Density

The two figures below present the density of the log percentage difference between the listing price and each of the reference prices. The left-hand side figure uses the appraised price as the reference price, while the right-hand side figure uses the initial purchase price as the reference price. Both figures use 200 bins.



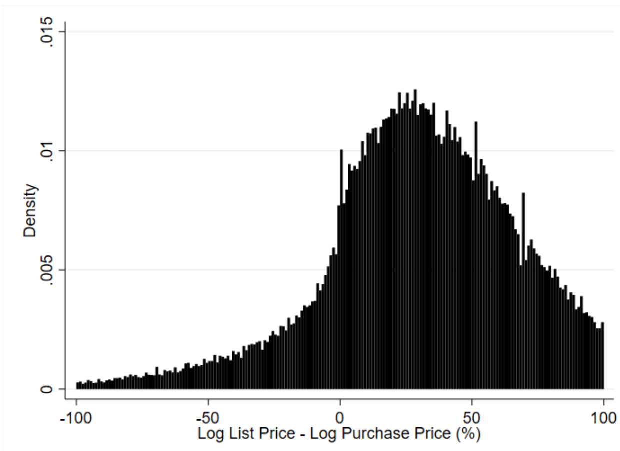
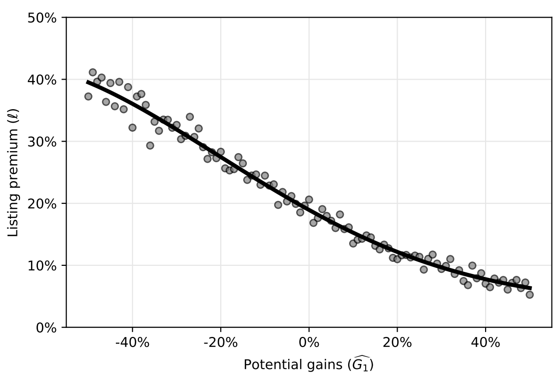


Figure 5: Listing Premium and Potential Gains

The figures presented below each display a binned-scatter plot and a fourth-order polynomial curve fit between the listing premium (denoted as ℓ) and the measure of potential gains (rounded to 1 percentage point). The solid line represents the fitted curve. The left figure uses the initial purchase price as the reference price, while the right figure uses the appraised price as the reference price.



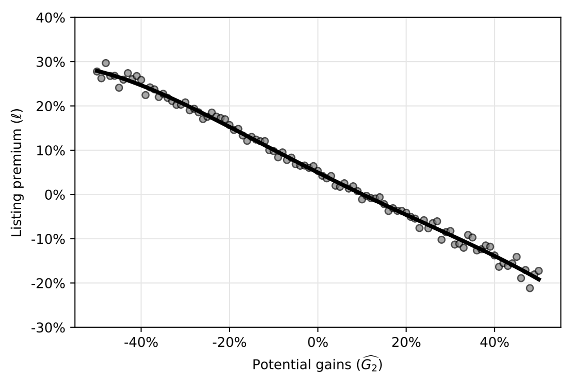
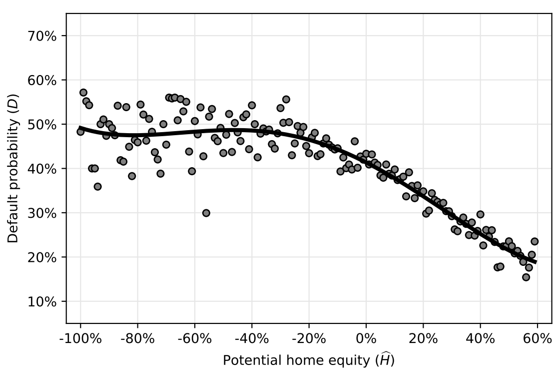


Figure 6: Home Equity, Default Decision and Listing Premium

Each figure below presents a binned-scatter plot and a polynomial curve fit between the variables of interest. The left figure displays the average default probability since listing across the potential home equity ($\hat{H}$) domain (rounded to 1 percentage point), with the solid line representing the fourth-order polynomial curve fit. The right figure displays the average listing premium across the potential home equity ($\hat{H}$) domain (rounded to 1 percentage point), with a solid line as the fourth-order polynomial curve fit.



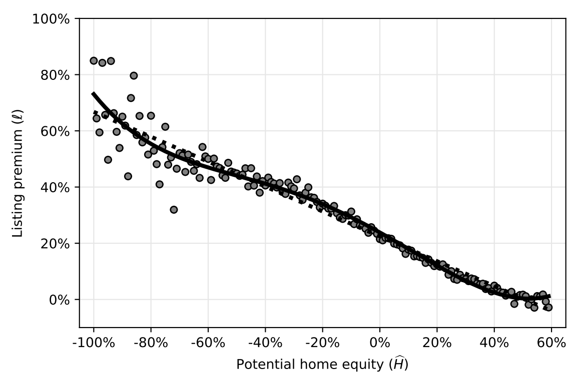


Figure 7: Home Equity and Listing Premium: Heterogeneity in Default Cost "c"

The two figures below present binned-scatter plots and fourth-order polynomial curve fits between the listing premium ℓ and potential home equity $\hat{H}$ (rounded to 1 percentage point) for different sample subsets, illustrating the role of default cost in affecting the listing premium. The left figure displays the average listing premium ℓ across each 1-percentage point bin of potential home equity ($\hat{H}$), differentiated by owner-occupied status as self-reported by the households. The green scatters and line correspond to the sample where properties are the primary residence of the households, while the orange scatters and line represent the sample where properties are not the primary residence location of the households. The right figure shows the average listing premium ℓ across each 1-percentage point bin of potential home equity ($\hat{H}$), depending on the borrower's credit quality at loan origination. The green scatters and line correspond to the sample where borrowers' FICO Scores are between 620 and 660 (inclusive), while the orange scatters and line represent the sample where borrowers' FICO Scores are either below 620 or above 660 (exclusive).

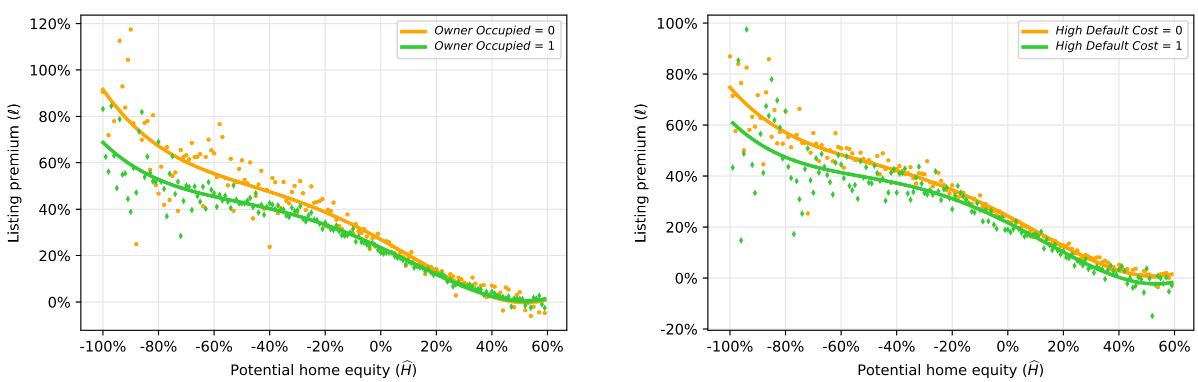


Figure 8: Concave Demand

The figure presents the average probability of sale (i.e., $\alpha(\ell)$) within 26 weeks across each 1-percentage point bin of the listing premium. The dashed line indicates a linear fit and the sold line indicates a generalized logistic function fit.

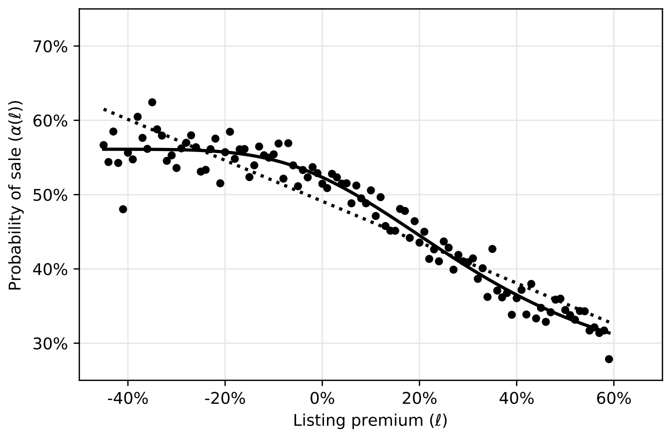


Figure 9: Price Realization

The figure presents the average realized price premium (i.e., $\beta(\ell)$) for sold listings across each 1-percentage point bin of the listing premium. The dashed line indicates a linear fit.

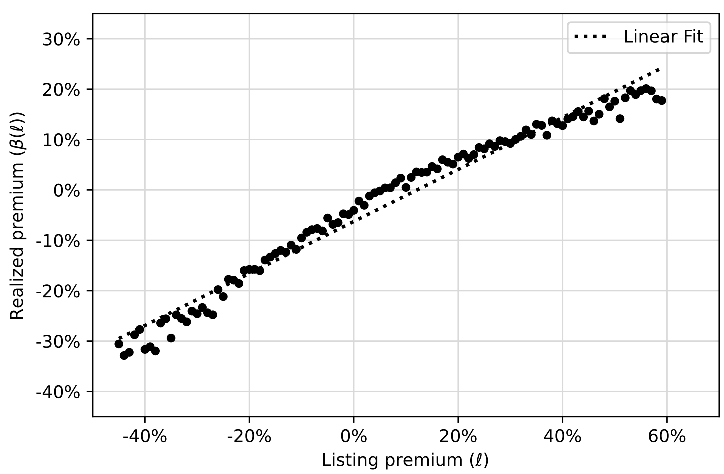


Figure 10: Extensive Margin

The figure presents the average annual probability of listing a property for sale across each 1-percentage point bin of potential home equity for the properties in the final listing sample. I first expand the listing sample to a property-year panel, with the panel starting year being 2007 and the last year being the lesser of 2015 or the year of the second sale. Then, I calculate the potential home equity for each property-year by imputing the difference between the sum of the property base value and the census tract price index averaged in a given year, and the outstanding mortgage balance. Finally, I compute the ratio of the number of properties listed for sale to the number of housing stock in this property-year panel sample for each 1-percentage point bin of potential home equity.

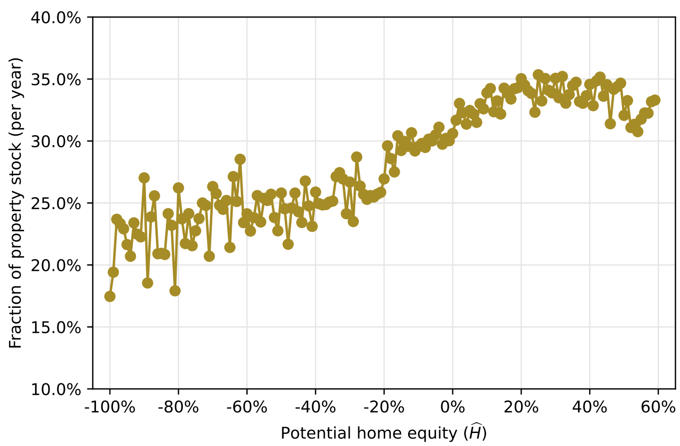


Figure 11: Model Fit Overview

The figures reports the set of data and model-implied moments. The model moments are evaluated at the estimated structural parameters in Table 7. The lines correspond to fourth-order polynomial curve fits.

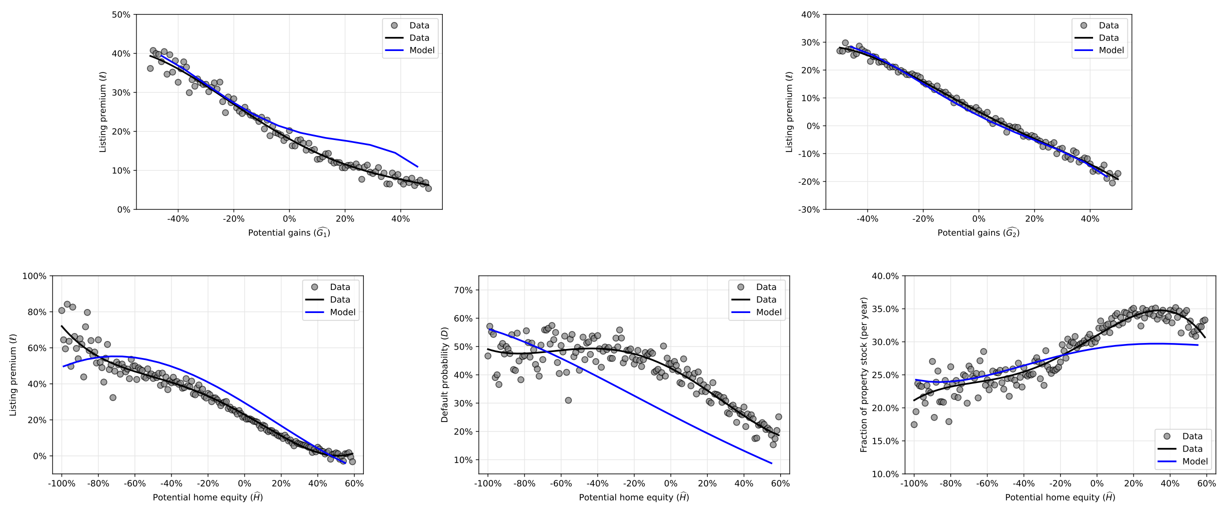


Figure 12: Interaction between Loss Aversion and Financial Constraints

The figures report the estimated slope of the listing premium along the potential gains and home equity dimension, conditioning on different levels of home equity and potential gains shown on the horizontal axis. The vertical intervals represent 99% confidence intervals for each point estimate.

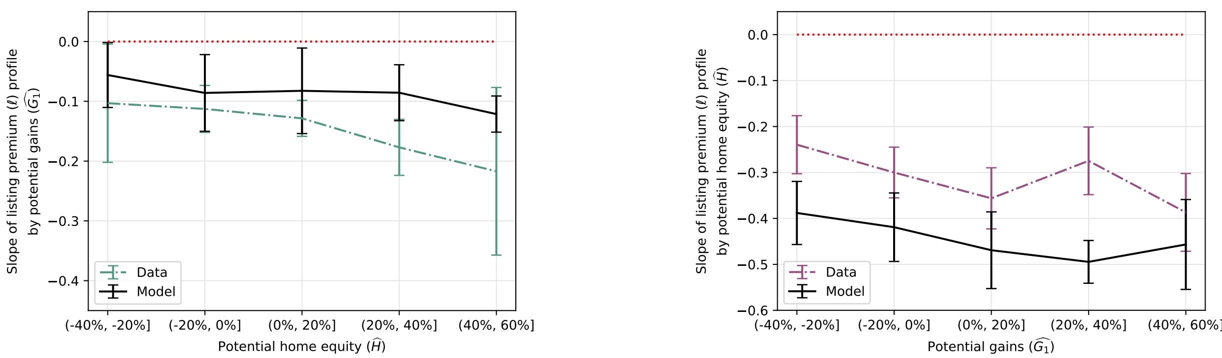
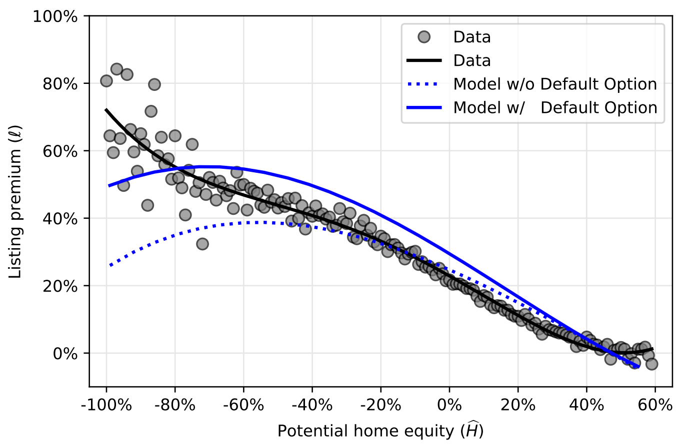


Figure 13: The Role of the Default Option

The figure presents the relationships between the listing premium ℓ and potential home equity $\hat{H}$, as seen in the data and models. The dashed and solid lines represent fourth-order polynomial curve fits. The dashed blue line illustrates the listing premium profile derived from a model that excludes a default option, whereas the solid blue line illustrates the listing premium profile using a model that includes a default option.



Appendix

A1: CoreLogic Mortgage, Owner Transfer, Multiple Listing Service and Property Assessment Merging Process

***Match between the CoreLogic Mortgage and Owner Transfer datasets*:** The goal of this matching process is to identify the mortgage contract associated with a sales transaction. Firstly, I filter the mortgage sample as described in Section 4.1 and require CoreLogic's proprietary property ID and loan origination date to be non-missing. Then, I filter the property transaction sample to include arms-length transactions, sales confirmed by the county assessor's office, and standard sales deeds. Additionally, I exclude short sales, real-estate-owned sales, foreclosure sales, and interfamily transfers. I also require the property types to be for residential use and all relevant variables to be non-missing, including CoreLogic's proprietary property ID and sales date. More specifically, for this matching exercise, I exclude cash purchases.

Upon completing the initial data cleaning, the matching procedure is straightforward. I match on: (1) CoreLogic's proprietary property ID; (2) property sales date and loan origination date, ensuring they are closest and within [-30, 30] days; and (3) retain unique matches. There are 52 million mortgages and 116 million property sales that can be matched. The match ratio in this exercise is approximately 91%.

***Match between the CoreLogic Multiple Listing Service and Owner Transfer datasets*:** The aim of this matching process is to associate listing information, such as listing date and time on the market, with a sales transaction. Firstly, I clean the owner transfer sample following the same cleaning procedure mentioned earlier, with the exception of including full-cash purchases for this match. For cleaning the MLS sample, I retain properties that are listed for sale (not for rent/lease) and exclude listings for short sales, real-estate-owned sales, and foreclosure sales. I also require the property types to be for residential use and all relevant variables to be non-missing, including CoreLogic's proprietary property ID and listing start and end dates. In light of the complexity of real listing episodes and the potential for multiple data entries corresponding to the same property with overlapping listing periods, I follow the methodology of Garriga and Hedlund 2020 to further refine the MLS data. This is achieved by consolidating overlapping listing spells associated with the same property, and subsequently retaining only listings identified as closed sales for the matching process.

The matching process proceeds as follows: (1) match based on CoreLogic's proprietary property ID; and (2) ensure the property sales date and sold listing's closing date are closest and within [-30, 30] days. There are 48 million closed transactions from MLS and 116 million property sales that can be matched with a sold listing. The match ratio in this exercise is 72.4%.

***Match between the CoreLogic Property Assessment and Owner Transfer/Mortgage datasets*:** The objective of this matching process is to acquire the property characteristics associated with a sale or mortgage transaction. To clean the property assessment data, I exclude foreclosed properties and retain those identified as for residential use by CoreLogic. Additionally, I require CoreLogic's proprietary property ID and the tax assessment year of the property to be non-missing.

The matching process involves: (1) CoreLogic's proprietary property ID; and (2) ensuring the sale/financing year and property tax assessment year are the closest. The match ratios for both the property transaction data and the mortgage data are above 99%.

A2. CoreLogic-BlackBox Match Evaluation

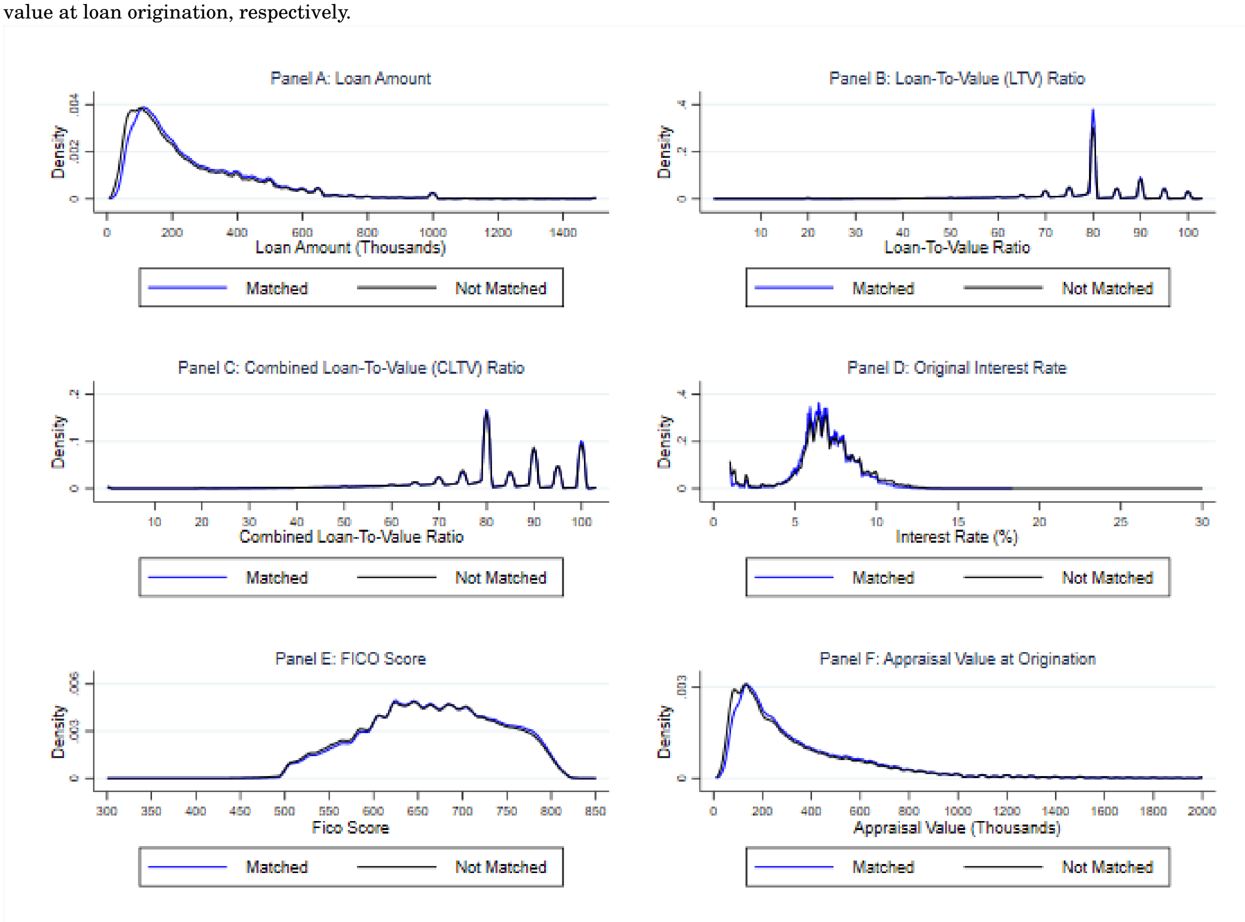
Although I have made every effort to be as careful as possible in my matching process, potential mismatches are still unavoidable and may bias the statistical inference of my empirical analysis. There are several ways to assess the accuracy and representativeness of my match.

First, CoreLogic classifies mortgages based on their eligibility for agency mortgages. Although I do not specifically aim to incorporate this information during the matching procedure, it is remarkable that none of the BlackBox loans are matched with loans eligible for GSEs from CoreLogic. This untargeted moment serves as a validation of the quality of my matching process.

Second, to ensure that the final matched sample is not a selected sample, I compare the density distribution between the matched and unmatched loans in the BlackBox sample. Figure A1 displays the comparisons across six observed loan, borrower, and property characteristics, including original loan amount, LTV, CLTV, interest rate at origination, borrower FICO score, and property appraised value at origination. As illustrated in this figure, the distributions are relatively balanced between the two groups, suggesting that my match does not result in a selected sample and my statistical inference can be extrapolated to the broader non-agency market.

Figure A1: Epanechnikov kernel densities for matched and non-matched BlackBox loans

This figure provides a comparison of the Epanechnikov kernel densities of observable variables for BlackBox loans that are matched with mortgages in the CoreLogic Mortgage dataset and the BlackBox loans that remain unmatched. Panels A through F display the densities of the loan amount (in thousands of dollars), LTV, combined-LTV, interest rate at loan origination, FICO Score, and property appraisal value at loan origination, respectively.



A3: Loss Aversion under Two Reference Points - Alternative Specifications and Robustness

Alternative specifications: The analytical solution of Equation (6) under a simplified version of the model indicates that the home equity level at listing predicts the listing premium and drives household default decisions. In Table A1, I explore a set of alternative specifications by replacing LTV with a linear spline function with two knots for home equity to examine the effects of down-payment constraints and default options.

The results consistently demonstrate that sellers exhibit a salient effect of loss aversion when using appraised price as the anchoring point. Moreover, the negative relationship between potential home equity and listing price, with steeper slopes at 20% (i.e., down-payment constraints) and -80% (i.e., default option) home equity levels, aligns with the model's predictions.

Multicollinearity: Another potential concern arises from the high correlation between the two nominal loss measures[38] by construction, as they are both measured relative to the hedonic predicted value. This correlation could lead to one measure absorbing the variation of the other, resulting in differences in significance levels and magnitudes of the coefficients. To address this multicollinearity concern, I estimate the effect of loss aversion for the two reference prices separately in Table A2. The results prove to be robust. In the linear loss specifications, the appraised price exhibits a larger impact on the listing price, as indicated by both the larger coefficients and higher R-squared values

Sample selection: On one hand, the refinancing decision is primarily driven by the aggregate economic environment, as discussed by Dunn and McConnell 1981, Schwartz and Torous 1989, Stanton 1995, Deng, Quigley, and Van Order 2000, and Koijen, Van Hemert, and Van Nieuwerburgh 2009. On the other hand, Bennett, Peach, and Peristiani 2001 and Agarwal, Rosen, and Yao 2016 provide evidence that borrower sophistication and awareness of the refinancing option can also explain the refinancing decision. As a result, by focusing on households who have refinanced, I might be dealing with a selected sample. To address this sample selection issue, I include three indicator variables (refinanced once, twice, and more than twice, respectively) as proxies for borrower sophistication in Table A3. The results remain robust under these alternative specifications.

Repeat sales approach: To address prediction errors in the hedonic model stemming from unobserved, time-invariant housing quality, I utilize a repeat sales model approach. This requires amending equation (8) to incorporate property fixed effects. Subsequently, I re-estimate equation (7) using

38. The correlation of the two variables is 0.67.

these adjusted loss metrics. Results depicted in Table A4 underscore the enduring significance of loss aversion effects when employing the appraised price, rather than the purchase price, as the reference point under this alternative hedonic model.

Instrumental variables: In this robustness check, I explore an institutional feature of the appraisal process that may produce potentially credible exogenous variation in the final appraised price. Specifically, appraisers rely on "comps", nearby properties with similar attributes that are sold recently, during the appraisal process. Therefore, if potential candidates are significantly different from the subject property, the appraised price is more likely to exhibit greater biases. On the contrary, when properties in the surrounding neighborhood are more homogeneous, it provides limited scope for appraisers to deviate from the current market value, leading to reduced pricing dispersion (Jiang and Zhang 2022).

I measure the heterogeneity of properties transacted in a local housing market ("comps") using the coefficient of variation (CV) (Van Nieuwerburgh and Weill 2010), which is determined by the ratio of the standard deviation to the mean, providing a scale-neutral measure of dispersion. Appraisers typically take into account seven hedonic attributes: property age, land size, living area, number of bedrooms and bathrooms, and the property's geolocation. These attributes are subsequently used in calculating the CV. Specifically, the CV for a particular hedonic attribute is calculated using all properties sold in census tract n in month $m-1$ (the month before the refinancing occurs), as follows:

$$CV^{\Theta}_{nm-1} = \frac{\sigma^{\Theta}_{nm-1}}{\mu^{\Theta}_{nm-1}}, \quad \forall \Theta \in \{age, landsize, livingsize, bed, bath, geo\}$$

Unlike Genesove and C. Mayer 2001, one of the measurement errors (ω_{im}) in the loss measured by the appraised price in this study is from idiosyncratic appraisal bias at the point of refinancing, rather than from overpayment or underpayment by the current homeowner at the point of purchase. Thus, the identification assumptions are that appraisers can fully capture the "unobserved" housing quality without errors and the remaining exogenous variation comes from the appraisal bias. Moreover, it is reasonable to argue that the heterogeneity of the potential comps sold one month before refinancing should affect the listing price solely through the appraisal bias (exclusion restriction).

The following Two-Stage Least Squares (2SLS) specification is then estimated:

Stage 1: $\text{Loss\_AppraisalPrice}_{inmt} = \beta_1 CV^{age}_{nm-1} + \beta_2 CV^{landsize}_{nm-1} + \beta_3 CV^{livingsize}_{nm-1} + \beta_4 CV^{bed}_{nm-1} + \beta_5 CV^{bath}_{nm-1} + \beta_6 CV^{geo}_{nm-1}$

$+ \beta LTV_{int} + \delta \widehat{Log\ Price}_{int} + \alpha_0 + \epsilon_{inmt}$

Stage 2: $\log(Y_{insmt}) = m_1 \text{Loss\_PurchasePrice}_{inst} + m_2 \text{Loss\_AppraisalPrice}_{inmt}$

$+ \beta LTV_{int} + \delta \widehat{Log\ Price}_{int} + \alpha_0 + \epsilon_{insmt}$

I further confirm that the baseline results remain qualitatively unchanged and even yield a larger coefficient estimate for m_2 as displayed in Table A5. The statistics from the weak identification test suggest that the instruments used are strong predictors of the loss measured by the appraised price.

Table A1: Loss Aversion under Two Reference Points: Robustness - Alternative Specifications

This table presents the regression results examining the effect of loss aversion relative to two reference prices, using the final listing dataset and an alternative specification. The dependent variable is the log of the original listing price. Loss: Appraised Price is the maximum of the difference between the log appraised price of the last refinance mortgage and the hedonic predicted market price, and 0. Loss-squared: Appraised Price is the squared term. Loss: Purchase Price is the maximum of the difference between the log purchase price and the hedonic predicted market price, and 0. Loss-squared: Purchase Price is the corresponding squared term. H represents the potential home equity, measured as the difference between the hedonic predicted price and the log of (1 + Outstanding Mortgage Balance) at the month of listing. H if ≤ 20% is the minimum of the difference between H and 0.2, and 0. H if ≤ -80% is the minimum of the difference between H and -0.8, and 0. Estimated Value refers to the base value of the property estimated from the hedonic model without including the census tract-specific monthly price index. Estimated Price Index is the census tract-specific monthly price index. Residual is the error term of the previous purchase. Months since last refinance measures the difference in months between the last refinance and the current listing, while Months since last purchase measures the difference in months between the last purchase and the current listing. Bootstrapped standard errors are displayed in parentheses.

	(1)	(2)	(3)	(4)
VARIABLES	Log (Original Listing Price)			
Loss: Appraised Price	0.409***	0.404***	0.551***	0.515***
	(0.007)	(0.007)	(0.015)	(0.015)
Loss-squared: Appraised Price			-0.136***	-0.103***
			(0.016)	(0.015)
Loss: Purchase Price	0.080***	-0.023**	0.127***	-0.010
	(0.010)	(0.009)	(0.015)	(0.013)
Loss-squared: Purchase Price			-0.007	0.015
			(0.012)	(0.013)
H	-0.003***	-0.003***	-0.004***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
H if ≤ 20%	-0.169***	-0.147***	-0.127***	-0.119***
	(0.008)	(0.009)	(0.009)	(0.007)
H if ≤ -80%	-0.126	-0.255***	-0.467***	-0.489***
	(0.095)	(0.071)	(0.083)	(0.073)
Estimated Value	1.043***	1.050***	1.042***	1.050***
	(0.003)	(0.003)	(0.003)	(0.003)
Estimated Price Index	1.023***	1.006***	1.023***	1.006***
	(0.002)	(0.002)	(0.002)	(0.002)
Residuals		0.271***		0.264***
		(0.006)		(0.005)
Months since	-0.000***	-0.000***	-0.000**	-0.000***
last purchase	(0.000)	(0.000)	(0.000)	(0.000)
Months since	-0.001***	-0.000***	-0.001***	-0.000***
last refinance	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.472***	-0.527***	-0.472***	-0.529***
	(0.036)	(0.032)	(0.034)	(0.031)
Observations	97,635	97,635	97,635	97,635
R-squared	0.863	0.870	0.864	0.870

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A2: Loss Aversion under Two Reference Points: Robustness - Alternative Specifications

This table presents the regression results in separate specifications to examine the effect of loss aversion relative to two reference prices using the final listing dataset. The dependent variable is the log of the original listing price. Columns (1) - (4) display the regression results for the expected nominal loss measure using the purchase price, while columns (5) - (8) display the regression results for the expected nominal loss measure using the appraised price of the last refinance mortgage. Loss-squared is the corresponding term squared. LTV if ≥ 80 represents the maximum value between the difference of the current LTV and 0.8, and 0. Estimated Value refers to the base value of the property estimated from the hedonic model without including the census tract-specific monthly price index. Estimated Price Index is the census tract-specific monthly price index. Residual is the error term from the previous purchase. Months since last refinance measures the difference in months between the last refinance and the current listing, while Months since last purchase measures the difference in months between the last purchase and the current listing. Bootstrapped standard errors are displayed in parentheses

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Log (Original Listing Price)							
	Purchase price				Appraised price			
Loss	0.408***	0.299***	0.570***	0.425***	0.482***	0.430***	0.588***	0.519***
	(0.010)	(0.010)	(0.020)	(0.018)	(0.005)	(0.005)	(0.015)	(0.012)
Loss-squared			-0.143***	-0.110***			-0.093***	-0.077***
			(0.022)	(0.017)			(0.016)	(0.013)
LTV if ≥ 80%	0.002***	0.002***	0.002***	0.002***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Estimated Value	1.014***	1.022***	1.013***	1.021***	1.047***	1.052***	1.047***	1.052***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Estimated Price Index	0.990***	0.972***	0.991***	0.974***	1.023***	1.008***	1.022***	1.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Residuals		0.293***		0.286***		0.274***		0.272***
		(0.007)		(0.006)		(0.006)		(0.005)
Months since last purchase	-0.000***	-0.000***	-0.000***	-0.000***				
	(0.000)	(0.000)	(0.000)	(0.000)				
Months since last refinance					-0.001***	-0.001***	-0.001***	-0.001***
					(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.064*	-0.136***	-0.063	-0.133***	-0.524***	-0.562***	-0.535***	-0.571***
	(0.036)	(0.036)	(0.041)	(0.040)	(0.030)	(0.032)	(0.033)	(0.034)
Observations	97,635	97,635	97,635	97,635	97,635	97,635	97,635	97,635
R-squared	0.844	0.852	0.845	0.853	0.859	0.867	0.860	0.867

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A3: Loss Aversion under Two Reference Points: Robustness - Borrower Sophistication

This table presents the robustness of regression results examining the effect of loss aversion relative to two reference prices using the final listing dataset. In addition to the main explanatory variables, all columns include a set of dummy variables (not shown) to control for the number of originated refinancing mortgages between the initial purchase and the current listing as a proxy for borrower sophistication. The dependent variable is the log of the original listing price. Loss: Appraised Price represents the maximum value between the difference of the log appraised price from the last refinance mortgage and the hedonic predicted market price, and 0. Loss-squared: Appraised Price is the squared term of this value. Loss: Purchase Price denotes the maximum value between the difference of the log purchase price and the hedonic predicted market price, and 0. Loss-squared: Purchase Price is the corresponding squared term. LTV if ≥ 80 represents the maximum value between the difference of the current LTV and 0.8, and 0. Estimated Value refers to the base value of the property estimated from the hedonic model without including the census tract-specific monthly price index. Estimated Price Index is the census tract-specific monthly price index. Residual is the error term from the previous purchase. Months since last refinance measures the difference in months between the last refinance and the current listing, while Months since last purchase measures the difference in months between the last purchase and the current listing. Bootstrapped standard errors are displayed in parentheses.

	(1)	(2)	(3)	(4)
VARIABLES	Log (Original Listing Price)			
Loss: Appraised Price	0.446***	0.436***	0.567***	0.527***
	(0.006)	(0.006)	(0.013)	(0.011)
Loss-squared: Appraised Price			-0.124***	-0.089***
			(0.016)	(0.011)
Loss: Purchase Price	0.094***	-0.008	0.147***	0.008
	(0.010)	(0.009)	(0.019)	(0.014)
Loss-squared: Purchase Price			-0.010	0.012
			(0.016)	(0.012)
LTV if ≥ 80%	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Estimated Value	1.047***	1.053***	1.046***	1.053***
	(0.003)	(0.003)	(0.003)	(0.003)
Estimated Price Index	1.027***	1.009***	1.026***	1.009***
	(0.002)	(0.002)	(0.002)	(0.002)
Residuals		0.275***		0.268***
		(0.005)		(0.006)
Months since	-0.000*	-0.000***	-0.000	-0.000***
last purchase	(0.000)	(0.000)	(0.000)	(0.000)
Months since	-0.001***	-0.001***	-0.001***	-0.001***
last refinance	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.518***	-0.569***	-0.523***	-0.574***
	(0.033)	(0.035)	(0.040)	(0.039)
Observations	97,635	97,635	97,635	97,635
R-squared	0.860	0.867	0.861	0.868

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A4: Loss Aversion under Two Reference Points: Robustness - Repeat Sales Model

This table presents the regression results that examine the effect of loss aversion relative to two reference prices using the final listing dataset. The hedonic model employs the repeat sales approach and add the property fixed effects as additional controls. The dependent variable is the log of the original listing price. Loss: Appraised Price represents the maximum value between the difference of the log appraised price from the last refinance mortgage and the hedonic predicted market price, and 0. Loss-squared: Appraised Price is the squared term of this value. Loss: Purchase Price denotes the maximum value between the difference of the log purchase price and the hedonic predicted market price, and 0. Loss-squared: Purchase Price is the corresponding squared term. LTV if ≥ 80 represents the maximum value between the difference of the current LTV and 0.8, and 0. Estimated Value refers to the base value of the property estimated from the hedonic model without including the census tract-specific monthly price index. Estimated Price Index is the census tract-specific monthly price index. Residual is the error term from the previous purchase. Months since last refinance measures the difference in months between the last refinance and the current listing, while Months since last purchase measures the difference in months between the last purchase and the current listing. Bootstrapped standard errors are displayed in parentheses.

	(1)	(2)	(3)	(4)
VARIABLES	Log (Original Listing Price)			
Loss: Appraised Price	0.419***	0.415***	0.502***	0.500***
	(0.007)	(0.008)	(0.022)	(0.018)
Loss-squared: Appraised Price			-0.067***	-0.070***
			(0.020)	(0.018)
Loss: Purchase Price	-0.118***	-0.102***	-0.209***	-0.186***
	(0.011)	(0.012)	(0.036)	(0.031)
Loss-squared: Purchase Price			0.086**	0.082***
			(0.034)	(0.032)
LTV if ≥ 80%	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Estimated Value	0.982***	0.981***	0.981***	0.981***
	(0.002)	(0.002)	(0.002)	(0.002)
Estimated Price Index	0.808***	0.813***	0.809***	0.814***
	(0.008)	(0.008)	(0.007)	(0.008)
Residuals		-0.082***		-0.080***
		(0.012)		(0.014)
Months since	-0.001***	-0.001***	-0.001***	-0.001***
last purchase	(0.000)	(0.000)	(0.000)	(0.000)
Months since	-0.001***	-0.001***	-0.001***	-0.001***
last refinance	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.385***	0.385***	0.387***	0.386***
	(0.025)	(0.024)	(0.021)	(0.021)
Observations	70,773	70,773	70,773	70,773
R-squared	0.869	0.870	0.870	0.870

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A5: Loss Aversion under Two Reference Points: Robustness - Instrumental Variables

This table presents the regression results that examine the effect of loss aversion relative to two reference prices using the final listing dataset. Columns 1-2 report OLS results from Table 2, and columns 3-5 report 2SLS results. The dependent variable is the log of the original listing price. Loss: Appraised Price represents the maximum value between the difference of the log appraised price from the last refinance mortgage and the hedonic predicted market price, and 0. Loss-squared: Appraised Price is the squared term of this value. Loss: Purchase Price denotes the maximum value between the difference of the log purchase price and the hedonic predicted market price, and 0. Loss-squared: Purchase Price is the corresponding squared term. LTV if ≥ 80 represents the maximum value between the difference of the current LTV and 0.8, and 0. Estimated Value refers to the base value of the property estimated from the hedonic model without including the census tract-specific monthly price index. Estimated Price Index is the census tract-specific monthly price index. Months since last refinance measures the difference in months between the last refinance and the current listing, while Months since last purchase measures the difference in months between the last purchase and the current listing. Standard errors are displayed in parentheses.

	OLS		2SLS		
	(1)	(2)	(3)	(4)	(5)
VARIABLES	Log (Original Listing Price)				
Loss: Appraised Price	0.444***	0.565***	0.764***	0.706***	1.284***
	(0.006)	(0.013)	(0.034)	(0.052)	(0.186)
Loss-squared: Appraised Price		-0.124***			-0.652***
		(0.015)			(0.125)
Loss: Purchase Price	0.095***	0.150***		-0.118***	-0.094
	(0.008)	(0.015)		(0.040)	(0.065)
Loss-squared: Purchase Price		-0.011			0.150***
		(0.010)			(0.040)
LTV if ≥ 80%	0.001***	0.001***	-0.001***	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Estimated Value	1.046***	1.045***	1.062***	1.057***	1.052***
	(0.003)	(0.003)	(0.005)	(0.005)	(0.006)
Estimated Price Index	1.025***	1.025***	1.072***	1.053***	1.047***
	(0.002)	(0.002)	(0.005)	(0.005)	(0.005)
Months since	-0.000***	-0.000***		-0.001***	-0.000***
last purchase	(0.000)	(0.000)		(0.000)	(0.000)
Months since	-0.001***	-0.001***	-0.001***	-0.000***	-0.001***
last refinance	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.505***	-0.511***	-0.783***	-0.675***	-0.672***
	(0.037)	(0.039)	(0.060)	(0.067)	(0.085)
Observations	97,635	97,635	58,283	58,283	58,283
R-squared	0.860	0.861	0.848	0.852	0.847
Underidentification t-stat			1166.403	701.606	182.990
Weak identification t-stat			169.997	101.426	26.217

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

A4. Alternative Functional Form of Down-payment Penalty Function

As noted by Andersen et al. 2022, the estimated structural parameters might be sensible to the assumption regarding the functional form of the down-payment constraints. To address this concern, I confirm that the key household preference parameters are virtually unaffected when a convex penalty function is employed in this section. Specifically, the penalty function is formulated as follows:

$$\kappa(P(\ell)) = \bar{\mu}(\gamma\hat{P} - H(\ell))^2 \tag{13}$$

where:

$$\bar{\mu} = \begin{cases} \mu & \text{if } H(\ell) < \gamma\hat{P} \\ 0 & \text{if } H(\ell) \geq \gamma\hat{P} \end{cases}$$

Figure A2 depicts the overall fit of this alternative model. It suggests that this model can adequately capture the levels and relationships observed in the data, with a slight underprediction of the default probability profile. Furthermore, Tables A6 and A7 report the prediction errors and the estimated parameters of this alternative model, respectively. The overall prediction error slightly increases, but still falls within an acceptable range. As shown in Table A7, this model produces very robust results with reference dependence parameters and loss aversion parameters of η_1 = 0.022, η_2 = 0.393, λ_1 = 1.322, and λ_2 = 3.735. Again, these estimates indicate that sellers strongly anchor to the most recent property valuations and exhibit heightened aversion to the consequent nominal loss. Quantitatively, sellers are about 2.8 times more loss averse to this updated valuation than to the original purchase price. It is worth noting that the financial constraints parameters μ and the disutility associated with listing ϕ are sensitive to the functional form of the penalty function. Different functional forms would result in differences in the disutility from the down-payment penalty for the same home equity position, thus leading to differences in how responsive a seller is to the down-payment constraints and also the extensive margin selling decision, which is reflected by the search and listing cost.

Figure A2: Model Fit Overview - Alternative Penalty Function

The figures reports the set of data and model-implied moments using a convex down-payment penalty function. The model moments are evaluated at the estimated structural parameters in Table A6. The lines correspond to fourth-order polynomial curve fits.

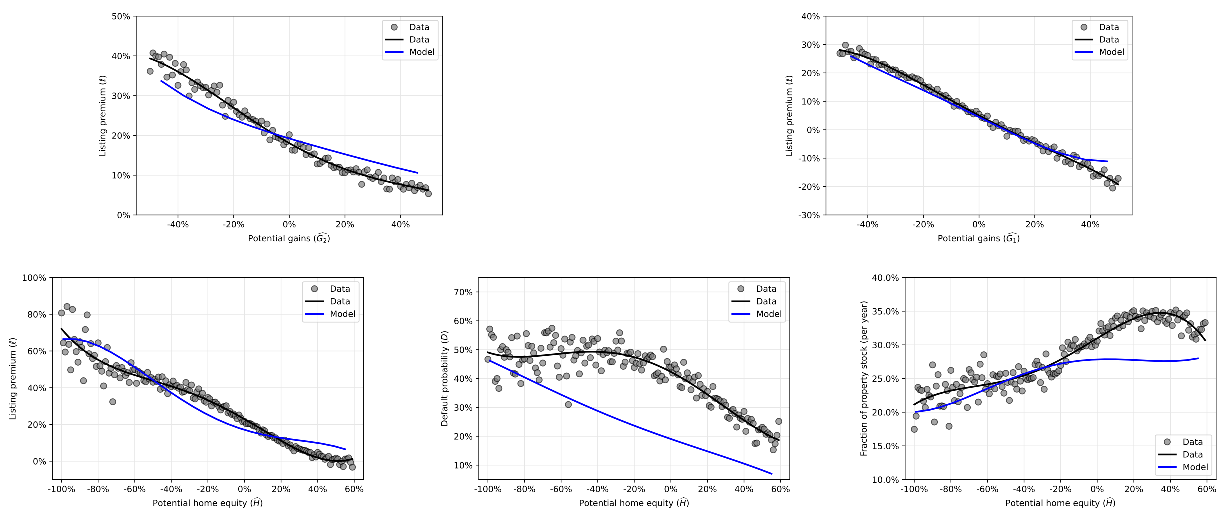


Table A6: Goodness of Fit - Alternative Penalty Function

This table reports root mean squared differences between the level of each moment in the model and the data, normalized by the mean value of each data moment. The model is estimated using a model variant with a convex down-payment penalty function. This prediction error is explainable in percentage terms, as reported below for each moment separately, as well as jointly for the full set of moments shown in the last column.

	$\ell(\hat{G}_1)$	$\ell(\hat{G}_2)$	$\ell(\hat{H})$	$f_{default}(\hat{H})$	$f_{list}(\hat{H})$	Overall Fit
RMSE	20.1%	38.4%	23.3%	43.8%	14.1%	28.0%

Table A7: Structural Parameter Estimates - Alternative Penalty Function

This table reports structural parameter estimates resulting from the method of moments estimation using a convex down-payment penalty function.

	Reference Dependence η_1	Reference Dependence η_2	Loss Aversion λ_1	Loss Aversion λ_2	Financial Constraints μ	Average Motivation Shock θ_m	St. dev. of Motivation Shock σ_m	Average Default Cost θ_c	St. dev. of Default Cost σ_c	Search and Listing Cost ϕ
Parameter Estimates	0.022	0.393	1.322	3.735	0.578	2.984	0.916	0.473	0.550	0.058