FEDERAL RESERVE BANK *of* ATLANTA WORKING PAPER SERIES

Mortgage Lock-in, Lifecycle Migration, and the Welfare Effects of Housing Market Liquidity

Kristopher Gerardi, Franklin Qian, and David Zhang

Working Paper 2024-15
October 2024

Abstract: We use a search and matching model to study the heterogeneous welfare effects of housing market illiquidity due to mortgage lock-in over the lifecycle. We find that younger home buyers are disproportionately affected by mortgage lock-in, which disrupts their typical pattern of moving to higher-quality neighborhoods. We estimate a model with heterogeneous seller-buyers bargaining within markets defined by CBSA-income terciles and with endogenous migration across markets. We find that on average mortgage lock-in reduces household listing probabilities by 21 percent to 23 percent, increases time on the market by 52 percent to 142 percent, increases house prices by 3 percent to 8 percent, and decreases match surplus by 3 percent to 29 percent through its effects on the search and matching process. The pricing and match surplus effects are larger for younger households and for households in lower-income areas, due to a higher idiosyncratic dispersion in buyer valuation leading to larger match surplus variation in those areas. Our study highlights the welfare benefits of market thickness in real estate markets.

JEL classification: G18, G21, E52

Key words: mortgage lock-in, moving to opportunity, housing market liquidity, idiosyncratic dispersion in house prices, FRMs

https://doi.org/10.29338/wp2024-15

The authors thank Elliot Annenberg, John Campbell, Jack Liebersohn, Lu Liu, and Gregor Schubert for helpful discussions. They also thank Hansen Pen and Payten Werbowsky for excellent research assistance. The views expressed here are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Kris Gerardi is with the Federal Reserve Bank of Atlanta. Franklin Qian is with the University of North Carolina at Chapel Hill. David Zhang is with Rice University. Please address questions regarding content to Kristopher Gerardi, Federal Reserve Bank of Atlanta, 1000 Peachtree Street NE, Atlanta, GA 30309, kristopher.gerardi@atl.frb.org.

1 Introduction

The liquidity of housing markets plays a pivotal role in shaping both the buyer-seller match surplus and overall market outcomes under idiosyncratic buyer valuations. A thicker market, characterized by a greater number of potential sellers and buyers in the market, could enhances the match surplus and promote household mobility. We examine the effect of mortgage lock-in, including due to the recent rise in interest rates commencing in the second quarter of 2022, on these dynamics. Furthermore, we explore the heterogeneous effects across various household demographics and geographic markets. To this end, we employ a search and matching model that integrates endogenous seller listing behaviors with buyer-seller decision-making across a large number of U.S. markets.

We begin our study by identifying a life-cycle pattern of migration among mortgage borrowers. We observe that younger borrowers are more likely to relocate and, upon moving, generally transition to higher-quality locations. Specifically, our regression estimates indicate that mortgage borrowers with primary applicants under the age of 35 exhibit a significantly higher propensity to move: 56% above the baseline moving rate compared to the reference group of borrowers aged 55 to 65. In contrast, borrowers aged 35 to 45 show a smaller differential of 39%. Our findings also reveal that primary applicants between 45 to 55 years and those over 65 do not exhibit an increased likelihood to move when compared to the reference group. These results underscore pronounced differences in mobility across the life cycle, with notably higher mobility observed earlier in life.

We further document that, upon relocating, younger borrowers tend to choose areas with higher location quality. Utilizing a measure from the Opportunity Atlas—which estimates the likelihood of a child reaching the 20th income percentile when the parent's income is in the 75th percentile (Chetty, Friedman, Hendren, Jones, and Porter, 2018)–our analysis shows that borrowers under the age of 35, as well as those aged 35 to 45, move to locations that improve their intergenerational mobility measure by 0.17 standard deviations. This pattern holds consistently across various indicators of location quality, including additional metrics from the Opportunity Atlas, log-adjusted IRS income, and the Health Resources and Services Administration (HRSA)'s area deprivation index. These findings highlight the substantial value of such migrations: back-of-the-envelope calculations, based on estimates from Chetty et al. (2018), suggest that each relocation could potentially enhance the lifetime earnings of each child growing up in a lower-income family by approximately $45,000.

In environments of rising interest rates, the mobility of households with existing mortgages secured at lower rates diminishes, a phenomenon known as the "lock-in effect" of higher interest rates. We document how this effect disrupts the typical life-cycle migration patterns in terms of location quality. Specifically, using microdata on household mortgage debt throughout the life cycle of borrowers, we find that, all else being equal, younger households are disproportionately affected by this mortgage lock-in. Their likelihood of selling their home is approximately twice as sensitive to interest rate incentives compared to older households. This heightened sensitivity can be attributed to their greater financial illiquidity and the fact that the net present value (NPV) of their favorable mortgage rates may constitute a larger proportion of their overall household wealth. Thus, we expect the liquidity effects of mortgage lock-in to have heterogeneous impacts across households and markets.

To quantify the heterogeneous welfare effects of mortgage lock-in across households and markets, we develop a search and matching model that incorporates endogenous listing decisions, duration on the market, price negotiations, and migration of sellers and buyers across multiple markets. Each market within our model is defined by a Core-Based Statistical Area (CBSA) and income terciles within Census tracts. According to our model, mortgage lock-in significantly diminishes the likelihood of homeowners listing their properties for sale, with variations in the extent of this reduction across different age groups of households. This decrease in listings leads to diminished liquidity within the housing market, consequently lowering the expected match surplus for buyers. As some sellers transition to become buyers, the diminished match surplus further discourages selling while also reducing the supply of buyers in each market. Our model, therefore, captures both the direct effects of mortgage lock-in on housing market liquidity and its varied impacts across different markets and household demographics.

Our model builds upon the extensive body of research on search and matching models in housing markets. This literature is quite large, and includes Wheaton (1990), Krainer (2001), Albrecht, Anderson, Smith, and Vroman (2007), Novy-Marx (2009), Piazzesi and Schneider (2009), Genesove and Han (2012b), Ngai and Tenreyro (2014), Head, Lloyd-Ellis, and Sun (2014), Gabrovski and Ortego-Marti (2019), Ouazad and Rancière (2019), Piazzesi, Schneider, and Stroebel (2020), Anenberg and Bayer (2020), Guren and McQuade (2020), Buchak, Matvos, Piskorski, and Seru (2020), Gabrovski and Ortego-Marti (2021), Sagi (2021), and Jiang, Kotova, and Zhang (2024). Our contribution to this literature, beyond the mortgage lock-in setting, is to add rich heterogeneity in seller listing behavior as an endogenous input to the model along with selling-rebuying behavior across many markets, which significantly increases the dimensionality of the model as is used in the literature but is still solvable through the strategic use of matrix inversion techniques. Our model thus captures the notion of "selling pressure" in real estate markets which affects market thickness and match surplus. It may also apply to assessing the general equilibrium market and welfare impact of policies in other search and matching settings.

In our model, homeowners have the option to list their homes for sale. This decision is influenced by several factors including the combined value of the seller's value and expected rebuy value from the search and matching process as well as the household's age groups, the effects of mortgage lock-in, mortgage status, time since initial purchase, its square, and idiosyncratic mobility shocks. This decision is estimated using micro-BLP techniques (Petrin, 2002; Berry, Levinsohn, and Pakes, 2004). Once the decision to list is made, homeowners can choose either to rebuy within the same or a different market or opt not to rebuy at all as the outside option. These choices are modeled as endogenous to various factors such as the distance between origin and destination markets, whether they are the same market and location quality considerations.

The proportion of potential buyers in the market is determined by the mix of those looking to rebuy and first-time home buyers. Although the number of first-time buyers is allowed to be influenced by mortgage interest rates and market-specific time fixed effects, it is assumed to be largely unaffected by lock-in effects. This observation aligns with a substantial body of research indicating that household demand for mortgages is relatively inelastic to monetary incentives, aside from debt-to-income (DTI) constraints (DeFusco and Paciorek, 2017; Bhutta and Ringo, 2021; Bosshardt, Di Maggio, Kakhbod, and Kermani, 2023).

Within our model, homeowners who decide to list their properties enter into a search and matching process with potential buyers. The frequency of these meetings and the resulting expected match surplus are significantly influenced by the degree to which mortgage lock-in curtails sellers' listing activities, as well as the idiosyncratic dispersion in buyers' valuations of the properties. Specifically, the impact of mortgage lock-in on match surplus intensifies when it more substantially reduces the sellers' tendency to list and when there is greater idiosyncratic dispersion in how buyers value the properties. This latter scenario illustrates the benefits that buyers experience in a thicker market, where a broader array of housing options is available.

We calibrate the parameters of our search and matching model based on empirical data moments including listing probabilities, duration on the market, average transaction prices, and the standard deviation of these prices within each market of income tercile by CBSA for each quarter. Using the model, we conduct counterfactual analysis by removing the influence of interest rate incentives on homeowners' decisions to list their properties. This approach specifically isolates the direct impact of mortgage lock-in on listing reductions while maintaining the effect of higher interest rates on both the flow utility of homeownership and market demand.

Quantitatively, our model reveals significant and heterogeneous consequences of mortgage lock-in. Specifically, it indicates that listing probabilities decrease by 21-23%, time on market rises by 52-142%, the match surplus for households declines by 3-29%, and house prices increase by 3-8% The large effect on time on the market combined with the more muted but still significant effect on house prices implies that sellers are willing to wait longer rather than accepting a lower price. Interestingly, the effect of mortgage lock-in on match surplus rises significantly over time, from 3-7% in 2022 Q2 to 17.7-28.8% in 2022 Q4, despite the listing, time on the market, and pricing effects being stable between 2022 Q2 to 2022 Q4. This suggests that the additional value of listings to match surplus is non-linear, and is higher when the number of listings is reduced due to higher interest rates.

Our model results exhibit considerable heterogeneity along several dimensions. First, they vary across geographic regions and income terciles. In general, we find larger price and time-on-market effects in lower-income areas despite similar effects on listings, which aligns with their higher idiosyncratic value dispersion. The higher idiosyncratic value dispersion in lower income terciles is also documented in Jiang and Zhang (2022), but we find a greater price and time-on-market effect of mortgage lock-in as a novel consequence. Second, our findings indicate that younger households are disproportionately affected by mortgage lock-in compared to other demographic groups. This outcome aligns with our reduced-form results, which demonstrate that mortgage lock-in interrupts the typical life-cycle migration pattern based on location quality, thereby bearing potential intergenerational implications.

In addition to the literature on search and matching, our paper is related to a growing literature on mortgage lock-in. Fonseca and Liu (2023) estimates the mortgage lock-in effect as a function of the interest rate differential, Δr, and emphasizes the resulting labor market frictions. Liebersohn and Rothstein (2024) uses a control hazard function approach to study the mobility effects of mortgage lock-in, using cash buyers as a control group. Batzer, Coste, Doerner, and Seiler (2024) studies the effect of mortgage lock-in on transactions and house prices, Amromin and Eberly (2024) examines the house price effect of mortgage lock-in relative to a macroeconomic model where house prices are affected by interest rates, and Fonseca, Liu, and Mabille (2024) uses a macroeconomic life-cycle model with a housing ladder to study the effect of mortgage lock-in as well as policies targeting first-time home-buyers. None of these models examines the effect of mortgage lock-in on liquidity and market thickness. A literature studying the lock-in effects of earlier episodes of interest rate increases include Quigley (1987), Quigley (2002), Ferreira (2010), and Ferreira, Gyourko, and Tracy (2010). Relative to this literature, we use a search and matching model to isolate the liquidity effect of mortgage lock-in.

More broadly, our paper develops a general framework for endogenizing rich seller heterogeneity and seller-buyer flows across many real estate markets in a search-and-matching setting to study the welfare effects of market thickness. We find that market thickness has sizeable welfare implications, which is discussed in the earlier literature but not quantified with respect to a given market phenomenon (Genesove and Han, 2012a; Jiang et al., 2024). Our framework may then be useful for studying the liquidity and market thickness implications of other market phenomena, such as the effect of rising interest rates more generally (Bosshardt et al., 2023).

2 Data

We combine several large datasets in our analysis. First, we use CRISM (Equifax Credit Risk Insight Servicing McDash Database) data from June 2005–December 2023. CRISM is an anonymous credit file match from the Equifax consumer credit database to Intercontinental Exchange's McDash loan-level mortgage dataset. The CRISM data contains detailed information on borrower and loan characteristics including the interest rate on the loan, the time of origination, the FICO score, the loan-to-value ratio, and the ZIP code of the property. It is a monthly dataset that also tracks mortgage performance over time from the month of origination to the month of termination. In addition, it includes information on each borrower's credit profile six months before the mortgage is originated and six months after the mortgage is terminated. Finally, the CRISM database also contains information about changes in the borrower's address over time and the borrower's other mortgage debt.

To measure longer-run location outcomes, we also link the CRISM data with the NY Fed Consumer Credit Panel (CCP) database, which is a 5% random sample of all individuals with a credit file in the US. We create this CRISM-CCP sample by taking the full universe of the intersection between the two datasets. Furthermore, we supplement the CRISM-CCP data with various measures of location quality. This includes Opportunity Atlas data from Chetty et al. (2018), ZIP code adjusted gross income data from the IRS Statistics of Income (SOI), and area deprivation index (ADI) data from Kind and Buckingham (2018).

To measure housing market outcomes, we combine data from CoreLogic MLS from 2021 and onwards, and CoreLogic mortgage and deeds data from 2000 and onwards. We merge the CoreLogic MLS data with CoreLogic deeds data using the vendor-provided property ID ("CLIP"). The MLS data includes information on listings and time on the market for properties with and without existing mortgage debt. We further limit our sample to single-family properties. The fraction of homeowners without an existing fixed rate mortgage is sizable, at 42% of homeowners based on Table 2. In our model, we leverage the CoreLogic deeds-CoreLogic MLS and CRISM-CCP samples to create moments for estimation.

Summary statistics of our CRISM-CCP data, supplemented with our measures of location quality change conditional on moving, are presented in Table 1. Summary statistics from our CoreLogic deeds-CoreLogic MLS data, as well as the statistics we computed for calibration, are presented in Table 2.

3 Reduced-form Analysis

In this section, we conduct two sets of reduced-form, econometric analysis and document several stylized empirical facts. First, we test whether homeowner mobility rates differ by age and whether conditional on moving, there is heterogeneity by age in terms of the location quality of homeowners' destinations. In a second set of empirical exercises, we test to see if mortgage lock-in reduces household mobility rates and whether there are heterogeneous effects across the age distribution.

3.1 Life-cycle Migration Decisions of Households

We begin by estimating linear probability models of homeowner mobility on our CRISM-CCP sample to determine how mobility differs over the life cycle. The linear probability specifications take the following form:

$$\mathbb{1}(\text{Move}) = \sum_a \beta_a \mathbb{1}(\text{age group} = a) + \xi_t^o + \xi_t + \text{additional controls} + \epsilon, \tag{1}$$

where the dependent variable, $\mathbb{1}(\text{Move})$, is equal to 100 if the borrower paid off their mortgage and sold their home in a given month, and 0 otherwise. β_a are age group specific coefficients, ξ_t^o are origination month fixed effects, and ξ_t are current month fixed effects. In some specifications, we include county or county-by-month fixed effects as well as additional controls for borrower and loan characteristics.[1]

Table 3 displays the estimation results. In column (1), we limit our fixed effects to the origination year-month ξ_t^o and current year-month ξ_t, so that we are comparing borrowers who obtained their mortgage in the same month and are making their mobility decisions in the same month. The coefficient associated with the under-35 age group suggests that the youngest borrowers in our sample are 0.187 percentage points more likely to move in a given month compared to the omitted group, which is borrowers between the ages of 55 and 65. This is an economically large difference, as it is approximately 56% of the average monthly mobility rate across all borrowers in our sample, which is 0.336 percentage points (panel B of Table 1). Similarly, borrowers in the 35-45 age group are 0.114 percentage points more likely to move in a given month than the omitted group, a difference that is equal to 34% of the average monthly mobility rate in our sample. On the other hand, older borrowers between the ages of 45-55 and borrowers over the age of 65 do not have significantly different mobility rates compared to the omitted group. In columns (2)-(3) of Table 3 we include additional fixed effects including county and county-by-month, and in column (4) we add numerous controls for borrower and loan characteristics. The estimation results do not materially change with these additions. In Figure 1 we plot the borrower mobility coefficient estimates from the specification in column (4) of Table 3. It is clear from the figure that household mobility declines significantly with age until borrowers reach 45-55 years old at which point mobility does not decline further with age.

Now that we have established younger homeowners are significantly more likely to move, we

[1]The controls include the term of the loan, indicator variables for whether the mortgage is a jumbo loan or insured by the Federal Housing Administration (FHA), an indicator for whether the mortgage is a refinance, indicators for the type of property (single-family detached, townhouse, or condominium), indicators for the level of documentation in the underwriting process (full, low, or none), indicators for the LTV ratio rounded to the nearest 5%, indicators for 20-point credit score bins, and the amount of the loan rounded to the nearest $50,000.

investigate whether conditional on moving, they are more likely to move to better locations. To do so, we estimate OLS regressions of the following form:

$$\Delta\text{Location Quality} = \sum_a \beta_a \mathbb{1}(\text{age group} = a) + \xi_t^o + \xi_t^c + \text{additional controls} + \epsilon, \tag{2}$$

where we use several different measures of location quality from the literature. Table 4 displays the estimation results. In columns (1)-(3) we use measures of location quality from Chetty and Hendren (2018). In column (1) the dependent variable is the probability of a child born in the 75th percentile of the household income distribution in a given county moving up to the 20th percentile. In column (2), it is the county-level probability of a household born in the 25th percentile of the household income distribution moving into the 1st percentile. In column (3), it is the county-level probability of a household born in the 25th percentile of the household income distribution. The dependent variable in column (4) of Table 4 is the logarithm of average gross county income, which comes from the IRS SOI. Finally, in column (5) the dependent variable is the area deprivation index (ADI) from Kind and Buckingham (2018). All specifications include origination year-month fixed effects ξ_t^o and county-by-year-month fixed effects ξ_t^c, as well as additional control variables corresponding to the most saturated specification of Table 3. Furthermore, we standardize all dependent variables to have a mean of 0 and a standard deviation of 1 before taking the difference (Δ) pre-and-post move, so that the regression coefficients are interpretable in terms of standard deviation.

In all of the specifications in Table 4, we find that conditional on moving younger households are more likely to move to higher-quality areas and the differences are economically large in magnitude. Taking column (1) as an example, the coefficient estimate associated with households in the youngest age group (≤35) is 0.165, which implies that, on average, they move to counties where the probability of a child moving from the 75th to the 20th percentile of the income distribution is 0.165 standard deviations higher than counties where homeowners aged 55-65 move (the omitted group). The coefficient estimate is similar for the group of homeowners aged 35-45 years (0.168), but is significantly lower for the group aged 45-55 years (0.0835), and close to zero for the oldest group (>65). A similar pattern holds across different location quality measures in columns (2)-(5).[2] Figure 2 plots these coefficient estimates and shows that homeowners in the three youngest age groups are significantly more likely to move to higher-quality locations.

3.2 The Heterogeneous Effect of Mortgage Lock-In

In the previous section, we documented that younger households are both more mobile and more likely to move to higher-quality locations when they do move. In this section, we estimate the effect of mortgage lock-in on homeowner mobility and test whether the magnitude of the effect is stronger for younger households.

We estimate linear probability models similar to equation (1) above that take the following form:

$$\mathbb{1}(\text{Move}) = \sum_a \beta_a \mathbb{1}(\text{age group} = a) \times \Delta r + \xi_a + \xi_t^c + \text{additional controls} + \epsilon \qquad (3)$$

where the dependent variable, $\mathbb{1}(\text{Move})$, is equal to 100 if the borrower paid off their mortgage and sold their home in a given month, and 0 otherwise, and Δr is a variable that measures mortgage lock-in and is defined as the difference between the household's mortgage interest rate at origination and the interest rate that is currently available to them if they were to originate a new mortgage, which we impute based on their mark-to-market LTV and current credit score as in Gerardi, Willen, and Zhang (2023). As Δr becomes negative, the interest rate that the borrower is paying on their current loan is lower than the rate that they would pay on a new loan, and lock-in effects should increase. We infer moving from the CRISM data based on the methodology of Lambie-Hanson and Reid (2018), which was also used in Gerardi et al. (2023). Specifically, we assume that a household moves when they terminate their mortgage and then subsequently change their address.[3] Finally,

[2]Note that the negative coefficient estimates in column (5) are consistent with the estimates in columns (1)-(4) because an increase in the area deprivation index corresponds to a lower quality location.

[3]Note, we exclude from our sample mortgages associated with non-owner-occupied properties.

ξ_a corresponds to age group fixed effects, ξ_t^c corresponds to county-by-month fixed effects, and the additional controls are identical to those in Tables 3 and 4.

Table 5 displays the estimation results. In columns (1) and (3), we present results from specifications that do not include interactions between Δr and the age group dummies and thus capture the effect of mortgage lock-in on average mobility. Column (1) shows results from an OLS specification, while column (3) shows results from an instrumental variables (IV) specification where we instrument for Δr using the change in the Freddie Mac Primary Mortgage Market Survey (PPMS) 30-year FRM interest rate since origination. Our construction of the Δr measure and the instrumental variables approach is similar to that of Fonseca and Liu (2023), except our Δr further captures the time-varying market-to-market LTV and credit scores of borrowers. The instrumental variables strategy, using changes in the PPMS 30-year FRM rate, is designed to address potential endogeneity bias. Such bias may arise if borrowers, who secure lower mortgage rates due to greater financial sophistication or superior credit quality, are also inherently more likely to relocate. This IV approach isolates variation in Δr that comes from differences in the timing of mortgage origination.

The OLS estimate of 0.075 in column (1) implies that a 1 percentage point increase in Δr increases the likelihood that a borrower moves in a given month by 0.075 percentage points, which is approximately 22% of the average monthly mobility rate in our sample. The IV estimate in column (3) is almost twice as large and suggests that a 1 percentage point increase in Δr increases mobility by 0.14 percentage points. These effects are large, compared to an average mobility rate of 0.336 percentage points overall in Table 1. These results suggest that prepay sales, which is a stringent measure of mobility, are more affected by mortgage lock-in than the zip code measure we examine subsequently.

Columns (2) and (4) of Table 5 display results from specifications that interact Δr with the age groups. We see significant heterogeneity in the effect of mortgage lock-in across the age groups. In column (2) the OLS estimates fall from 0.131 for the $\leq$35 age group to 0.045 in the 55-65 age group, while the IV estimates in column (4) fall from 0.233 for the $\leq$35 age group to 0.101 in the

55-65 age group. Thus, the negative effect of mortgage lock-in on mobility is much stronger for younger homeowners compared to older homeowners.

Some households who experience lock-in may decide to move without selling their current home. This could confound our CRISM-based heterogeneity results by age if younger borrowers are disproportionately likely to do so. To investigate this possibility in the data, we construct an alternative mobility indicator that uses ZIP code changes in the consumer credit panel (CCP) to identify neighborhood changes that do not require borrowers to pay off their mortgages. We then re-estimate the specifications in Table 5 with this alternative measure. The results, which are displayed in Table A.2 in the online appendix, show largely similar patterns by age though with smaller average magnitudes. In particular, the OLS estimate of 0.0407 in column (1) of Table A.2 implies that a 1 percentage point increase in Δr increases the likelihood that a borrower moves in a given month by 0.0407 percentage points, which is approximately 6.7% of the average monthly zip code mobility rate in our sample. The IV estimate in column (3) is almost twice as large and suggests that a 1 percentage point increase in Δr increases zip code mobility by 0.0781 percentage points or 12.8% of the average monthly mobility rate in our sample, a similar finding compared to the 9–14% estimate in Fonseca and Liu (2023). Taken together, these results suggest that zip code mobility is less sensitive to mortgage lock-in compared to prepay sales, though the heterogenous effects by age remain robust to both measures.

Our measure of lock-in thus far, Δr, is the simple difference between a borrower's current mortgage interest rate and the market rate that the borrower could obtain on a new loan. We investigate whether the results in Table 5 are similar to alternative lock-in measures.

We investigate a net present value (NPV) based measure that better captures the percentage gain in NPV due to the borrower having a different interest rate at origination than their current market rate. We compute the NPV change in percentage terms to account for the fact that borrowers with larger mortgage balances may also have higher moving costs and larger potential benefits from moving. We compute this % NPV measure as:

$$\% \text{ NPV} = 100 * \frac{V_{i,m} - V_{i,r}}{V_{i,m}}$$

where

$$V_{i,m} = \sum_{s=1}^{TM_i - k_i} \frac{P_i}{(1+m_{it})^s}$$
$$V_{i,r} = \sum_{s=1}^{TM_i - k_i} \frac{P_i}{(1+r_i)^s}$$

and r_i is borrower i's mortgage rate, TM_i is the mortgage term, k_i is the age/seasoning of the mortgage, m_{it} is the prevailing market rate which depends on the borrower-specific mark-to-market LTV and credit score, and P_i is the mortgage payment. This measures the % change in NPV of the remaining mortgage payment between when it is discounted at the origination interest rate, r_i, and the market interest rate, m_{it}. This measure is also used to measure the call option value on the mortgage in Deng, Quigley, and Van Order (2000).

Table 6 presents estimation results that use the % NPV measure to measure mortgage lock-in. In columns (1) and (2) of the table, we present linear probability results from regressions of the mobility indicator on % NPV with various control variables. In columns (3) and (4), we instrument for % NPV with an alternative version of % NPV that uses the Freddie Mac Primary Mortgage Market Survey rate at origination and over time to construct r_i and m_{it}, respectively. The IV approach helps to address potential endogeneity bias due to borrower selection of upfront interest rates and their time-varying LTV and credit scores. The results in Table 6 are qualitatively similar to our results with the Δr measure in Table 5, which suggests that using the % NPV measure, which better captures the remaining term of the mortgage as well as the non-linear impact of interest rate on mortgage payments, has little impact on our heterogeneity effect of mortgage lock-in by age. In Appendix Table A.3 we re-do the analysis in Table 6 using CCP ZIP code changes to construct the mobility indicator and find similar results.

Thus far, our empirical analyses have shown that homeowners have life-cycle migration patterns whereby younger households are both more mobile and, conditional on moving, on average relocate to higher-quality areas. In addition, younger households' mobility decisions are also more affected by mortgage lock-in. This suggests that mortgage lock-in may also adversely impact the life-cycle migration patterns of homeowners.

One potential mitigating factor is if households that are more affected by mortgage lock-in move to better locations conditional on moving so that the households who derive the most benefits from moving are also less affected by mortgage lock-in. If true, this would imply that the life-cycle migration pattern is less disrupted than implied by our reduced-form analyses. Table 7 shows that this is not the case. The table displays the results of regressions of the change in location quality conditional on moving on homeowner age groups interacted with our % NPV measure of mortgage lock-in. We see little evidence that younger households in the ≤35, 35-45, and 45-55 age groups, who experience greater mortgage lock-in, are moving to more favorable locations, as the coefficients are economically small and generally not statistically significant.

4 Model

In this section, we develop a structural model of mortgage lock-in and its heterogeneous impact across markets and households through the liquidity channel. Our model has two stages. First, potential sellers decide to enter the market by listing their homes for sale. They do this with rational expectations of the gains from trade, including the utility of buying another house either in the same or in a different market. The sellers meet with a pool of potential buyers, where the size of the pool varies depending on the fraction of sellers that choose to re-buy. Second, sellers and buyers negotiate prices based on a search and matching model with Nash bargaining. The resulting price effects of mortgage lock-in and the implied value of market liquidity come from the idiosyncratic variation in buyers' valuations. In the model, a thicker market generates a higher match surplus. Furthermore, buyer-seller imbalances affect prices and time-on-market.

4.1 Potential seller and buyers

Potential Sellers. Housing markets are indexed by h. We define each housing market h as comprising of a CBSA and a Census tract-level income tercile. Homeowners make a market entry decision at the Poisson rate of 1 per quarter[4] and gain expected utility u^s_{it} from selling in market h and then potentially rebuying in market h':

$$u^s_{iath} = \underbrace{-\beta_a \kappa(M, r_i, \tau, \bar{r}_{it})}_{\text{lock-in effect}} + \underbrace{\gamma^s \left(V^s_{th} + E\tilde{V}^b_{iath} \right)}_{\text{seller value and rebuy option value}} + \delta^s X^s_{it} + \alpha^s_h + \eta_t + \underbrace{\xi^s_{ht} + \epsilon^s_{it}}_{\text{error terms}}, \tag{4}$$

where β_a is a coefficient that varies by borrower age a, $\kappa(M, r_i, \tau, \bar{r}_{it})$ is the percent NPV loss from taking out a new mortgage at rate $\bar{r}_{it}$ instead of r_i at origination for a mortgage balance M and remaining term τ, V^s_{th} is the expected gains from selling in time t and market h, $E\tilde{V}^b_{iath}$ is the expected utility from the option to re-buy after selling taking into account all other markets h' as described in the following two paragraphs, X^s_{it} are a set of seller control variables including an indicator for whether the seller has a mortgage, time since home purchase, its square, and seller age, α^s_h is a market fixed effect, η_t is a time fixed effect, ξ^s_{ht} is a market-by-time residual, and ϵ^s_{it} is an i.i.d. shock to the value of selling which we assume to be Type I extreme value.

The utility of re-buying in market h' after selling in market h is:

$$u^b_{iathh'} = \underbrace{\gamma^b V^b_{th'}}_{\text{buyer value in market } h'} + \delta^b X^b_{iahh'} + \underbrace{\xi^b_{athh'} + \epsilon^b_{iathh'}}_{\text{error terms}} \tag{5}$$

where $V^b_{th'}$ is the expected buyer value at time t in market h', $X^b_{iahh'}$ is a set of control variables including an indicator variable for the selling market h and the buying market h' being in the same CBSA, an indicator variable for whether h and h' are the same CBSA by income tercile market, the log of the distance between market h and market h', age fixed effects, income tercile fixed effects, an indicator for h' being a higher income tercile than h, and the indicator interacted with the income

[4] As explained in Section 5.1 of Arciadiacono, Bayer, Blevens, and Ellickson (2016), the Poisson rate parameter is the continuous-time analog of the choice of time interval (e.g. monthly, quarterly) in discrete time choice modeling.

terciles. $\xi^b_{athh'}$ is a set of age groups by time by selling market h by buying market h' residuals, and $\epsilon^b_{iathh'}$ is a logit error term. The utility from the outside choice of the seller not re-buying in any market, or $h' = 0$, is normalized to zero, and $E\tilde{V}^b_{iath} = E \max_{h'} u_{iathh'}$ is the expected value of the re-buy option.

The probability of observing a listing by a homeowner with age a in a given quarter t in market h, and with control variables X^s_{it} is then:

$$p^s_{ath}(X^s_{it}) = \Pr\left(u^s_{iath} \geq 0 | X^s_{it}\right) \tag{6}$$

Potential Buyers. In each market, there is a potential pool of buyers of size M^b_{th}. This pool is comprised of both the pool of seller-buyers and other buyers. Specifically,

$$M^b_{th} = \sum_{h'} \sum_{ia} \underbrace{K_{iath'} \Pr\left(u^s_{iah'} \geq 0 | X^s_{it}\right) \Pr(u^b_{iath'h} = \max_{\tilde{h}} u^b_{iath'\tilde{h}})}_{\text{number of seller-rebuyers from } h' \text{ to } h} + \underbrace{M^{bo}_{th}}_{\text{other potential buyers}} \tag{7}$$

where $K_{iath'}$ is the number of potential sellers of type i, age a, at time t and in market h', $\Pr\left(u^s_{iah'} \geq 0 | X^s_{it}\right)$ is the probability of the agent selling in market h', and $\Pr(u^b_{iath'h} = \max_{\tilde{h}} u^b_{iath'\tilde{h}})$ is the probability that the seller re-buys in market h. M^{bo}_{th} denotes the mass of other potential buyers in the market. Our model assumes that mortgage lock-in, as a distinct effect from higher interest rates, affects the number of seller-rebuyers from h' to h, but not the mass of other potential buyers M^{bo}_{th}.

4.2 Matching

Potential buyers and sellers play a continuous time-matching game. The game largely follows the setup in Jiang et al. (2024) and those of the earlier search and matching literature. Specifically, in each market h, a mass of sellers M^s_{th} and buyers M^b_{th} match with flow probability:

$$m\left(M^b_{th}, M^s_{th}\right) = \alpha \left(M^b_{th}\right)^{\phi} \left(M^s_{th}\right)^{1-\phi}, \tag{8}$$

Once matched, buyers draw an idiosyncratic match surplus $\epsilon \sim G(\epsilon)$ if they were to purchase the house. This captures the idiosyncratic value that buyers may place on a specific house. To the extent that mortgage lock-in reduces market thickness by reducing both the number of sellers and buyers in the market, it consequently lowers the expected value of the idiosyncratic match surplus. This reduction leads to a discrepancy between buyers' ideal housing preferences and the properties they acquire.

The bilateral surplus from trade is then:

$$V_{th}^{m}(\epsilon) - V_{th}^{b} - V_{th}^{s} \tag{9}$$

where $V_{th}^{m}(\epsilon)$ is the value of buyers who got matched. The option values of staying in the market and remaining active buyers and sellers are V_{th}^{b} and V_{th}^{s}, respectively. Sellers get surplus $P - V_{th}^{s}$, buyers get surplus $V_{th}^{m}(\epsilon) - V_{th}^{b} - P$. The sum of the two is Equation (9).

Given Nash bargaining, the equilibrium price is:

$$P_{th}(\epsilon) = V_{th}^{s} + \theta\left(V_{th}^{m}(\epsilon) - V_{th}^{b} - V_{th}^{s}\right) \tag{10}$$

where $\theta \in [0, 1]$ governs the split of the surplus between sellers and buyers. Note that in equilibrium trade will only occur if there is positive surplus, such that $V_{th}^{m}(\epsilon) - V_{th}^{b} - V_{th}^{s} \geq 0$, or $\epsilon \geq \epsilon^*$ for some ϵ^* such that $V_{th}^{m}(\epsilon^*)$ exceeds $V_{th}^{b} + V_{th}^{s}$.

We assume a stationary equilibrium in each period, such that buyers and sellers expect the current flow rates and volumes in market h to persist across time t. This assumption is justified to the extent that the higher interest rate environment is expected to continue and that the existing listing probabilities and volume of buyers will persist for the relevant horizon of the sellers and potential buyers playing this matching game.

The equilibrium of the matching game is similar to that of Jiang et al. (2024), with a few modifications. First, we abstract from seller heterogeneity in impatience. Second, the mass of potential buyers M_{th}^{b} is no longer a primitive of the model but rather determined by Equation (7).

Third, we add a flow equality constraint so that the seller flow in each market is pinned down by Equation (6). These allow us to extend the search and matching model to a multi-market setting with rich heterogeneity in terms of the seller listing decision. Specifically, the equilibrium of the matching game can be written as:

Proposition 1. *Given primitives:*

$$r, \alpha, \phi, \theta, \lambda^m_{th}, p^0_{th}, \sigma^\epsilon_{th}$$

A stationary equilibrium of the model is described by buyer and seller masses M^b_{th}, M^s_{th}*, matching rates* $\lambda^s_{th}, \lambda^b_{th}$*, listing rates* λ^m_{th}*, value functions* $V^s_{th}, V^m_{th}(\epsilon), V^b_{th}$*, and a trade cutoff function* ϵ^*_{th}*, which satisfy the following conditions:*

Buyer, seller, and matched owner Bellman equations:

$$rV^b_{th} = \lambda^b_{th} \int_{\epsilon>\epsilon^*} (1-\theta) \left(V^m_{th}(\epsilon) - V^b_{th} - V^s_{th} \right) dG(\epsilon) \tag{11}$$

$$rV^s_{th} = \lambda^s_{th} \int_{\epsilon>\epsilon^*} \theta \left(V^m_{th}(\epsilon) - V^b_{th} - V^s_{th} \right) dG(\epsilon) \tag{12}$$

$$rV^m_{th}(\epsilon) = p^0_{th} + \epsilon + \lambda^m_{th} \left(V^s_{th} - V^m_{th}(\epsilon) \right) \tag{13}$$

Trade cutoffs:

$$V^m_{th}(\epsilon^*) = V^s_{th} + V^b_{th} \tag{14}$$

Matching rates:

$$M^s_{th}\lambda^s_{th} = M^b_{th}\lambda^b_{th} = \alpha \left(M^b_{th} \right)^\phi \left(M^s_{th} \right)^{1-\phi} \tag{15}$$

Flow equality:

$$(1 - M^s_{th})\lambda^m_{th} = \lambda^s_{th} M^s_{th} \left(1 - G(\epsilon^*)\right) = \int_{a, X^s_{it}} p^s_{ath}(X^s_{it}) \tag{16}$$

Buyer mass:

$$M^b_{th} \text{ satisfies Equation (7)} \tag{17}$$

Proof. Based on Jiang et al. (2024). □

Note that Equation (13) combined with Equation (16) implies that matched buyers expect to become sellers with probability equal to the market average λ^m_{th} implied by flow equality, as opposed to an individual-specific λ^m_{ith} which may vary by household characteristics. This assumption is made for the purposes of computational tractability given that the model has numerous markets and endogenous market-to-market flows.[5] However, the effect of this assumption is likely small since λ^m_{th} is small at quarterly frequencies, and the flow utility of housing for matched homeowners is dominated by $p^0_{th} + \epsilon$ rather than the term involving λ^m_{th}. Effectively, this assumption implies that expected seller values become homogeneous conditional on market and time, after entering the search and matching stage.

4.3 Equilibrium

To compute counterfactuals without mortgage lock-in ($\beta_a = 0$), we find the counterfactual listing propensities and buyer masses $p^s_{ath}(X^s_{it}), M^b_{th}$ such that the following equilibrium conditions are satisfied:

1. Sellers' listing probability follows Equation (6),
2. Buyers' mass follows Equation (7),
3. Matching between sellers and active searchers follows the game in Proposition 1.

This equilibrium can be computed for each market h and time t iteratively. First, we conjecture a set of V^b_{th}, V^s_{th} and calculate the flows p^s_{ath} for all markets h and time t based on Equation (6). Second, we calculate the implied M^b_{th} based on these flows and the sellers' rebuy demand in Equations (5) and (7). Third, we compute the implied counterfactual V^b_{th}, V^s_{th} by computing the

[5]Another way we increase the computational tractability of our model after discretizing ϵ as in Jiang et al. (2024) is to solve the system of linear equations defined by Equations (11) to (13) via matrix inversion rather than fixed-point approaches, which leads to a significant speed-up over solving Equations (11) to (13). We discretize ϵ over 1,000 grid points.

matching game equilibrium in Proposition 1 for each market h and time t. Iterating on this procedure allows us to compute a counterfactual equilibrium.

5 Estimation

Estimation proceeds in three steps. First, we calibrate $\{p^0_{th}, \sigma^\epsilon_{th}, M^b_{th}, \lambda^m_{th}\}$ for each market by matching on each CBSA-income tercile's average log sale price in each listing quarter, the residual standard deviation of the log sale price after controlling for extensive observables, the average time on the market, and listing probability. This allows us to extract the implied seller and buyer values V^s_{th}, V^b_{th}. Second, using the extracted V^s_{th}, V^b_{th} we estimate the seller-rebuy utility in Equation (5) using Berry (1994). Third, we estimate the seller listing decision in Equation (4) using the micro-BLP approach of Petrin (2002) and Berry et al. (2004). We now describe each of these steps in detail.

First, our calibration of the search and matching model and the resulting seller and buyer values V^s_{th}, V^b_{th} proceeds as follows. For each market h consisting of a CBSA-income tercile pair and each quarter t, we estimate the average log sale price, the residual standard deviation of the log sale price after controlling for an extensive set of observable property and market characteristics, the time on the market, and the listing probability. Summary statistics of these estimates are presented in Panel B of Table 2 and form the basis of our calibration.

In terms of model parameters, we allow $\{p^0_{th}, \sigma^\epsilon_{th}, M^b_{th}, \lambda^m_{th}\}$ to vary across markets and time. We further set $r = 0.05, \phi = 0.84, \theta = 0.5$, based on Genesove and Han (2012b) and Jiang et al. (2024). In Jiang et al. (2024), the search efficiency α is set to 1 due to the lack of an empirical target to calibrate to. We present results with both $\alpha = 1$ and $\alpha = 10$, with the former implying that in balanced markets with $M^b_{th} = M^s_{th}$, a potential buyer matches with a potential seller at a rate of once per year, and the latter implying a potential buyer matches with a potential seller at a rate of 10 times per year. We view these as plausible bounds on this parameter. Our calibration of the search and matching model matches the target moments quite well. We present summary statistics on our calibrated parameters in Appendix Table A.1 and our model fit in Appendix Figure A.1, which shows that the model largely replicates the targeted moments. In turn, the model produces the implied buyer and seller values V^s_{th}, V^b_{th} for each market h and quarter t.

Second, using the model implied buyer and seller values, we estimate the buyer-seller utility in Equation (5) à la Berry (1994), which established that this class of models may be estimated using the OLS relationship:

$$\log(s^b_{athh'}) - \log(s_0) = \underbrace{\gamma^b V^b_{th'}}_{\text{buyer value in market h'}} + \delta^b X^b_{ahh'} + \xi^b_{athh'} \tag{18}$$

where $s^b_{athh'}$ is the share of sellers in market h re-buying in market h', with h' possibly being equal to h, and s_0 is the outside option of not re-buying at all. In this model, the choice of which market to re-buy and whether to re-buy at all depends on the buyer value in the target market h', $V^b_{th'}$, as well as a set of other variables $X^b_{ahh'}$. We estimate Equation (18) via OLS and present the results for the $\alpha = 1$ case in Table 8.

The first row of Table 8 shows that sellers tend to re-buy in the "same_mkt", an indicator variable for h and h' being the same market. On the other hand, in the second row of Table 8 we see a negative coefficient associated with "same_cbsa", which suggests that although sellers are more likely to re-buy in the same market, they are less likely to re-buy in the other income tercile markets in the same CBSA. This likely reflects strong inertia in terms of the income tercile markets that households select. The third row of Table 8 suggests that sellers prefer to re-buy in closer markets. The fourth row of Table 8 suggests that v_b, which is our variable name for $V^b_{th'}$, is not significantly correlated with re-buying decisions, which suggests that re-buy demand is relatively inelastic to the match value. This is consistent with the literature showing that home purchase decisions are relatively inelastic to interest rate incentives other than through DTI constraints (DeFusco and Paciorek, 2017; Bhutta and Ringo, 2021; Bosshardt et al., 2023). The fourth and fifth rows of Table 8, combined, show that households in the first initial income tercile tend to move to a higher income tercile, although the same is not true for households in the second initial tercile. We include age group and income tercile fixed effects. In Appendix Table A.4 we explore interacting the "higher_tercile" variables by age group, but do not find significant results, which suggests that although younger borrowers are more likely to move up in location quality, the force is not strong enough to push those borrowers to move up in terms of income terciles.

Third, we estimate the seller listing decision in Equation (4) via micro-BLP (Petrin, 2002; Berry et al., 2004). This consists of drawing a set of parameters $\beta_a, \gamma^s, \delta^s$, extracting $\alpha_h^s + \eta_t + \xi_{ht}$ via a nested-fixed point algorithm and then using a linear projection to separately identify $\alpha_h^s, \eta_t, \xi_{ht}$. We match moments via GMM, with weights equal to one over the number of moments in each set such that each set of moments is weighted equally. We use six sets of moments, the first five of which are:

- From CoreLogic-MLS data:
    1. Probability of listing a home for sale, by quarters since purchase,
    2. Relative probability of listing a home for sale, for owners with and without a FRM,
- From CRISM-CCP data:
    3. Relative probability of selling a home by borrower age group,
    4. Relative probability of selling a home by quintile of %NPV lock-in effect,
    5. Relative probability of selling a home, by borrower age group interacted with quintile of %NPV lock-in effect,

Note that new construction is included in the CoreLogic-MLS data. Therefore, new construction observations are included in our seller demand estimation as a sale with a relatively low time since purchase. The CoreLogic-MLS data also includes owners without a fixed-rate mortgage, which we include as an indicator variable in our estimation of Equation (4). For owners with a fixed-rate mortgage, we assume that our CRISM-CCP data is representative of that sub-population. We construct the relative probabilities of listing or selling a home relative to the first category of the set of moments 2–5. Finally, as in Berry et al. (2004) we add a sixth moment condition:

$$E\left(\left(V_{th}^{s} + E\tilde{V}_{iath}^{b}\right)\xi_{ht}\right) = 0 \tag{19}$$

to identify the coefficient γ^s.

Results of this third estimation step, along with GMM standard errors, for the $\alpha = 1$ case are shown in Table 9. In particular, rows 1-2 of Table 9 suggest that listing propensity is negatively related to time since purchase, although the relationship weakens over time. The age group results in rows 3-6 of Table 9, with the omitted group being the ≤35 group, suggest a decreasing propensity to list their home for sale as the borrower increases in age from ≤35 to 35-45 and again from 45-55, although the results flatten thereafter and increases slightly for the >65 group. Rows 7-12 of Table 9 suggest that the lock-in effect is strong, although its magnitude cannot be directly compared with the estimates in Table 6 since it is a logit functional form rather than linear. Further examination of rows 7-12 of Table 9 suggests that the lock-in effect is similar across age groups in logit form, although it is larger for younger borrowers in linear form. This suggests that mortgage lock-in has a similar proportional effect across age groups, although it has a larger magnitude effect on younger borrowers due to their higher propensity to relocate in general. Row 13 of Table 9 suggests that sellers are sensitive to the expected value from selling their homes and re-buying in a different market. Finally, row 14 of Table 9 suggests that households with a fixed-rate mortgage are significantly more likely to list their homes for sale compared to households without fixed-rate mortgages.

We present analogous results of calibration fit and the second and third step estimation results for the $\alpha = 10$ case in Appendix Figure A.2, and Appendix Table A.5, respectively. In general, we do not find sizable differences in either the calibration fit or the second and third-step estimated parameters between the $\alpha = 1$ and $\alpha = 10$ cases.

6 Model results

After estimating the model, we compute counterfactuals without mortgage lock-in by setting the lock-in effect on sellers β_a in Equation (4) to zero. This allows us to isolate the liquidity effect of mortgage lock-in in reducing sellers' tendency to sell and their endogenous re-buy propensities in different markets. This liquidity effect then impacts house prices and match surplus by altering the buyers' equilibrium cut-off in terms of their ϵ draw.

Importantly, our counterfactual holds fixed other market condition changes that may affect the mass of potential sellers and buyers, such as the effect of elevated interest rates. By holding fixed such factors, we isolate the heterogeneous effect of mortgage lock-in through the liquidity channel.

6.1 Housing Market Liquidity, House Prices, and Welfare

A summary of our model counterfactual is presented in Table 10. The first and fifth row of Table 10 suggests a large effect of mortgage lock-in on listing probabilities relative to a counterfactual without mortgage lock-in in the same market and quarter. With no lock-in effects, listing probabilities are approximately 21-23% higher in both the $\alpha = 1$ and $\alpha = 10$ cases after the second quarter of 2022. This is consistent with a large listing effect among mortgage borrowers as well as the substantial presence of homeowners without a fixed-rate mortgage who are not directly impacted by mortgage lock-in. The second and sixth row of Table 10 suggests that mortgage lock-in has had a large effect on time on the market and that removing the effect of mortgage lock-in would increase time on the market by 53-143%. The third and seventh row suggests that house prices without mortgage lock-in would fall by between 3 to 8%.

The large change in time on the market with smaller, though still significant changes, in house prices, is due to the search and matching process, whereby sellers wait longer for a sale rather than accepting a lower price. In Toronto, Canada, a prominent North American market in which most households have shorter-term mortgages and are less affected by mortgage lock-in, time on the market went up 110.4% while prices fell by 9.2% in December 2022 relative to a year prior, which is in the range of our model outcomes.[6]

The fourth and eighth row of Table 10 looks at match surplus, which we define as $E(V^m_{th}(\epsilon) - P_{th}(\epsilon)|\epsilon > \epsilon^*_{ht})$, or the utility received by buyers net of the house price. A positive match surplus exists when buyers have an idiosyncratic valuation over housing and capture part of that value above the selling price of the house. The results suggest that without mortgage lock-in, the match surplus would grow by between 3-7.2% in 2022Q2, 6.5-12.9% in 2022Q3, and 17.7-28.8% in 2022Q4. The increasing effect of removing mortgage lock-in on match surplus is determined by the decreasing transaction volume due to factors such as higher interest rates in Q3 and Q4 of 2022. When transaction volume is already thin, the additional liquidity taken up by mortgage lock-in becomes exceedingly valuable in terms of generating match surplus.

In addition to average effects, our model reveals substantial heterogeneity in the effect of mortgage lock-in across markets. We illustrate this heterogeneity using violin plots. In Figure 3, we present the violin plots of the counterfactual change in listing probabilities by quarter. On average, consistent with Table 10, we find a roughly 21-23% increase starting in Q2 of 2022. However, there is also substantial heterogeneity, with increases of over 30% in some markets and even a decline in listing probabilities in some markets due to cross-market buyer-seller flows. These results highlight how the inter-connectedness of markets affects the equilibrium outcomes.

Figure 4 presents analogous violin plots for time on the market. As in Figure 3, we find substantial heterogeneity in time on market effects across markets in addition to a large average increase starting in 2022Q2. In Figure 5 we display price effects, which are also quite dispersed across markets. Figure 6 presents results on the idiosyncratic variation in price or the price dispersion observed in each market. Interestingly, price dispersion falls on average in our counterfactual without mortgage lock-in, reflecting a thicker market that generates more homogeneous match values. Figure 7 presents results on buyer expected values V^b_{th} and Figure 8 presents results on seller expected values V^s_{th}, with buyers seeing an increase in expected values without mortgage lock-in and sellers seeing a decrease in expected values, consistent with more sellers in the market.

[6]See: https://justo.ca/blog/2022-gta-housing-market-year-in-review.

6.2 Heterogeneous Effects across Households and Markets

We now turn to a closer examination of heterogeneous effects across households and markets. Figure 9 shows the change in the expected utility of potential sellers, Eu^s_{iath}, by age. In general, the removal of mortgage lock-in leads to a negative expected utility effect on the potential sellers due to house price declines. Younger households are especially affected by mortgage lock-in, as they become more mobile and experience a greater fall in seller utility without mortgage lock-in. The reason is that markets with more younger borrowers are experience larger equilibrium effects from mortgage lock-in. Hence, younger homeowners derive more implicit benefit from mortgage lock-in in equilibrium, even though it disrupts their life-cycle migration patterns.

Another dimension of heterogeneity we examine is by terciles of Census tract income. Figure 10 presents the results for listing probabilities and suggests that on average the effect of mortgage lock-in on listing probabilities is similar across Census tract income terciles, though there is greater dispersion in the lowest income tercile. Examining time on the market in Figure 11 and log price differences in Figure 12, however, reveals that the lowest income tercile would experience significantly larger effects on time on the market and greater negative effects on prices without mortgage lock-in. This effect is attributable to the higher relative idiosyncratic value dispersion in the lower income tercile, as measured by $\sigma^{\epsilon}_{th}/p^0_{th}$ and can be computed in Appendix Table A.1. Because housing in the lower income areas experiences more idiosyncratic value dispersion, an effect also documented in Jiang and Zhang (2022), prices and time on the market in those areas are more sensitive to the liquidity effect of mortgage lock-in.

Figure 13 shows the results on price dispersion by Census tract income. We find that price dispersion in lower-income areas is less affected by mortgage lock-in on average, which is also related to the higher idiosyncratic value dispersion that persists in those areas, regardless of mortgage lock-in. Figures 14 and 15 show the heterogeneous effect of mortgage lock-in for buyer and seller expected values and suggest that the effects are generally larger for lower-income terciles.

7 Conclusion

In this paper, we find that mortgage lock-in not only disrupts the life-cycle migration pattern of households but also has substantial equilibrium effects on market liquidity and market thickness. We develop a framework for endogenizing the equilibrium effects of market liquidity by extending a search and matching model to incorporate rich seller heterogeneity and buyer-seller flows across many markets. In doing so, we find large and heterogeneous effects of mortgage lock-in on equilibrium market outcomes, particularly in the time-on-market and match surplus dimensions but also on prices and transaction volumes.

We find that the match surplus effects of mortgage lock-in grew non-linearly over time, as higher interest rates reduced market liquidity in general making the additional reduction in liquidity from mortgage lock-in particularly salient for buyers with idiosyncratic tastes in housing. Thus, our paper highlights the importance of real estate market thickness for matching surplus and welfare, which has been discussed but not highlighted in the earlier literature (Genesove and Han, 2012a; Jiang et al., 2024). The welfare effects of market thickness in other settings then serve as a promising direction for future research.

Table 1: Summary statistics for the CRISM-CCP sample

This table presents summary statistics for our CRISM-CCP sample. The sample is at the borrower-by-month frequency. Panel A presents summary statistics for our mortgage origination variables, Panel B presents summary statistics for time-varying variables, and Panel C presents summary statistics for our location change variables conditional on moving. All location quality measures were standardized and winsorized at the 1st and 99th percentiles. In Panel C, the location quality changes are measured in units of standard deviation of location quality.

	Mean	SD	P25	Median	P75	N
Panel A: Mortgage Origination Variables						
log(loan amount)	12.144	0.628	11.724	12.126	12.555	2655948
term	324.067	71.708	360.000	360.000	360.000	2655948
borrower age	48.935	13.378	38.000	48.000	58.000	2655948
SATO	0.019	0.564	-0.305	0.040	0.330	2655948
credit score	743.034	85.386	702.000	769.000	809.000	2655948
LTV ratio	75.524	18.501	65.000	79.170	90.810	2655948
FHA	0.221	0.415	0.000	0.000	0.000	2655948
jumbo	0.055	0.229	0.000	0.000	0.000	2655948
refi	0.556	0.497	0.000	1.000	1.000	2655948
missing purpose	0.055	0.229	0.000	0.000	0.000	2655948
condo	0.131	0.337	0.000	0.000	0.000	2655948
manufactured home	0.013	0.113	0.000	0.000	0.000	2655948
missing property type	0.063	0.243	0.000	0.000	0.000	2655948
low documentation	0.201	0.400	0.000	0.000	0.000	2655948
missing documentation	0.281	0.449	0.000	0.000	1.000	2655948
Panel B: Time-Varying Variables						
move_crism	0.336	5.784	0.000	0.000	0.000	2655948
move_ccp	0.607	7.766	0.000	0.000	0.000	2381836
Δr	0.420	0.976	-0.246	0.351	1.019	2655948
$\%NPV$	-3.686	9.071	-9.523	-3.089	2.173	2655948
Panel C: Location Quality Change						
p20_p75	0.025	0.729	-0.127	0.000	0.221	94954
p01_p25	0.034	1.052	-0.137	0.000	0.191	94954
p25	0.017	0.794	-0.166	0.000	0.230	94954
log_agi	0.077	1.031	-0.142	0.000	0.364	94954
rating	0.052	0.906	0.000	0.000	0.036	94954
adi	-0.021	0.675	-0.198	0.000	0.115	94954

Table 2: Summary statistics for the CoreLogic deeds and MLS sample

This table presents summary statistics for our CoreLogic deeds and MLS sample. Panel A presents summary statistics for the variables used in our seller regressions, and Panel B presents summary statistics for the market-quarter level statistics we computed for calibration. Markets are defined at the CBSA by income tercile level, where income is measured at the Census tract level using the 2010 Census where available, and imputed using the zip code level income when unavailable. Market-quarters with no prepay sales in either the CoreLogic-MLS data or the CRISM-CCP data are dropped.

Variable Name	Mean	SD	P25	Median	P75	N
Panel A: CoreLogic-MLS Variables						
Listing	0.0119	0.108	0	0	0	5891978
Has FRM	0.581	0.493	0	1	1	5891978
Quarters since purchased	35.540	26.919	12	28	58	5891978
Quarters since purchased sq	1987.702	2389.547	144	784	3364	5891978
Panel B: Market-Level Statistics for Model Calibration						
Log(Price)	12.620	0.170	12.493	12.623	12.749	642
Price dispersion	0.178	0.060	0.139	0.171	0.213	642
TOM	51.801	3.827	48.985	51.860	54.561	642
Listing Probability	0.0131	0.00561	0.00898	0.0124	0.0165	642

Table 3: Mobility estimates over the life-cycle

This table presents linear regression coefficients of moving on age group. Moving is coded as an indicator variable equals to 0 if the borrower did not prepay and move and 100 if the borrower moved in a given month. The age 55–65 group is omitted. The sample is our CRISM-CCP sample. Additional controls include the term of the loan, indicators for whether the mortgage is a FHA or Jumbo loan, an indicator for whether the mortgage is a refinance or is missing purpose, indicators for the type of the property, indicators for borrower documentation, indicators for the LTV of the borrower rounded to the nearest 5, indicators for the credit score of the borrower rounded to the nearest 20, and the amount of the loan rounded to the nearest $50,000. Robust standard errors are clustered at the county and year-month level. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

	(1)	(2)	(3)	(4)
<=35	0.187***	0.181***	0.185***	0.179***
	(0.0120)	(0.0121)	(0.0131)	(0.0136)
35-45	0.114***	0.111***	0.109***	0.0873***
	(0.00969)	(0.00958)	(0.0105)	(0.0109)
45-55	0.00805	0.00752	0.00703	-0.00540
	(0.00793)	(0.00806)	(0.00890)	(0.00914)
>65	0.00228	0.00470	-0.000628	0.00813
	(0.0130)	(0.0131)	(0.0142)	(0.0146)
Orig Month FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	No	No
County FE	No	Yes	No	No
CountyXMonth FE	No	No	Yes	Yes
Add'l Controls	No	No	No	Yes
Observations	2932550	2932537	2825032	2706246

Table 4: Location quality change conditional on moving, CRISM-CCP sample

This table presents linear regression coefficients of changes in location quality conditional on moving. Location quality variables are standardized with a mean of 0 and standard deviation of 1 and are held fixed within a zipcode over time. The age 55–65 group is omitted. The sample is our CRISM-CCP sample. Additional controls include the term of the loan, indicators for whether the mortgage is a FHA or Jumbo loan, an indicator for whether the mortgage is a refinance or is missing purpose, indicators for the type of the property, indicators for borrower documentation, indicators for the LTV of the borrower rounded to the nearest 5, indicators for the credit score of the borrower rounded to the nearest 20, and the amount of the loan rounded to the nearest $50,000. Robust standard errors are clustered at the county and year-month level. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

	(1)	(2)	(3)	(4)	(5)
	p20_p75	p01_p25	p25	log_agi	adi
<=35	0.165***	0.126***	0.185***	0.225***	-0.164***
	(0.0147)	(0.0235)	(0.0161)	(0.0217)	(0.0128)
35-45	0.168***	0.157***	0.180***	0.257***	-0.170***
	(0.0123)	(0.0194)	(0.0133)	(0.0172)	(0.00960)
45-55	0.0835***	0.0827***	0.0939***	0.147***	-0.0936***
	(0.0119)	(0.0193)	(0.0121)	(0.0173)	(0.00971)
>65	0.00718	-0.00481	0.0198	-0.00239	0.0145
	(0.0125)	(0.0241)	(0.0145)	(0.0224)	(0.0143)
Orig Month FE	Yes	Yes	Yes	Yes	Yes
CountyXMonth FE	Yes	Yes	Yes	Yes	Yes
Add'l Controls	Yes	Yes	Yes	Yes	Yes
Observations	54225	54225	54225	55275	55619

Table 5: Mortgage lock-in effects and its heterogeneity

This table presents linear regression and instrumental variables coefficients of the lock-in effect. The sample is our CRISM-CCP sample. The dependent variable is coded as an indicator variable equals to 0 if the borrower did not prepay and move and 100 if the borrower moved in a given month. Additional controls include the term of the loan, indicators for whether the mortgage is a FHA or Jumbo loan, an indicator for whether the mortgage is a refinance or is missing purpose, indicators for the type of the property, indicators for borrower documentation, indicators for the LTV of the borrower rounded to the nearest 5, indicators for the credit score of the borrower rounded to the nearest 20, and the amount of the loan rounded to the nearest $50,000. A robustness check where mobility is defined using zip code changes in the CCP data is in Appendix Table A.2. Robust standard errors are clustered at the county and year-month level. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

	OLS		IV	
	(1)	(2)	(3)	(4)
	move_crism	move_crism	move_crism	move_crism
Δr	0.0751***		0.140***	
	(0.00796)		(0.0116)	
<=35 $\times$ Δr		0.131***		0.233***
		(0.0174)		(0.0250)
35-45 $\times$ Δr		0.0919***		0.167***
		(0.0125)		(0.0175)
45-55 $\times$ Δr		0.0620***		0.119***
		(0.0115)		(0.0155)
55-65 $\times$ Δr		0.0445***		0.101***
		(0.0125)		(0.0156)
>65 $\times$ Δr		0.0626***		0.115***
		(0.0128)		(0.0167)
Age Group FE	Yes	Yes	Yes	Yes
CountyXMonth FE	Yes	Yes	Yes	Yes
Add'l Controls	Yes	Yes	Yes	Yes
Observations	1922088	1922088	1922088	1922088

Table 6: Mortgage lock-in effects and its heterogeneity, % NPV changes

This table presents linear regression and instrumental variables coefficients of the lock-in effect. The measurement of lock-in is the % change in the NPV of the mortgage discounted at the mortgage rate. The sample is our CRISM-CCP sample. Additional controls include the term of the loan, indicators for whether the mortgage is a FHA or Jumbo loan, an indicator for whether the mortgage is a refinance or is missing purpose, indicators for the type of the property, indicators for borrower documentation, indicators for the LTV of the borrower rounded to the nearest 5, indicators for the credit score of the borrower rounded to the nearest 20, and the amount of the loan rounded to the nearest $50,000. A robustness check where mobility is defined using zip code changes in the CCP data is in Appendix Table A.3. Robust standard errors are clustered at the county and year-month level. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

	OLS		IV	
	(1)	(2)	(3)	(4)
	move_crism	move_crism	move_crism	move_crism
$\%NPV$	0.00759***		0.0175***	
	(0.000902)		(0.00138)	
<=35 × $\%NPV$		0.0114***		0.0232***
		(0.00173)		(0.00248)
35-45 × $\%NPV$		0.00889***		0.0191***
		(0.00130)		(0.00189)
45-55 × $\%NPV$		0.00605***		0.0153***
		(0.00135)		(0.00191)
55-65 × $\%NPV$		0.00474***		0.0138***
		(0.00150)		(0.00195)
>65 × $\%NPV$		0.00707***		0.0158***
		(0.00144)		(0.00211)
Age Group FE	Yes	Yes	Yes	Yes
CountyXMonth FE	Yes	Yes	Yes	Yes
Add'l Controls	Yes	Yes	Yes	Yes
Observations	1922057	1922057	1922057	1922057

Table 7: Mortgage lock-in and location quality changes conditional on moving

This table presents linear regression coefficients of changes in location quality conditional on moving. Location quality variables are standardized with a mean of 0 and standard deviation of 1 and are held fixed within a zipcode over time. The variable "p20_p75" refers to the probability of a child born in the 75th percentile of household income in a given county who moves to the 20th percentile from Chetty and Hendren (2018). The variable "p01_p25" refers to the probability of a child born in the 25th percentile of household income in a given county who moves to the 1th percentile from Chetty and Hendren (2018). The variable "p25" refers to the probability of a child born in the 25th percentile of household income from Chetty and Hendren (2018). The variable "log_agi" refers to the log of the annual gross income from IRS SOI. The variable "adi" refers to the area deprivation index from (ADI) from Kind and Buckingham (2018). The sample is our CRISM-CCP sample. Additional controls include the term of the loan, indicators for whether the mortgage is a FHA or Jumbo loan, an indicator for whether the mortgage is a refinance or is missing purpose, indicators for the type of the property, indicators for borrower documentation, indicators for the LTV of the borrower rounded to the nearest 5, indicators for the credit score of the borrower rounded to the nearest 20, and the amount of the loan rounded to the nearest $50,000. Robust standard errors are clustered at the county and year-month level. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

(6)	(1) p20_p75	(2) p01_p25	(3) p25	(4) log_agi	(5) adi
<=35 × $\%NPV$	-0.000820	-0.000589	-0.000985	-0.00102	0.000447
	(0.00105)	(0.00138)	(0.00111)	(0.00135)	(0.000840)
35-45 × $\%NPV$	0.00130	0.00126	0.000617	0.00255**	-0.00184**
	(0.000851)	(0.00103)	(0.000901)	(0.00105)	(0.000749)
45-55 × $\%NPV$	0.00134	0.000811	0.00144	0.00158	-0.00173**
	(0.000815)	(0.00136)	(0.000872)	(0.00142)	(0.000843)
55-65 × $\%NPV$	0.00552***	0.00333**	0.00372***	0.00742***	-0.00328***
	(0.00109)	(0.00167)	(0.00125)	(0.00124)	(0.000776)
>65 × $\%NPV$	0.00541***	0.00606***	0.00328**	0.00906***	-0.00328**
	(0.00147)	(0.00182)	(0.00161)	(0.00228)	(0.00132)
Age Group FE	Yes	Yes	Yes	Yes	Yes
CountyXMonth FE	Yes	Yes	Yes	Yes	Yes
Add'l Controls	Yes	Yes	Yes	Yes	Yes
Observations	43787	43787	43787	44634	44890

Table 8: Berry logit for sell-rebuy values

This table presents results of a Berry (1994) logit on estimating the utility of sellers in market h re-buying in market h', with not re-buying in any market as the outside option. The data used is the 2021 CCP data. "same_mkt" is an indicator variable for h and h' being the same market, "same_cbsa" is an indicator variable for whether h and h' are in the same CBSA, "log_dist" is the log distance in miles between h and h' and is set to 0 when h and h' are equal to the same market, "v_b" is the value of buying in the target market, "higher_tercile" is an indicator variable for whether the h' is in a higher income tercile in h and has variation only for sellers in the bottom two terciles. A version of this table where "higher_tercile" is interacted with age groups is in Appendix Table A.4. A version of this table for the $\alpha = 10$ case is in Appendix Table A.5. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

	(1) berry_logit
same_mkt	0.670*** (0.129)
same_cbsa	-0.388* (0.197)
log_dist	-0.232*** (0.0770)
v_b	-0.00226 (0.0482)
higher_tercile	0.631*** (0.197)
higher_tercile#tercile_init=2	-0.611** (0.260)
Age Group FE	Yes
Income Tercile FE	Yes
N	298

Table 9: BLP results

This table presents results of the seller listing problem model estimation following the methodology of Berry, Levinsohn, and Pakes (1993). Age groups include <= 35, 35 − 45, 45 − 55, 55 − 65, and > 65. A version of this table for the $\alpha = 10$ case is in Appendix Table A.7. GMM standard errors are presented. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Variable	Estimate	Std Dev
Time since purchase	-0.1314***	0.0061
Time since purchase sq	0.0056***	0.0004
35-45	-0.0316***	0.0154
45-55	-0.215***	0.0242
55-65	-0.1896***	0.0245
>65	-0.087***	0.0256
<=35*%NPV	0.0299***	0.0006
35-45*%NPV	0.0367***	0.0008
45-55*%NPV	0.0393***	0.001
55-65*%NPV	0.0349***	0.0011
>65*%NPV	0.0335***	0.0012
Seller + Rebuy Value	0.0092***	0.0043
Has FRM	0.2315***	0.0108

Table 10: Summary of model counterfactuals

This table presents the results of our model counterfactuals for the $\alpha = 1$ and $\alpha = 10$ cases. The counterfactuals are computed by setting the lock-in $\beta_a = 0$ and then re-computing the equilibrium in Section 4.3. Violin plots of counterfactuals across markets are shown in Figures 3 to 8.

Variable (average % change)	2022 Q1	2022 Q2	2022 Q3	2022 Q4
		$\alpha = 1$		
Listing probability	4.9	20.8	21.6	22.5
Time on market	12.9	53.4	52.1	56.4
House prices	-0.8	-3.5	-3.4	-2.8
Match surplus	2.5	7.2	12.9	28.8
		$\alpha = 10$		
Listing probability	4.9	20.8	21.6	22.9
Time on market	20.2	132.8	142.9	142.9
House prices	-1.7	-6.0	-6.6	-8.3
Match surplus	3.3	3.0	6.5	17.7

Figure 1: Mobility estimates over the life-cycle

This figure plots the probability that a household in the CRSIM-CCP data prepays their mortgage and sells their home over the household's life-cycle. The x-axis is the age group of the borrower. The y-axis plots the coefficient estimates and their 95% confidence intervals from column (4) of Table 3.

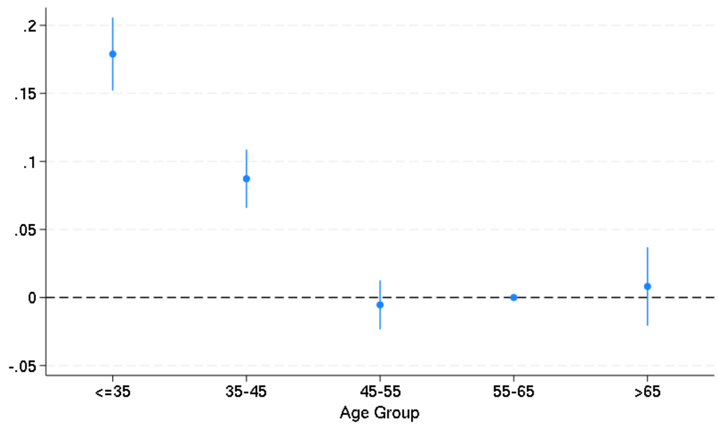
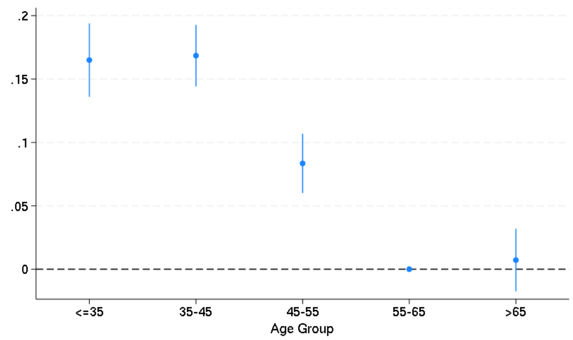
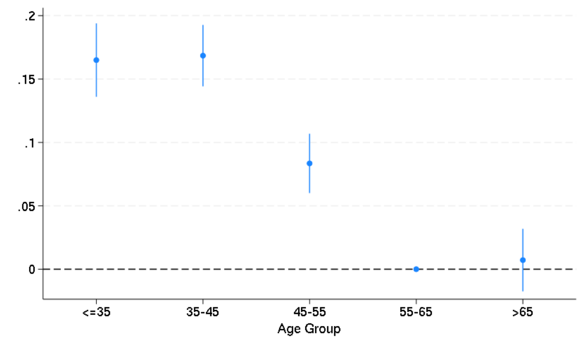


Figure 2: Location quality change conditional on moving, CRISM-CCP sample

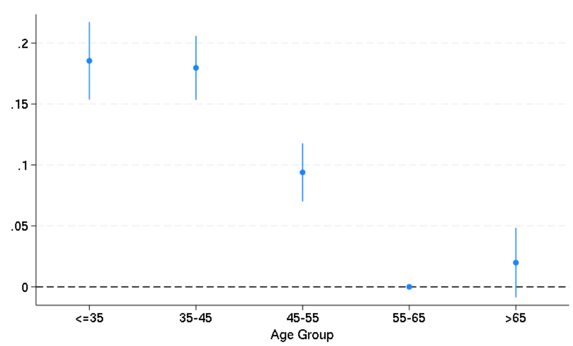
This figure plots the borrowers' location quality change conditional on moving. The x-axis is the age group of the borrower. The y-axis plots the coefficient estimates and their 95% confidence intervals from Table 4. The variable "p20_p75" refers to the probability of a child born in the 75th percentile of household income in a given county who moves to the 20th percentile from Chetty and Hendren (2018). The variable "p01_p25" refers to the probability of a child born in the 25th percentile of household income in a given county who moves to the 1th percentile from Chetty and Hendren (2018). The variable "p25" refers to the probability of a child born in the 25th percentile of household income from Chetty and Hendren (2018). The variable "Log(agi)" refers to the log of the annual gross income from IRS SOI. The variable "ADI" refers to the area deprivation index from (ADI) from Kind and Buckingham (2018).



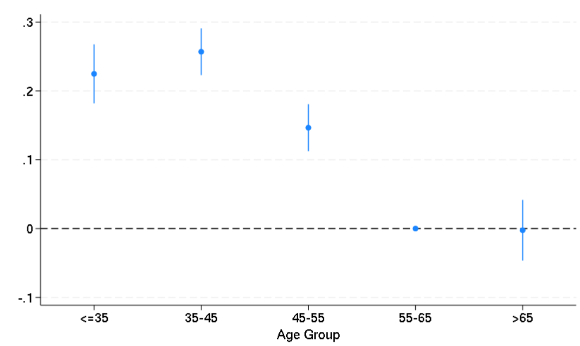
(a) p20_p75 change



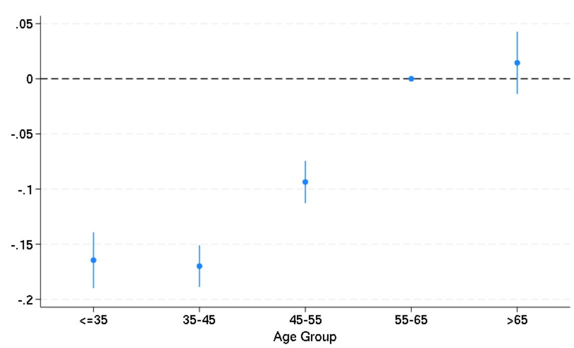
(b) p01_p25 change



(c) pooled_p25 change



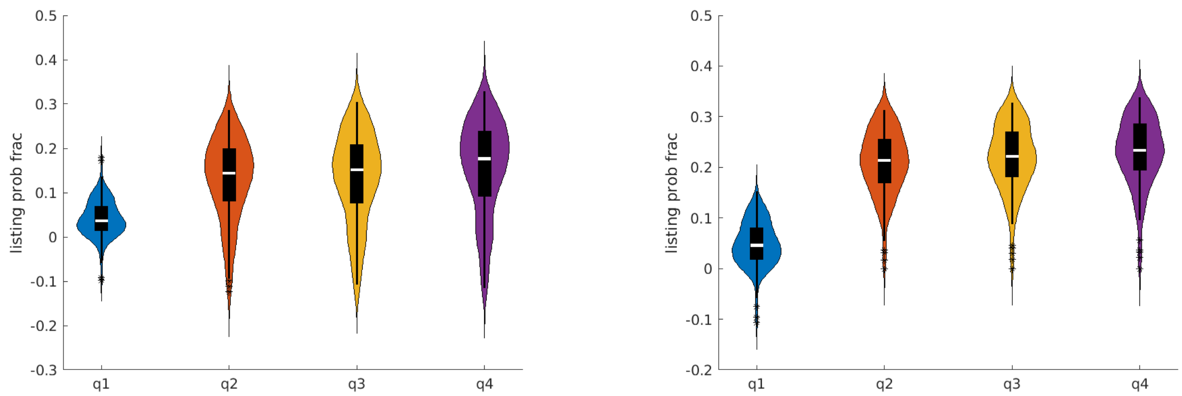
(d) Log(agi) change



(e) ADI change

Figure 3: Listing probability

This figure shows violinplots the counterfactual fraction change in listing probabilities without mortgage lock-in. The x-axis consists of quarters in 2022. The counterfactuals are computed by setting the lock-in $\beta_a = 0$ and then re-computing the equilibrium in Section 4.3. Panel (a) shows the results for the $\alpha = 1$ case, and panel (b) shows the results for the $\alpha = 10$ case.

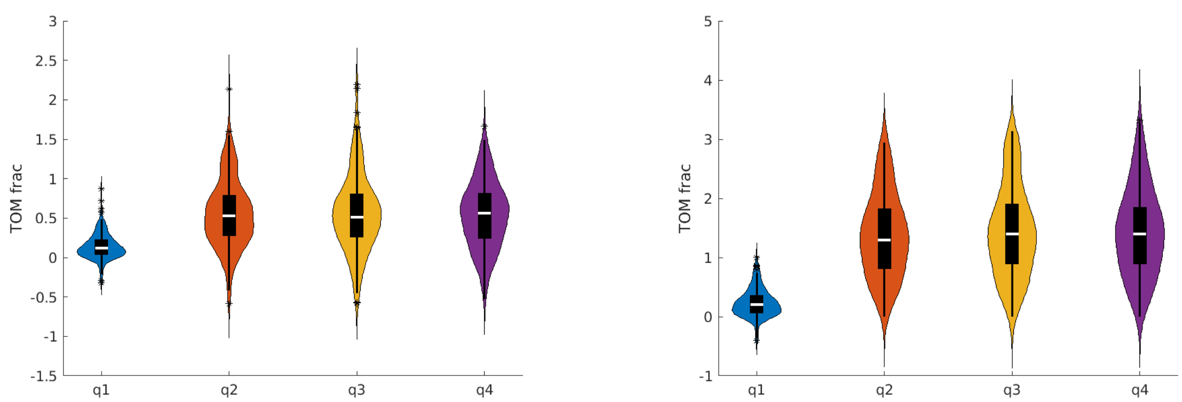


(a) Listing probability, $\alpha = 1$

(b) Listing probability, $\alpha = 10$

Figure 4: Time on Market

This figure shows violinplots the counterfactual fraction change in time on market without mortgage lock-in. The x-axis consists of quarters in 2022. The counterfactuals are computed by setting the lock-in $\beta_a = 0$ and then re-computing the equilibrium in Section 4.3. Panel (a) shows the results for the $\alpha = 1$ case, and panel (b) shows the results for the $\alpha = 10$ case.

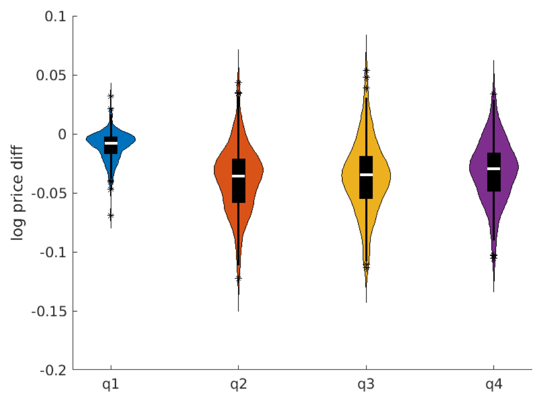


(a) TOM, $\alpha = 1$

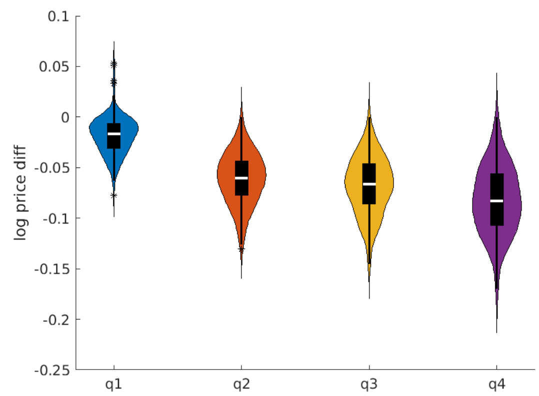
(b) TOM, $\alpha = 10$

Figure 5: Log Price Difference Between Counterfactual and Actual

This figure shows violinplots the log price difference between the actual and counterfactual market outcomes without mortgage lock-in. The x-axis consists of quarters in 2022. The counterfactuals are computed by setting the lock-in $\beta_a = 0$ and then re-computing the equilibrium in Section 4.3. Panel (a) shows the results for the $\alpha = 1$ case, and panel (b) shows the results for the $\alpha = 10$ case.



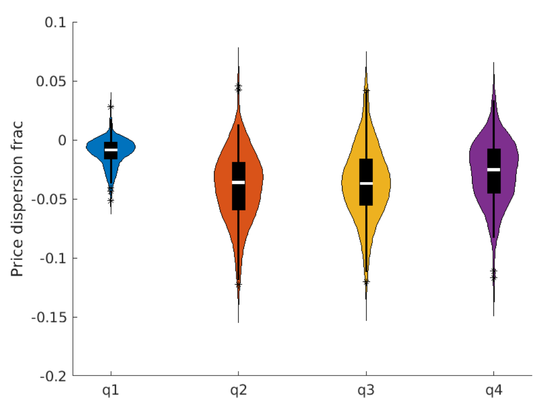
(a) Price difference, $\alpha = 1$



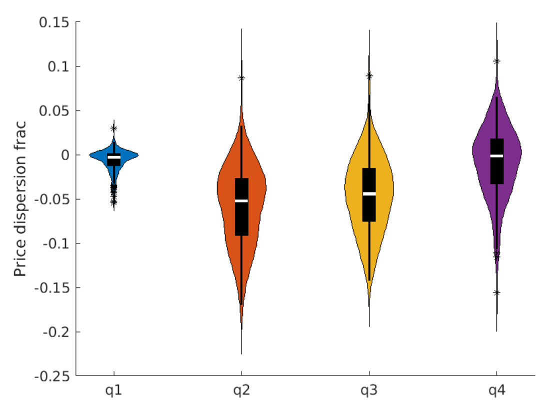
(b) Price difference, $\alpha = 10$

Figure 6: Price dispersion

This figure shows violinplots the fraction change in idiosyncractic variation in price (“price dispersion”), measured as the standard deviation of log prices, between the actual and counterfactual market outcomes without mortgage lock-in. The x-axis consists of quarters in 2022. The counterfactuals are computed by setting the lock-in $\beta_a = 0$ and then re-computing the equilibrium in Section 4.3. Panel (a) shows the results for the $\alpha = 1$ case, and panel (b) shows the results for the $\alpha = 10$ case.



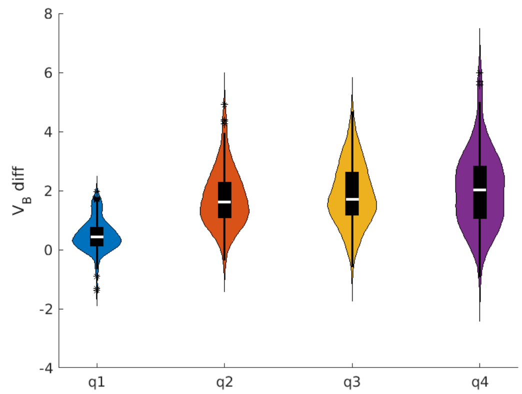
(a) Price dispersion, $\alpha = 1$



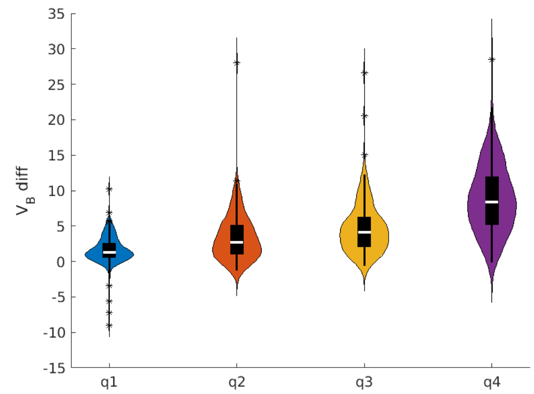
(b) Price dispersion, $\alpha = 10$

Figure 7: Buyer Expected Match Values

This figure shows violinplots the fraction change in buyer values V^b_{th} between the actual and counterfactual market outcomes without mortgage lock-in. The x-axis consists of quarters in 2022. The counterfactuals are computed by setting the lock-in $\beta_a = 0$ and then re-computing the equilibrium in Section 4.3. Panel (a) shows the results for the $\alpha = 1$ case, and panel (b) shows the results for the $\alpha = 10$ case.



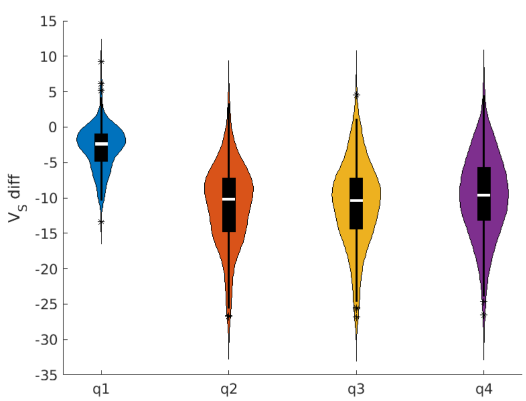
(a) Buyer value, $\alpha = 1$



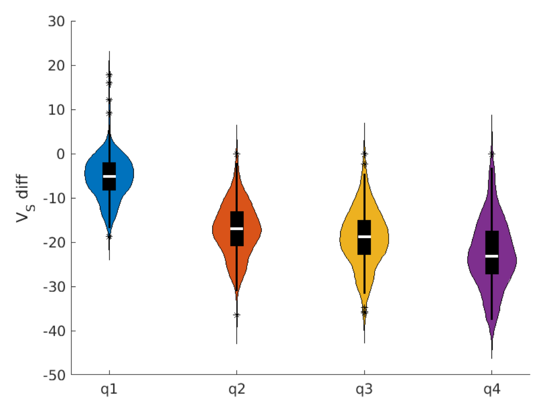
(b) Buyer value, $\alpha = 10$

Figure 8: Seller Expected Match Values

This figure shows violinplots the fraction change in seller values V^s_{th} between the actual and counterfactual market outcomes without mortgage lock-in. The x-axis consists of quarters in 2022. The counterfactuals are computed by setting the lock-in $\beta_a = 0$ and then re-computing the equilibrium in Section 4.3. Panel (a) shows the results for the $\alpha = 1$ case, and panel (b) shows the results for the $\alpha = 10$ case.



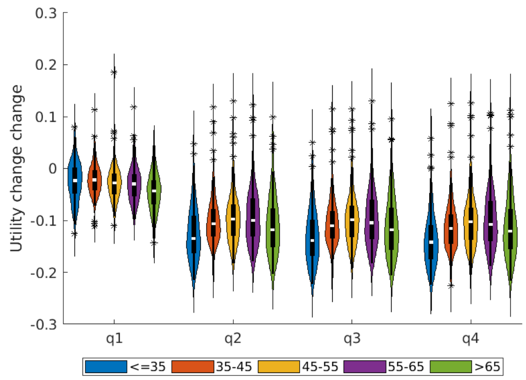
(a) Seller value, $\alpha = 1$



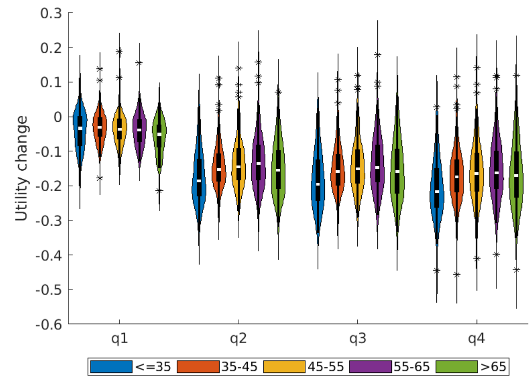
(b) Seller value, $\alpha = 10$

Figure 9: Utility difference of potential sellers by age

This figure shows violinplots the fraction change in utility of potential sellers, Eu^s_{iaht}, between the actual and counterfactual market outcomes for various age groups a without mortgage lock-in. The x-axis consists of quarters in 2022. The counterfactuals are computed by setting the lock-in $\beta_a = 0$ and then re-computing the equilibrium in Section 4.3. Panel (a) shows the results for the $\alpha = 1$ case, and panel (b) shows the results for the $\alpha = 10$ case.



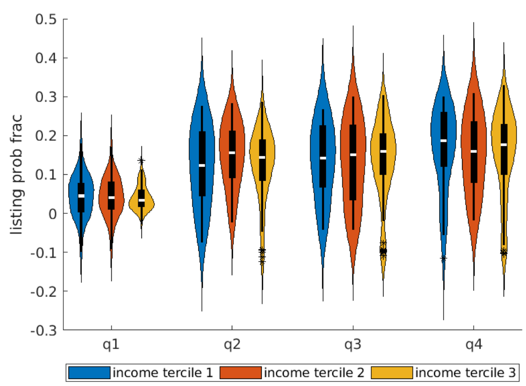
(a) Utility difference, $\alpha = 1$



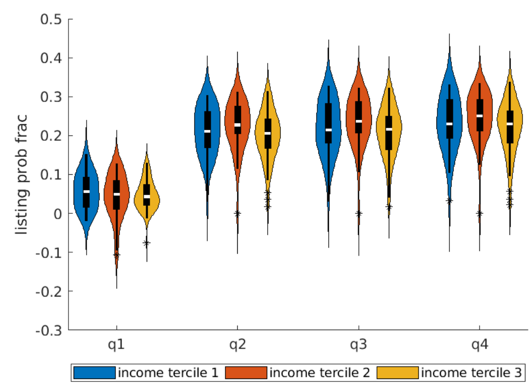
(b) Utility difference, $\alpha = 10$

Figure 10: Listing probability by Census tract income

This figure shows violinplots the counterfactual fraction change in listing probabilities without mortgage lock-in by Census tract income terciles. The x-axis consists of quarters in 2022. The counterfactuals are computed by setting the lock-in $\beta_a = 0$ and then re-computing the equilibrium in Section 4.3. Panel (a) shows the results for the $\alpha = 1$ case, and panel (b) shows the results for the $\alpha = 10$ case.



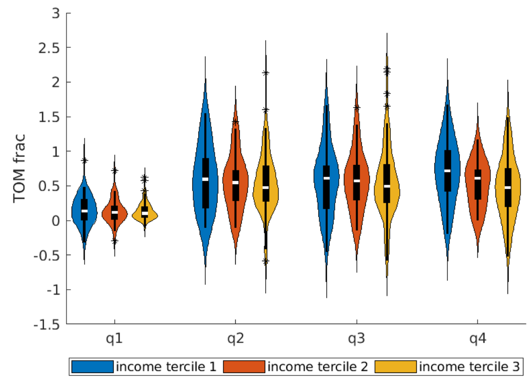
(a) Listing probability, $\alpha = 1$



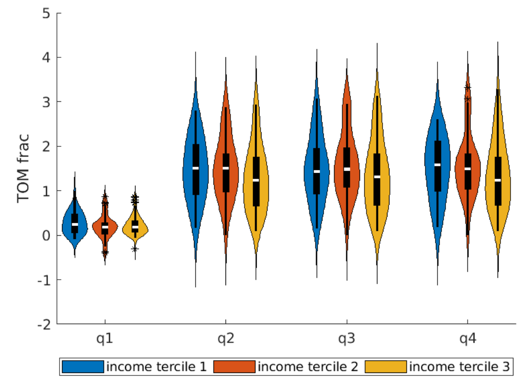
(b) Listing probability, $\alpha = 10$

Figure 11: Time on market by Census tract income

This figure shows violinplots the counterfactual fraction change in time on market without mortgage lock-in by Census tract income terciles. The x-axis consists of quarters in 2022. The counterfactuals are computed by setting the lock-in $\beta_a = 0$ and then re-computing the equilibrium in Section 4.3. Panel (a) shows the results for the $\alpha = 1$ case, and panel (b) shows the results for the $\alpha = 10$ case.



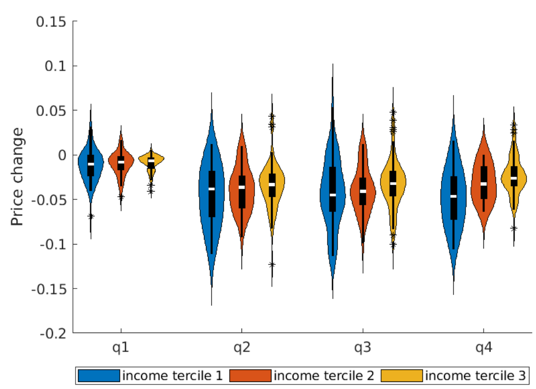
(a) TOM, $\alpha = 1$



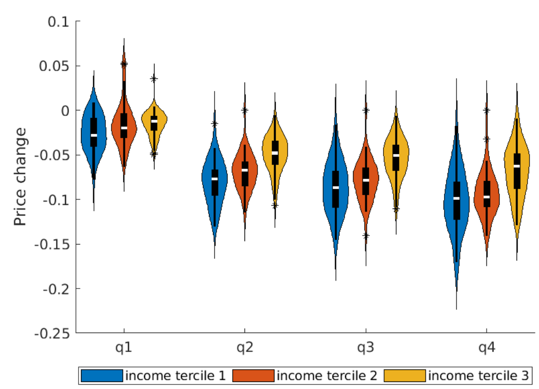
(b) TOM, $\alpha = 10$

Figure 12: Log Price Difference Between Counterfactual and Actual, by Census tract income

This figure shows violinplots the log price difference between the actual and counterfactual market outcomes without mortgage lock-in by Census tract income terciles. The x-axis consists of quarters in 2022. The counterfactuals are computed by setting the lock-in $\beta_a = 0$ and then re-computing the equilibrium in Section 4.3. Panel (a) shows the results for the $\alpha = 1$ case, and panel (b) shows the results for the $\alpha = 10$ case.



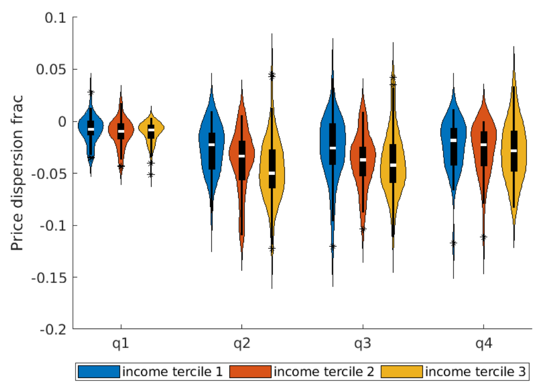
(a) Price difference, $\alpha = 1$



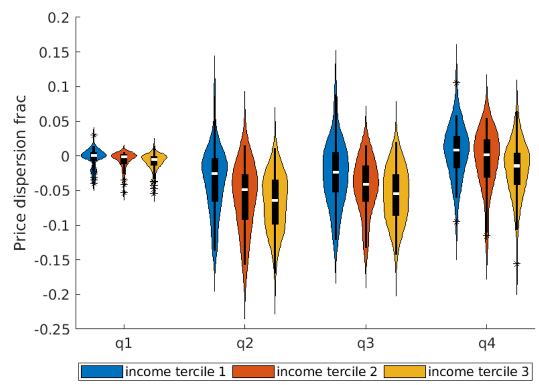
(b) Price difference, $\alpha = 10$

Figure 13: Price dispersion by Census tract income

This figure shows violinplots the fraction change in idiosyncractic variation in price ("price dispersion"), measured as the standard deviation of log prices, between the actual and counterfactual market outcomes without mortgage lock-in by Census tract income terciles. The x-axis consists of quarters in 2022. The counterfactuals are computed by setting the lock-in $\beta_a = 0$ and then re-computing the equilibrium in Section 4.3. Panel (a) shows the results for the $\alpha = 1$ case, and panel (b) shows the results for the $\alpha = 10$ case.



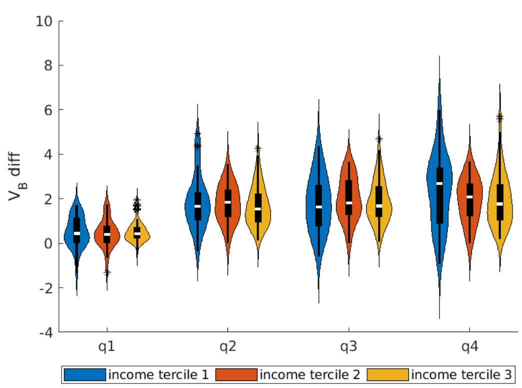
(a) Price dispersion, $\alpha = 1$



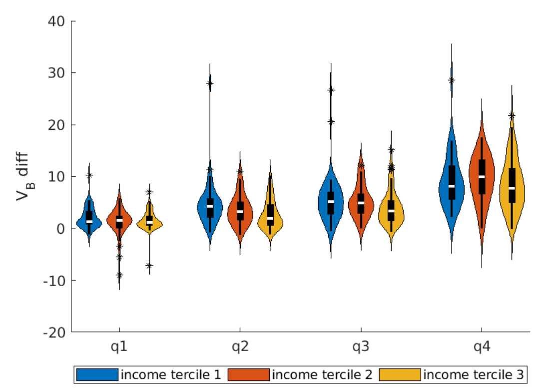
(b) Price dispersion, $\alpha = 10$

Figure 14: Buyer expected match value by Census tract income

This figure shows violinplots the fraction change in buyer values V^b_{ht} between the actual and counterfactual market outcomes without mortgage lock-in by Census tract income terciles. The x-axis consists of quarters in 2022. The counterfactuals are computed by setting the lock-in $\beta_a = 0$ and then re-computing the equilibrium in Section 4.3. Panel (a) shows the results for the $\alpha = 1$ case, and panel (b) shows the results for the $\alpha = 10$ case.



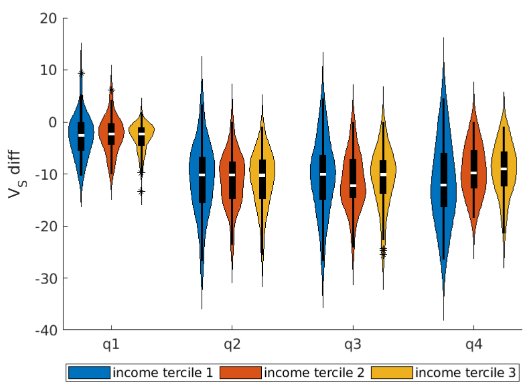
(a) Buyer value, $\alpha = 1$



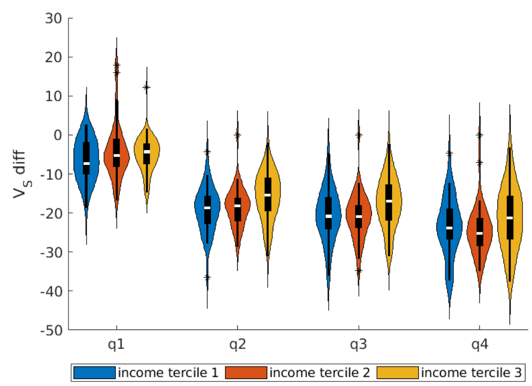
(b) Buyer value, $\alpha = 10$

Figure 15: Seller expected match value by Census tract income

This figure shows violinplots the fraction change in seller values V^s_{ht} between the actual and counterfactual market outcomes without mortgage lock-in by Census tract income terciles. The x-axis consists of quarters in 2022. The counterfactuals are computed by setting the lock-in $\beta_a = 0$ and then re-computing the equilibrium in Section 4.3. Panel (a) shows the results for the $\alpha = 1$ case, and panel (b) shows the results for the $\alpha = 10$ case.



(a) Seller value, $\alpha = 1$



(b) Seller value, $\alpha = 10$

Internet Appendix

This appendix supplements the empirical analysis of this paper. Below is a list of the sections contained in this appendix.

Table of Contents

A Additional Calibration Results

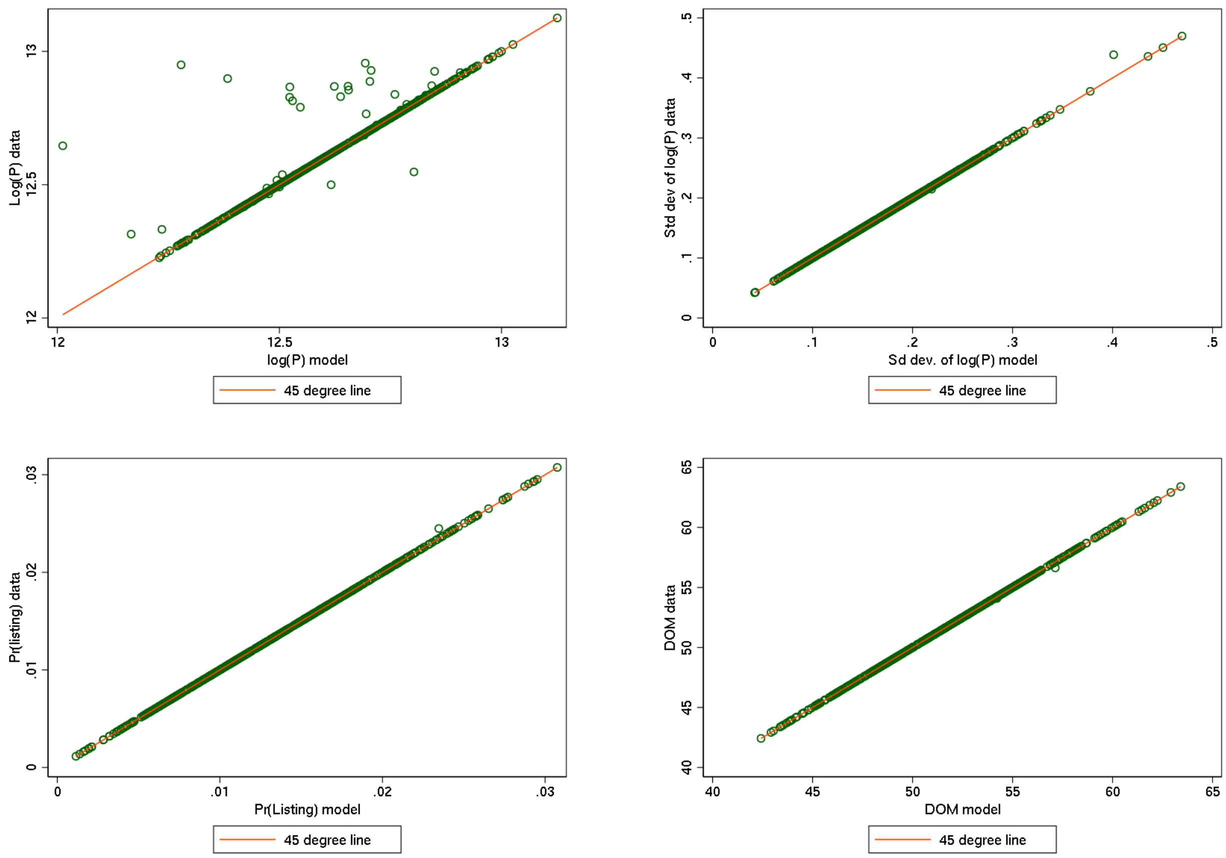
Table A.1: Calibration values

This table presents our calibrated parameters for each market h and time t following our estimation methodology in Section 5. p^0_{th} refers to the minimum value of ϵ, σ^{ϵ}_{th} refers to the idiosyncratic variation in house values, M^b_{th} refers to the mass of potential buyers, λ^m_{th} refers to the rate at which existing homeowners lists their home for sale.

Variable Name	Mean	SD	P25	Median	P75	N
Panel A: $\alpha = 1$						
p^0_{th}	11.089	2.757	9.014	10.879	13.063	642
Income tercile 1	8.383	1.515	7.340	8.418	9.385	161
Income tercile 2	9.844	1.739	8.639	9.609	10.842	148
Income tercile 3	12.951	2.124	11.448	12.849	14.313	333
σ^{ϵ}_{th}	11.122	4.063	8.238	10.929	13.683	642
Income tercile 1	11.076	4.069	7.724	11.382	13.690	161
Income tercile 2	10.142	3.346	7.884	10.392	12.605	148
Income tercile 3	11.579	4.278	8.587	11.122	14.305	333
M^b_{th}	.131	.058	.095	.124	.157	642
Income tercile 1	.136	.078	.096	.124	.160	161
Income tercile 2	.135	.048	.099	.137	.168	148
Income tercile 3	.128	.050	.095	.122	.155	333
λ^m_{th}	.053	.023	.036	.050	.067	642
Income tercile 1	.052	.023	.037	.049	.066	161
Income tercile 2	.051	.021	.035	.050	.065	148
Income tercile 3	.054	.024	.036	.051	.067	333
Panel B: $\alpha = 10$						
p^0_{th}	11.085	2.766	9.036	10.879	13.060	641
Income tercile 1	8.324	1.557	7.201	8.400	9.374	160
Income tercile 2	9.920	1.704	8.822	9.620	10.876	152
Income tercile 3	12.965	2.126	11.460	12.930	14.320	329
σ^{ϵ}_{th}	11.095	4.048	8.238	10.910	13.627	641
Income tercile 1	10.938	4.174	7.421	11.228	13.670	160
Income tercile 2	10.377	3.340	8.101	10.452	12.744	152
Income tercile 3	11.503	4.240	8.573	11.026	14.222	329
M^b_{th}	.133	.068	.096	.124	.157	641
Income tercile 1	.143	.105	.095	.123	.162	160
Income tercile 2	.133	.046	.099	.134	.165	152
Income tercile 3	.128	.051	.096	.122	.155	329
λ^m_{th}	.053	.023	.036	.050	.067	641
Income tercile 1	.052	.022	.036	.048	.066	160
Income tercile 2	.052	.021	.036	.050	.068	152
Income tercile 3	.054	.024	.036	.051	.067	329

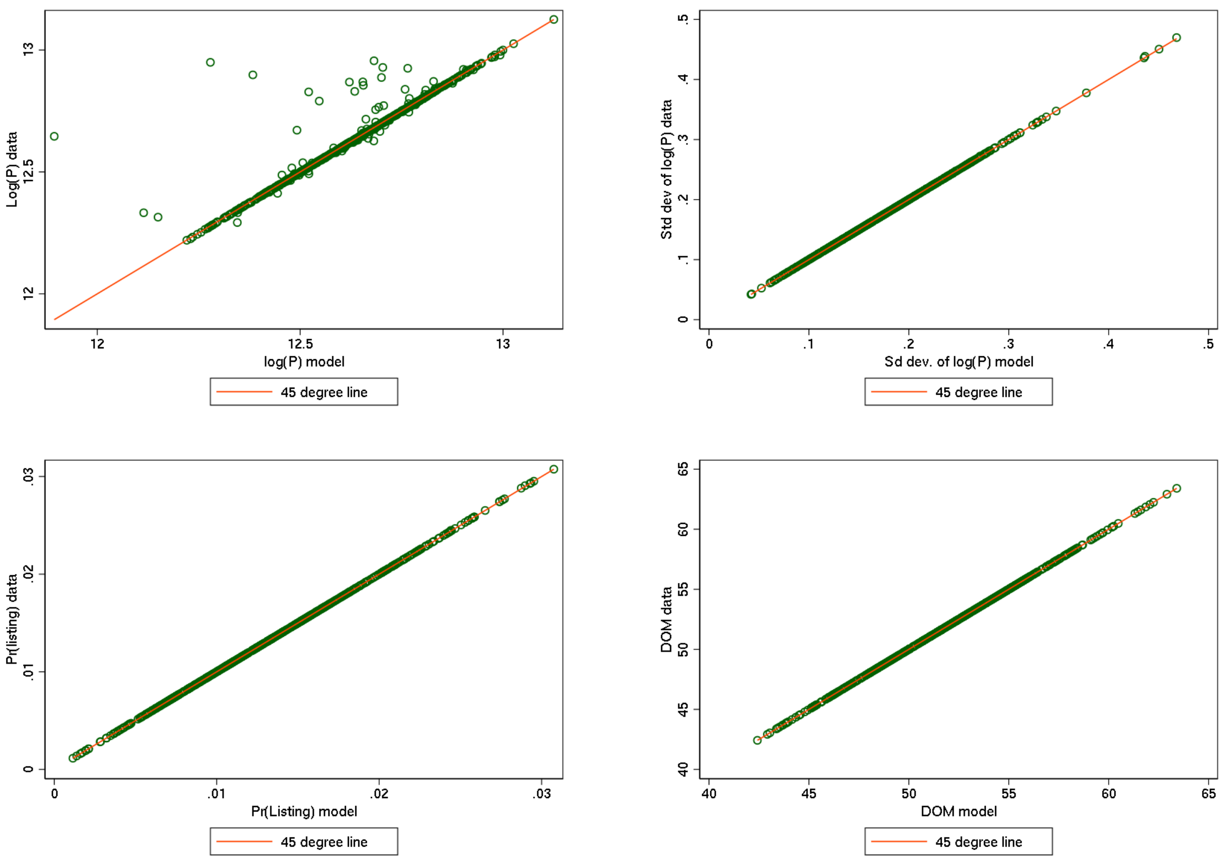
Figure A.1: Calibration fit, $\alpha = 1$

This figure presents our calibation fit for each market h and time t following our estimation methodology in Section 5 for the $\alpha = 1$ case. The x-axis consists of the model-implied values, whereas the y-axis consists of the observed data. "Log(P)" refers to the average log price of housing transactions. "Sd dev. of log(P)" refers to the standard deviation of log prices across transactions, after residualizing against property and time fixed effects and a rich set of hedonic regression variables. "Pr(Listing)" refers to the probability of listing a home, and "DOM" refers to the number of days on the market.



This figure presents our calibation fit for each market h and time t following our estimation methodology in Section 5 for the $\alpha = 10$ case. The x-axis consists of the model-implied values, whereas the y-axis consists of the observed data. "Log(P)" refers to the average log price of housing transactions. "Sd dev. of log(P)" refers to the standard deviation of log prices across transactions, after residualizing against property and time fixed effects as well as a rich set of hedonic regression variables. "Pr(Listing)" refers to the probability of listing a home, and "DOM" refers to the number of days on market.

Figure A.2: Calibration fit, $\alpha = 10$



B Additional Empirical Results

Table A.2: Mortgage lock-in effects and its heterogeneity, CCP zip code changes

This table presents linear regression and instrumental variables coefficients of the lock-in effect. Moving is defined as a change in zip code in the CCP sample or a prepay sale. The sample is our CRISM-CCP sample. Additional controls include the term of the loan, indicators for whether the mortgage is a FHA or Jumbo loan, an indicator for whether the mortgage is a refinance or is missing purpose, indicators for the type of the property, indicators for borrower documentation, indicators for the LTV of the borrower rounded to the nearest 5, indicators for the credit score of the borrower rounded to the nearest 20, and the amount of the loan rounded to the nearest $50,000. Robust standard errors are clustered at the county and year-month level.

	OLS		IV	
	(1)	(2)	(3)	(4)
	move_ccp	move_ccp	move_ccp	move_ccp
Δr	0.0407***		0.0781***	
	(0.0121)		(0.0166)	
<=35 $\times$ Δr		0.0874***		0.144***
		(0.0242)		(0.0318)
35-45 $\times$ Δr		0.0439**		0.0944***
		(0.0173)		(0.0229)
45-55 $\times$ Δr		0.0397**		0.0698***
		(0.0166)		(0.0215)
55-65 $\times$ Δr		0.0115		0.0443**
		(0.0163)		(0.0197)
>65 $\times$ Δr		0.0380*		0.0647***
		(0.0202)		(0.0245)
Age Group FE	Yes	Yes	Yes	Yes
CountyXMonth FE	Yes	Yes	Yes	Yes
Add'l Controls	Yes	Yes	Yes	Yes
Observations	1735775	1735775	1735775	1735775

Table A.3: Mortgage lock-in effects and its heterogeneity, % NPV changes, CCP zip code changes

This table presents linear regression and instrumental variables coefficients of the lock-in effect. The measurement of lock-in is the % change in the NPV of the mortgage discounted at the mortgage rate. Moving is defined as a change in zip code in the CCP sample or a prepay sale. The sample is our CRISM-CCP sample. Additional controls include the term of the loan, indicators for whether the mortgage is a FHA or Jumbo loan, an indicator for whether the mortgage is a refinance or is missing purpose, indicators for the type of the property, indicators for borrower documentation, indicators for the LTV of the borrower rounded to the nearest 5, indicators for the credit score of the borrower rounded to the nearest 20, and the amount of the loan rounded to the nearest $50,000. Robust standard errors are clustered at the county and year-month level.

	OLS		IV	
	(1)	(2)	(3)	(4)
	move_ccp	move_ccp	move_ccp	move_ccp
$\%NPV$	0.00474***		0.0109***	
	(0.00126)		(0.00199)	
<=35 × $\%NPV$		0.00704***		0.0137***
		(0.00233)		(0.00326)
35-45 × $\%NPV$		0.00428**		0.0114***
		(0.00186)		(0.00258)
45-55 × $\%NPV$		0.00472***		0.0102***
		(0.00179)		(0.00258)
55-65 × $\%NPV$		0.00247		0.00794***
		(0.00197)		(0.00250)
>65 × $\%NPV$		0.00625***		0.0115***
		(0.00226)		(0.00306)
Age Group FE	Yes	Yes	Yes	Yes
CountyXMonth FE	Yes	Yes	Yes	Yes
Add'l Controls	Yes	Yes	Yes	Yes
Observations	1735749	1735749	1735749	1735749

Table A.4: Berry logit for sell-rebuy values, interacted

	(1)
	berry_logit
same_mkt	0.672***
	(0.135)
same_cbsa	-0.420**
	(0.202)
log_dist	-0.237***
	(0.0779)
v_b	0.00208
	(0.0501)
higher_tercile	0.495
	(0.397)
higher_tercile#tercile_init=2	-0.610
	(0.614)
higher_tercile#35.age_grp	0.00830
	(0.430)
higher_tercile#45.age_grp	0.277
	(0.545)
higher_tercile#65.age_grp	0.163
	(0.524)
higher_tercile#75.age_grp	0.270
	(0.463)
higher_tercile#tercile_init=2#35.age_grp	0.256
	(0.734)
higher_tercile#tercile_init=2#45.age_grp	-0.207
	(0.806)
higher_tercile#tercile_init=2#65.age_grp	0.0865
	(0.804)
higher_tercile#tercile_init=2#75.age_grp	0.0515
	(1.059)

Age Group FE	Yes
Income Tercile FE	Yes
Age Group X Income Tercile FE	Yes
N	298

Table A.5: Berry logit for sell-rebuy values, $\alpha = 10$

This table presents results of a Berry (1994) logit on estimating the utility of sellers in market h re-buying in market h', with not re-buying in any market as the outside option, for the $\alpha = 10$ case. “same_mkt” is an indicator variable for h and h' being the same market, “same_cbsa” is an indicator variable for whether h and h' are in the same CBSA, “log_dist” is the log distance in miles between h and h', “v_b” is the value of buying in the target market, “higher_tercile” is an indicator variable for whether the h' is in a higher income tercile in h and has variation only for sellers in the bottom two terciles. An version of this table where “higher_tercile” is interacted with age groups is in Appendix Table A.6.

	(1) berry_logit
same_mkt	0.672***
	(0.129)
same_cbsa	-0.391**
	(0.198)
log_dist	-0.234***
	(0.0777)
v_b	-0.00209
	(0.00950)
higher_tercile	0.636***
	(0.196)
higher_terciletercile_init=2	-0.614**
	(0.259)
higher_terciletercile_init=3	0
	(.)
Age Group FE	Yes
Income Tercile FE	Yes
N	298

Table A.6: Berry logit for sell-rebuy values, interacted, $\alpha = 10$

	(1)
	berry_logit
same_mkt	0.673***
	(0.136)
same_cbsa	-0.420**
	(0.204)
log_dist	-0.238***
	(0.0788)
v_b	-0.000344
	(0.00968)
higher_tercile	0.496
	(0.397)
higher_tercile#tercile_init=2	-0.609
	(0.615)
higher_tercile#35.age_grp	0.00864
	(0.430)
higher_tercile#45.age_grp	0.277
	(0.545)
higher_tercile#65.age_grp	0.163
	(0.523)
higher_tercile#75.age_grp	0.272
	(0.460)
higher_tercile#tercile_init=2#35.age_grp	0.254
	(0.737)
higher_tercile#tercile_init=2#45.age_grp	-0.208
	(0.806)
higher_tercile#tercile_init=2#65.age_grp	0.0848
	(0.806)
higher_tercile#tercile_init=2#75.age_grp	0.0504
	(1.060)

Age Group FE	Yes
Income Tercile FE	Yes
Age Group X Income Tercile FE	Yes
N	298

Table A.7: BLP results with $\alpha = 10$

This table presents results of a Berry (1994) logit on estimating the utility of sellers in market h re-buying in market h', with not re-buying in any market as the outside option, for the $\alpha = 10$ case. The data used is the 2021 CCP data. "same_mkt" is an indicator variable for h and h' being the same market, "same_cbsa" is an indicator variable for whether h and h' are in the same CBSA, "log_dist" is the log distance in miles between h and h' and is set to 0 when h and h' are equal to the same market, "v_b" is the value of buying in the target market, "higher_tercile" is an indicator variable for whether the h' is in a higher income tercile in h and has variation only for sellers in the bottom two terciles. A version of this table where "higher_tercile" is interacted with age groups is in Appendix Table A.4. A version of this table for the $\alpha = 10$ case is in Appendix Table A.5. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Variable	Estimate	Std Dev
Time since purchase	-0.1324***	0.0061
Time since purchase sq	0.0056***	0.0004
35-45	-0.0325***	0.0161
45-55	-0.2167***	0.0245
55-65	-0.1873***	0.0245
>65	-0.0932***	0.0257
<=35*%NPV	0.0303***	0.0006
35-45*%NPV	0.0361***	0.0008
45-55*%NPV	0.0395***	0.001
55-65*%NPV	0.0350***	0.0011
>65*%NPV	0.0334***	0.0012
Seller + Rebuy Value	0.0067***	0.0045
Has Mortgage	0.2200***	0.0101

References

Albrecht, J., A. Anderson, E. Smith, and S. Vroman (2007). Opportunistic matching in the housing market. *International Economic Review 48*(2), 641–664.

Amromin, G. and J. Eberly (2024). Macro shocks and housing markets. Technical report, Working Paper.

Anenberg, E. and P. Bayer (2020). Endogenous sources of volatility in housing markets: The joint buyer–seller problem. *International Economic Review 61*(3), 1195–1228.

Arciadiacono, P., P. Bayer, J. R. Blevens, and P. B. Ellickson (2016). Estimation of dynamic discrete choice models in continuous time with an application to retail competition. *The Review of Economic Studies 83*(3 (296)), 889–931.

Batzer, R. M., J. Coste, W. M. Doerner, and M. J. Seiler (2024). The lock-in effect of rising mortgage rates. Technical report, Federal Housing Finance Agency.

Berry, S., J. Levinsohn, and A. Pakes (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of political Economy 112*(1), 68–105.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242–262.

Berry, S. T., J. A. Levinsohn, and A. Pakes (1993). Automobile prices in market equilibrium: Part i and ii.

Bhutta, N. and D. Ringo (2021). The effect of interest rates on home buying: Evidence from a shock to mortgage insurance premiums. *Journal of Monetary Economics 118*, 195–211.

Bosshardt, J., M. Di Maggio, A. Kakhbod, and A. Kermani (2023). The credit supply channel of monetary policy tightening and its distributional impacts. Technical report, National Bureau of Economic Research.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2020). Why is intermediating houses so difficult? evidence from ibuyers. Technical report, National Bureau of Economic Research.

Chetty, R., J. N. Friedman, N. Hendren, M. R. Jones, and S. R. Porter (2018). The opportunity atlas: Mapping the childhood roots of social mobility. Technical report, National Bureau of Economic Research.

Chetty, R. and N. Hendren (2018). The impacts of neighborhoods on intergenerational mobility ii: County-level estimates. *The Quarterly Journal of Economics 133*(3), 1163–1228.

DeFusco, A. A. and A. Paciorek (2017). The interest rate elasticity of mortgage demand: Evidence from bunching at the conforming loan limit. *American Economic Journal: Economic Policy 9*(1), 210–240.

Deng, Y., J. M. Quigley, and R. Van Order (2000). Mortgage terminations, heterogeneity and the exercise of mortgage options. *Econometrica 68*(2), 275–307.

Ferreira, F. (2010). You can take it with you: Proposition 13 tax benefits, residential mobility, and willingness to pay for housing amenities. *Journal of Public Economics 94*(9-10), 661–673.

Ferreira, F., J. Gyourko, and J. Tracy (2010). Housing busts and household mobility. *Journal of urban Economics 68*(1), 34–45.

Fonseca, J. and L. Liu (2023). Mortgage lock-in, mobility, and labor reallocation. *Working paper*.

Fonseca, J., L. Liu, and P. Mabille (2024). Unlocking mortgage lock-in: Evidence from a spatial housing ladder model. *Working paper*.

Gabrovski, M. and V. Ortego-Marti (2019). The cyclical behavior of the beveridge curve in the housing market. *Journal of Economic Theory 181*, 361–381.

Gabrovski, M. and V. Ortego-Marti (2021). Search and credit frictions in the housing market. *European Economic Review 134*, 103699.

Genesove, D. and L. Han (2012a). Measuring the thinness of real estate markets. *manuscript, University of Toronto*.

Genesove, D. and L. Han (2012b). Search and matching in the housing market. *Journal of urban economics 72*(1), 31–45.

Gerardi, K., P. S. Willen, and D. H. Zhang (2023). Mortgage prepayment, race, and monetary policy. *Journal of Financial Economics 147*(3), 498–524.

Guren, A. M. and T. J. McQuade (2020). How do foreclosures exacerbate housing downturns? *The Review of Economic Studies 87*(3), 1331–1364.

Head, A., H. Lloyd-Ellis, and H. Sun (2014). Search, liquidity, and the dynamics of house prices and construction. *American Economic Review 104*(4), 1172–1210.

Jiang, E., N. Kotova, and A. L. Zhang (2024). Search frictions and idiosyncratic price dispersion in the us housing market. *Available at SSRN 3386353*.

Jiang, E. X. and A. Zhang (2022). Collateral value uncertainty and mortgage credit provision. *University of Chicago, Chicago*.

Kind, A. and W. Buckingham (2018). Making neighborhood disadvantage metrics accessible: The neighborhood atlas. *New England Journal of Medicine* (378), 2456–2458.

Krainer, J. (2001). A theory of liquidity in residential real estate markets. *Journal of urban Economics 49*(1), 32–53.

Lambie-Hanson, L. and C. Reid (2018). Stuck in subprime? examining the barriers to refinancing mortgage debt. *Housing Policy Debate 28*(5), 770–796.

Liebersohn, J. and J. Rothstein (2024). Household mobility and mortgage rate lock.

Ngai, L. R. and S. Tenreyro (2014). Hot and cold seasons in the housing market. *American Economic Review 104*(12), 3991–4026.

Novy-Marx, R. (2009). Hot and cold markets. *Real Estate Economics 37*(1), 1–22.

Ouazad, A. and R. Rancière (2019). Market frictions, arbitrage, and the capitalization of amenities. Technical report, National Bureau of Economic Research.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of political Economy 110*(4), 705–729.

Piazzesi, M. and M. Schneider (2009). Momentum traders in the housing market: Survey evidence and a search model. *American Economic Review 99*(2), 406–411.

Piazzesi, M., M. Schneider, and J. Stroebel (2020). Segmented housing search. *American Economic Review 110*(3), 720–759.

Quigley, J. M. (1987). Interest rate variations, mortgage prepayments and household mobility. *The Review of Economics and Statistics*, 636–643.

Quigley, J. M. (2002). Homeowner mobility and mortgage interest rates: New evidence from the 1990s.

Sagi, J. S. (2021). Asset-level risk and return in real estate investments. *The Review of Financial Studies 34*(8), 3647–3694.

Wheaton, W. C. (1990). Vacancy, search, and prices in a housing market matching model. *Journal of political Economy 98*(6), 1270–1292.