Mortgage Rate Lock and House Prices *

Justin Katz †

Robert Minton ‡

[CLICK FOR MOST RECENT VERSION]

This version: March 2026

First version: December 2024

Abstract

When interest rates rise, fixed-rate mortgages generate a financial incentive for owners to keep their homes, creating "rate lock". Does rate lock dampen the negative impact of rising interest rates on house prices? To estimate rate lock's causal effect on market-level house prices, we instrument for the rate lock incentive in the outstanding local mortgage stock using unexpected family size shocks that induce moves at times with different mortgage rates. We find that when interest rates increased over 2021-23, a one standard deviation increase in the rate lock incentive, corresponding to a 0.3pp lower average outstanding mortgage rate, raised nominal house price growth by 2.6pp. To understand the mechanism, we compare mobility of owners who purchased homes just before versus just after sharp mortgage rate increases. A 1pp lower outstanding mortgage rate reduces moves from owning to renting by 33%, a force increasing the price-to-rent ratio, and reduces overall moves by 42%. Using these mobility estimates, we calibrate a dynamic structural model to quantify how much rate lock dampens the aggregate price effects of a higher cost of capital. Model simulations indicate that the 2021-23 tightening would have reduced the price-to-rent ratio by 12.3% with adjustable-rate mortgages, and hence no rate lock, versus only 4.4% with fixed-rate mortgages.

Keywords: fixed-rate mortgages, monetary policy, house prices, household finance

JEL classification: G51, E21, R21, R31, L78

* For valuable comments and discussions, we thank Josh Abel, John Campbell, Gabe Chodorow-Reich, Mark Egan, Kris Gerardi, Ed Glaeser, Benny Goldman, Adam Guren, Robin Greenwood, Sam Hanson, Myrto Kalouptsidi, David Laibson, Max Miller, Ariel Pakes, Sarah Robinson, Maya Roy, Peleg Samuels, Jesse Shapiro, Andrei Shleifer, Jeremy Stein, Adi Sunderam, Paul Willen, Ryan Xia, and participants at the Harvard workshops in financial economics, industrial organization, public and labor, and macroeconomics. Katz gratefully acknowledges support from The Institute of Consumer Money Management (ICMM) Pre-doctoral Fellowship on Consumer Financial Management, awarded through the NBER, and from the John Meyer Dissertation Fellowship of the Harvard Joint Center for Housing Studies. The views expressed in this paper are solely those of the authors and not necessarily those of the Federal Reserve Bank of Boston, the Federal Reserve Board, nor the Federal Reserve System. All errors are ours.

† Harvard University and Federal Reserve Bank of Boston. Email: jkatz@g.harvard.edu.

‡ Federal Reserve Board. Email: robert.j.minton@frb.gov

Acknowledgements.

Justin Katz acknowledges support from The Institute of Consumer Money Management (ICMM) Pre-doctoral Fellowship on Consumer Financial Management, awarded through the NBER, and from the John Meyer Dissertation Fellowship of the Harvard Joint Center for Housing Studies. He has no other funding sources or other conflicts of interest to disclose.

Robert Minton has no funding sources or conflicts of interest to disclose.

The views expressed in this paper are solely those of the authors and not necessarily those of the Federal Reserve Bank of Boston, the Federal Reserve Board, nor the Federal Reserve System.

From 2021 to 2023, the ratio of house prices to rents rose by 17%, while the average rate on new 30-year fixed-rate mortgages rose by 3.9pp.[1] The surge in house prices alongside this rapid mortgage rate increase is puzzling, given a long literature showing tighter financial conditions and a higher cost of capital reduce housing demand and valuation ratios.[2] One explanation is that rising rates reduce the supply of existing homes due to "rate lock", where owners with fixed-rate mortgages at below-market rates have a financial incentive to keep their current homes.[3] Jerome Powell expressed this view at a 2023 press conference: "people who have low-rate mortgages [...] [are] not going to [sell] [...] which means that supply of existing homes is really, really tight, which is keeping prices up" (Federal Reserve Board 2023).

Can rate lock explain why a house price boom coincided with rapid financial tightening? While rate lock implies lower mobility (Quigley 1987, Fonseca and Liu 2024, Liebersohn and Rothstein 2024), it does not obviously impact prices. If rate lock reduces sales for owners who would have bought a new home in the same housing market, the number of buyers and sellers falls equally, leaving prices unchanged. But if rate lock reduces sales for owners who would have moved to a different market, it shifts relative demand and affects relative prices. In particular, rate lock can increase aggregate house prices relative to rents by reducing moves into the rental market. Quantifying rate lock's price effect is challenging, because recent house price dynamics reflect both higher rates and fundamental shocks (Van Nieuwerburgh 2023).

This paper quantifies rate lock's effect on housing market equilibrium. First, we estimate a large effect of rate lock on house prices, using family size shocks that cause moves at times with different mortgage rates as instruments for the rate lock incentive in the outstanding mortgage stock. Second, we find that rate lock reduces own-to-rent moves, a key mechanism for effects on aggregate price-to-rent ratios, by comparing homeowners who buy just before and after a sharp mortgage rate increase. Finally, we develop and estimate a model of dynamic housing demand to simulate a rate increase under existing fixed-rate mortgages and alternative mortgage designs, and find that rate lock offsets much, but not all, of the resulting negative price effects. Other fundamental shocks are thus required to explain recent price growth.

Section 1 describes a theory of how rate lock offsets the negative house price effects of higher interest rates. Owners who sell a home with a fixed-rate mortgage at below-market

[1] From 2021Q1 to 2023Q4, single-family house prices rose 35.2% (FHFA all-transaction index); rents rose 18.4% (BLS); and the average rate on new 30-year fixed-rate mortgages rose from 2.7% to 6.6% (Freddie Mac).

[2] Papers emphasizing the role of financial conditions in explaining house prices include Henderson and Ioannides (1983), Poterba (1984), Himmelberg *et al.* (2005), Mayer and Sinai (2007), Hubbard and Mayer (2009), Favilukis *et al.* (2017), Garriga and Hedlund (2020), Mian and Sufi (2022), Justiniano *et al.* (2022), Foote *et al.* (2021), Samuels (2024), Greenwald and Guren (2024), and papers reviewed in Kuttner (2014) and Časta (2025), among others. Shiller (2006), Glaeser *et al.* (2012), and Kaplan *et al.* (2020), among others, emphasize other factors to explain price fluctuations. Appendix D.1 shows 2021-23 rate hikes reduced prices internationally.

[3] Conventional mortgage "due-on-sale" clauses require repayment at face value on home sale, prohibiting buyers from assuming mortgages. USDA and FHA loans allow assumptions, with∼0.5% assumed a year (Park 2022).

interest rates incur a "prepayment cost": the gap between the mortgage balance repaid on sale and the present-value of remaining payments at the locked-in rate.[4] We add prepayment costs to a user-cost model of house prices (Poterba 1984, Himmelberg *et al.* 2005, Glaeser *et al.* 2012), where households move between partially-segmented markets with owner-occupied or rental units. When interest rates rise, the cost of capital increases, reducing demand for owner-occupied units relative to rentals. Prepayment costs also rise, which reduces moves from owning to renting by incentivizing sellers to keep their current home. This decreases the supply of owner-occupied units and incrementally raises the price-to-rent ratio.

Sections 2 and 3 then empirically estimate the impact of prepayment costs on price growth in local housing markets. We use two administrative datasets, provided to the Federal Reserve System, that cover most post-2005 US housing market activity: a credit bureau panel merged to mortgage terms and repayments data, and a property transactions panel. We calculate prepayment costs as the gap between a mortgage's balance and its estimated market value, scaled by property value. We study the period from the end of 2021 to the end of 2023, the only period in our data with significant prepayment cost changes. We compare zip codes within Core-Based Statistical Areas (CBSAs) to ensure markets have similar fundamentals.

Regressing market-level price growth on the change in average prepayment cost likely yields biased estimates. Markets with more inflows during low-rate periods, with higher average prepayment costs, may have unobservable differences related to 2021-23 house price growth.[5] Supporting this concern, average prepayment costs strongly correlate with population growth, age, income, and other demographics. Existing papers instrument for *individual household* mortgage rates using the average rate at origination, and demographic balance when comparing individuals *within* a local market (e.g., Fonseca and Liu 2024). However, at the *market level*, we find that instruments using average origination timing correlate strongly with market-level demographics. For unbiased estimates, we thus need *market-level* instruments.

We develop three new market-level instruments for the rate lock incentive in the outstanding mortgage stock. The instruments use increasingly broad variation, making estimates potentially more representative but introducing potential biases. The three approaches give similar results. All are quite different from estimates using endogenous origination timing.

The first instrument is based on past moves due to having children, a main non-financial reason for mortgage origination. We focus on unexpected moves due to pre-2006 twin births. We estimate that having twins rather than one child increases moves when twins are born and

[4] "Prepayment costs" are distinct from "prepayment penalties", fees some lenders charge for early repayment.

[5] Lower interest rates increase housing demand, inducing moves and more mortgage originations. This could cause bias in either direction. People move to where housing units are available – e.g., places with elastic housing supply – causing downward bias, as such places have lower price growth. People also move to where they expect price growth, causing upwards bias if expectations are correct.

enter elementary school. As a result, variation in the timing of twin births across markets generates twin-driven household moves at different times that happen to coincide with different Treasury yields. We use *changes* in the twin birth rate[6] from 1995-2006 to predict local moves from 1995-2021, and construct the instrument using co-movement between predicted moves and the prepayment cost of an average mortgage originated in each year. We restrict to pre-2006 twin births so that twin-driven moves predate 2021, avoiding direct effects on housing markets over 2021-23 that would violate the exclusion restriction.

Consistent with the exclusion restriction, variation in the year-to-year *change* in the local twin birth rate appears driven by finite sample variation, conditional on CBSA. This implies induced moves are quasi-random. As further support, we find that the instrument does not correlate with local demographics like age, income, or population growth. The first stage is strong (F-stat=75): twin birth rate fluctuations from 1995-2006 lead to small cross-sectional differences in annual origination *flow*, which accumulate to substantial differences in the outstanding mortgage *stock* by 2021.[7]

The twins instrument uses relatively narrow variation in moves, and so may identify a non-representative local average treatment effect. Two other instruments, using broader variation, yield similar estimates, suggesting the twins estimate is representative. Our "first birth" instrument follows the same methodology as the twins instrument but predicts moves with pre-2006 changes in the number of births to first-time mothers, a family formation proxy. Our third instrument uses observed, rather than predicted, origination timing, but strips out variation we think is most prone to bias. We remove all variation from recent originations in 2019-2021, and also control for the share of mortgages originated in half-year bins from 2002-2019. We thus exploit differences in the timing of moves within half-year intervals across markets.

Estimates using the twins instrument imply a one standard deviation increase in prepayment cost, equivalent to a 0.3pp decrease in 2021 average outstanding mortgage rates, causes a 2.6pp increase in 2021-23 nominal house price growth. Scaling this estimate linearly implies that eliminating 2023 prepayment costs would reverse a third of realized 2021-23 house price growth, and that prepayment cost variation explains 33% of the cross-sectional variation in price growth. Robustness checks show that estimates are not driven by in-vitro fertilization

[6]The twin birth rate is the fraction of births that are twins.

[7]For example, in our sample, a market with 10,000 owner-occupied households with a mortgage, with a population around 70,000, has 900 births/year, 3.6% of which are twins. If twin births are iid binomial, the standard deviation of the change in the twin birth rate is $\sqrt{2} \times \sqrt{3.6\% \times 96.4\%/900} = 0.88\%$. Over 25 years, the standard deviation of the gross change is $\sqrt{25} \times 0.88\% = 4.4\%$. Moving 2 standard deviations above and below the mean affects the timing of moves for $4 \times 4.4\% \times 900 = 158$ households, 1.6% of the housing stock. If 5% of households sell in a given year, twin-affected households will be 1.6% of potential sales in the average market, but due to variation in the set of households who move, will be >2% of potential sales in 25% of markets. Appendix D.7.2 applies constructs the instrument using simulated iid binomial twin births. In the median market, ±2 standard deviations around the simulated instrument mean changes the timing of moves for 1.3% of households.

births nor affected by controlling for the *expected* twin birth rate change, ensuring estimation only uses random exposure from sampling variation. The twin and first birth instruments can be constructed with public data as general-purpose mobility instruments.

To understand the mechanisms driving the price effects of rate lock, we study the household-level response to the rate lock incentive in Section 4. Since we focus on individual rather than market outcomes, we use variation in mortgage origination timing for exogenous rate changes, following Fonseca and Liu 2024 and Liebersohn and Rothstein 2024. We differ from these papers in two ways. First, we only use differences in origination timing due to moves (not refinancing), since moves are frictional and hard to time around high-frequency rate changes. Second, we compare households who purchased a home just before and after the sharp rate increase around November 2016, the largest month-to-month increase in the 10-year Treasury rate over 2009-2019, which we find did not induce selection on observable characteristics.

A 1pp increase in outstanding fixed mortgage rates raises the four-year sale probability by 42% (5.2pp). Forward-looking wealth effects, rather than liquidity effects, appear to drive lower mobility. Estimates are higher for households with higher loan-to-value ratios, where mortgage rate variation impacts wealth more, but do not vary with debt-to-income ratios, a proxy for liquidity. Furthermore, owners with lower rates sell at higher prices, with the price difference implying compensation for many years of higher payments at the new rate.

We also find that a 1pp increase in fixed mortgage rates increases the four-year probability of an own-to-rent move by 33% (1.7pp). An effect on own-to-rent moves implies that the cross-sectional price effects described above have implications for the aggregate price-to-rent ratio. The potential for aggregate effects is large: Chodorow-Reich *et al.* (2025) estimate that one-third of moves by owners are from owning to renting, a finding we replicate in our administrative data and the National Longitudinal Survey of Youth (NLSY79).

We use these results to estimate a structural model of dynamic housing demand. The model quantifies how much rate lock offsets the negative price effects of a higher user cost of capital when rates rise, and clarifies how cross-sectional price estimates aggregate. We also evaluate recent policy proposals to reduce rate lock: making mortgages portable (kept by borrowers upon move, as in Canada and the UK), which the Trump administration is "actively evaluating" (CNN 2025), or giving borrowers a "buyback option" to repurchase their mortgages at market, rather than face value (as in Denmark, discussed in Berger *et al.* 2026 and Campbell and Ramadorai 2026, and proposed in a Wall Street Journal op-ed by Pollock and Kupiec 2026).

The model embeds a dynamic discrete housing choice in an incomplete-markets lifecycle consumption-saving model with fixed-rate mortgages and stochastic riskfree rates. This captures both consumption and investment components of dynamic housing demand (Henderson and Ioannides 1983), connecting models with rich housing preferences but stylized mortgages

(e.g., Bayer *et al.* 2016) and those focused on home equity as investment (e.g., Landvoigt *et al.* 2015). Households vary by age, wealth, homeownership, and mortgage characteristics. Each period, they choose whether to remain in their current unit, move to a new owner-occupied unit, or move to a new rental unit. Equilibrium house prices equate home sales and purchases.

Rising interest rates impact housing demand through standard user-cost channels: home purchases become expensive for renters, and bonds become more attractive relative to home equity. But fixed-rate mortgages introduce a rate lock channel. Own-to-own moves fall due to the increase in new mortgage rates, and own-to-rent moves fall as owners both expect higher mortgage rates on future rent-to-own moves and exploit the spread between low-rate borrowing and high-rate saving. Rate lock raises house prices by reducing own-to-rent moves.

We estimate the key parameters that control the effect of rate lock on home sales by matching causal mobility moments from the 2016 event study. Crucially, the estimated model matches the untargeted IV estimate of the effect of prepayment costs on 2021-23 price growth, even though we only estimate using effects on mobility, not prices. The model's key mechanism—effects on own-to-rent moves—can thus generate the price effects observed in the data.

We then simulate the 2021-23 rate hikes, holding fixed the 2021 mortgage stock, age and wealth distribution, household preferences, and rents. With fixed-rate mortgages and rate lock, house prices fall by 4.4%. With new and existing owners forced into adjustable-rate mortgages and no rate lock, prices fall by 12.3%. We attribute 6.4pp of this 7.9pp difference to rate lock for existing owners, based on a counterfactual in which existing owners face adjustable-rate mortgages while renters continue to buy with fixed-rate loans.

Rate lock thus played an important role in recent aggregate house price changes. The 6.4pp price increase from rate lock explains one-third of the gap between the 12.3% decrease from higher user costs and the observed 5.6% increase in real house prices from year-end 2021-23. The model's 4.4% price decrease is closer to the data than reduced-form (Federal Reserve 2024) or structural (Amromin and Eberly 2023) work without rate lock that predicts 20% and 37% decreases, respectively. But rate lock has a smaller effect than higher user costs, so changes in preferences, wealth, or housing supply are needed to explain recent price *growth*.

In simulations with portable or buyback option mortgages, tightening has a similar impact on aggregate prices as with fixed-rate mortgages. With portable mortgages, rate lock no longer impacts own-to-own moves, but still reduces own-to-rent moves, and hence raises aggregate prices. Buyback option mortgages avoid rate lock, but give owners a realized capital gain that reduces own-to-rent moves by increasing their wealth. Fixed rates, not features like non-portability, are thus the key friction attenuating the price effects of tightening.

Related literature. Rising interest rates have revived interest in the impact of rate lock on mobility (Bonanno 1971, Quigley 1987, Goodman and Bai 2017, Fount and Oundee 2020, Fonseca and Liu 2024, Batzer *et al.* 2024, Abel 2024, Liebersohn and Rothstein 2024). We use a well-identified event study to provide new facts on how rate lock affects moves between housing market segments, such as from owning to renting. With especially rich data on mortgage contracts, we identify the key mechanism as the negative net worth effect of losing low-rate mortgages. More broadly, our analysis of the price impact of existing-home supply choices gives a framework to study equilibrium effects of other housing lock-in effects, such as "tax lock" due to asymmetric treatment of capital gains (Newman and Reschovsky 1987, Cunningham and Engelhardt 2008); "equity lock" when falling house prices put owners underwater on their mortgages (Ferreira *et al.* 2010, Andersson and Mayock 2014, Bernstein and Struyven 2022), and lock-in due to other transactions costs (Quigley 2002).[8]

Several contemporaneous papers study the price effects of rate lock using structural models, emphasizing lifecycle dynamics (Amromin and Eberly 2023), relative prices across housing market quality segments (Fonseca *et al.* 2024), and the price effects of market liquidity (Gerardi *et al.* 2024, Aladangady *et al.* 2025). These papers study perfect foresight, steady-state equilibria where interest rates are constant over time and outstanding mortgages have identical rates, and typically model rate lock as a reduced-form transaction cost. Our paper puts more emphasis on estimating the reduced-form impact of rate lock on local house prices, using new instruments to reveal a large causal effect. Our structural model features uncertain, time-varying interest rates. Rate lock emerges endogenously when the riskfree rate rises, and so coincides with an increased user cost of housing.[9] These unique features make it possible to simulate a rate hike given the observed 2021 mortgage rate distribution, quantify aggregate price effects attributed to rate lock and higher user costs, and evaluate counterfactual mortgage contracts. Because the model captures dynamic responses to rate changes, we can estimate it with causal mobility effects and validate it with IV estimates.

We also contribute to a large literature on how the financial environment and the mortgage market influence housing demand and house prices. One strand of this literature studies effects of interest rates on house prices (e.g., Himmelberg *et al.* 2005, Smith and Smith 2006, Glaeser *et al.* 2012, Adelino *et al.* 2025) and rents (Abramson *et al.* 2025, De la Roca *et al.* 2026). Estimates of the house price semi-elasticity to long-term rates range from 0 to -15 in reviews by Kuttner (2014) and Časta (2025). One reason for this wide range could be that the rate lock incentive, which offsets the negative impact of tightening, varies over time. Moreover, the rate

[8] Broker commissions are an especially policy-relevant example, given recent litigation surrounding National Association of Realtors price fixing (Akcam *et al.* 2024, Kim 2024). Papers such as Andersen *et al.* (2022) study the effects of nominal loss aversion on sales, but the equilibrium effects differ from our framework.

[9] Abel (2024)'s model has these features, but "is silent on home prices" (p. 4).

lock incentive depends on past conditions, implying path dependence monetary policy's house price effects. Low rates in the 2010s increased the share of borrowers with historically cheap mortgages; while Berger *et al.* (2021) argue this reduced room for future monetary stimulus, our results imply it also dampened the house price effects of tightening.

Another strand focuses on collateral constraints, using applied theory (e.g., Stein 1995, Campbell and Hercowitz 2005, Kaplan *et al.* 2020, Favilukis *et al.* 2017) or quasi-experimental variation in credit supply (e.g., Loutskina and Strahan 2015, Di Maggio and Kermani 2017, Mian and Sufi 2022). In our context, fixed rates on existing mortgages grant owners the ability to borrow cheaply, similar to a housing collateral expansion tied to owners' current units. That this effect influences house prices is an economic application of the Fostel and Geanakoplos (2016) notion of an asset's "collateral value," where asset prices increase if their ownership facilitates borrowing on favorable terms. While we focus on housing markets, the supply and demand forces we study apply to equilibrium in any durable goods market where ownership with fixed-rate financing is tied to use.

Sections 1-6, respectively, present the conceptual model to organize our empirical analysis; data; market-wide equilibrium price effects; reduced-form evidence on household behaviors; the empirical model, its estimation, and counterfactuals; and the conclusion.

1 Valuing houses with fixed-rate mortgages

This section adds fixed-rate mortgages to a "user-cost" model of the price-to-rent ratio to show how rate lock impacts equilibrium house prices and transaction volume. The results guide reduced-form analysis in Sections 3 and 4. For quantitative exercises in Section 5, we add household heterogeneity, uncertainty, non-stationarity, credit constraints, and other frictions.

Section 1.1 provides intuition with a static, two-market model, and delivers modified versions of the formulas in Poterba (1984), Himmelberg *et al.* (2005), and Glaeser *et al.* (2012) Section 1.2 summarizes results from an extension in Appendix A.3 with an arbitrary number of markets and housing demand microfounded as the solution to a dynamic saving-consumption problem. Proofs are in Appendix A.

1.1 Simple two-market framework.

A city has two housing market segments: one in which all units are owner-occupied, and one with rented units. Owner-occupied units are financed with fixed-rate mortgages originated at the current riskfree interest rate r. Renters pay per-period spot rents to absentee landlords. Converting a rental unit to owner-occupied and vice-versa is costly, so the number of units in each market segment is not fully flexible. Spot rents are the numeraire.

Households and housing supply. A fixed mass of households each initially starts as an owner-occupant or a renter. Owner-occupants have fixed-rate mortgages with balance M and interest rate r_0. Households then make a once-and-for-all choice of whether to stay in their current unit or move to a alternative unit. This choice depends on the present-value cost of housing, where the discount rate for future costs correlates with the riskfree rate r.[10] The more a move increases present-value costs, the fewer households choose to move.

Renters who purchase an owner-occupied unit increase net demand for owner-occupied units. This move changes the present-value cost of housing by $P - PV[\text{Rent}]$, the difference between owner-occupied house price P and the present-value cost of renting. Denote the mass of households moving from renting to owning as $D(P - PV[\text{Rent}])$, with $D' < 0$. An increase in house prices relative to the present-value cost of rents reduces net demand.

Owners who sell their home to become renters increase the net supply of owner-occupied units. The change in present value from this move is $(P - M) + PV[\text{Pmt}(r_0)] - PV[\text{Rent}]$: revenue from the home sale, plus the (avoided) present-value cost of fixed-rate mortgage payments, less the present-value cost of renting. Rearranging, denote the mass of households moving from renting to owning as $S(P - (M - PV[\text{Pmt}(r_0)]) - PV[\text{Rent}])$, with $S' > 0$. An increase in house prices relative to the present-value cost of rents increases net supply.

Due to fixed-rate mortgages, an increase in market interest rates create a "prepayment cost": $PC := M - PV[\text{Pmt}(r_0)]$. The prepayment cost is a financial penalty for moving, and equals gap between the mortgage balance owners must repay upon move and the present-value cost of mortgage payments they would have paid if they stayed: $PC := M - PV[\text{Pmt}(r_0)]$. If $r_0 = r$, mortgages are actuarially fair, and lenders and borrowers discount costs at the same rate, then $PC = 0$. Prepayment costs rise with an *increase* in the market risk-free rate ($\frac{\partial PC}{\partial r} > 0$) or due to *lower* outstanding mortgage rates, which reduce payments ($\frac{\partial PC}{\partial r_0} < 0$).

Prepayment costs arise because of two frictions: first, mortgages have fixed interest rates, and second, mortgages must be repaid at face value. If instead r_0 adjusted to match r, or borrows could repay the present-value cost of payments given r, then $PC = 0$.

Finally, a mass of contractors may purchase owner-occupied units and pay a cost to convert them to rental units and vice versa. Rental units trade at the present-value cost of rents. Denote the mass of units converted from rentals to owner-occupied by $Q_S(P - PV[\text{Rent}])$, with $Q_S' > 0$.

Equilibrium. Equilibrium owner-occupied house prices equate demand and supply:

$$D(P - PV[\text{Rent}]) = S(P - PV[\text{Rent}] - PC) + Q_S(P - PV[\text{Rent}]) \tag{1}$$

We highlight two key equilibrium properties. First, markets with higher prepayment costs

[10] Appendix A.2 shows results are identical if moves instead depend on per-period user costs.

have higher house prices and lower sales:

$$\frac{\partial P}{\partial PC} = \frac{S'(\cdot)}{S'(\cdot) - D'(\cdot) + Q'_S(\cdot)} := \omega \geq 0, \quad \frac{\mathrm{d}S}{\mathrm{d}PC} \leq 0 \tag{2}$$

Intuitively, higher prepayment costs reduce own-to-rent moves, and hence the supply of owner-occupied units. There is no offsetting effect on rent-to-own moves, because renters do not have fixed-rate mortgages. Equilibrium prices thus rise.

In general, $\frac{\partial P}{\partial PC}$ is higher when prepayment costs have a larger impact on own-to-rent moves (high $\frac{\partial S}{\partial PC} = S'$). If markets are fully segmented, then $S'(\cdot) = 0$, and $\omega = 0$. Existing owners also have a larger impact on total supply when housing supply is less price elastic. As housing supply becomes more flexible (Q'_S increases), ω falls.

Interestingly, $\frac{\partial P}{\partial PC}$ does not decrease if households view owner-occupied and rental markets as closer substitutes. Suppose that $S(\cdot) = \alpha \cdot S_0(\cdot)$, and $D(\cdot) = \alpha D_0(\cdot)$. Higher α implies moves are more sensitive to cost differences, reflecting a higher elasticity of substitution. In this case, $\frac{\partial \omega}{\partial \alpha} \geq 0$, with equality holding when $Q'_S = 0$. Intuitively, a higher elasticity of substitution means greater own-rent arbitrage from current renters, which should reduce the price effects of PC. However, a higher elasticity of substitution *reduces* own-rent arbitrage for current *owners*, since prepayment cost becomes relatively more important. Higher α also makes choices from current owners more important for prices compared to unit conversions, so $\frac{\partial \omega}{\partial \alpha}$ is weakly *positive*.

Second, prepayment costs reduce the price effect of an increase in interest rates:

$$\frac{\partial P}{\partial r} = \underbrace{\frac{\partial PV[\text{Rent}]}{\partial r}}_{\text{User cost: } <0} + \underbrace{\omega \frac{\partial PC}{\partial r}}_{\text{Rate lock: } >0} \tag{3}$$

The first term is negative, and reflects the fact that a higher risk-free rate makes renting more attractive by increasing the opportunity cost of capital. With $\Delta PC = 0$, this would predict that the price impact of higher rates equals the impact on the present-value cost of rents, familiar from standard user-cost analysis (e.g., Poterba 1984, Himmelberg *et al.* (2005)). The second term is positive, and reflects the price impacts of increased prepayment cost. The size of this offset is equal to the cross-sectional effect of prepayment costs on prices, and depends on the impact of prepayment costs on own-to-rent moves.

Equations (2) and (3) imply that the capital structure of housing impacts equilibrium prices. There are two frictions that explain this apparent Modigliani-Miller violation. First, due to fixed-rate mortgages, capital structure impacts households' cash flow – a mortgage with a below-market rate generates an implicit flow dividend based on the gap between the market rate of saving and the rate of borrowing. Second, we implicitly assume limits to arbitrage, since conversions between owner-occupied and rental units are costly. As $Q'_S \to \infty$ and contractors

can freely convert rental units to owner-occupied, $\omega \to 0$.

1.2 Extension: Multi-market model.

In Appendix A.3, we show that the key results from the simple two-market framework still apply with many partially-segmented markets. We analyze an extended model with an arbitrary number of owner-occupied and rented housing market segments, and micro-found moves across segments as the solution to a dynamic saving-consumption problem with embedded discrete housing choice. Equilibrium prices equate total *inflows* into each market with *outflows* from each market. We study the price effect of interest rate and prepayment cost increases, approximated by a first-order Taylor expansion around a steady-state equilibrium.[11]

Proposition 2 extends the cross-sectional predictions of equation (2) to the case of multiple owner-occupied housing markets. Indexing housing market segments by j, we consider the regression $\Delta P_j = a + \beta \cdot \Delta PC_j + \epsilon_j$. In general, β may be positive or negative. But if prepayment shocks ΔPC_j are exogenous to local housing market fundamentals, then $\beta > 0$. Intuitively, markets with a greater increase in prepayment costs have a larger decrease in outflows, both to other markets and a numeraire rental market. This reduces the supply of available units and increases prices.

Proposition 2.2 shows how considering arbitrarily similar market segments does not attenuate β. The intuition is the same as for why $\frac{\partial \omega}{\partial \alpha} \geq 0$ in the simple framework. When markets are more similar, cross-market arbitrage from inflows leads to price compression and lower β. However, this is exactly offset by *reduced* cross-market arbitrage from *outflows*. Proposition 2.3 and 2.4 show β is higher in places where housing supply is less flexible or where prepayment cost has a larger effect on local outflows, with intuition similar to the simple framework.

Proposition 3 extends the aggregate predictions of equation (3). Specifically, we study the effects of an increase in interest rates and average prepayment cost on average owner-occupied house prices, relative to the numeraire rental market. As in the simple framework, there is a negative price effect due to a higher opportunity cost of capital that is attenuated by an increase in prepayment costs. The effect of average prepayment costs is *increasing* in β, when β is estimated using exogenous shocks to ΔPC_j (under mild conditions that rule out pathological reasons for increased β). Intuitively, β partly captures the price effects of lower outflows to the numeraire rental market, which is what is relevant for aggregates.

2 Data

We use several administrative datasets provided to the Federal Reserve system and public data.

[11] A steady-state equilibrium has constant interest rates, house prices, and population in each market segment.

Equifax Credit Risks Insight Servicing (CRISM). CRISM is a borrower-level monthly panel starting in 2005 that merges mortgage servicer data from ICE, McDash®, covering about two-thirds of residential installment loans, with credit bureau data on other consumer debts from Equifax. ICE, McDash®provides details about mortgage contract terms and property features that are unavailable in standard credit bureau data, such as mortgage interest rates, whether a loan is fixed or variable, appraised property value, origination loan-to-value ratios, and origination debt-to-income ratios. Fields from Equifax's credit bureau files add information on other mortgage and non-mortgage debt, as well as demographics such as household age, monthly zip codes of residence, and address changes.

We construct two datasets from CRISM: a borrower panel to study household behaviors, and a local housing market panel at the zip-code level to study market-wide outcomes. The borrower panel is based on a 10% sample of the full monthly panel. In this dataset, we infer existing-home sales and relate the decision to sell to mortgage contract features, borrower demographics, and the financial environment. The local housing market panel is derived from a 50% sample of year-end snapshots. In this dataset, we estimate average market-wide prepayment cost and relate it to local house price growth. Appendix Table C.6, panel (a) provides summary statistics; see Appendix B.1 for dataset construction details.

ICE, McDash®. To measure pre-2005 prepayment costs, we also extract a 0.5% sample from the full ICE, McDash®dataset. This provides a mortgage-level monthly panel with information on loan characteristics and repayment behavior. The panel covers 1992-present.

CoreLogic. To study housing market outcomes, we use property-level real estate public records data provided by CoreLogic. For most US residential properties, we observe the history of transactions dating back until around 2005, including purchase price, purchase date, and the identity of buyers and sellers. We also see limited mortgage information connected with properties, including mortgage balance, origination date, term, and whether the mortgage has a fixed or variable interest rate. CoreLogic also provides property characteristics such as the number of bedrooms and bathrooms. A separate dataset has information on building permits filed in connection with individual parcels.

We use these data to construct a dataset relating the sale price of a housing unit to the inferred characteristics of outstanding mortgages on the property. We also use building permits data to infer local construction of single-family units, multi-family units, and renovations. Appendix Table C.7 presents summary statistics, and Appendix B.2 provides details.

Other sources. We also use publicly-available data on local demographics, housing market characteristics, and construction, introduced where relevant and described in Appendix B.3.

2.1 Estimating prepayment costs.

Section 1 showed that the rate lock incentive depends on outstanding mortgages' prepayment costs: the present-value cost of repaying a mortgage at face value, rather than through fixed-interest installments. In this section, we empirically estimate "scaled prepayment costs": a mortgages' prepayment costs as a fraction of property value. Unlike "rate gaps" used in previous work, the measure is forward-looking and scales incentives by household wealth.

Define $r_{ft,q}$ as the q-period risk-free interest rate, r_{mt} as the average interest rate on a 30-year fixed-rate mortgage originated in period t, and Q as a risk-neutral measure. The "scaled prepayment cost" of mortgage i, secured by a property appraised at P_{ia}, depends on its balance M_i, fixed payments m_i, remaining term n_i, and origination spread Δ_i:

$$SPC_{it} := \left(M_i - E_t^Q \sum_{0 \leq q \leq n} \frac{m^*_{t+q}(M_i, m_i, n_i, \Delta_i)}{(1 + r_{ft,q})} \right) / P_{ia} \tag{4}$$

with m^*_{t+q} the uncertain $t+q$ payment. We assume $SPC_{it} = 0$ for adjustable-rate mortgages.[12]

The numerator in SPC_{it} is the gap between a mortgages' face value and its market value, equal to the present-value cost of future payments. Estimating market value is complex due to the option to prepay. In baseline analysis, we assume period-t households expect to hold their mortgage to maturity, but consider refinancing in the current period. The optimal refinancing threshold thus follows a break even rule, and:

$$E_t^Q \sum_{0 \leq q \leq n} \frac{m^*_{t+q}(\cdot)}{(1 + r_{ft,q})} := \min \left[E_t^Q \sum_{0 \leq q \leq n} \frac{m}{(1 + r_{ft,q})},\ c_R(M) + E_t^Q \sum_{0 \leq q \leq n} \frac{u(r_{mt} + \Delta, M, n)}{(1 + r_{ft,q})} \right] \tag{5}$$

This expression takes the minimum hold-to-maturity costs of the current mortgage (first term) and a rate refi mortgage (second term) with origination rate $r_{mt} + \Delta$, balance M, term n, and "updated" monthly payments $u(\cdot)$, accounting for closing costs $c_R(M)$. Following Agarwal *et al.* (2013) and Keys *et al.* (2016), we assume $c_R(M) := 2000 + 0.01 \cdot M$.[13]

The hold-to-maturity assumption has advantages and drawbacks. On the one hand, it delivers a straightforward optimal refinancing threshold, and by assuming m and $u(\cdot)$ are constant, it allows estimation of $E_t^Q(1 + r_{ft,q})^{-1}$ with the period-t price of a zero-coupon bond with maturity q. On the other hand, the assumption may over-state future payments by ignoring the future option to refinance and move. Appendix B.4 describes two alternatives for estimating market value of the option to prepay in future periods: (i) using calibrations of a model of optimal refinancing from Agarwal *et al.* (2013), and (ii) using mortgage-backed securities prices in the TBA market. We find similar results when using alternative measurement approaches.

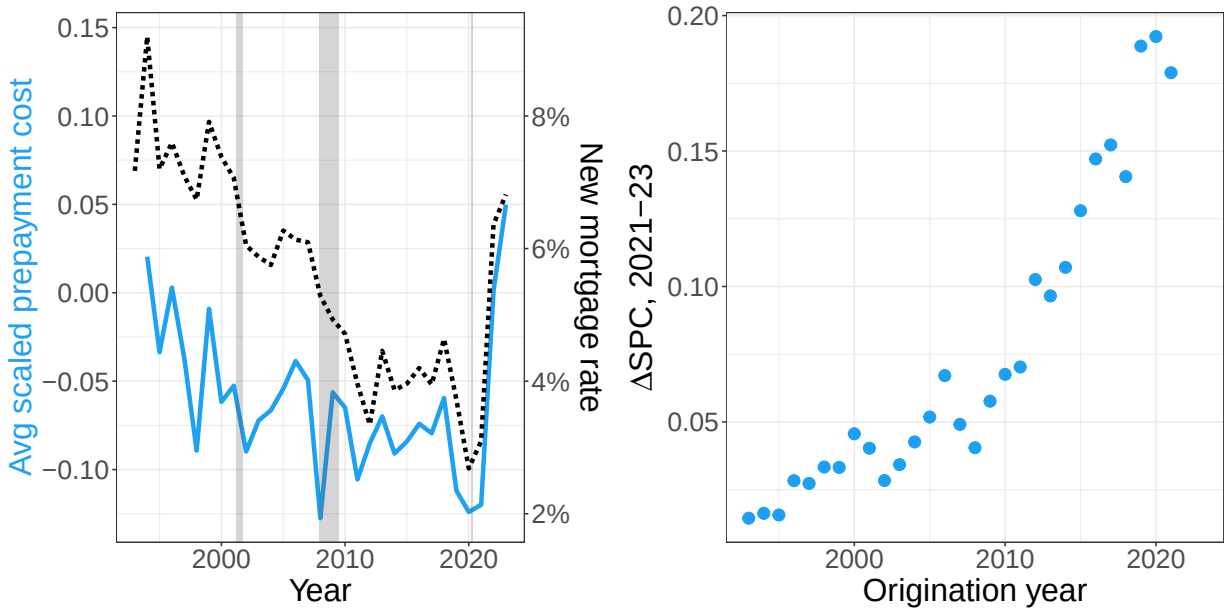
[12] This assumption abstracts from prepayment costs during fixation periods and between rate resets.

[13] We deflate the $2,000 fixed closing cost to 2010 dollars, so technically $c_R(\cdot)$ is indexed by time.

Scaled prepayment costs have intuitive comparative statics. For a given mortgage, an increase in rates after origination increases SPC_{it}, as the present-value cost of mortgage payments fall relative to the mortgage balance. This also implies that in a given year, SPC_{it} is higher for mortgages with lower mortgage rates. These differences are of a larger magnitude for mortgages with higher loan-to-value ratios and a longer remaining term.

Figure 1: Average scaled prepayment costs over time.

(a) By year. (b) By origination year.



Source: ICE, McDash® and PMMS. (a). The blue solid series (left scale) plots $SPC_t(I_{t-2})$ by t. The black dashed series (right scale) plots the Freddie Mac PMMS. Grey regions indicate NBER recessions. (b). The figure plots $SPC_{2023}(I_y) - SPC_{2021}(I_y)$ on the vertical axis against origination year y on the horizontal axis, where I_y is the set of mortgages originated in y.

Implementation and results. We estimate the time-varying zero coupon bond price using estimated parameters from the Fed's daily nominal yield curve (Gürkaynak *et al.* 2007). We estimate SPC_{it} for each mortgage and year t in CRISM and ICE, McDash®. We use the former to estimate the distribution of scaled prepayment cost at the zip code level, and the latter to estimate scaled prepayment cost in the time series stretching back to 1992. Define the year-t average scaled prepayment cost for a set of mortgages I as $SPC_t(I) := |I|^{-1} \sum_{i \in I} SPC_{it}$.

Figure 1 shows how aggregate scaled prepayment cost varies over time. We study effects of the initial mortgage stock on subsequent equilibrium outcomes, with a focus on two-year price growth. Therefore, we examine the initial mortgage stock with a two-year lag, before price growth happened. Panel (a) plots $SPC_t(I_{t-2})$, where I_{t-2} is the set of outstanding mortgages two years before t,[14] along with the average new prime 30-year fixed mortgage rate

[14] Appendix Figure B.1 compares results from I_{t-2} to I_{t-1} and I_t.

from the Freddie Mac Primary Mortgage Market Survey (PMMS). Prepayment costs are mostly stable until 2021. Tightening from 2021-23 causes the largest year-over-year increase in scaled prepayment costs.

Panel (b) shows that the year of mortgage origination is an important driver of the 2021-23 increase in scaled prepayment cost. The horizontal axis groups mortgages by origination year y, and the vertical axis plots $SPC_{2023}(I_y) - SPC_{2021}(I_y)$, where I_y is the set of mortgages originated in year y. Prepayment costs rise from 2021-23 for all origination cohorts, but the increase is greatest for mortgages originated in the 2010s. Older cohorts have higher mortgage rates at origination, and hence have an in-the-money refinancing option given low rates in 2021. Their prepayment costs are thus unaffected by much of the rate increase from 2021-23. Recent cohorts have low mortgage rates at origination and so are affected by the full rate increase. Variation in origination timing creates large differences in the increased financial cost of giving up a low-rate mortgage from 2021-23. An owner with a mortgage originated in 2016 would see an increase in prepayment costs equal to 15% of home value (around $47,000 for the typical home), whereas an owner with a mortgage originated in 2001 would see an increase of only around 5% (around $16,000).

3 Effect of rate lock on local housing market equilibrium

This section estimates the pass-through of prepayment costs into equilibrium house prices, outflows, and residential investment, using variation across local housing markets over the 2021-23 financial tightening.

3.1 Estimation framework.

We define local housing markets as zip codes indexed by j. Motivated by a first-order Taylor expansion of the market-clearing equation (1) from Section 1, we estimate the following cross-sectional specification, where changes are from the end of 2021 to the end of 2023:

$$\%\Delta P_j = \beta \cdot \Delta SPC_j + \eta'_{c(j)} \cdot X_j + \omega_{c(j)} + \epsilon_j \tag{6}$$

where $\%\Delta P_j$ is house price growth; ΔSPC_j is the change in average scaled prepayment cost from 2021-23; X_j is a vector of observable housing market characteristics in 2021, including demographics and price-rent ratios, and $\omega_{c(j)}$ is a CBSA fixed effect.

The coefficient of interest is β: the impact of an increase in scaled prepayment cost on price growth from 2021-23. As discussed in Section 1, higher β implies that aggregate prepayment costs have a larger positive impact on average house prices, relative to rents.

Specification (6) has four key features that help isolate the price impacts of rate lock. First, we focus on the period from the end of 2021 to the end of 2023, because this is the period

with the largest prepayment cost shock in our data. Second, we estimate the impact of prepayment costs in *changes*, so that β picks up variation in exposure to the aggregate shock, rather than persistent housing market characteristics. Third, we include CBSA fixed-effects to absorb confounds such as local labor market shocks that may both explain moves during low-rate periods and price growth.[15] And finally, we measure prepayment costs at a lag relative to prices, to capture how the *initial* mortgage stock impacts *subsequent* price growth. Where I_{jt} is the set of mortgages outstanding in zipcode j at the end of year t, we estimate $\Delta SPC_j := SPC_{2023}(I_{j,2021}) - SPC_{2021}(I_{j,2019})$. The lag is crucial: originations in year t shade average SPC for mortgages outstanding in t towards zero, biasing estimates of β.

We estimate $\%\Delta P_j$ with the zip-level FHFA repeat-sales house price index. This ensures we estimate effects on average property value, not on the composition of properties sold. The index is equal-weighted, which aligns with our equal-weighted ΔSPC_j estimate.

Need for instruments. The estimate of β in equation (6) with CBSA fixed effects, $\hat{\beta}_{FE}$, is likely biased. Within-CBSA variation in $SPC_{j,2023}$ drives variation in ΔSPC_j, because as explained in Section 2.1, most outstanding mortgages had in-the-money refi options in 2021. $SPC_{j,2023}$ is higher in places where outstanding mortgages have high leverage and low mortgage rates: that is, where mortgages have low origination spreads, or where origination volumes are high during periods with low long-term rates. However, leverage, origination spreads, and origination timing all likely correlate with unobserved drivers of local house price growth.

The estimate of β could be biased upwards or downwards. It could be biased upwards if, for example, people take advantage of low aggregate mortgage rates to move to places with unobserved positive amenity shocks, or places where they expect house price appreciation. This creates spurious positive correlation between increased prepayment costs and price growth.

The estimate could also be biased downwards for several reasons. First, outstanding mortgages will have higher mortgage spreads, and hence lower prepayment costs, in markets that are attractive to younger and more financially-precarious borrowers. This demand makes these markets *more* exposed to aggregate housing demand and credit supply growth that increase prices (e.g., Landvoigt *et al.* 2015). Over 2021-23, there were positive shocks to aggregate housing demand and credit supply, (Delgado and Gravelle 2023, Gamber *et al.* 2023), implying downwards bias.

Second, and more perniciously, places with high origination flow during low-rate periods, and hence higher prepayment costs, likely have more elastic housing supply. Lower rates increase housing demand, and additional demand causes people to move where units are avail-

[15] Restricting to within-CBSA variation: (i) removes variation from cross-CBSA moves, which rate lock does not appear to affect (see Section 4); and (ii) does not attenuate estimates of β, as discussed in Section 1.

able. Units are available in places where housing supply is more price elastic, such as suburban markets. These markets have lower price growth, causing downwards bias.

Finally, places with high prepayment costs are more exposed to negative demand shocks due to Covid-19. In high-density, urban housing markets, the typical mortgage was recently-originated, because urban markets have high churn (Henning-Smith *et al.* 2023). But for the 2021 outstanding mortgage stock, a recently-originated mortgage has higher prepayment costs, due to falling mortgage rates over the prior decade. Urban markets suffered negative demand shocks due to Covid-19, leading to lower price growth over 2021-23 (Frost 2023).

Validating these concerns about bias, ΔSPC_j correlates with many demographics that are important for prices. The top panel of Figure 2 reports t-statistics from a regression of ΔSPC_j on local housing market demographics, and shows correlation with characteristics like such as population density, age, and income.

IV strategy. Section 3.2 presents benchmark instruments that address endogeneity from spreads and leverage, but that still use endogenous variation due to origination timing. We then use three strategies to address the endogeneity of origination timing. Section 3.3, describes two instruments that use exogenous shocks to predict moves. Section 3.4 describes an approach for controlling for origination timing directly. The three approaches produce similar estimates, all larger than the benchmarks that are endogenous to origination timing.

3.2 Benchmark: Addressing endogeneity from spreads and leverage.

The first instrument, $\Delta SPC_j^{\text{Avg rate x LTV}}$, strips ΔSPC_j of variation from borrower mortgage origination spreads. For this instrument, variation depends only on borrower leverage and the average mortgage rate at origination. For each outstanding mortgage in 2019 and 2021, we set mortgage spread $\Delta_i = 0$ and re-calculate remaining mortgage balance M_i and time remaining n_i, fixing origination mortgage balance (see Appendix Section D.2 for the formula).

The second instrument, $\Delta SPC_j^{\text{Avg rate}} := SPC_{j,2023}^{\text{Avg rate}} - SPC_{j,2021}^{\text{Avg rate}}$, only uses variation in origination timing, and is constructed to capture nonlinearities in prepayment costs. We calculate: $SPC_{jt}^{\text{Avg rate}} = \sum_{\tau \leq t} w_{j\tau,t} \cdot SPC_{\tau,t}$, where $w_{j\tau,t}$ is the share of mortgages outstanding as of year $t-2$ in j that were originated in year-month τ; and $SPC_{\tau,t}$ is the scaled prepayment cost of a synthetic loan outstanding as of year $t-2$ and originated in year-month τ with average characteristics. The synthetic loan has mortgage rate $r_{m\tau}$, and origination LTV and time remaining equal to the national average for outstanding mortgages at $t-2$.

Variation in $\Delta SPC_j^{\text{Avg rate}}$ comes from variation in "origination timing shares" $w_{j\tau,t}$: places with more originations during periods with low mortgage rates have higher prepayment costs when market rates rise. This is similar to the household-level mobility instruments in Liebersohn and Rothstein (2024), Fonseca and Liu (2024), and Batzer *et al.* (2024). However, the

instrument does not appear balanced on demographics at the *market level*. The second panel of Figure 2 shows that $\Delta SPC_j^{\text{Avg rate}}$ has many of the same correlations as the endogenous variable. This suggests that origination timing, which is the key driver of variation, may relate to unobserved local housing market characteristics that cause price growth.

3.3 Addressing origination timing endogeneity by predicting moves.

We construct market-level origination timing instruments by identifying shocks that cause household moves but are otherwise unrelated to 2021-23 price growth. We use these shocks to predict local housing market churn in each year before 2021. Markets where these shocks cause a high *flow* of households to move when market mortgage rates are low accumulates to a large *stock* of households with low outstanding mortgage rates when rates rise in 2021. These markets are exogenously exposed to higher prepayment cost growth.

To construct instruments, $\Delta SPC_j^S := SPC_{j,2023}^S - SPC_{j,2021}^S$, using shock S, we *predict* origination timing shares $w_{j\tau,t}^S$, rather than using *endogenous* $w_{j\tau,t}$. The shares $w_{j\tau,t}^S$ predict the share of mortgages outstanding in year $t-2$ in market j that were originated in year τ due to churn from past shocks. Using $w_{j\tau,t}^S$, we predict period-t scaled prepayment cost due to shock-driven churn: $SPC_{jt}^S := \sum_{\tau \le t} w_{j\tau,t}^S \cdot SPC_{\tau,t}$.

We estimate $w_{j\tau,t}^S$ in three steps. First, we estimate the impact of shocks to households in year τ' on moves in subsequent years $\tau \ge \tau'$. Second, using these estimates, we predict the year-τ churn in each market j caused by the sequence of shocks in past years $\tau' \le \tau$. Third, we estimate $w_{j\tau,t}^S$ based on how churn in year τ will impact the mortgage distribution $t \ge \tau$. Due to moves and refinancing, the mortgage distribution in t depends more on recent τ; we therefore down-weight churn in τ as $t-\tau$ grows.

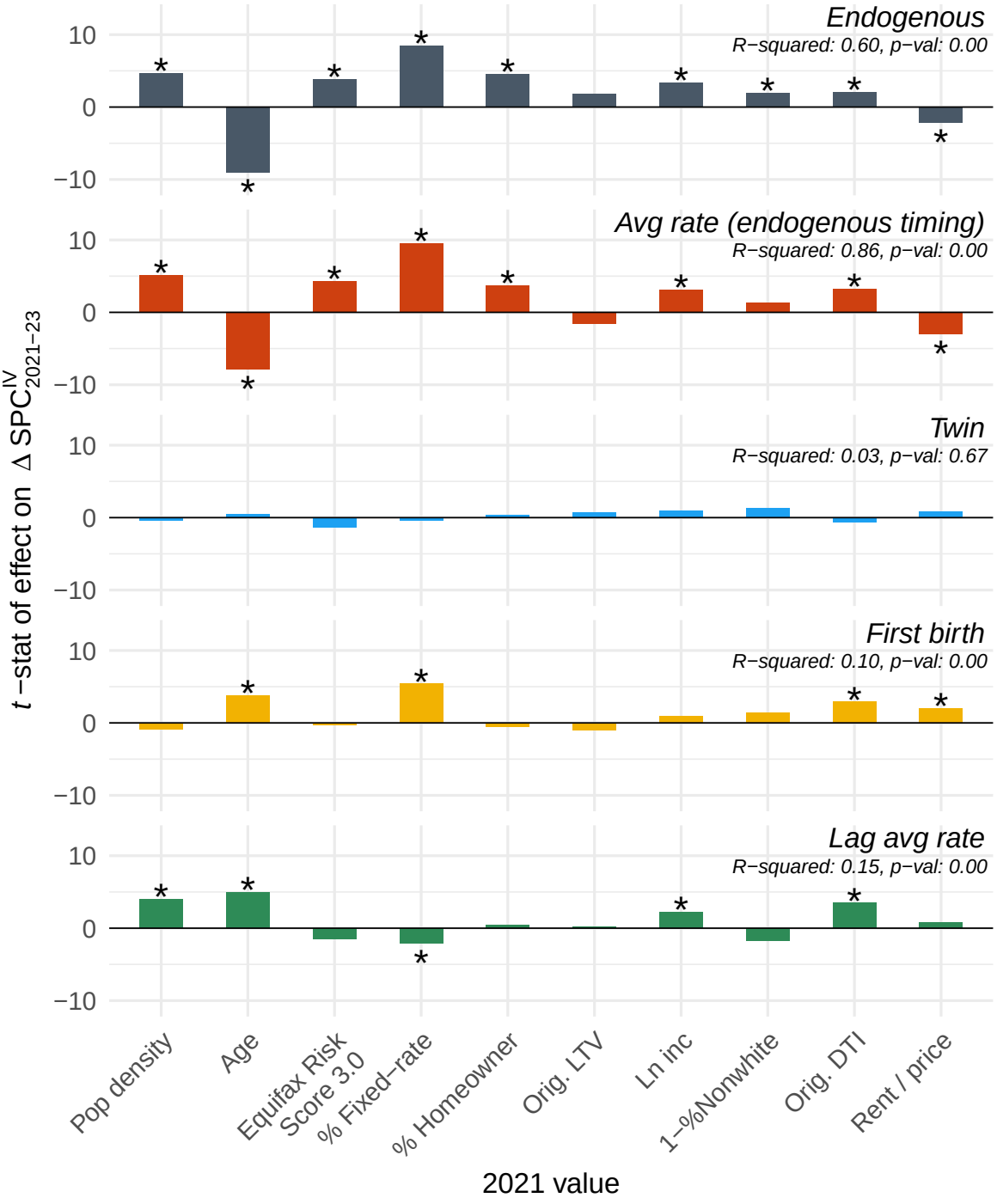
Formally, define the shock experienced in year τ' by households in j as $S_{j\tau'}$; predicted churn in year τ from a linear combination of past shocks as $\nu_{j\tau} := \nu_\tau^S\left(\{S_{j\tau'}\}_{\tau'<\tau}\right)$; and the a-year mortgage survival probability as $K(a)$. We estimate $w_{j\tau,t}^S := \frac{K(t-\tau)\cdot \nu_{j\tau}^S}{\sum_{\tau \le t} K(t-\tau)\cdot \nu_{j\tau}^S}$.

The instrument's variation comes from the path of *changes* in the shock. The identifying assumption is therefore an exclusion restriction: *changes* in the shock are conditionally mean independent of unobserved drivers of house price growth from 2021-23. Changes, rather than levels, drive variation because we effectively *correlate* predicted moves in each year with average prepayment costs of mortgages originated in each year.[16] Formally:

Proposition 1. *The two-stage least squares estimate of β in equation (6) using ΔSPC_j^S is consistent if $\ln S_{j\tau} - \ln S_{j,\tau-1}$ is mean-independent of ϵ_j, conditional on $\omega_{c(j)}$, X_j.*

[16] The definition of $w_{j\tau,t}^S$ makes this clear: level differences in S scale all $\nu_{j\tau}^S$ proportionately in the numerator and denominator, which cancel out.

Figure 2: Instrument balance on area demographics.



Source: Equifax Credit Risks Insight Servicing and ICE, McDash®. The figure reports the t-statistics on coefficient estimates from a multiple regression of ΔSPC_j and instruments for ΔSPC_j on local area demographics: see Appendix Table D.20 for all coefficient estimates. Stars indicate statistical significance at a 5% threshold. The outcome variable varies by panel, and is, respectively: $\Delta SPC_j, \Delta SPC_j^{\text{Avg rate}}, \Delta SPC_j^{\text{Twin}}, \Delta SPC_j^{\text{First birth}}$, and $\Delta SPC_j^{\text{Lag avg rate}}$. All specifications include CBSA fixed effects and controls for 2019 demographics. The specification with $\Delta SPC_j^{\text{Lag avg rate}}$ includes controls for the fraction of 2021 mortgages originated in each pre-2019 half-year bin, as described in Section 3.4. Each panel reports the within R-squared for all regressors, as well as the p-value of the Wald test of joint nullity of the 2021 regressors plotted in the figure.

As shown in Appendix Section C.1, having kids is a major driver of household moves: a first child causes moves on impact, while additional children cause moves with a lag. We form origination timing instruments using two shocks to the number of kids in a household.

We first use *unexpected* family size shocks from having twins (hereafter, the "twins" instrument). Around 3% of births are twins, so around 6% of families ever have twins.[17] Below, we find that having twins locks in a different pattern of moves over childhood than having one child, and that within a CBSA, the *change* in the fraction of births that are twins is consistent with finite sample variation. Changes in the twin birth rate thus cause quasi-random variation in moves, which accumulate over decades to impact the 2021 outstanding mortgage stock.

We also consider moves due to family formation, proxied by the number of births to first-time mothers (hereafter, the "first births" instrument), conditional on the local population and age distribution. This instrument has higher power, but raises identification concerns, because economic conditions impact family formation and fertility (Autor *et al.* 2019).

Moves driven by twin births or family formation could have direct housing market effects. We therefore construct instruments *only* with pre-2006 variation in the twin birth rate and number of first births. This also alleviates identification concerns, as confounds driving pre-2006 variation are unlikely to correlate with 2021-23 house price growth.

3.3.1 Moving shock 1. Twin births.

We first construct an instrument for ΔSPC_j using the moving shock $p_{j\tau}$, the fraction of births in j that are twins in year τ.

Relevance. We first estimate the effect of $p_{j\tau}$ on moves. For a family, having twins rather than one child is a large family size shock that causes moves around birth and school transitions. An especially large twin birth cohort therefore generates greater market churn around these ages. We model the effects of changes in twin birth rates on housing market churn as:

$$\Pr(Move)_{jt} = \alpha_j + \omega_{c(j),t} + \sum_{a \in A} \beta_a p_{j,t-a} + \varepsilon_{jt} \tag{7}$$

$\Pr(Move)_{jt}$ is the year-t move rate in j, α_j is a zip code fixed effect, and $\omega_{c(j),t}$ is a year-by-CBSA effect. The coefficient β_a is the period-t effect on moves of the twin birth rate a years prior, when a twin birth cohort is aged a.

Figure 3a plots first-difference estimates of β_a from equation (7).[18] Lead effects are statistically zero, implying twin births are unexpected. Having twins increases moves when twins are

[17] The typical woman has 2.1 children by the end of childbearing age (Human Fertility Database 2025), implying around $100\% - 97\%^{2.1} = 6.2\%$ of women eventually have twins.

[18] Public data on twin births from the CDC is available by county. We use data from 1995-2006. Appendix Section B.3 describes our mapping from counties to zip codes.

born and parents realize their family is larger than expected; when twins enter kindergarten at age 5; and when twins turn 9, around the start of middle school. The effect dissipates after 15 years, once twins are in high school and less likely to influence moves directly. To avoid direct effects of twin births on housing markets, we restrict to pre-2006 twin birth rate variation.

Estimates reflect not just moves by households with twins, but the broader market *churn* as these moves generate vacancies filled by households creating additional vacancies. Appendix Section D.7.1 shows that coefficient magnitudes align with the estimated effect of additional kids on moves and recent estimates of vacancy spillover effects by French and Gilbert (2023) and Anenberg and Ringo (2022).

We follow the above steps to construct the instrument, visualized in Appendix Figure D.17. We predict twin-driven moves in year τ by projecting the history of twin birth rates on coefficient estimates from equation (7). Using the results, we estimate $w^{\text{twin}}_{j\tau,t}$: the twin-driven probability that a mortgage outstanding in t was originated in τ.[19] The instrument is $\Delta SPC^{\text{twin}}_j := SPC^{\text{twin}}_{j,2023} - SPC^{\text{twin}}_{j,2021}$, where $SPC^{\text{twin}}_{jt} := \sum_{\tau \leq t} w^{\text{twin}}_{j\tau,t} \cdot SPC_{\tau,t}$.

The instrument $\Delta SPC^{\text{twin}}_j$ predicts actual scaled prepayment costs. Figure 3b shows the positive relationship between ΔSPC_j and $\Delta SPC^{\text{twin}}_j$, residualizing out CBSA fixed effects and baseline controls. The instrument is strong, with a first-stage F-stat of 75.

The instrument is strong because twin-driven origination timing shares, $w^{\text{twin}}_{j\tau,t}$, predict actual origination timing, which is a key driver of outstanding mortgage rates. Appendix Table D.22 shows in column (1) that twin-driven origination shares, $w^{\text{twin}}_{j\tau,t}$, predict actual origination shares, $w_{j\tau,t}$ (R-squared = 1.3%); and in column (2) that the twin-driven average outstanding mortgage rate in 2021 predicts the actual average outstanding mortgage rate in 2021.

Exclusion. From Proposition 1, the identifying assumption is that the pre-2006 change in the local twin birth rate is mean-independent of unobserved drivers of 2021-23 house price growth, conditional on CBSA.[20]

Consistent with this assumption, the third panel in Figure 2 shows that $\Delta SPC^{\text{twin}}_j$ is uncorrelated with observable local housing market characteristics.

The instrument does not correlate with demographics because the within-CBSA year-to-year change in the twin birth rate can be explained by finite sample variation. Whether a maternity produces twins is stochastic, so otherwise identical markets will have different *realized* twin birth rates each year. The law of large numbers means this variation falls with

[19] Formally, $w^{\text{twin}}_{j\tau,t} := \frac{K(t-\tau)\nu^{\text{twin}}_{j\tau}}{\sum_{\tau<t} K(t-\tau)\nu^{\text{twin}}_{j\tau}}$, where $\nu^{\text{twin}}_{j\tau} = \sum_{\tau' \leq 2006} \hat{\beta}_{\tau-\tau'} \cdot p_{j\tau'}$. $K(\cdot)$ is estimated as the pooled probability that a mortgage originated a years prior remains outstanding using ICE, McDash® (see Appendix Section D.5.3). To improve power, we parameterize β_a as a fifth-order polynomial with fixed effects for years 1, 5, and 9 (see Appendix Section D.5.2). Appendix D.9.3 shows results are similar to using estimates from Figure 3a directly.

[20] Note that $\ln p_{j\tau} - \ln p_{j,\tau-1} \approx p_{j\tau} - 1 - (p_{j,\tau-1} - 1) = p_{j\tau} - p_{j,\tau-1}$.

the number of births. But the variation is still meaningful because even markets with a large *population* have a relatively small number of *births*.

There is considerable year-to-year variation in the twin birth rate. The blue solid line in Figure 3c plots the distribution of the year-to-year change in the twin birth rate. The standard deviation is 0.64%. The variation exactly aligns with expectations due to sampling variation. The red dashed line plots a distribution based on simulations from a theoretical benchmark that satisfies Proposition 1, where county-level twin births are iid draws from CBSA-by-year specific binomial distributions.[21] Empirical and simulated distributions overlap almost exactly (two-sample Kolmogorov-Smirnov test p-value=0.34).

Variation declines in larger markets, as expected from the law of large numbers. Figure 3d groups county-years into deciles based on number of births, and plots the standard deviation of the change in the twin birth rate, both in the data (solid blue) and based on theoretical finite sample variation from the above benchmark (hollow red). Larger markets have less variation, but the level and rate of decline in the data aligns with expected finite sample variation.[22]

As further evidence, within-CBSA twin birth rate changes are not predictable from local housing market or household characteristics, such as average income, age, and property value. Appendix Table D.21 regresses the level and change in the twin birth rate on local characteristics, and finds they are jointly insignificant (p-value of Wald test=0.91).

These findings align with the public health literature about drivers of twin births. Twin births are random for the 25-30% of twin births that are monozygotic ("identical").[23] Remaining twin births are dizygotic ("fraternal"), and naturally occur due to excess secretion of follicle stimulating hormone (FSH) controlling the release of eggs. Fraternal twinning runs in families due to heritability of FSH secretory drive: sisters and daughters of twins are twice as likely to have twins. Fraternal twinning rises with age as follicle density declines (Hoekstra *et al.* 2008).

Fraternal twins can also result from in-vitro fertilization (IVF) if multiple embryos are transferred. However, fewer than 1.2% of births use IVF in our study period [24] and IVF explains under 0.13% of twin birth rate variation across markets.[25] Over 1995-2006, the twin birth rate rose by 0.08pp per year, which Hoekstra *et al.* (2008) and Tandberg *et al.* (2007) attribute to

[21] See Appendix Section D.4, Example 1 to show this satisfies Prop 1.

[22] Theoretical variance is slightly higher than the data for two reasons. First, the CBSA-by-year twin birth rate is estimated, slightly increasing dispersion. Second, as shown in Table D.21, column (3), the twin birth rate is mean-reverting. This makes biological sense: if genetic drivers of twinning are iid across the population, then the remaining population has a low twinning rate in a year after an unusually high number of high-twinning mothers gives birth. Appendix Section D.9.2 presents mean reversion-robust estimates.

[23] "There are no clear associations between m[ono]z[ygotic] twinning and maternal, environmental or genetic factors and the mechanisms have not been identified" (Hoekstra *et al.* 2008).

[24] IVF rates were 0.5% in 1995 (Guyer *et al.* 1997, Wright *et al.* 2006) and 1.2% in 2004 (CDC 2004).

[25] This could be because the CDC recommends single embryo transfer (Practice Committee of the Society for Assisted Reproductive Technology 2006). See Appendix Section D.5.1 for details.

healthier and older mothers (e.g., decline in teen pregnancies) rather than IVF.

Two robustness exercises described in Appendix Section D.9 verify that variation in maternal age and IVF do not drive our results. First, we exclude twin births to older women, which strips out 75% of IVF births. Second, we control for the local change in the expected twin birth rate directly, which we prove ensures only finite sample variation drives estimates.

Discussion. Does our methodology correctly capture the impact of twin birth rates on prepayment costs? Appendix D.7.3 verifies this via an over-identification test. We regress the twin birth rate history on $\Delta SPC_j^{\text{twin}}$ and ΔSPC_j directly. If the instrument's model of mortgage origination is valid, then coefficients on each lag of the twin birth rate should align across specifications. A bootstrapped J-test rejects that coefficients are different (p-value=0.95).

Why are twin-driven moves so relevant to the 2021 mortgage stock? In a given year, cross-market variation in twin-driven moves produces small differences in the *flow* of new originations. Crucially, the outstanding *stock* many years later depends on how these small differences *accumulate* over many years. Appendix Figure D.19 plots the difference in $w_{j\tau,2023}^{\text{twin}}$ between zip codes with $SPC_{j,2023}^{\text{twin}}$ 1.96 standard deviations above and below the average. In a given year, the difference in origination probability is at most 0.1%. Over 25 years, twin-driven moves reshuffle origination timing for 1.7% of households.

Is variation in $\Delta SPC_j^{\text{twin}}$ large enough to plausibly impact prices? Previous research on housing supply and speculative activity shows that variation impacting 1% of housing units moves equilibrium prices.[26] Appendix Figure D.18 shows that instrument variation is driven by changes in the fraction of households with a rate lock incentive that is very strong (mortgage rate around 4%) and very weak (mortgage rate above 6%). Combining these two results, variation in the implied twin-driven average outstanding mortgage rate spans about 15 basis points across zip codes in 2021 (see Appendix Figure D.20a), similar in magnitude as previous research estimating the equilibrium price effects of interest rate variation.[27]

3.3.2 Moving shock 2. First births.

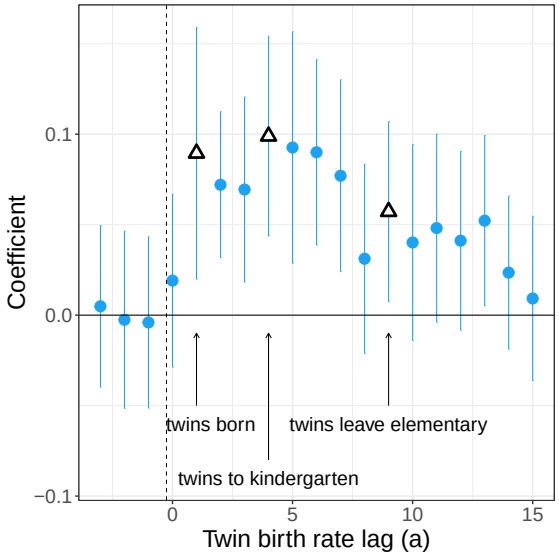
We next construct an instrument for ΔSPC_j using the number of births to first-time mothers. We give an overview below, and present details in Appendix D.8.

[26] Mian and Sufi (2022) find that activity by the 1.4% of the population that are housing speculators explained 100% of increased lending financed by non-core deposits in the lead-up to the 2008 housing crisis, which caused a 19.1% increase in MSA-level prices.Moreover, single-family housing unit completions are typically 0.5-1% of the residential housing stock (`COMPU1USA/ ETOTALUSQ176N` from the St. Louis Fed), and variation in supply elasticities across geography affects house prices (e.g., Baum-Snow and Han 2024).

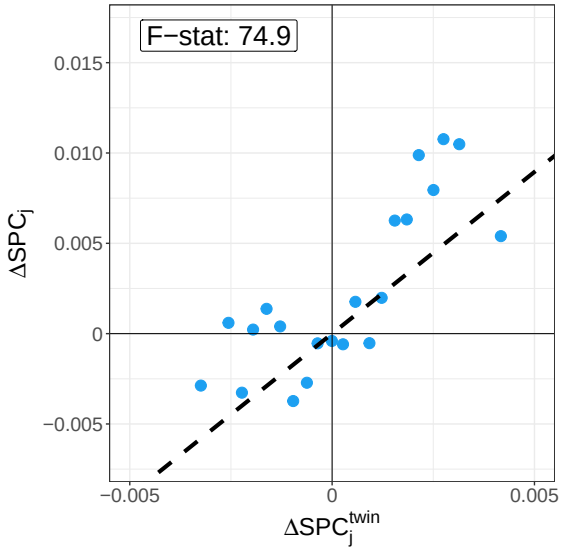
[27] Adelino *et al.* (2025) identify price effects from a change in average borrowing costs of 6-8 basis points at the GSE non-conforming loan threshold. Loutskina and Strahan (2015) estimate that a 1pp increase the fraction of local borrowers receiving this 6-8 bps mortgage subsidy causes 0.5% higher price growth at the CBSA level.

Figure 3: Twins instrument: Exclusion and relevance

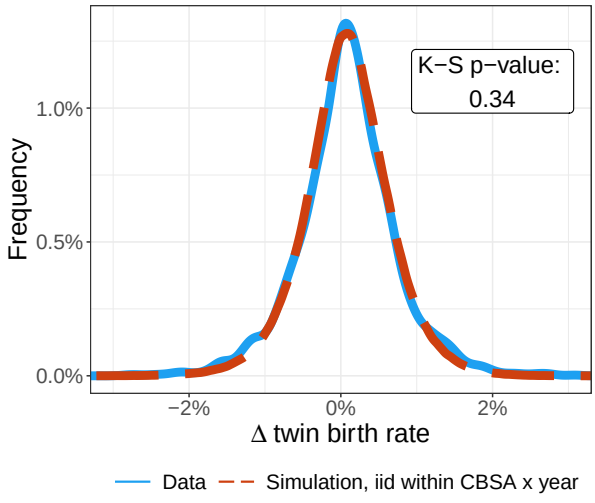
(a) Δ twin birth rate predicts moves.



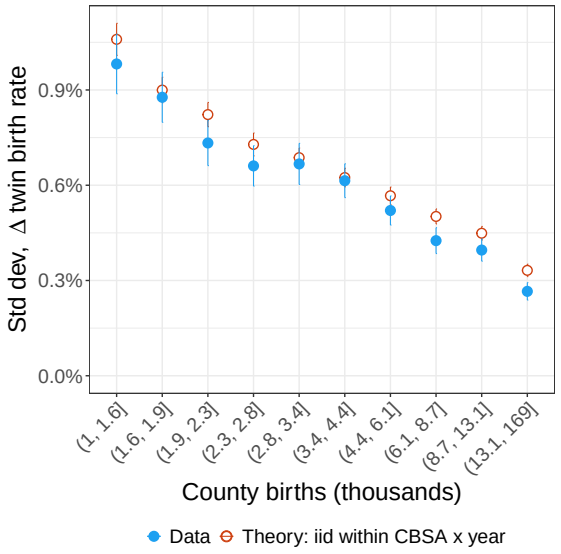
(b) Twin instrument predicts ΔSPC.



(c) Distribution of Δ twin birth rate.



(d) Δ twin birth rate by population.



Source: Author's calculations using CDC natality data from 1995-2006, Equifax Credit Risks Insight Servicing, and ICE, McDash®. (a) The figure plots $\hat{\beta}$ from specification (7), estimated in first differences. Spikes plot 95% confidence intervals calculated from standard errors clustered by county and calendar year. (b) The figure is a binscatter of ΔSPC_j against $\Delta SPC_j^{\text{twin}}$, residualized using CBSA fixed effects and CBSA-specific slopes for the 2021 rent-price ratio and the 2021-23 change in rents over 2021 price. The black dashed line shows the slope. The box shows the projected F-stat assuming county-level clustering. (c) The figure plots the distribution of the year-over-year change in the twin birth rate (blue solid) and 500 simulations that assume twin births are iid draws from CBSA-by-year binomial distributions. We restrict to markets in the estimation sample for equation (6). A two-sample Kolmogorov-Smirnov test fails to reject the null that observations are drawn from the same distribution at the 5% level (p-value = 0.34). (d) The figure plots the standard deviation of year-over-year twin birth rate change in percent within county birth deciles, in the data (blue solid) and assuming twin births are iid draws from CBSA-by-year binomial distributions (red hollow). Spikes plot 95% confidence intervals calculated using the Delta method (for data) or with 500 bootstraps (for theory). Appendix Section D.6 derives the formula for the benchmark.

Relevance. We first estimate the effect of first births on moves. We control for the local age distribution, because the number of first births depends on the ages of local residents, which independently affect moves. Where B_{jt} is the number of first births, Pop_{jt} the local population, and $\Pr(Age_{it} = l | i \in j)$ the fraction of households in zip j of age l in year t, we estimate:[28]

$$Moves_{jt} = \alpha_j + \omega_{c(j),t} + \sum_{a \in A} \pi_a B_{j,t-a} + \delta_t Pop_{jt} + \sum_{l \in L} \pi_{lt} \Pr(Age_{it} = l | i \in jt) + \varepsilon_{jt} \quad (8)$$

Appendix Figure D.14 shows fixed effects estimates of π_a. There are positive lead effects, suggesting couples move expecting a first child. Moves increase in the year after childbirth, before kindergarten, and around early elementary school, and die out after 15 years.

Using coefficient estimates from equation (8) and the local history of the number of first births, we predict moves in year τ.[29] With these results, we estimate $w^{\text{births}}_{j\tau,t}$ and $\Delta SPC^{\text{births}}_{j}$ as with the twins instrument.

Exclusion. In the framework of Proposition 1, the exclusion restriction is that the pre-2006 change in first births, as a fraction of average moves, is mean independent of ε_j (see Appendix Section D.8.1). There are a few reasons to think this assumption is reasonable. First, we restrict to pre-2006 variation, so even if the *current* economic environment impacts *current* fertility, it is less likely that these effects persist to affect 2021-23 house price growth. Second, over 40% of pregnancies are unplanned, and hence likely unrelated to local conditions (Rossen *et al.* 2023). And finally, although fertility decisions relate to individual economic circumstances, they may have a weaker relationship with aggregate economic conditions. Autor *et al.* (2019) find that negative shocks to men's employment *lower* fertility, but shocks to women's employment *raise* fertility. For a local housing market, positive and negative shocks might roughly offset.

We evaluate the identifying assumption by regressing $\Delta SPC^{\text{births}}_{j}$ on local demographics, and plot coefficient estimates in the fourth panel of Figure 2. The first birth instrument is much more balanced on demographics than the average rate benchmark, but has statistically significant loadings on covariates such as age and debt-to-income ratios.

Although the first birth instrument raises some identification concerns relative to the twins instrument, it uses more variation, and hence has more power. To understand how much variation there is in $\Delta SPC^{\text{births}}_{j}$, we predict a first-birth driven outstanding mortgage rate in 2021: $\sum_{\tau \leq 2021} w^{\text{births}}_{j\tau,2023} \cdot r_{m\tau}$. Appendix Figure D.21a shows its distribution, which spans roughly a percentage point, compared to 15 basis points for the twins instrument. Appendix Table D.23

[28] We estimate in levels, not rates, because the number of first births mechanically affects the local population.

[29] Formally, we predict moves as: $\nu^{\text{births}}_{j\tau} = \sum_{\tau' \leq 2006} \hat{\pi}_{\tau-\tau'} B_{j\tau'} + \overline{Moves}_j$, where $\overline{Moves}_j$ is the 2006-2019 zip-level average. The average re-centers level differences across location. Given $\nu^{\text{births}}_{j\tau}$, we form w^{births}_{jt} as with the twins instrument. As with the twins instrument, we slightly increase power by estimating π_a as a high-dimensional polynomial with indicators at meaningful milestones. See Appendix Section D.8.2 for details.

regresses the actual average outstanding interest rate in 2021 against its first-birth predicted counterpart. The within-CBSA R-squared is 6.2%, compared to 1.3% for the twins instrument.

3.4 Controlling for origination timing.

Our final approach estimates equation (6) by removing the sources of origination timing variation that are most likely to be endogenous. First, we instrument for ΔSPC_j using a "lagged average rate" instrument, ΔSPC_j^L, that only uses pre-2019 variation in origination timing. This removes confounds from the correlation between recent demand shocks and outstanding mortgage rates. Additionally, we control for the fraction of outstanding mortgages in 2021 that were originated in each half-year intervals h from 2002-2019. This ensures we only use variation across markets in the fraction of mortgages originated in different year-months *within* relatively narrow time intervals.

ΔSPC_j^L is constructed in the same way as $\Delta SPC_j^{\text{Avg rate}}$, except we use "lagged" origination shares $w_{j\tau,t}^L$ in place of $w_{j\tau,t}$. To estimate $w_{j\tau,2023}^L$, for each j and $\tau' \leq 2019$, we calculate the probability that a mortgage outstanding at the end of year τ' was originated in each of the previous 24 year-months. With these estimates, we predict the probability that a mortgage outstanding at the end of 2021 was originated in each year-month from 2019-2021. We calculate $w_{j\tau,2023}^L$ using actual origination shares for τ in pre-2019 year-months, and predicted origination shares for τ in 2019-2021, re-normalizing so shares sum to one. To estimate $w_{j\tau,2019}^L$, we follow the same procedure shifted backwards two years.

ΔSPC_j^L still uses variation from pre-2019 origination timing. We thus add controls for the the fraction of outstanding mortgages in 2021 originated in each pre-2019 half-year h: $\left\{\zeta_h \cdot \sum_{\tau:\, h(\tau)=h} w_{j\tau,2023}\right\}_h$. Appendix Section D.3 shows that this approach accounts for an unobserved origination half-year fixed effect for property-level price growth.

The identifying assumption is that pre-2019 variation in the share of mortgages originated within six-month intervals is uncorrelated with unobserved drivers of 2021-23 house price growth. To evaluate this assumption, the bottom panel of Figure 2, plots t-statistics from a regression of ΔSPC_j^L on local demographics, residualizing out origination timing controls. As with the first births instrument, the lagged average rate instrument is more balanced than $\Delta SPC_j^{\text{Avg rate}}$, but is less balanced than the twins instrument, due to correlation with demographics like population density and debt-to-income ratios.

3.5 Main Results.

We now present IV estimates of the effect of scaled prepayment cost on house price growth and moves out of local housing markets.

House price growth. Panel (a) of Table 1 presents estimates of β in equation (6) using fixed effects (column (1)) and the benchmark instruments (columns (2)-(3)).[30] All include endogenous origination timing variation. As expected, point estimates progressively increase: from column (1) to column (2), as we strip out variation from borrower-specific spreads by instrumenting with $\Delta SPC_j^{\text{Avg rate x LTV}}$, and from column (2) to column (3), as we strip out variation from borrower-specific leverage by instrumenting with $\Delta SPC_j^{\text{Avg rate}}$. Estimates in Column (3), which are only based on endogenous origination timing variation, indicate that a 1pp increase in scaled prepayment cost increases local price growth by a statistically-insignificant 0.25pp.

Table 1, panel (b) reports TSLS estimates of β using the twins, first birth, and lagged average rate instruments.[31] These approaches address the endogeneity of origination timing. Instruments are strong, with first-stage F-stats and Wald test stats above 74 and 23, respectively.[32]

The new instruments reveal a large effect of the rate lock incentive on house price growth. Column (4), using the twins instrument, shows a 1pp increase in average scaled prepayment cost causes a 1.75pp increase in house price growth from 2021-23. This implies that a one-standard deviation (0.015) increase in $SPC_{j,2023}$, which corresponds to a 0.33pp decrease (from 4% to 3.67%) in the average outstanding 2021 mortgage rate, causes a 2.6pp increase in nominal house price growth. The 2.6pp increase is about 11% of average nominal house price growth over 2021-23. Rate lock is also quite important in explaining the cross-section of house price growth during the recent tightening. Variation in scaled prepayment cost can account for 30% of cross-sectional variation in house price growth.[33]

The twins instrument uses somewhat small variation in prepayment costs, and so could estimate a non-representative local average treatment effect. The estimates in columns (5) and (6), which use broader variation from the first births and lagged average rate instruments, yield similar estimates. This builds confidence that the magnitudes in Panel (b) are both unbiased and representative. As further evidence, Panel (b) of Appendix Table D.24 shows these estimates are robust to including additional CBSA-by-demographic controls.

Overall, instruments that address endogenous origination timing endogeneity uncover a much greater price impact of the rate lock incentive than instruments that do not. Thus, on

[30] Estimates include rent-price controls suggested by the model from Section 1. Appendix Table D.24 adds other CBSA-by-demographic controls, including for 2021 log income, log population, log unit value, homeownership rate, population density, average age, average origination LTV ratio, average origination DTI ratio, the fixed-rate mortgage share, the nonwhite share, the rent-price ratio, and the ratio of 2023 rents to 2021 prices.

[31] See Appendix Table D.28 for the first-stage and reduced-form estimates separately.

[32] With one endogenous regressor and one instrument, the Wald test stat is equivalent to the Kleibergen (2007) Wald statistic and the Olea and Pflueger (2013) effective F-stat. With one endogenous regressor, one instrument, heteroskedastic errors, and TSLS estimation, the effective F-stat can be compared to Stock and Yogo (2005) critical values, typically around the rule-of-thumb value of 10. See Andrews *et al.* (2019), pp. 12-14.

[33] This comes from $\sqrt{\beta^2 \text{Var}(\Delta SPC_j)/\text{Var}(\%\Delta P_j)} = 0.30$.

net, origination timing endogeneity biases coefficient estimates downwards. This implies the previously-discussed forces for downwards bias, such as high scaled prepayment costs in places with unobserved negative demand shocks or positive housing supply shocks, outweigh the forces for upwards bias, such as households anticipating where price growth will be strong.

Housing market outflows. Section 1 explained that the rate lock incentive increases house prices by decreasing local home sales by households who would have otherwise moved to a different local market. Table 2 presents specifications where the outcome is the probability that a household moves out of the local owner-occupied housing market from 2022-2023, with Appendix Table D.25 adding controls. Coefficients are negative in all specifications, but are much more negative in panel (b) when estimated with instruments accounting for endogeneity of origination timing.

Estimates in panel (b) indicate that rate lock significantly reduces local outflows. Based on estimates in Column (4), a one standard-deviation increase in scaled prepayment costs, which corresponds to a 0.33pp decrease in the average outstanding 2021 mortgage rate, causes a 1.4pp decrease in outflows. As detailed in Section 4 below, these mobility estimates align with estimates from a separate, household-level natural experiment. Prepayment costs are an important driver of cross-sectional variation in mobility over 2022-23: variation in prepayment costs can account for 71% of cross-sectional variation in 2022-23 outflows.

Rate lock mostly reduces within-CBSA moves, but not across-CBSA moves. This gives another reason to include CBSA fixed effects in our specifications, as cross-CBSA variation does not reflect changes in existing-home supply and demand due to rate lock. In Appendix Table D.26a, we estimate β, where the outcome is the probability of a cross-CBSA move. Coefficients are non-negative. This is unsurprising, given that cross-CBSA moves are rare (>20% of all moves by owner-occupants, see Appendix Figure C.4), and are probably less discretionary, and so less elastic to the rate lock incentive. Aladangady *et al.* 2025 and Gerardi *et al.* (2024) report similar results, using individual-level, rather than market-level designs.

Appendix Table D.26b shows that rate lock significantly reduces own-to-rent moves. This implies rate lock affects moves between owner and renter market segments so as to increase aggregate house prices, as described in Section 1. This aligns with survey data showing around 25-40% of moves out of owner-occupied units are own-to-rent (see Appendix Figure C.4a).

Aggregate implications. We now consider what our estimates imply about the impact of aggregate scaled prepayment costs on aggregate house prices, based on a naive linear extrapolation of cross-sectional estimates. This is motivated by our analysis in Section 1, which shows that the impact of aggregate scaled prepayment costs on aggregate house prices are increasing in the cross-sectional estimate from specification (6).

Table 1: Effect of rate lock on local equilibrium prices

	Panel (a). Endogenous origination timing. %Δ Price, 21-23			Panel (b). Origination timing instruments. %Δ Price, 21-23		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	0.10	0.09	0.25	1.75***	2.37***	1.41**
	(0.07)	(0.07)	(0.17)	(0.64)	(0.64)	(0.59)
Instrument		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
CBSA FEs	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls	Yes	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls	No	No	No	No	No	No
Orig timing controls	None	None	None	None	None	Half-yr
Observations	8,000	8,000	8,000	5,975	6,377	6,377
F-stat (1st stage)		117,326.6	1,610.0	74.9	97.2	188.6
Wald (1st stage)		24,639.5	166.8	23.7	30.2	92.9

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

Source: Author's calculations using Equifax Credit Risks Insight Servicing, ICE, McDash®, and CDC natality data. Each column reports an estimate of β in equation (6) using a different methodology. Panel (a) reports fixed effects and IV estimates with endogenous origination timing as described in Section 3.2. Panel (b) reports IV estimates that address endogenous origination timing, as described in Section 3.3. Reported F-stats are the projected F-stat of the first stage. Reported Wald stats are calculated from county-clustered standard errors, and in these specifications with a single instrument for a single endogenous regressor, are equivalent to both the Kleinbergen and Paap ("KP") robust test statistic and the Olea and Pflueger (2013) effective F-stat.

Table 2: Effect of rate lock on local equilibrium outflows

	Panel (a). Endogenous origination timing. Pr(Move), 2022			Panel (b). Origination timing instruments. Pr(Move), 2022		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	-0.12***	-0.10***	-0.22***	-0.94***	-0.69***	-0.58***
	(0.02)	(0.02)	(0.05)	(0.24)	(0.15)	(0.17)
Instrument		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
CBSA FEs	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls	Yes	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls	No	No	No	No	No	No
Orig timing controls	None	None	None	None	None	Half-yr
Observations	8,000	8,000	8,000	5,975	6,377	6,377
F-stat (1st stage)		117,326.6	1,610.0	74.9	97.2	188.6
Wald (1st stage)		24,639.5	166.8	23.7	30.2	92.9

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

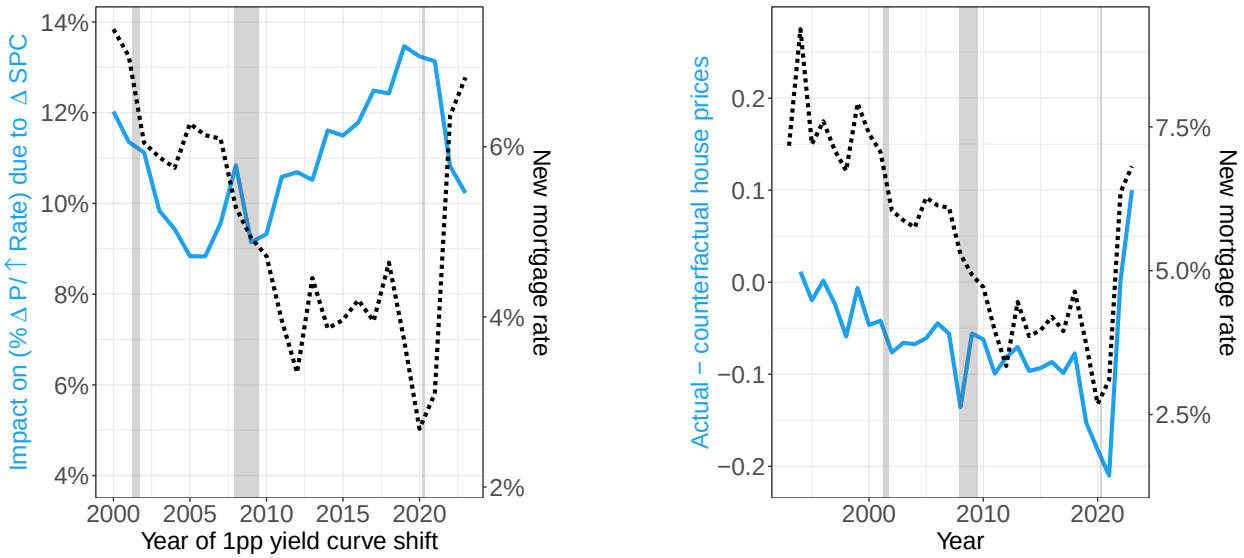
Source: Author's calculations using Equifax Credit Risks Insight Servicing, ICE, McDash®, and CDC natality data. The table has the same format as Table 1, except the outcome variable is the probability of moving from 2022-23. See notes to Table 1 for details.

Recall that a 33 basis-point decrease in average outstanding rates, implied by a 1 standard-deviation increase in $SPC_{j,2023}$, causes a 2.6pp increase in 2021-23 nominal house price growth – a semi-elasticity of 8. Static user cost formulas that do not consider fixed-rate mortgages, as in Himmelberg *et al.* 2005, often predict the semi-elasticity of price-rent ratios to long-term interest rates of around -20. Based on this benchmark, over 2021-23, the rate lock incentive offset about 40% of the negative user-cost effects of higher rates.

This offset to the semi-elasticity of house prices with respect to interest rates varies over

Figure 4: Effect of rate lock on house prices over time

(a) Offset to price-rate semi-elasticity. (b) Price effect of removing rate lock.



Source: Equifax Credit Risks Insight Servicing and ICE, McDash®. In each figure, the black dashed line plots the average mortgage rate from the Freddie Mac PMMS on the right axis. (a) For the year given on the horizontal axis, we calculate ΔSPC_t due to a 1pp parallel shift in the yield curve, and plot $\hat{\beta} \cdot \Delta SPC_t$, where $\hat{\beta}$ is the twins IV estimate from Column (4) of Table 1. (b) For each year, we calculate $SPC_{t-2,t}$, scaled prepayment costs in year t for the mortgage distribution outstanding two years prior. We plot $\hat{\beta} \cdot SPC_{t-2,t}$, the effect on nominal house prices of removing outstanding prepayment cost.

time, depending on the previous path of mortgage rates. Using the 0.5% ICE, McDash®dataset, we calculate the aggregate ΔSPC_t for mortgages outstanding in t due to a 1pp parallel shift in the yield curve. We multiply ΔSPC_t by the twins estimate of β to estimate the offset to the house price semi-elasticity.

Figure 4a plots results from 2000-2023. The offset rose from 9% in 2005 to over 13% by 2020. Persistently low rates in the 2010s increased the fraction of borrowers with low-rate mortgages. This made the outstanding mortgage distribution increasingly sensitive to price effects of rate lock. Fixed-rate mortgages thus create path dependence in the house price effects of interest rates: after a persistent period of low rates, a rate hike has a more muted effect on house prices. This is the flip-side of the argument in Berger *et al.* (2021), where the stimulus effects of rate *cuts* are weaker after persistent low-rate periods due to the high fraction of borrowers with low-rate mortgages.

What do estimates imply about the contribution of fixed-rate mortgages to overall 2021-23 price growth? If outstanding mortgage rates tracked the policy rate, then scaled prepayment cost would decrease from 4.5% in 2023 to zero. This is analogous to the price effects of forcing all borrowers into adjustable-rate mortgages with continuous adjustment and no fixation periods. Eliminating scaled prepayment cost would reduce house prices by 7.9pp, or 33% of

house price growth from 2021-23.

Looking over time, we find that prepayment costs lead to higher house prices when mortgage rates deviate from recent trends. Figure 4b shows the percent difference between actual nominal house prices and counter-factual prices if outstanding scaled prepayment costs were eliminated, and borrowers were implicitly forced into adjustable-rate mortgages, in each year since 1992. Until 2021, mortgage rates were on a downward trajectory. Over this period, prepayment costs due to fixed-rate mortgages led to slightly *lower* house prices, consistent with papers showing that adjustable-rate mortgages make house prices more responsive to monetary stimulus (e.g., Guren *et al.* 2021). Over 2021-23, mortgage rates rose, and prepayment costs led to *higher* house prices.

3.6 Robustness checks and additional outcomes.

This section summarizes several robustness checks and presents additional results on the effect of scaled prepayment costs on residential investment and rents.

Robustness: Alternative prepayment cost measurement. Appendix Section D.11 estimates price and outflow effects with a re-calculated ΔSPC_j which values the option to refinance in future periods, using either the model of optimal refinancing behavior from Agarwal *et al.* (2013) or mortgage-backed securities prices in the TBA market, as described in Appendix Section B.4. The results are essentially the same, up to rescaling by differences in average ΔSPC_j.

Robustness: Twins instrument. Appendix Section D.9 presents two robustness checks for the twins IV. Section D.9.1 excludes twin births to older mothers, where IVF use is more common. Section D.9.2 uses a method adapted from Borusyak and Hull (2023) that controls for market-specific expected changes in the twin birth rate. This ensures identifying variation is only finite sample variation relative to the expectation. Estimates align with the main text.

Robustness: November 2016 event study. Appendix D.10 presents a robustness check for the lagged average rate instrument. We only use variation from originations in the 8 months surrounding November 2016: due to Donald Trump's unexpected election, the 10-year Treasury jumped by 70 basis points in the largest month-over-month increase outside of a recession for which we have data. Section 4 describes this event study in more detail. Estimates align with results in Panel (b) in Tables 1 and 2.

Robustness: Placebo test. Appendix Table D.27a shows results of a placebo test, where we estimate specification (6) with the outcome as house price growth from 2019-2021. If instruments only detect the price impacts of 2021-23 scaled prepayment costs, then coefficient estimates should be zero. In specifications with instruments endogenous to origination timing,

estimates are nonzero: more evidence that estimates in Panel (a) of Table 1 are confounded. Estimates are statistically zero in specifications with instruments addressing origination timing endogeneity in Panel (b).

Robustness: Price growth timing. Appendix Table D.27b shows results where we measure price growth over a longer period. Specifically, we regress price growth from 2019-23 on scaled prepayment cost growth from 2019-23. Results are unchanged relative to Table 1, indicating that results are not distorted by basing initial prices during the post-Covid housing market.

Effects on residential investment. We estimate the effect of ΔSPC_j on local single-family construction, multi-family construction, and remodeling. We measure construction activity using the Census Building Permits Survey and new CoreLogic data on permits for renovation.

Appendix D.13 reports two findings. First, scaled prepayment costs increase single-family construction. We scale coefficient estimates by the estimated effect of prepayment costs on price growth from Table 1, so estimates can be interpreted as a price elasticity of supply. Using instruments that address origination timing endogeneity, implied supply elasticities align with external estimates from Baum-Snow and Han (2024). When estimated using instruments that are endogenous to origination timing, implied supply elasticities are much too large. This suggests the endogenous origination timing instruments predict high ΔSPC_j in supply-elastic markets, as expected. Second, we show weak negative effects on remodeling activity. Rate lock could increase renovations for households who stay put rather than move to a more suitable unit, but could also decrease remodeling that happens when new owners purchase a unit by reducing housing market churn.[34] Our results suggest the latter effect dominates.

Effects on rents. Appendix Section D.14 present estimates of β in equation (6) where the outcome is the change in nominal rents from 2021 Q4-2023 Q4. Measuring rents on a constant-quality basis with spatial and temporal granularity is a major challenge. To overcome this challenge, we construct a zip-level constant-quality rent index using monthly administrative building-level data from RealPage, a leading property management software company. We broadly find evidence of a zero-to-negative impact on rents, consistent with our findings that rate lock increases relative demand for owner-occupied housing.[35]

[34] As explained by a former Senior Manager for Merchandising Strategy at The Home Depot: "unless it's still really new, when you buy an existing house, you're going to want to go put your own spin on. You're going to buy paint, you're going to buy flooring, and you might buy new light fixtures" (AlphaSense 2025).

[35] In recent work, **?** compare properties within Census tracts in LA county, and find that higher prepayment costs in owner-occupied "starter homes" increase contract rents in units within a 0.5-mile radius. Their methodology is most similar to our "average rate" instrument.

3.7 Mechanisms: Rate lock affects prices where housing supply is scarce.

Section 1 predicts that rate lock should impact prices most in places where lower existing-home supply has a larger impact on overall housing unit availability. This section gives empirical support for this prediction: rate lock has a larger price effect in markets with inelastic housing supply, and higher predicted outflows absent rate lock. We also find higher prepayment costs in supply-elastic markets and places with low expected outflows, helping to explain the downwards bias in fixed-effects estimates of the price effect of rate lock.

We estimate the following equation:

$$\%\Delta P_j = \beta_1 \cdot \Delta SPC_j + \beta_2 \cdot \gamma_j + \beta_3 \cdot \gamma_j \times \Delta SPC_j + \eta'_{c(j)} \cdot X_j + \omega_{c(j)} + \epsilon_j \tag{9}$$

where γ_j is a measure of either supply elasticity or predicted outflows. The coefficient of interest is β_3: if positive, rate lock has a greater price effect when γ_j is high. We instrument for ΔSPC_j and $\gamma_j \times \Delta SPC_j$ with $\Delta SPC_j^{\text{Avg rate x LTV}}$ and its interaction with γ_j. We use the average rate-by-LTV instrument because it has the greatest power, and here we aim to test what relationships hold directionally rather than assess magnitudes.

Elasticity of housing supply. We first estimate β_3 where γ_j is the elasticity of housing supply in standardized deviations from the mean. We estimate γ_j as the housing-unit weighted average of Census tract-level long-run new unit supply elasticities reported by Baum-Snow and Han (2024).

Table 3, column (1) presents estimates. The interaction effect is negative and quite large: a one standard deviation decrease in the supply elasticity more than doubles the marginal effect of scaled prepayment cost on price growth. Column (2) regresses ΔSPC_j on the standardized supply elasticity. As hypothesized in Section 3.1, markets with more elastic housing supply experienced more scaled prepayment cost growth from 2021-23. Thus, markets with higher scaled prepayment cost growth are also those less likely to see price growth due to elastic supply, leading to downwards bias in fixed effects estimates of β in equation (6).

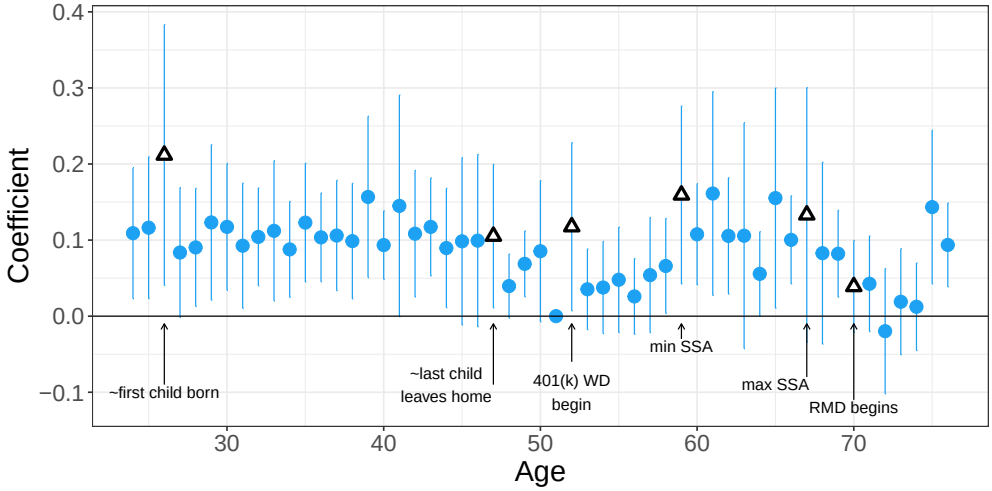
Expected outflows. We next estimate β_3 where γ_j is expected outflows. We proxy for expected outflows using the probability that a household moves out of j from 2022-23 in standard deviation units.

A challenge is that outflows are higher in markets that face negative economic or amenity shocks that create persistently lower price growth, potentially biasing estimates of β_3. To overcome this challenge, we instrument for outflows using the local age distribution, due to the fact that moves vary predictably with age. We use the 2019 local age distribution and the historical relationship between the age distribution and three-year ahead mobility to predict 2022-23

outflows, and use this prediction as the instrument for γ_j. We estimate the following equation using pre-2019 data:

$$\Pr(Move)_{j,t+3} = \alpha_j + \omega_{c(j),t} + \sum_{l \in L} \pi_l \Pr(Age_{it} = l | i \in jt) + \varepsilon_{j,t+3} \tag{10}$$

Figure 5: Effect of the local age distribution on moves.



Source: Equifax Credit Risks Insight Servicing data. The figure plots fixed effects estimates of π_l in equation (10). We label age 26 (three years before age 29, the median parental age of first child); 47 (21 years later, the age by which most children leave home); 52 (three years before age 55 1/2, when non-penalized 401(k) withdrawals begin); 59 (three years before the minimum age of claiming social security); 67 (three years before the maximum age of claiming social security); and 70 (three years before minimum distribution requirements for retirement accounts).

Figure 5 plots coefficient estimates. Places with many households in early adulthood and early retirement see higher outflows three years later, with inflection points around meaningful lifecycle events. We instrument for γ_j by predicting 2022-23 outflows using the 2019 age distribution: $\sum_{l \in L} \pi_l \Pr\big(Age_{i,2019} = l | i \in j, 2019\big)$.[36]

Table 3, column (3) shows TSLS estimates of equation (9) where γ_j is standardized predicted outflows. We control for $\{\Pr(Age_{it} = \ell | i \in jt)\}_{\ell \in L}$; thus, estimates are driven only by the correlation between the age distribution and outflows, rather than the through the direct effect of the age distribution on prices. The coefficient on the interaction term is positive, meaning rate lock has a larger price effect where predicted outflows are high. The magnitude is large: a one standard deviation increase in predicted outflows increases the effect of scaled prepayment cost growth on price growth by about 45%.

Column (4) regresses ΔSPC_j on predicted outflows. Markets with higher predicted outflows have lower scaled prepayment cost growth. This gives another channel through which fixed effects estimates are biased downwards. Markets with high predicted outflows have low ΔSPC_j, perhaps because high outflows – and hence high churn – means that households have

[36] Appendix Figure D.22 shows that there significant dispersion in predicted outflows.

not held onto low-rate mortgages. But those markets are where rate lock has scope to impact prices, meaning average effects are biased downwards.

Table 3: Effect of rate lock on prices: Heterogeneity by supply elasticity and outflows.

Dependent Variables:	%Δ Price, 21-23 (1)	Δ SPC, 2021-23 (2)	%Δ Price, 21-23 (3)	Δ SPC, 2021-23 (4)
Δ SPC, 2021-23	0.09		0.40***	
	(0.09)		(0.10)	
New unit supply elast (std)	0.04***	0.01***		
	(0.01)	(0.00)		
%Δ SPC, 2021-23 x supply elast (std)	-0.19**			
	(0.08)			
Pr(Move), 2022 (std)			-0.03**	-0.01***
			(0.01)	(0.00)
%Δ SPC, 2021-23 x Pr(Move), 2022 (std)			0.18**	
			(0.08)	
Instrument	Avg rate x LTV		Avg rate x LTV	
CBSA FEs	Yes	Yes	Yes	Yes
Rent/price x CBSA controls	Yes	Yes	Yes	Yes
Demo x CBSA controls	Yes	Yes	Yes	Yes
Age dist controls	No	No	Yes	Yes
Observations	6,194	6,194	7,950	7,950
Within Adjusted R^2	0.018	0.083	0.023	0.092

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

Source: Equifax Credit Risks Insight Servicing and ICE, McDash®. The table presents coefficient estimates for equation (9) using the LTV-by-Treasury rate instrument. Column (1) proxies for $-\hat{\gamma}_{j,2023}$ using physical supply elasticities. Column (2) proxies for $\hat{\gamma}_{j,2023}$ using predicted outflows based on the three-period lagged age distribution.

4 Effect of rate lock on individual households

This section studies how rate lock affects existing-home sales decisions at the household-level. We do so to analyze the mechanisms driving results in Section 3 in more detail, and inform the structural model in Section 5. We quantify how much rate lock disrupts overall existing-home sales and moves across different housing market segments. We target these causal moments in structural estimation.

4.1 Empirical approach.

An ideal experiment would randomly assign mortgage rates to borrowers immediately before a large increase in long-term risk-free rates. But in the data, there are two challenges leading borrowers' outstanding mortgage rates to correlate with unobserved drivers of mobility. First, outstanding mortgage rates are not random, but depend on borrower risk and long-term rates at origination. Second, long-term rates have trended downwards over time, meaning most mortgage rate movements do not create variation in the rate lock incentive, but in the incentive to refinance.

The correlation between mobility and outstanding mortgage rates could be positive or negative, but either way creates bias when estimating the effect of the rate lock incentive on existing-home sales. There could be positive correlation because more mobile borrowers tend to be younger and riskier, and hence have higher mortgage spreads, while less mobile borrowers expect to stay in their homes for longer, and so time moves or refinance in low-rate periods. The correlation could also be negative. Due to falling long-term rates, more mobile borrowers mechanically end up with lower outstanding mortgage rates. Moreover, more mobile borrowers could be more interest-rate sensitive, and so tend to move or refinance in low-rate periods.

To overcome the first challenge, we compare borrowers who initially purchase a home just before and just after sharp increases in long-term interest rates. Using aggregate rate movements removes endogeneity due to borrower mortgage spreads. Focusing on sharp rate changes reduces borrower selection: moving is frictional, and exact dates result from months of planning that are hard to time around high-frequency rate fluctuations.[37]

To overcome the second challenge, we only use variation from rate *increases* followed by persistently high rates. This rules out most rate variation in CRISM from 2006-present. Figure 6, panel (d) plots the month-over-month change in the new 30-year fixed-rate mortgage rate, measured in the PMMS, against the average difference with future mortgage rates over the next four years. The top-right quadrant shows candidate periods, when a sharp mortgage rate increase preceded a period of high rates.

We focus on the largest month-over-month increase that is followed by persistently high rates, which occurred in November 2016. During a period of otherwise stable conditions, 10-year Treasury yields jumped in response to Donald Trump's surprise presidential election victory.[38] The change was sharp, but not driven by current economic conditions or credit market stress. It is thus unlikely to affect the mortgage market except through the mortgage rate, unlike other candidate episodes — e.g., the 2013 "taper tantrum", which created credit market distress that could have impacted mortgage origination, or the 2005 tightening just before the housing market crash.[39]

We use variation from the November 2016 increase to study effects of rate lock on existing-

[37] Most online guides recommend at least 3-6 months for finding a new house, selling an existing one, and arranging moving logistics (example here).

[38] From the Fed's December 2016 meeting minutes: "[s]urveys of market participants indicated that revised expectations for government spending and tax policy following the U.S. elections in early November were seen as the most important reasons, among several factors, for the increase in longer-term Treasury yields."

[39] When the Fed announced the end of quantitative easing in 2013, mortgage spreads jumped by 50 bps (Gordon 2023), partly due to lower liquidity supplied by Treasury broker-dealers to the MBS market (Goldberg 2020). Changes in 2005 happened around house price peaks; falling house prices over the next few years could confound mobility effects due to "equity lock" (Ferreira *et al.* 2010). Changes in 2015 are small, and driven by bond market speculation about the the start of federal fund rate increases.

home sales within an event study framework.[40] For households who initially moved in a narrow window around November 2016, we estimate the impact of the November 2016 rate increase on outstanding mortgage rates as the coefficient π_1 in the regression:

$$r_{i0} = \pi_1 \cdot T_i + \pi_2 \cdot d_i + \pi_3 \cdot d_i \times T_i + \gamma_{z(i)} + \eta_{g(i)} + u_i \tag{11}$$

where r_{i0} is borrower i's mortgage rate at move-in; T_i is an indicator for origination in the period after November 7, 2016; d_i is the number of days between origination and the increase; $\gamma_{z(i)}$ are property zip code fixed effects; and $\gamma_{g(i)}$ are fixed effects for group $g(i)$, which interact age, income, and appraised property value quintile.

We then use T_i as the excluded instrument to estimate β_1 in the structural equation:

$$y_i = \beta_1 \cdot r_{i0} + \beta_2 \cdot d_i + \beta_3 \cdot d_i \times T_i + \gamma_{z(i)} + \eta_{g(i)} + \varepsilon_i \tag{12}$$

where y_i is an indicator for existing-home sales over a fixed horizon. The coefficient of interest is β_1, which compares move rates for borrowers with mortgages originated just before and just after the increase. Doing so ensures that estimates flexibly hold fixed the effect on existing-home sales of the path of financial environment, time since initial home purchase, and any interaction between them. The identifying assumption is that unobserved drivers of y_i in ε_i do not shift discontinuously after the interest rate increase at the event date.

We also estimate equation (12) where y_i is the log resale price for borrowers who sell their home. This gives a revealed-preference measure for the financial value borrowers place on their fixed-rate mortgages: a higher value raises sellers' willingness-to-accept and therefore the resale price. However, the estimate of β_1 is a *lower bound* on willingness-to-accept. If housing markets were fully integrated, then within a market, variation in seller willingness-to-accept would not affect transaction prices. In reality, search-and-matching frictions allow sellers to partly capitalize their willingness-to-accept (Kotova and Zhang 2020), meaning we expect to estimate an attenuated (but still positive) effect.

Our methodology builds on Fonseca and Liu (2024), Gerardi *et al.* (2024), Batzer *et al.* (2024), and Aladangady *et al.* (2025). These papers regress moves on the gap between r_{i0} and current market rates in a specification with calendar time and origination cohort fixed effects, instrumenting for r_{i0} with market mortgage rates at origination. This addresses the endogeneity of mortgage spreads, but not confounds due to origination timing or falling rates.

Our approach has two main innovations. First, we restrict to variation from move-in dates, stripping out originations due to refinancing. Move-in dates are less manipulable at high frequencies, and hence are less likely to correlate with unobserved drivers of mobility. Second,

[40] This design is sometimes called fuzzy regression discontinuity in time (see Hausman and Rapson 2018).

we focus on variation from rate fluctuations most relevant for rate lock. This matters: while Fonseca and Liu (2024) show that conditional on origination quarter, origination mortgage rates are *on average* balanced on borrower characteristics, Appendix C.2 shows the instrument *does* correlate with borrower income, risk, and leverage for originations during periods *that produce rate lock*, outside of the November 2016 event study.

Where y_i is moves, we estimate equation (12) using the CRISM borrower panel. We restrict to first-lien, fixed-rate, owner-occupied purchase originations that align with a borrower move.[41] We restrict to borrowers with mortgage origination dates within 4 months of the event window. Where y_i is sales price, we use CoreLogic data with analogous sample restrictions. Appendix Tables C.6b and C.7b show summary statistics for the restricted samples.

4.2 Results.

Visual evidence. Figure 6, panel (a) shows the first-stage of the November 2016 event study. The 10-year treasury yield increased about 70 bps in November, immediately passing through to new mortgage rates as measured in PMMS. Rates on mortgages in CRISM increased with a lag of 4-12 weeks, because many lenders allow borrowers to lock rates for up to 90 days.[42]

The main threat to identification is selection on unobserved borrower characteristics: borrowers who buy just after the rate increase have unobservably different mobility rates for reasons unrelated to the rate lock incentive. Additionally, higher rates could cause borrowers to purchase smaller homes, leading to lower match quality and higher mobility. Appendix Figure C.7b gives visual evidence against these concerns: there is no discontinuous change in age, income, credit score, LTV, or appraised property value at the boundary. Appendix Table C.8 estimates equation (12) where y_i is each of these characteristics, and finds that β_1 is statistically zero. Inconsistent with selection, Appendix Figure C.7a shows that 2016 origination volume closely tracks previous and subsequent years.

Panel (b) of Figure 6 shows the primary effect of higher origination rates: higher mortgage payments, which decrease the willingness-to-accept for owners. Given average appraised value of $330,000, annual mortgage payments increase by around $1,650 around the threshold.[43]

[41] This screens out mortgages made to real estate investors who may time origination around high-frequency rate fluctuations and exhibit different sales choices than owner-occupants.

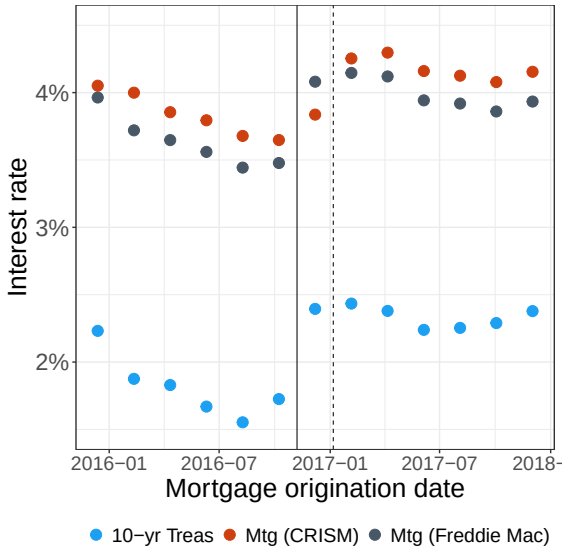
[42] Freddie Mac rates are slightly lower than in CRISM because Freddie Mac collects interest rates on prime 80% LTV mortgage. This is lower risk than the average mortgage, so the average mortgage rate is higher.

[43] Fonseca and Liu (2024) and Batzer *et al.* (2024) control for this variation in their estimates. All of their specifications control for monthly payments, mortgage balance, and remaining term: effectively, a first-order approximation to the mortgage rate. For a fixed-rate mortgage, the remaining balance after the nth payment is: $B_n = B_0 \cdot r_{m0} - m \cdot \frac{(1+r_{m0})^n - 1}{r_{m0}}$, where B_0 is origination balance, r_{m0} is origination mortgage rate, and m is fixed mortgage payment. To a first-order, $B_n \approx a_0 + a_1 B_0 + a_2 m + a_3 n + a_4 r_{m0} \implies r_{m0} \approx a_3^{-1} \cdot (B_n - a_0 - a_1 B_0 - a_2 m - a_3 n)$. Conditional on B_n, m, and n, variation in r_{m0} only comes from origination balance B_0. Fonseca and Liu (2024) control for local house prices, which closely correlates with B_0, and Batzer *et al.* (2024) control for B_0 explicitly.

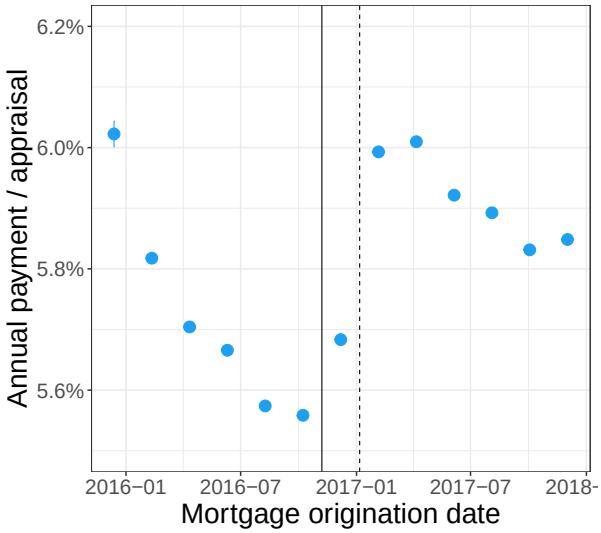
As predicted by the model in Section 1, panel (c) shows this decreased willingness-to-accept increases the probability that a borrower sells their home within four years by about 30%, and reduces the probability of an own-to-rent move by 20%.[44]

Figure 6: Effect of November 2016 rate increase on existing-home supply.

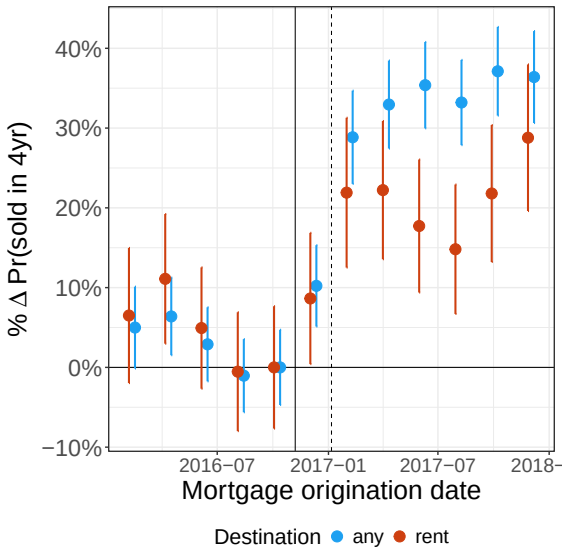
(a) Rates by mortgage origination date



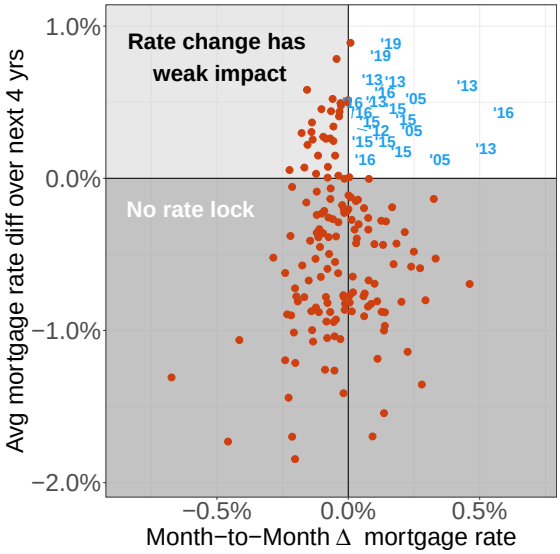
(b) Mortgage payment by origination date



(c) Existing-home sales by origination date



(d) Valid sources of variation



Source: Equifax Credit Risks Insight Servicing and FRED. Each figure plots outcomes in 60-day bins relative to November 10, 2016, when long-term interest rates started to increase following the 2016 presidential election. Panel (a) plots the average 10-year Treasury yield, the Freddie Mac Primary Mortgage Market Survey 30-year fixed rate, and the average rate in CRISM for first-lien purchase mortgages originated at move-in. Panel (b) plots monthly principal and interest mortgage payments. Panel (c) plots the percent change in probability of an existing-home sale relative to the period just before November 2016, both unconditionally (blue) and conditional on an own-to-rent move (red).

[44] Appendix Figure C.8a shows the change in levels, rather than percentage changes.

Effects on existing-home sales. Table 4 presents IV estimates of β_1 from equation (12). To account for the lag in pass-through, we omit originations in the 60 days after the event date, and compute event time d_i relative to 60 days after the event date for $T_i = 1$.

Panel (a) shows effects on existing-home sales. Column (1) shows a 1pp increase in the mortgage rate at origination increases the four-year sales probability by 5.2pp (42%). Crucially, Column (2) estimates that rate lock disrupts own-to-rent moves: a 1pp increase in the origination mortgage rate increases the four-year probability of an own-to-rent move by 1.7pp (34%). This key result implies that rate lock increases the aggregate price-to-rent ratio, as explained in Section 1.

Our finding that moving from owning to renting is an active margin aligns with data points from several other sources. Table C.6 shows that 30% of moves by owner-occupants within four years of move-in are from owning to renting. This aligns with survey data from the NLSY79, where depending on the age of an owner-occupant, around 25-39% of total moves are from owning to renting (Appendix Figure C.4a). It also aligns with separate survey analysis in the Survey of Income and Program Participation (SIPP) by Chodorow-Reich *et al.* (2025), where 34% of moves by owner-occupants are from owning-to-renting (Table E.1), and evidence from the American Housing Survey evidence over 1985-2007, where 39% of owner-occupant sales were for own-to-rent moves (Wheaton and Lee 2009).

These individual-level mobility estimates align with market-level cross-sectional estimates in Section 3.[45] Recall that the estimates from the twins IV in Table 2, panel (b) imply that a one standard-deviation increase in 2023 prepayment costs, which corresponds to a 0.33pp decrease in average outstanding mortgage rates in 2021, causes a 1.4pp decrease in local outflows — an 18.5% decrease relative to average outflows from 2006-2019. Using estimates in Table 4, a 0.33pp decrease in the average outstanding mortgage rate decreases individual-level moves by 17.3% of the average four-year move rate from 2006-2019. Individual-level estimates in Table 4 are *not* consistent with market-level estimates that use instruments endogenous to origination timing in 2, panel (a). For example, the estimates in column (3) imply that the same rate lock incentive considered above reduces outflows by 4.3% of the 2006-2019 average.

What types of existing-home sales does rate lock disrupt? Column (3) shows a statistically-insignificant effect on cross-CBSA moves, suggesting that rate lock mostly reduces within-CBSA moves. This aligns with results from Section 3.

Column (4) splits by age, and shows that rate lock has a larger impact for younger and older borrowers. This could reflect an endogenous feature of lifecycle wealth accumulation: young households are more sensitive to monthly payment differences due to low wealth, while older

[45] Note that the results are not directly comparable: we consider moves over different time periods, for households with different tenures in their current homes, using different samples and a different source of variation.

households, with high wealth and low future income, discount costs at close to the risk-free rate, and are more sensitive to the present-value cost of face value prepayment. These forces matter less for middle-aged borrowers. These features endogenously emerge in the empirical model in Section 5.

Table 4: November 2016 event study: Main estimates.

(a) Effects on existing-home sales.

Pr(Moved in 4yr):	all (1)	to rent (2)	diff CBSA (3)	(4)	all (5)	(6)
Rate	0.052***	0.017**	0.006		-0.102*	0.066
	(0.012)	(0.007)	(0.005)		(0.058)	(0.067)
Rate x Age quartile 1				0.068**		
				(0.028)		
Rate x Age quartile 2				0.056**		
				(0.026)		
Rate x Age quartile 3				0.040		
				(0.028)		
Rate x Age quartile 4				0.046**		
				(0.023)		
Rate x LTV					0.181***	
					(0.066)	
Rate x DTI						0.004
						(0.181)
Trend width	4mo	4mo	4mo	4mo	4mo	4mo
Fixed-effects						
Zip-ever refi	Yes	Yes	Yes		Yes	Yes
Age-Income-Appraised val.	Yes	Yes	Yes	Yes	Yes	Yes
Zip-ever refi-Age quartile				Yes		
Observations	68,941	68,941	68,941	55,903	68,719	38,386
R^2	0.225	0.206	0.162	0.330	0.225	0.290
Within R^2	0.014	0.005	0.003	0.015	0.013	0.014

****: 0.01, **: 0.05, *: 0.1. Standard errors clustered by person and orig. date in parentheses.*

(b) Effects on resale price.

	Ln(Resale price) (1)	(2)
Rate	-0.0238***	-0.0221***
	(0.0081)	(0.0076)
Trend width	6mo	6mo
Fixed-effects		
Zip-Resale qtr	Yes	
Zip-Bed x bath bin		Yes
Resale qtr-Bed x bath bin		Yes
Observations	605,057	605,057
F-test (1st stage), Rate	1,705,548.0	1,602,739.6
Wald (1st stage), Rate	463,431.2	436,988.0

Standard errors clustered by zip and resale month.

Source: Equifax Credit Risks Insight Servicing (a) and CoreLogic (b). The table presents TSLS estimates of β_1 in equation (12) along with interaction effects. In panel (a), the outcome is existing-home sales in 4 years. Column (1) shows effects on the probability of existing-home sales within 4yrs of move-in. Column (2) conditions on existing-home sales resulting in an own-to-rent move. Column (3) conditions on existing-home sales resulting in a within-CBSA move. Column (4) interacts r_{i0} with quartiles of age at move-in. Column (5) and (6) interact r_{i0} with origination LTV and DTI, in specifications that control for LTV and DTI directly (coefficients suppressed in table). Age-inc-appraised value fixed effects bin by age, income, and property appraised value at move-in. In panel (b), the outcome is log existing-home sales price.

Robustness checks. Appendix C.3 presents four robustness checks. First, we verify that seasonal patterns do not drive observed patterns by running a placebo test on originations around November 2014. Second, we present estimates from an alternative "difference-in-differences" design that uses origination date-of-year trends in 2014-2015 as a control group for the treated 2016-2017 sample. Point estimates are similar, but the design has more power. Third, we show that our main estimates are unchanged if we exclude moves to areas to and from the Washington, DC area, ensuring moves related to political churn around the November 2016 presidential election do not drive results. Finally, we show that main estimates are unchanged if we restrict to variation where the *initial* move was across-CBSA. These moves are likely less elastic at high frequencies than within-CBSA moves, and hence less likely to be subject to selection.

Comparing to other estimates. Estimates of the semi-elasticity of existing-home sales with respect to a 1pp change in households' mortgage rates range from 7-9% (Liebersohn and Roth-

stein 2024, Fonseca and Liu 2024) to 34% (Gerardi *et al.* 2024). While methodologies are conceptually similar to ours, these papers use different variation over different time periods.

Our estimates align closely with Gerardi *et al.* (2024). If we adjust for differences in average LTV in our sample and that in Gerardi *et al.* (2024) using estimates in Table 4, panel (a), column (5), we estimate an average semi-elasticity of 34%.

We suspect our estimates are on the higher range because we restrict to variation that is most relevant for rate lock. By using wider variation, other papers may have more measurement error in the rate lock incentive and hence estimate an attenuated effect.

Effects on resale price. We estimate equation (12) where the outcome is log resale price, and report results in Table 4, panel (b).[46] Column (1) includes zip code by resale quarter fixed effects, and so compares properties sold in the same zip code in the same quarter where owners initially purchased the home just before and just after the November 2016 mortgage rate increase. Estimates imply that a 1pp increase in mortgage rates reduces resale price by 2.4%. This indicates a lower fixed-rate mortgage increases seller willingness-to-accept. Column (2) addresses the possibility that the price difference is instead driven by different characteristics of homes sold by controlling for quarter-specific fixed effects for the number of bedrooms (0-2, 3-4, 4+) interacted with number of bathrooms (0-2, 3-4, 4+). Estimates are almost unchanged.

Wealth effects, not liquidity constraints, drive behavior. Why does rate lock reduce existing-home sales? In the Section 1 model, forward-looking borrowers don't sell to avoid the negative wealth effect of losing the capital gain on their low-rate mortgage. Alternatively, borrowers could make static choices based on current liquidity, where the rate lock effect reflects that households cannot afford higher monthly payments on new mortgages when interest rates rise.

Four pieces of evidence indicate wealth effects drive behavior. First, rate lock is higher when the capital gain on low-rate mortgages is a higher fraction of borrower wealth. Table 4, panel (a), Column (5) interacts T_i with a mortgage's loan-to-value (LTV) ratio, a proxy for the forgone capital gain. Rate lock reduces existing-home sales by more for borrowers with higher LTV. Second, rate lock does not vary by borrower liquidity. Column (6) interacts T_i with the ratio of debt payments to income (DTI), a proxy for the payment's liquidity impact. The interaction term is statistically zero.

Third, rate lock reduces own-to-rent moves. Our design compares households who face the same path of rental costs. The fact that borrowers with different mortgages make different own-to-rent choices therefore cannot be driven by different post-move housing costs.

[46] This *within*-market price effect depends on market clearing frictions, and so is different than the *equilibrium* impact of changing mortgages for the typical borrower; see Minton and Mulligan (2024) for further discussion.

Finally, the effects of rate lock on seller willingness-to-accept indicate that borrowers value their current mortgages in a way consistent with a forward-looking calculation. We calculate the risk-neutral discount rate to make a typical seller indifferent between a 2.3% higher resale price and the present-value of annual differences in payments due to a 1pp higher mortgage rate. We account for moving and prepayment using mortgage duration estimates from ICE, McDash®. Since resale price effects are a *lower* bound on willingness-to-accept effects, our procedure gives an *upper* bound on discount rates. Appendix Section C.4 gives details. Estimates imply that borrowers discount future costs at a risk-neutral rate of at most 4.6%.

Do households calculate these wealth effects of rate lock by discounting future costs at a rate related to the risk-free rate on saving? To investigate, we correlate the 2-year Treasury rate with a precise revealed-preference measure of relative borrower discounting: bunching at the 80% LTV threshold for conventional loans. Conventional loans with LTVs over 80% must buy private mortgage insurance (PMI), which adds an insurance premium to monthly payments. This discretely increases the present value of future mortgage costs. LTVs just above 80% are thus dominated by an LTV of exactly 80%, until the liquidity benefit from a smaller down payment today equals the present-value cost of discretely higher monthly payments. Borrowers who would have chosen an LTV in this dominated region without PMI requirements instead bunch at 80% exactly. The size of this region—and hence the amount of bunching—depends on borrower discount rates. When discount rates are low, bunching is high because future costs loom large relative to current liquidity.

Appendix Figure C.9 shows a strong negative relationship between the two-year Treasury yield and the amount of bunching at an 80% LTV (R-squared=17.8%). This is consistent with a *positive* correlation between borrower discount rates and the risk-free rate.

Implications for structural model. Three results from this section inform the structural model's design and estimation. First, we estimate structural parameters that determine the strength of the rate lock incentive by matching the causal effect of higher outstanding mortgage rates on overall and own-to-rent moves. Second, we ensure that households make forward-looking and dynamic mobility choices, so they value the rate lock incentive according to a forward-looking calculation. Third, we explicitly allow households to save at the riskfree rate. This endogenously ensures that households behave as though they discount future costs at a rate correlated with the riskfree rate.

5 Structural model: Housing demand with rate lock

Section 3 shows that markets more exposed to the rate lock incentive experienced relatively more price growth over 2021-23. But rising rates also affect housing demand by increasing the

user cost of capital. This section develops a structural model to estimate the net price effect of these forces, under fixed-rate mortgages and alternative mortgage contracts.

The model embeds a dynamic discrete housing choice in an incomplete-markets lifecycle saving-consumption model with fixed-rate mortgages and uncertain, time-varying interest rates. We capture important features of housing demand, such as uncertainty about the future economic environment, credit constraints, a lifecycle consumption profile, mobility frictions, and housing transactions costs. Sections 5.1-5.3 describe and parameterize the model; estimate parameters using causal moments from Section 4; and show the model matches important untargeted moments, including the cross-sectional effect of scaled prepayment cost on price growth from Section 3. Section 5.5 presents results from two sets of counterfacutal simulations. We first simulate the impact of the 2021-23 rate increase on equilibrium house prices, holding fixed other factors influencing housing demand. We then simulate tightening under counter-factual mortgage contracts, such as adjustable-rate and portable mortgages.

5.1 Model description.

A heterogeneous, fixed population chooses between otherwise undifferentiated owner-occupied and rental housing units. Households occupy one unit at a time, and differ by age, financial wealth, housing debt, and current unit type. Units are in fixed supply.

In each period, credit-constrained, risk-averse households make two dynamic choices under uncertainty. First, they choose non-housing consumption, financed with current income and financial wealth. Second, they make a discrete housing choice of whether to stay in their current unit, switch to a different owner-occupied unit, or switch to a different rental unit. Purchases of owner-occupied units are financed with non-assumable, amortizing fixed-rate mortgages. Households face uncertainty over the interest rate, rents, and house prices. Households are also uncertain over how preferences for housing units changes over time. Households live to age $\overline{A}$, and have preferences over bequests left after death.

In each period, house prices adjust to equate net moves between the owner-occupied and rental housing market. This price depends on the household state distribution, the risk-free rate, and rents. Changes in the risk-free rate affect housing demand by changing the rate on saving and borrowing. Outstanding fixed-rate mortgages are valuable when rates rise because they allow households to borrow at below-market rates, attenuating the negative effect of rising rates on housing demand. The model is partial equilibrium: while owner-occupied house prices are endogenous, interest rates, income, and rents are exogenous.

Household problem overview. In each discrete period t, households i make discrete housing choice d_{it} and continuous non-housing consumption choice c_{it}. Housing choice is between the current unit $d_{it} = 0$, a different rental unit $d_{it} = \ell$, and a different owner-occupied unit $d_{it} = o$.

Housing delivers flow utility $h(a_{it}, d_{it}, k_{it}, \boldsymbol{\varepsilon}_{it}, \chi_{it})$, a function of age a_{it}, housing tenure $k_{it} \in \{o, \ell\}$ ($o$wn or ℓease), and idiosyncratic household preferences $(\boldsymbol{\varepsilon}_{it}, \chi_{it})$, detailed below. Utility depends on both current status k_{it} and housing choice d_{it} to encode moving costs. Non-housing consumption delivers flow utility $u(c_{it})$. Households have preferences $v^b(W_{i,t+1})$ over bequests any financial wealth $W_{i,t+1}$ remaining after liquidation of any housing assets. The sequence of choices are assumed to maximize the lifetime expected present discounted value of additively-separable flow utility from housing and non-housing consumption:

$$E_t \left[\sum_{q \geq t}^{t+\bar{A}-a_{it}} \delta^{q-t} \left[h(a_{iq}, d_{iq}, k_{iq}, \boldsymbol{\varepsilon}_{iq}, \chi_{iq}) + u(c_{iq}) \right] + \delta^{\bar{A}-a_{iq}+1} v^b(W^b_{i,\bar{A}-a_{iq}+1}) \right] \tag{13}$$

Choices depend on the household state s_{it}, aggregate market conditions θ_t, and idiosyncratic household preferences $(\boldsymbol{\varepsilon}_{it}, \chi_i)$. The household state vector is $s_{it} := (a_{it}, k_{it}, W_{it}, M_{it}, m_{it})$ for mortgage balances M_{it} and mortgage payment m_{it}. The vector of aggregate market conditions is $\theta_t := (r_t, \ell_t, P_t)$, where r_t is the risk-free return on savings, ℓ_t is rents, and P_t is the purchase price of owner-occupied housing.

Households may place different value on the characteristics of owner-occupied and rental units. Some households might prefer owner-occupied units for their larger size, location in suburban markets with family-oriented amenities, and greater customization, while others might prefer the urban amenities and low sweat equity of rental units. Given that it is likely persistent, an ideal approach would model heterogeneous preferences for each choice as a Markov process, adding at least one unobserved state variable for each housing choice. We adopt a tractable alternative that retains some of the important features of the more realistic approach.

Specifically, the iid binomial variable χ_{it} encodes whether i considers a move in period t, while $\boldsymbol{\varepsilon}_{it} = \{\varepsilon_{idt}\}_d \overset{\text{iid}}{\sim} G_{k_{it}}$ describes the unobserved value household i places on choice d, conditional on considering a move. This conceptually represents the arrival of some change in household circumstances – for example, marriage, divorce, birth of a child, or health shocks – that leads households to consider a move. The formal structure produces persistence in household choice without having to track additional state variables. A high draw of ε_{idt} will have persistent effects in future periods so long as $\chi_{iq} = 0$ for $q > t$.

The model is in real dollars, with the exception of interest rates, which are nominal. We assume a constant rate of inflation ι.

Mortgage contracts, household state transitions, and constraints. Households finance home purchases using a fixed-payment mortgage with origination loan-to-value ratio of $1-\phi_a$.

For a mortgage originated in period t:

$$M_{it} = (1-\phi_a)P_t, \quad m_{it} = A(r_{mt}, (1-\phi_a)P_t) \tag{14}$$

where r_{mt} is the mortgage interest rate and $A(r, M) := M \cdot \frac{r_{mt}(1+r_{mt})^n}{(1+r_{mt})^n - 1}$ is the amortization formula for n the number of payments. Thereafter:

$$(1+\iota) \cdot M_{i,t+1} = M_{it} - \nu \cdot m_{it}, \quad (1+\iota) \cdot m_{i,t+1} = \mathbb{1}(M_{i,t+1} > 0) \cdot m_{it} \tag{15}$$

The mortgage balance transition depends on ν, rather than the origination interest rate. This eliminates a state variable, at the cost of slightly changing the convexity of the mortgage repayment schedule.[47] We pick ν so that a borrower with a fully-amortizing 30-year mortgage rate will repay within 30 years at the steady-state of the interest rate process, detailed below.

Households face credit constraints that prevent borrowing aside from mortgages: $W_{it} \geq 0$. Households spend financial wealth on housing and non-housing consumption, and save what remains in risk-free asset. Financial wealth after borrowers make their housing choice is:

$$W'_{it} = \underbrace{W_{it} + y_a}_{\text{resources}} - \underbrace{\mathbb{1}_{\text{own},t} \cdot (m_{it} + \tau_{ft}) - \mathbb{1}_{\text{lease},t} \cdot \ell_t}_{\text{per-period housing costs}} + \underbrace{\mathbb{1}_{\text{sell},t} \cdot ((1-\tau_s)P_t - M_{it}) - \mathbb{1}_{\text{buy},t} \cdot \phi_a P_t}_{\text{capital gains net down pmt}} \tag{16}$$

Household resources include financial wealth and age-specific income, y_a. Per-period housing costs include property taxes and maintenance τ_{ft}. If a household decides to sell their current house, they receive the market price of their home less an ad valorem transactions cost τ_s representing fees paid to real estate agents, and must repay their outstanding mortgage balance at face value. Finally, a household makes a down payment when they buy a new home. Financial wealth earns r_t after households choose consumption:

$$W_{i,t+1} = (1 + r_{t+1} - \iota) \cdot \left(W'_{it} - c_{it}\right) \tag{17}$$

Define s'_{it} as equal to s_{it}, with W_{it} replaced with W'_{it}. In the terminal period when $a_{it} = \bar{A}$, bequests are $W^b_{it} := \max(0, W_{i,t+1} + (1-\tau_s)P_{t+1} - M_{i,t+1})$.

Location k_{it} transitions based on d, with $k_{i,t+1} = k_{it}$ if $d_{it} = 0$ and $k_{i,t+1} = d_{it}$ otherwise.

Timing and recursive formulation of household problem. When the period begins, households observe state $(s_{it}, \theta_t, \boldsymbol{\varepsilon}_{it}, \chi_{it})$. They make housing choices if $\chi_{it} = 1$, taking expectations over the evolution of aggregate state variables, which households expect to follow a first-order Markov process (detailed below). θ_{t+1} is realized, and households make non-housing con-

[47] This change also impacts loan term, but in practice the change in term is less than our calibrated time step.

sumption choices, taking expectations over future realizations of idiosyncratic preferences.[48] Finally, the period advances and $\chi_{i,t+1}, \boldsymbol{\varepsilon}_{i,t+1}$ are drawn.

The optimal choice of d and c satisfy the following Bellman equation, where the value function $V(\cdot)$ represents the present-value of expected lifetime utility:

$$V(s_{it}, \theta_t, \boldsymbol{\varepsilon}_{it}, \chi_{it}) = \begin{cases} h(a_{it}, d_{it} = 0, k_{it}, \boldsymbol{\varepsilon}_{it}) + E_\theta \left[V^c(s'_{it}, \theta_{t+1}) | s_{it}, \theta_t \right] & \text{if } \chi_{it} = 0 \\ \max_d \; h(a_{it}, d, k_{it}, \boldsymbol{\varepsilon}_{it}) + E_\theta \left[V^c(s'_{it}, \theta_{t+1}) | s_{it}, \theta_t, d \right] & \text{if } \chi_{it} = 1 \end{cases} \tag{18}$$

$$\text{s.t. } W'_{it} \geq 0, \; (14), (15), (16);$$

$$V^c(s'_{it}, \theta_{t+1}) = \begin{cases} \max_{c \geq 0} \; u(c) + E_{\varepsilon,\chi} \left[V(s_{i,t+1}, \theta_{t+1}, \boldsymbol{\varepsilon}_{it}, \chi_{i,t+1}) | s'_{it}, c \right] & \text{if } a_{it} < \overline{A} \\ \max_c \; u(c) + v^b(W^b_{it}) & \text{if } a_{it} = \overline{A} \end{cases} \tag{19}$$

$$\text{s.t. } W_{i,t+1} \geq 0, \; a_{i,t+1} = a_{it} + 1, \; (17).$$

Let $d^*(s, \theta, \boldsymbol{\varepsilon}, \chi)$ denote the optimal location choice.

Equilibrium. The risk-free rate r_t and lease costs ℓ_t follow a correlated first-order Markov process. Owner-occupied house prices adjust in equilibrium to equalize total flows to and from owner-occupied and rental housing. Let N_o, N_ℓ denote the fixed number of owner-occupied and rental units; $\pi_{dk}(P; s, r, \ell) := \Pr_{\varepsilon,\chi}[d = d^*(s, \theta, \boldsymbol{\varepsilon}, \chi)]$ denote the probability of choosing alternative d from location k given state s, and G_{kt} the distribution of household states in location k. Equilibrium prices satisfy:

$$\underbrace{\sum_{k \in \{o,\ell\}} N_k \int_s \pi_{ok}(P_t; s, r_t, \ell_t) dG_{kt}}_{\text{owner-occupied inflows}} = \underbrace{\sum_{d \neq 0} N_o \int_s \pi_{do}(P_t; s, r_t, \ell_t) dG_{ot}}_{\text{owner-occupied outflows}} \tag{20}$$

Discussion. In our model, households react to a change in the risk-free rate for several reasons. First, higher rates increase the return on savings, making investment in owner-occupied housing less attractive. Second, higher rates increase per-period payments on new mortgages. This discourages new purchases. Third, higher rates reduce the present-value of future income, a negative wealth effect. And finally, higher rates increase the value of existing fixed-rate mortgages, by depressing the present-value of their payments.

Crucially, households care about these effects to the extent they impact the present value of future non-housing consumption. Non-housing consumption in our model endogenously depends on income, liquidity, age, and uncertainty over returns to housing. This makes the

[48] θ_{t+1} is realized after housing but before non-housing consumption choices are made so that households don't know their origination interest rate when making housing choices. This is so that the identifying assumption behind the November 2016 event study, which we will use to estimate preference parameters, holds in the model.

model able to realistically capture *how much* households care about impacts on future non-housing consumption across the state distribution and as financial conditions shift.

5.2 Model parametrization and solution.

Housing consumption. Flow utility from housing consumption is given by:

$$h(a,d,k,\boldsymbol{\varepsilon},\chi) = h_{k(d)} + \mathbb{1}_{d\neq 0}\cdot\lambda_0 + \mathbb{1}_{d\neq 0,k}\cdot\lambda_s + \mathbb{1}_{\chi=1}\cdot\varepsilon_d \tag{21}$$

where $h_{k(d)} \in \{h_o, h_\ell\}$ is the mean value of owner-occupied or rental housing units, which depends on destination $k(d)$; λ_0 is the net benefit of a move; and λ_s is the net benefit of switching housing tenure (e.g., from owning to renting). We assume that $\varepsilon_d \overset{\text{iid}}{\sim}$ nested logit, with $d = 0$ in one nest, $d \neq 0$ in another, and nesting parameter σ.[49] We assume that χ_{it} is distributed Bernoulli with success probability χ.

When $\chi_{it} = 0$, equation (21) equals $h_{k_{it}}$. Otherwise, the value of each housing alternative is state dependent, with λ_0 meant to capture any persistent value of a household's own unit relative to others of the same type k, and λ_s the persistent value of a household's current unit type k relative to the alternative. A positive value of λ indicates a persistent preference for the current relative to alternative type, while a negative value indicates a household is "unmatched" from their current arrangement when the option to move arrives.

Non-housing consumption. We assume that flow utility over non-housing consumption and bequest motives take the following constant relative risk aversion forms:

$$u(c) = \alpha\frac{c^{1-\gamma}}{1-\gamma}, \quad v^b(W) = \varphi_1\left(\frac{W+\varphi_2}{1-\gamma}\right)^{1-\gamma} \tag{22}$$

The functional form for v^b follows De Nardi (2004).

The parameter α controls the relative value of housing relative to non-housing consumption. It is the crucial parameter that determines how much the rate lock incentive and increased user costs impact housing demand, and hence equilibrium prices. When α is high, the decision to move is more sensitive to the relative price of different housing choices. This makes households' moving decisions more dependent on *both* the rate lock incentive *and* the cost of capital. When α is low, moving decisions are driven by demand for housing amenities, and the rate lock and user cost effects matter less.

Aggregate state variables and expectations. The nominal interest rate r_t follows an exogenous single-factor Cox, Ingersoll, and Ross (CIR) process. We assume that mortgage rates equal

[49] For $\sigma = 1$, ε is perfectly correlated across alternatives within nests; for $\sigma = 0$, ε is distributed type-I extreme value.

the implied 30-year fixed rate implied by CIR estimates, plus a fixed spread: $r_{mt} := r_{t,30} + \Delta$. Rental cost growth follows an exogenous process that correlates with the level of and shocks to the interest rate. Households further expect that house price growth follows a first-order Markov process correlated with the level of the interest rate as well as shocks both interest rates and rental cost growth. Appendix Section E.1 describes expectations in more detail.

Our approach to modeling expectations for future house prices departs from full information rational expectations, because as shown in the market clearing equation (20), house prices depend on the full household state distribution. As in Landvoigt *et al.* (2015), our model does not focus on expectations, and so we simplify here to enable more detailed modeling of more central pieces of the household problem. Given the high computational burden that processing the full state distribution would place on households, we view this simplification as behaviorally realistic. As in other papers that simplify by modeling expectations as first-order Markov (e.g. Kaplan *et al.* 2020, Guren *et al.* 2021, Jeon 2022), we ensure that expectations are consistent with the correlations that households have previously observed.

Model solution. We solve the household problem via backwards iteration. For each age a, we first solve for optimal $c(\cdot)$, which gives $V^c(\cdot)$ from equation (19). Given optimal consumption, we form choice-specific value functions $v_{do}(s, \theta_t) := h_{k(d)} + \mathbb{1}_{d\neq 0} \cdot \lambda_0 + \mathbb{1}_{d\neq 0,k} \cdot \lambda_s + E_\theta\left[V^c(s', \theta_{t+1})|d, s, \theta_t\right]$. Given choice-specific value functions, we can solve for equilibrium choice probabilities by integrating over the joint distribution of $\chi_{it}, \boldsymbol{\varepsilon}_{it}$, with $\pi_{do}(\cdot) = E[\chi_i] \cdot \Pr\left(d = \max_{d'} v_{d',o}(\cdot) + \varepsilon_{i,d',o}\right)$.

Key to our approach, we efficiently solve for optimal non-housing consumption using an Euler equation that accounts for future expected moves. Due to the embedded discrete choice over housing and liquidity constraints, the Euler equation is necessary, but not sufficient, for optimal policy. To overcome this, we use an extension of the endogenous grid point method described in Druedahl and Jørgensen (2017), which inverts the Euler equation at endogenous *segments* and takes the upper envelope to solve for optimal non-housing consumption at each state.

5.3 Model estimation.

We calibrate a number of parameters offline, and estimate key preference parameters using full-solution minimum distance (Rust 1987, Gourinchas and Parker 2002, Laibson *et al.* 2024). To capture household responses to rate lock, our estimation targets causal moments from the November 2016 event study from Section 4.

Offline calibration. Appendix Table E.29 summarizes calibrated parameters. We set the model time step to two years. Thus, r_t is the nominal 2-year Treasury yield. We calibrate pref-

erence parameters $\gamma, \delta, \varphi_1, \varphi_2$, and σ, along with the boundaries of the household age profile (with $a \in [22, 82]$), transactions costs τ_s, per-period owner-occupied housing costs τ_{ft}, and mortgage spreads Δ. We use the 2013-2019 Survey of Consumer Finances to estimate the age profile of real income $\{y_a\}$, the age profile of origination LTVs $\{\phi_a\}$, and the household state distribution G_{kt}, N_{kt}. Finally, we estimate aggregate state transition parameters using data on the nominal 2-year Treasury, rent growth, and house price growth. All nominal variables besides the 2019 interest rate are deflated to 2019 dollars with the CPI-U.

Appendix equation (E.2) shows estimates for how interest rates, rent growth, and house price growth correlate. The nominal 2-year Treasury has a steady-state value of 3.9% and autocorrelation coefficient of 0.9. Rent and house price growth correlate negatively with the interest rate, with a greater effect on house prices. Positive shocks to rental cost growth and nominal interest rates also increase house price growth. This is consistent with the identification concern from Section 3 that high rates coincide with high housing demand.

Minimum distance estimation. We estimate the five household preference parameters $\psi := (h_o, \lambda_0, \alpha, \lambda_s, \chi)$ using just-identified minimum distance (normalizing $h_\ell = 0$). We estimate ψ as the value that makes the model exactly match five empirical moments related to the November 2016 event study, taken directly from Section 4.

The first two moments target the probability that a household purchasing a home before November 2016 moves to either (i) another owner-occupied home within 4 years; or (ii) a rental unit within 4 years. The probability of an own-to-own move is most informative about λ_0, as the net benefit of a move is the only average non-financial differentiator across properties. The probability of an own-to-rent move is most informative about h_o, the average value of an owner-occupied relative to rental unit.

To estimate the strength of the rate lock incentive, we target the causal effect of the rate lock incentive on household-level moves. Specifically, we match the impact of the increase in mortgage rates in November 2016 on (i) all moves and (ii) own-to-rent moves within 4 years, where we scale effects by the overall 4-year move rate. The impact on all moves is most informative about α, which controls the relative importance of financial and non-financial factors in housing demand. The impact on own-to-rent moves has information about λ_s; conditional on the other parameters, a higher value of λ_s implies that substitution driven by changing costs of ownership is more likely to induce a change in tenure.

Finally, we target the variance of the total move rate from 2011-2019 to pin down χ. Conditional on other parameters, χ scales responsiveness of household behavior to changes in the economic environment. When χ is high, aggregate moves respond aggressively to changing house prices, interest rates, and rents, and vice versa when χ is low.

Formally, where Ω is the vector of empirical moments and $g(\psi, \Omega)$ a distance function where each element is the difference between the modeled and empirical moment given parameters ψ, the minimum-distance estimate is:[50]

$$\hat{\psi} := \arg\min_{\psi} \; g(\psi, \Omega)' g(\psi, \Omega) \tag{23}$$

Model estimates. Appendix Table E.30 presents $\hat{\psi}$, along with the value and description of the moment primarily used for identification, recognizing that all moments are relevant because parameters are estimated jointly. Estimates indicate that households prefer owner-occupied to rental housing on average, since $\hat{h}_o > 0$. Households also experience a utility cost when they change units, given that $\hat{\lambda}_0 < 0$. Our estimate of $\hat{\chi} = 0.39$ indicates that on average, households consider a move once every 5.2 years.

To quantify the value households place on moving costs and owner-occupied housing, we consider the wealth transfer that would make the average 2019 owner indifferent between their current situation, a forced move to a different owner-occupied unit, and a forced move to a renter-occupied unit. Appendix Figure E.23a plots required compensation by age as a fraction of total wealth (the sum of financial wealth, the present-value of future income, and home equity). First, required compensation for a forced own-to-rent move is generally higher than for an own-to-own move, reflecting the preference for owner-occupied units. Second, required compensation for own-to-own moves is rising in age. Young households, who are credit constrained, value liquidity from a home sale and so are willing to move. By age 60, a household would need a transfer equal to 0.25% of total wealth to compensate for an own-to-own move. Third, compensation required for own-to-rent moves is non-monotonic in age. It rises from around 0.1% at age 30, to 0.8% at age 64, before falling to around 0.5% by age 75. This reflects that older households want to liquidate housing wealth for consumption as they spend down savings in retirement.

5.4 Validation with untargeted moments.

To validate our estimates, we compare our model's predictions untargeted household behaviors and equilibrium house price effects. The predictions for untargeted price effects are especially important, both because our main outcome of interest is predicted price growth during 2021-23 tightening and because we did not directly target house prices or price growth anywhere in our estimation procedure. Price predictions depend on modeled impacts for how housing choice responds to economic conditions.

[50] Since our system is just-identified and we match targeted moments exactly, a weighting matrix is redundant.

Effect of rate lock on moves by age. Appendix Figure E.23b shows that the model reproduces the heterogeneous impact of rate lock on existing-home sales by age. We simulate the November 2016 event study for households that originally purchase a home at different ages, and compare model predictions to estimates from Table 4. The model endogenously replicates lifecycle patterns for the effects of rate lock, with greater impacts for younger households and persistent impacts across the lifecycle. This is both because of the lifecycle structure of the model – income depends on age, and households' decision horizon changes as they approach the end of their life – and because of the empirical correlation between age and financial wealth. Younger households, who are relatively more liquidity constrained, are sensitive to the liquidity effects of rate lock, while older households are sensitive to its net worth effects. This builds confidence that the model captures the main economic forces determining the impact of rate lock on housing choice.

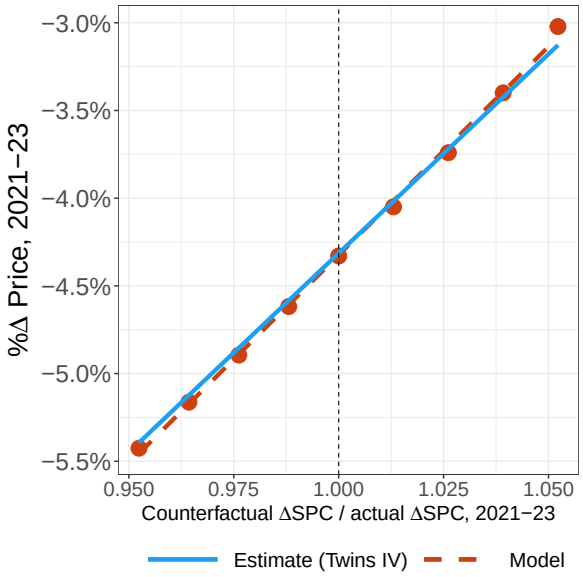
Cross-sectional impact of rate lock on price growth. Figure 7a shows the model predicts cross-sectional price effects of scaled prepayment cost on equilibrium price growth from 2021-23 that align with IV estimates from Section 3. We simulate the model's predictions for equilibrium house price growth due to the 2021-23 increase in the 2-year Treasury rate, given the household state distribution in 2021. We then perturb the 2021 distribution of outstanding mortgages to increase or decrease rate lock, and plot equilibrium house price growth from 2021-23 as a function of scaled prepayment cost growth. The modeled effect of scaled prepayment cost on house price growth aligns with the twin IV estimate from Table 1.

Aggregate price changes during last tightening. Appendix Figure E.24 compares model-implied to actual real growth in average real house prices during the last rate tightening. Given our model's 2-year time step, rates rose starting in 2015-2016, and peaked over 2017-2018. We compute the modeled price for for 2015-2016, and show how house prices incrementally change due to the 1.2pp increase in the 2-year Treasury rate, the 3.5% increase in market rents, and changes to the household state distribution (households grew older and wealthier).

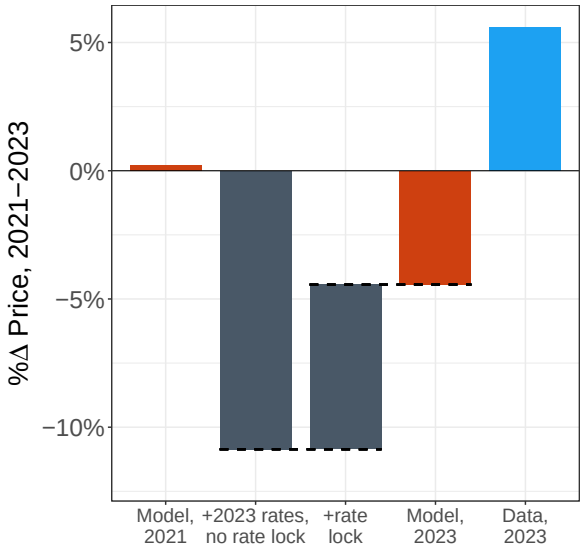
Over this period, real house prices rose by 6.8%. A static user cost calibration would have instead predicted a substantial decrease in house prices due to the changes over this period: for example, using parameters in Himmelberg *et al.* (2005), the elasticity of the real price-to-rent ratio with respect to a change in the 10-year real interest rate is around -19. This implies a 16.1% decrease in real house prices, accounting for modeled changes in the household state distribution, because given our estimates of the short rate process, a 1.2pp increase in the nominal 2-year Treasury yield corresponds to a 1.1pp increase in the real 10-year Treasury yield. Our model is much closer to the data, and predicts an increase of around 1%. It is not surprising that the model undershoots the data, because 2013-2019 includes recovery

Figure 7: Structural model results

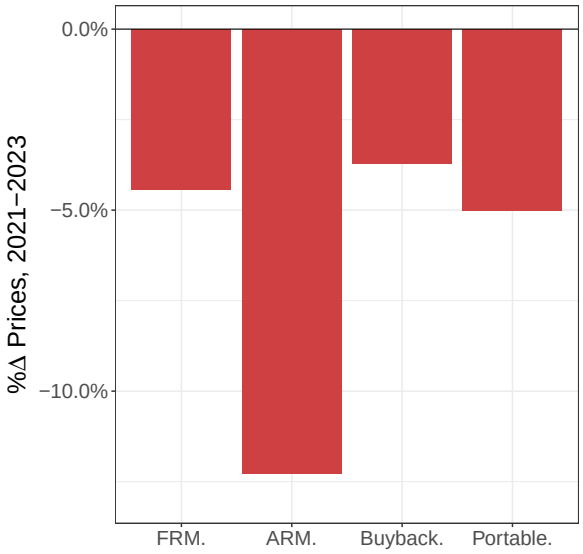
(a) Modeled ΔSPC vs price growth.



(b) Rate lock offsets price declines.



(c) Tightening with counterfactual contracts.



Source: Equifax Credit Risks Insight Servicing and ICE, McDash®. (a) Red dots plot the modeled impact of 2021-23 2-year Treasury increases on house price growth as a function of counterfactual ΔSPC fraction of modeled baseline ΔSPC. The red dashed line is the line of best fit. The blue line plots predictions based on the twins instrument estimates of the impact of scaled prepayment cost on 2021-23 house price growth from Table 1. (b) The figure decomposes the price effect of a 3.6pp increase in the 2-year Treasury rate. The first bar plots equilibrium prices where existing owners are all forced into adjustable-rate mortgages, where the ARM spread is calibrated to keep 2021 prices constant. The second bar shows the effect of the Treasury rate increase where existing owners have fixed-rate mortgages. The third bar shows the net effect. (c) The figure shows the price impact of 2021-23 tightening where all mortgages unexpectedly change from fixed-rate to (i) adjustable-rate, (ii) prepayable at present-value cost (with a buyback option), or (iii) portable.

from the aftermath of the 2008 housing crisis. At the beginning of this period, house prices were depressed for unmodeled reasons, such as foreclosures (Guren and McQuade 2020) and

especially tight leverage constraints (e.g., Di Maggio and Kermani 2017).

5.5 Model results.

The effect of 2021-23 tightening on equilibrium house prices. We use the estimated model to both simulate the effects of the 2021-23 increase in interest rates on house prices, and estimate the incremental impact of rate lock. We fix the household state distribution at 2021 levels, and simulate the price impact of a 3.6pp increase in the 2-year Treasury rate under two scenarios. In the first scenario, all households have fixed-rate mortgages (FRM), as above. In the second, we unexpectedly force all existing owners into adjustable-rate mortgages (ARM), where payments increase with the short rate and we set the ARM spread so that modeled 2021 prices align with those under FRM. Comparing the price effect of the two simulations gives the incremental effect of rate lock.

Figure 7b shows the results. Without rate lock, prices would fall by around 10.9%. Rate lock directly offsets 6.4pp of this decrease. As predicted in the Section 1 framework, this is a meaningful share of the 7.9pp offset predicted using cross-sectional estimates from the twins instrument in Section 3. The model predicts a price decrease of 4.4% due to 2021-23 rate hikes, all else equal. In the data, real house prices grew by 5.6% from 2021-23. Therefore, rate lock can explain around one-third of the gap between house prices if existing owners were fully exposed to user cost effects and observed price growth.

Our model predicts a much smaller price decrease than recent alternative approaches that consider tightening without rate lock. The Federal Reserve (2024)'s April Financial Stability report suggests that prices should have fallen 20% during tightening, based on a statistical model of the historical relationship between house prices, the real 10-year Treasury yield, and market rents to predict equilibrium valuations.[51] Amromin and Eberly (2023) calibrate the impacts of tightening on perfect-foresight steady-state equilibrium house prices in an overlapping generations lifecycle model with variable rate mortgages, and predict that 2021-23 tightening should have decreased house prices by 37%.[52] Therefore, accounting for rate lock is able to close 61-76% of the gap between actual and predicted price appreciation.

Figure 7b holds rents constant, and so can be interpreted as an impact on price-rent ratios. However, real rents grew during 2021-23. Is the combination of higher interest rates and rents enough to explain real price growth? Appendix Figure E.25 shows the answer is no: even accounting for rent growth, the model predicts a 3.7% decrease in owner-occupied house prices, lower than the 5.6% real house price growth observed on average over 2021-23. This

[51] Figure 1.18 shows that market over-valuation relative to the model increased from $\sim$ 5% in 2020-21 to $\sim$ 30% by 2024, a $\sim$ 25% increase. Since real house prices actually rose by $\sim$ 5% over this period, the Fed's study implicitly predicts a $\sim$ 20% drop in real house prices.

[52] See Table 6, row labeled "House price appreciation relative to peak."

implies other unmodeled factors – e.g., growth in real income, growth in real financial assets, or an increased nonfinacial preference for owner-occupied housing – are needed to explain the net *increase* in recent prices.

Tightening under alternative mortgage contracts. Finally, we simulate the impact of tightening under alternative mortgage designs. In 2021, we unexpectedly force all current owners and new buyers into alternative contracts to the FRM in our baseline model.

We consider three alternatives: ARMs described above, mortgages with a present-value repayment option, and portable mortgages. Mortgages with a present-value repayment option allow borrowers to repay their mortgage at current market values if the option to refinance is out-of-the-money (called the "buyback" or "delivery" option in the Danish mortgage system). Portable mortgages can be used to finance the purchase of a new home (precluded by due-on-sale clauses in the US, but standard in countries such as Canada and the UK). We hold fixed mortgage origination spreads Δ for these counterfactual mortgage contracts.[53]

Figure 7c compares the price effects of a 3.6pp increase in the 2-year Treasury from 2021-23 under each mortgage regime. With ARMs, prices fall by 12.3%. The rate lock effect disappears, causing existing owners to move out of their current homes.[54]

With mortgages prepayable at present-value costs, the price change is similar to under FRM. The reason is because a switch to the new contract has two offsetting effects on house prices. On the one hand, the contract eliminates the rate lock incentive, a force for relatively *lower* prices after an increase in interest rates. On the other, the delivery option allows owners to *realize* their otherwise-unrealized capital gain on their mortgage. This increases their wealth, which in our model, increases the demand for owner-occupied housing due to the diminishing marginal returns to non-housing consumption. Empirically, these forces roughly offset.

Portable mortgages have a similar impact on prices as fixed-rate mortgages. This is intuitive. Rate lock in FRMs creates an implicit cost to leaving the current unit. Portable mortgages create an implicit cost to leaving the owner-occupied housing market more broadly. Since exit from the owner-occupied market is what impacts prices in the aggregate, the rate lock incentive in portable fixed-rate mortgages has a similar impact on aggregate prices as non-portable fixed-rate mortgages.

[53] We make this assumption for parsimony and because our model does not endogenize mortgage pricing. But this assumption is likely a reasonable approximation. In the baseline equilibrium, lenders set lower mortgage rates because they anticipate a windfall from prepayment at par from movers with high prepayment costs. Berger *et al.* (2026) argues that because high prepayment costs are *ex-ante* low probability, and moves with high prepayment costs are endogenously low, this price impact is also low. In their model, switching to the Danish mortgage system in the US would only raise equilibrium mortgage rates by 3 basis points (3% of our calibrated spread) on average.

[54] The negative price effect is greater than in Figure 7b, where we only forced existing-owners, not buyers, into adjustable-rate contracts. Buyers prefer FRM to ARM, magnifying the negative price effect.

6 Conclusion

This paper provides theory and evidence to study how fixed-rate, non-assumable mortgages affect the housing market response to financial tightening. At the household level, we provide causal evidence that rate lock discourages existing owners from exiting the market for owner-occupied homes. At the local housing market level, we develop new instruments for household mobility to show a causal effect of these household behaviors on equilibrium house price growth during tightening. We present an empirical model of dynamic housing demand to understand the impact of rate lock on aggregate market equilibrium. We estimate model parameters using household-level substitution patterns, and verify that the model generates an effect of rate lock in the cross section that aligns with our IV estimates.

We show that rate lock has a significant impact on equilibrium house prices. Reduced-form estimates imply that the rate lock incentive can explain about 30% of the cross-sectional variation in market-level house price growth over 2021-23, and in the aggregate led to 8pp higher nominal house prices in 2023. Simulations from a model estimated with the impact of rate lock on household-level moves give a similar conclusion: all else equal, real house prices would have been 7.9pp lower in 2023 with adjustable-rate mortgages and no rate lock, compared to with fixed-rate mortgages. Although rate lock has a large impact on house prices, it does not fully offset the negative house price effects of higher user costs due to financial tightening. The model predicts that all else equal, 2021-23 tightening would reduce real house prices by 4.4%.

Our findings imply that fixed-rate mortgages reduce the impact of tighter monetary policy on house prices. This makes both monetary policy and housing affordability path dependent: a long era of low rates makes tightening less effective in rapidly curbing house price growth.

References

ABEL, J. (2024). Mortgage lock-in and home sales volume dynamics.

ABRAMSON, B., HAN, L. and DE LLANOS, P. (2025). Monetary policy and rents. *Available at SSRN 5353365*.

ADELINO, M., SCHOAR, A. and SEVERINO, F. (2025). Credit supply and house prices: Evidence from mortgage market segmentation. *Journal of Financial Economics (forthcoming)*, **163**, 103958.

AGARWAL, S., DRISCOLL, J. C. and LAIBSON, D. I. (2013). Optimal mortgage refinancing: a closed-form solution. *Journal of Money, Credit and Banking*, **45** (4), 591–622.

AKCAM, B. K., KIBIS, E. and AKCAM-KIBIS, Z. (2024). Analyzing the initial reactions to national association of realtors settlement on broker commissions in 2024. *Journal of Real Estate Practice and Education*, **26** (1), 2377873.

ALADANGADY, A., KRIMMEL, J. and SCHARLEMANN, T. (2025). Locked in: Mobility, market tightness, and house prices.

ALPHASENSE (2025). Former manager at home depot sees post-covid trends favoring floor & decor for tile projects over home depot. Interview, interview conducted by AlphaSense.

AMROMIN, G. and EBERLY, J. (2023). *Macro shocks and housing markets*. Tech. rep., Working paper.

ANDERSEN, S., BADARINZA, C., LIU, L., MARX, J. and RAMADORAI, T. (2022). Reference dependence in the housing market. *American Economic Review*, **112** (10), 3398–3440.

ANDERSSON, F. and MAYOCK, T. (2014). How does home equity affect mobility? *Journal of Urban Economics*, **84**, 23–39.

ANDREWS, I., STOCK, J. H. and SUN, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics*, **11** (1), 727–753.

ANENBERG, E. and RINGO, D. (2022). The propagation of demand shocks through housing markets. *American Economic Journal: Macroeconomics*, **14** (3), 481–507.

ATTOM (2020). U.s. property taxes levied on single family homes in 2019 increased to more than $306 billion. https://www.attomdata.com/news/market-trends/attom-data-solutions-2019-property-tax-report/, accessed: 2026-01-02.

AUTOR, D., DORN, D. and HANSON, G. (2019). When work disappears: Manufacturing decline and the falling marriage market value of young men. *American Economic Review: Insights*, **1** (2), 161–178.

BATZER, R., COSTE, J., DOERNER, W. and SEILER, M. (2024). *The Lock-In Effect of Rising Mortgage Rates*. Tech. rep., Federal Housing Finance Agency.

BAUM-SNOW, N. and HAN, L. (2024). The microgeography of housing supply. *Journal of Political Economy*, **132** (6), 1897–1946.

BAYER, P., MCMILLAN, R., MURPHY, A. and TIMMINS, C. (2016). A dynamic model of demand for houses and neighborhoods. *Econometrica*, **84** (3), 893–942.

BERGER, D., JEONG, J., MARX, J., OLESEN, M. V. and TOURRE, F. (2026). A danish fix for us mortgage lock-in?

—, MILBRADT, K., TOURRE, F. and VAVRA, J. (2021). Mortgage prepayment and path-dependent effects of monetary policy. *American Economic Review*, **111** (9), 2829–2878.

BERNSTEIN, A. and STRUYVEN, D. (2022). Housing lock: Dutch evidence on the impact of negative home equity on household mobility. *American Economic Journal: Economic Policy*, **14** (3), 1–32.

BONANNO, J. F. (1971). Due on sale and prepayment clauses in real estate financing in california in times of flucuating interest rates-legal issues and alternatives. *USFL Rev.*, **6**, 267.

BORUSYAK, K. and HULL, P. (2023). Nonrandom exposure to exogenous shocks. *Econometrica*, **91** (6), 2155–2185.

CAMPBELL, J. R. and HERCOWITZ, Z. (2005). The role of collateralized household debt in macroeconomic stabilization.

CAMPBELL, J. Y. and RAMADORAI, T. (2026). *Household finance in retrospect and prospect*. Tech. rep., National Bureau of Economic Research.

ČASTA, M. (2025). The impact of mortgage rates on the housing market. *Journal of Housing Economics*, p. 102079.

CDC (2004). Assisted reproductive technology success rates (2004).

CENSUS (2021). Census Bureau Announces Changes for 2020 American Community Survey 1-Year Estimates — census.gov. https://www.census.gov/newsroom/press-releases/2021/changes-2020-acs-1-year.html, [Accessed 29-10-2024].

CENTER, T. P. (2024). Who benefits from the mortgage interest deduction and who misses out? Accessed: 2025-10-19.

CHODOROW-REICH, G., GLAESER, E., JARAVEL, X. and LIPTON, A. (2025). Homeownership and the cost of living.

CNN (2025). Portable mortgages: What to know. *CNN*, accessed: 2026-03-24.

CUNNINGHAM, C. R. and ENGELHARDT, G. V. (2008). Housing capital-gains taxation and homeowner mobility: evidence from the taxpayer relief act of 1997. *Journal of urban Economics*, **63** (3), 803–815.

DE LA ROCA, J., GIACOLETTI, M. and LIU, L. (2026). Mortgage rates and rents: Evidence from local mortgage lock-in effects. *Available at SSRN 4968066*.

DE NARDI, M. (2004). Wealth inequality and intergenerational links. *The Review of Economic Studies*, **71** (3), 743–768.

DELGADO, M. and GRAVELLE, T. (2023). Central bank asset purchases in response to the covid-19 crisis. bis cgfs papers no 68.

DI MAGGIO, M. and KERMANI, A. (2017). Credit-induced boom and bust. *The Review of Financial Studies*, **30** (11), 3711–3758.

DRUEDAHL, J. and JØRGENSEN, T. H. (2017). A general endogenous grid method for multi-dimensional models with non-convexities and constraints. *Journal of Economic Dynamics and Control*, **74**, 87–107.

ESRI (2023). Usa zip code areas. https://www.arcgis.com/home/item.html?id=8d2012a2016e484dafaac0451f9aea24.

FANNIE MAE (2025). How to build your maintenance and repair budget. https://yourhome.fanniemae.com/own/how-build-your-maintenance-and-repair-budget, accessed: 2026-01-02.

FAVILUKIS, J., LUDVIGSON, S. C. and VAN NIEUWERBURGH, S. (2017). The macroeconomic effects of housing wealth, housing finance, and limited risk sharing in general equilibrium. *Journal of Political Economy*, **125** (1), 140–223.

FEDERAL RESERVE (2024). *Financial Stability Report*. Tech. rep., Board of Governors of the Federal Reserve System.

FEDERAL RESERVE BOARD (2023). Transcript of chair powells press conference, july 26, 2023. U.S. Federal Reserve, Media Center, accessed October 13, 2025. https://www.federalreserve.gov/mediacenter/files/FOMCpresconf20230726.pdf.

FERREIRA, F., GYOURKO, J. and TRACY, J. (2010). Housing busts and household mobility. *Journal of urban Economics*, **68** (1), 34–45.

FONSECA, J. and LIU, L. (2024). Mortgage lock-in, mobility, and labor reallocation. *The Journal of Finance*.

—, — and MABILLE, P. (2024). *Unlocking mortgage lock-in: Evidence from a spatial housing ladder model*. Tech. rep., Working paper.

FOOTE, C. L., LOEWENSTEIN, L. and WILLEN, P. S. (2021). Cross-sectional patterns of mortgage debt during the housing boom: evidence and implications. *The Review of Economic Studies*, **88** (1), 229–259.

FOSTEL, A. and GEANAKOPLOS, J. (2016). Financial innovation, collateral, and investment. *American Economic Journal: Macroeconomics*, **8** (1), 242–284.

FOUNT, H. and OUNDEE, O. S. (2020). *Homeowner Mobility and Impediments to Moving Housing Insights*. Tech. rep., Fannie Mae Research and Insights.

FRENCH, R. and GILBERT, V. (2023). Suvbburban housing and urban affordability: Evidence from residential vacancy chains.

FROST, R. (2023). Did more people move during the pandemic? *Harvard Joint Center for Housing Studies Working Paper Series*.

GAMBER, W., GRAHAM, J. and YADAV, A. (2023). Stuck at home: Housing demand during the covid-19 pandemic. *Journal of Housing Economics*, **59**, 101908.

GARRIGA, C. and HEDLUND, A. (2020). Mortgage debt, consumption, and illiquid housing markets in the great recession. *American Economic Review*, **110** (6), 1603–1634.

GERARDI, K., QIAN, F. and ZHANG, D. (2024). Mortgage lock-in, lifecycle migration, and the welfare effects of housing market liquidity. *Lifecycle Migration, and the Welfare Effects of Housing Market Liquidity (July 28, 2024)*.

GLAESER, E. L., GOTTLIEB, J. D. and GYOURKO, J. (2012). Can cheap credit explain the housing boom? In *Housing and the financial crisis*, University of Chicago Press, pp. 301–359.

GOLDBERG, J. (2020). Liquidity supply by broker-dealers and real activity. *Journal of Financial Economics*, **136** (3), 806–827.

GOODMAN, L. and BAI, B. (2017). The impact of higher interest rates on the mortgage market. *The Journal of Structured Finance*, **23** (3), 45–55.

GORDON, G. (2023). Mortgage spreads and the yield curve. *Richmond Fed Economic Brief*, **23** (27).

GOURINCHAS, P.-O. and PARKER, J. A. (2002). Consumption over the life cycle. *Econometrica*, **70** (1), 47–89.

GREENWALD, D. L. and GUREN, A. (2024). *Do credit conditions move house prices?* Tech. rep., National Bureau of Economic Research.

GUREN, A. M., KRISHNAMURTHY, A. and MCQUADE, T. J. (2021). Mortgage design in an equilibrium model of the housing market. *The Journal of Finance*, **76** (1), 113–168.

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

GÜRKAYNAK, R. S., SACK, B. and WRIGHT, J. H. (2007). The us treasury yield curve: 1961 to the present. *Journal of monetary Economics*, **54** (8), 2291–2304.

GUYER, B., MARTIN, J. A., MACDORMAN, M. F., ANDERSON, R. N. and STROBINO, D. M. (1997). Annual summary of vital statistics1996. *Pediatrics*, **100** (6), 905–918.

HAUSMAN, C. and RAPSON, D. S. (2018). Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics*, **10** (1), 533–552.

HENDERSON, J. V. and IOANNIDES, Y. M. (1983). A model of housing tenure choice. *The American Economic Review*, **73** (1), 98–113.

HENNING-SMITH, C., TUTTLE, M., SWENDENER, A., LAHR, M. and YAM, H. (2023). Differences in residential stability by rural/urban location and socio-demographic characteristics. *University of Minnesota Rural Health Research Center*.

HIMMELBERG, C., MAYER, C. and SINAI, T. (2005). Assessing high house prices: Bubbles, fundamentals and misperceptions. *Journal of Economic Perspectives*, **19** (4), 67–92.

HOEKSTRA, C., ZHAO, Z. Z., LAMBALK, C. B., WILLEMSEN, G., MARTIN, N. G., BOOMSMA, D. I. and MONTGOMERY, G. W. (2008). Dizygotic twinning. *Human reproduction update*, **14** (1), 37–47.

HUBBARD, R. G. and MAYER, C. J. (2009). The mortgage market meltdown and house prices. *The BE Journal of Economic Analysis & Policy*, **9** (3).

HUMAN FERTILITY DATABASE (2025). Max planck institute for demographic research (germany) and vienna institute of demography (austria).

HUTCHINSON, M. F. (1989). A stochastic estimator of the trace of the influence matrix for laplacian smoothing splines. *Communications in Statistics-Simulation and Computation*, **18** (3), 1059–1076.

III (2025). Average premiums for homeowners and renters insurance by state, 2022. https://www.iii.org/table-archive/21407, accessed: 2026-01-02.

JEON, J. (2022). Learning and investment under demand uncertainty in container shipping. *The RAND*

Journal of Economics, **53** (1), 226–259.

JUSTINIANO, A., PRIMICERI, G. E. and TAMBALOTTI, A. (2022). The mortgage rate conundrum. *Journal of Political Economy*, **130** (1), 121–156.

KAPLAN, G., MITMAN, K. and VIOLANTE, G. L. (2020). The housing boom and bust: Model meets evidence. *Journal of Political Economy*, **128** (9), 3285–3345.

KEYS, B. J., POPE, D. G. and POPE, J. C. (2016). Failure to refinance. *Journal of Financial Economics*, **122** (3), 482–499.

KIM, G. H. (2024). The equilibrium impacts of broker incentives in the real estate market.

KLADÍVKO, K. (2007). Maximum likelihood estimation of the cox-ingersoll-ross process: the matlab implementation. *Technical Computing Prague*, **7** (8), 1–8.

KLEIBERGEN, F. (2007). Generalizing weak instrument robust iv statistics towards multiple parameters, unrestricted covariance matrices and identification statistics. *Journal of Econometrics*, **139** (1), 181–216.

KOTOVA, N. and ZHANG, A. L. (2020). Search frictions and idiosyncratic price dispersion in the us housing market. *Available at SSRN 3386353*.

KUTTNER, K. N. (2014). Low interest rates and housing bubbles: still no smoking gun. *The role of central banks in financial stability: How has it changed*, **30**, 159–185.

LAIBSON, D., LEE, S. C., MAXTED, P., REPETTO, A. and TOBACMAN, J. (2024). Estimating discount functions with consumption choices over the lifecycle. *The Review of Financial Studies*, p. hhae035.

LANDVOIGT, T., PIAZZESI, M. and SCHNEIDER, M. (2015). The housing market (s) of san diego. *American Economic Review*, **105** (4), 1371–1407.

LIEBERSOHN, J. and ROTHSTEIN, J. (2024). *Household mobility and mortgage rate lock*. Tech. rep., National Bureau of Economic Research.

LOUTSKINA, E. and STRAHAN, P. E. (2015). Financial integration, housing, and economic volatility. *Journal of Financial Economics*, **115** (1), 25–41.

MAYER, C. J. and SINAI, T. (2007). *Housing and behavioral finance*. Wharton School, University of Pennsylvania.

MIAN, A. and SUFI, A. (2022). Credit supply and housing speculation. *The Review of Financial Studies*, **35** (2), 680–719.

MINTON, R. and MULLIGAN, C. B. (2024). *A Market Interpretation of Treatment Effects*. Tech. rep., National Bureau of Economic Research.

NEWMAN, S. and RESCHOVSKY, J. (1987). An evaluation of the one-time capital gains exclusion for older homeowners. *Real Estate Economics*, **15** (1), 704–724.

OLEA, J. L. M. and PFLUEGER, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, **31** (3), 358–369.

PARK, K. A. (2022). What happens when you assume. *Cityscape*, **24** (3), 317–338.

PLEMMONS, R. J. (1977). M-matrix characterizations. inonsingular m-matrices. *Linear Algebra and its applications*, **18** (2), 175–188.

POLLOCK, A. J. and KUPIEC, P. H. (2026). How to end mortgage lock-in and get americans moving again. *Wall Street Journal*, accessed: 2026-03-24.

POTERBA, J. M. (1984). Tax subsidies to owner-occupied housing: an asset-market approach. *The quarterly journal of economics*, **99** (4), 729–752.

PRACTICE COMMITTEE OF THE SOCIETY FOR ASSISTED REPRODUCTIVE TECHNOLOGY (2006). Guidelines on number of embryos transferred. *Fertility and Sterility*, **86** (5), S51–S52.

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

— (2002). Transactions costs and housing markets.

ROSSEN, L. M., HAMILTON, B. E., ABMA, J. C., GREGORY, E. C., BERESOVSKY, V., RESENDEZ, A., CHANDRA, A. and MARTIN, J. A. (2023). Updated methodology to estimate overall and unintended pregnancy rates in the united states. *Centers for Disease Control*.

RUST, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, pp. 999–1033.

SAMUELS, P. (2024). Local real estate investors and rent-price dispersion, working Paper.

SHILLER, R. J. (2006). Long-term perspectives on the current boom in home prices. *The Economists' Voice*, **3** (4).

SMITH, M. H. and SMITH, G. (2006). Bubble, bubble, where's the housing bubble? *Brookings Papers on Economic Activity*, **2006** (1), 1–67.

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

STOCK, J. and YOGO, M. (2005). Asymptotic distributions of instrumental variables statistics with many instruments. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, **6**, 109–120.

SUNDERAM, S. (2019). Assisted reproductive technology surveillanceunited states, 2016. *MMWR. Surveillance summaries*, **68**.

SVENSSON, L. E. (1994). Estimating and interpreting forward interest rates: Sweden 1992-1994.

TANDBERG, A., BJØRGE, T., BØRDAHL, P. E. and SKJAERVEN, R. (2007). Increasing twinning rates in norway, 1967–2004: the influence of maternal age and assisted reproductive technology (art). *Acta obstetricia et gynecologica Scandinavica*, **86** (7), 833–839.

VAN NIEUWERBURGH, S. (2023). The remote work revolution: Impact on real estate values and the urban environment: 2023 areuea presidential address. *Real Estate Economics*, **51** (1), 7–48.

WHEATON, W. C. and LEE, N. J. (2009). The co-movement of housing sales and housing prices: Empirics and theory.

WRIGHT, V. C., CHANG, J., JENG, G. and MACALUSO, M. (2006). Assisted reproductive technology surveillance–united states, 2003. *Morbidity and Mortality Weekly report. Surveillance Summaries (Washington, DC: 2006)*.

Online appendix to: "Mortgage Rate Lock and House Prices"

Justin Katz and Robert Minton

This version: March 2026

Table of Contents

A Derivations for Section 1 62

A.1 Derivations for simple framework. . . . 62
A.2 Simple framework based on one-period user costs. . . . 63
A.3 Extension: Multiple market segments. . . . 64
A.4 Once-and-for-all location choice is optimal. . . . 68
A.5 Formulas for equilibrium flows. . . . 69
A.6 Steady-state equilibrium details. . . . 69
A.7 Derivation of equation (A.20). . . . 71
A.8 Proof of Proposition 2. . . . 72
A.9 Proof of Proposition 3. . . . 79

B Data construction details 83

B.1 CRISM and ICE, McDash®details. . . . 83
B.2 CoreLogic details. . . . 84
B.3 Supplemental data sources. . . . 85
B.4 Alternative prepayment cost pricing models. . . . 87

C Section 4 details. 90

C.1 Facts about existing-home sales across the lifecycle. . . . 90
C.2 Confounds from generic origination timing instruments. . . . 91
C.3 November 2016 event study: Robustness checks 92
C.4 Implicit borrower discount rate 95
C.5 Additional figures 96
C.6 Additional tables 98

D Section 3 details. 100

D.1 House price effects of tightening internationally. . . . 100
D.2 Calculating $SPC(\cdot)$ with counterfactual mortgage terms. . . . 101
D.3 Origination timing fixed effects controls. . . . 102
D.4 Proof of Proposition 1 102
D.5 Twin birth rate instrument construction. . . . 104

D.6 Standard deviation calculations in Figure 3. . . . 105
D.7 Calibrating twin instrument relevance and power. . . . 106
D.8 First births instrument construction details. . . . 110
D.9 Twins robustness. . . . 110
D.10 November 2016 event study. . . . 113
D.11 Alternative prepayment cost measurement. . . . 115
D.12 Interpreting estimates. . . . 118
D.13 Effects on residential investment. . . . 118
D.14 Effects on rents. . . . 119
D.15 Additional figures. . . . 123
D.16 Additional tables. . . . 126

E Section 5 details. 134
E.1 Household expectations . . . 134
E.2 Offline calibration. . . . 134
E.3 Additional figures and tables. . . . 136

A Derivations for Section 1

A.1 Derivations for simple framework.

Equation (2). To derive ω, differentiate equation (1) with respect to PC:

$$D'(\cdot)\frac{\partial P}{\partial PC} = S'(\cdot)\left(\frac{\partial P}{\partial PC} - 1\right) + Q'_S(\cdot)\frac{\partial P}{\partial PC} \tag{A.1}$$

$$\implies \frac{\partial P}{\partial PC} = \frac{\partial S'(\cdot)}{\partial S'(\cdot) - D'(\cdot) + Q'_S(\cdot)} = \omega \tag{A.2}$$

Totally differentiating $S = S(\cdot)$ with respect to PC:

$$\frac{\mathrm{d}}{\mathrm{d}PC}S\left(P - PV[\text{Rent}] - PC\right) = S'(\cdot)\left(\frac{\partial P}{\partial PC} - 1\right) = S'(\cdot)\underbrace{(\omega - 1)}_{\leq 0} \leq 0 \tag{A.3}$$

where the sign follows because $\omega \leq 1$ due to the fact that $S'(\cdot) - D'(\cdot) \geq S'(\cdot)$ and $Q'_S(\cdot) \geq 0$.

$\frac{\partial P}{\partial PC}$ unaffected by greater cross-market substitutability. Substitute $S'(\cdot) = \alpha \cdot S_0(\cdot)$ and $D'(\cdot) = \alpha \cdot D_0(\cdot)$ into equation (2):

$$\omega = \frac{\alpha \cdot S'_0(\cdot)}{\alpha \cdot S'_0(\cdot) - \alpha \cdot D'_0(\cdot) + Q'_S(\cdot)} = \frac{S'_0(\cdot)}{S'_0(\cdot) - D'_0(\cdot) + Q'_S(\cdot)/\alpha} \implies \frac{\partial \omega}{\partial \alpha} \geq 0 \tag{A.4}$$

Equation (3). Differentiate equation (1) with respect to r:

$$D'(\cdot)\left(\frac{\partial P}{\partial r}-\frac{\partial PV[\text{Rent}]}{\partial r}\right)=S'(\cdot)\left(\frac{\partial P}{\partial r}-\frac{\partial PV[\text{Rent}]}{\partial r}-\frac{\partial PC}{\partial r}\right)+ Q_S'(\cdot)\left(\frac{\partial P}{\partial r}-\frac{\partial PV[\text{Rent}]}{\partial r}\right) \tag{A.5}$$

$$\implies \frac{\partial P}{\partial r}=\frac{\partial PV[\text{Rent}]}{\partial r}+\frac{S'(\cdot)}{S'(\cdot)-D'(\cdot)+Q_S'(\cdot)}\cdot\frac{\partial PC}{\partial r} \tag{A.6}$$

$$=\frac{\partial PV[\text{Rent}]}{\partial r}+\omega\frac{\partial PC}{\partial r} \tag{A.7}$$

A.2 Simple framework based on one-period user costs.

Suppose that household's decision to stay or move depends not on the relative present-value cost of housing, but on the user cost of housing relative to spot rents. For owners, user costs equal the opportunity cost of capital invested in home equity plus mortgage payments. Assume that mortgages are fixed-rate perpetuities, so a mortgage originated with interest rate r and balance M has monthly payment rM. Denote rents as ℓ (for "lease" costs), and assume they are constant over time.

Net demand for owner-occupied units comes from renters who purchase a home. The difference in the user cost of owning and renting for households who initially rent is $rP-\ell$. Denote the mass of households moving from renting to owning as $D_u(rP-\ell)$, with $D'<0$.

Net supply of owner-occupied units comes from owners moving to rent. The difference in the user cost of owning and renting is $r(P-M)+r_0M-\ell$, the sum of the opportunity cost of capital on home equity and fixed-rate mortgage payments less rents. Denote the mass of households moving from owning to renting as $S_u(rP-(r-r_0)M-\ell)$, with $S'>0$.

Rearranging the difference in user costs:

$$r(P-M)+r_0M-\ell=rP-r\left(M-\frac{r_0M}{r}\right)-\ell=rP-r\cdot PC-\ell \tag{A.8}$$

with "prepayment cost" $PC:=M-PV[\text{Pmt}(r_0)]=M-\frac{r_0M}{r}$. As in the main text, the prepayment cost is the gap between the mortgage balance households must repay upon move and the present-value cost of fixed-rate mortgage payments they would make if they stayed. Prepayment costs reduce the user cost of owning for current owner-occupants by $r\cdot PC=(r-r_0)M$: when $r_0<r$, the current fixed-rate mortgage delivers an implicit flow dividend equal to the gap between the outstanding and current market rate $r-r_0$.

As in the main text, the mass of conversions from rental to owner-occupied units is given by $Q_s(P-PV[\text{Rent}])$, where $Q_S'>0$ and $PV[\text{Rent}]=\ell/r$.

Equation (1) becomes:

$$D(rP-\ell)=S(rP-r\cdot PC-\ell)+Q_S(P-\ell/r) \tag{A.9}$$

Higher PC increases equilibrium house prices:

$$\frac{\partial P}{\partial PC}=\frac{S_u'(\cdot)}{S_u'(\cdot)-D_u'(\cdot)+Q_S'(\cdot)/r}:=\omega_u\geq 0 \tag{A.10}$$

As in the main text, (i) $\frac{\partial P}{\partial PC}$ is increasing in the effect of prepayment costs on own-to-rent moves; and (ii) $\frac{\partial P}{\partial PC}$ is decreasing in the elasticity of supply Q_S'.

Furthermore, prepayment cost attenuates the negative effect of higher mortgage rates on house

prices:

$$\frac{\partial P}{\partial r} = \frac{\partial PV[\text{Rent}]}{\partial r} + \omega_u \frac{\partial PC}{\partial r} + \epsilon \tag{A.11}$$

$$\text{where } \epsilon := \frac{S_u'(\cdot)PC - (S_u'(\cdot) - D_u'(\cdot))(P - \ell/r)}{rS_u'(\cdot) - rD_u'(\cdot) + Q_S'(\cdot)} \tag{A.12}$$

This is identical to equation (3) when $\epsilon = 0$ – that is, for changes in r around $r = r_0$ where house prices initially equal dividend yields, with $P = \ell/r$ and $PC = 0$. Otherwise, the formula is conceptually similar, with a slight difference given by ϵ to reflect differences in functional form.

A.3 Extension: Multiple market segments.

We extend the framework from Section 1.1 to the case with an arbitrary number of owner-occupied and rental market segments.

Section A.3.1 sets up a general framework to connect fixed-rate mortgages to dynamic housing demand. For this section, we make simplifying assumptions to derive a closed form solutions, and use results to guide reduced-form analysis in Sections 4 and 3. Proofs start following this section in Appendix A.

A.3.1 General setup.

Time is discrete and indexed by t. A city has $J+1$ markets $j = 0, \ldots J$ with N_{jt} units that provide housing services to households who live for T periods. Units in the first J_ℓ markets are rented, and require periodic payments ℓ_{jt} to landlords. Units in the remaining J_o markets are owner-occupied and trade at price P_{jt}. Bold typeface indicates vectors – e.g., $\boldsymbol{\mathcal{P}}_t := (\boldsymbol{\ell}_t, \boldsymbol{P}_t)$ collects rents and owner-occupied prices across markets. The risk-free interest rate is r_t.

Households. Households are indexed by i and each occupy one unit. Their utility in period τ is defined over non-housing consumption c and housing services h. The value of housing services depends on household-specific preferences for amenities in the market of residence and whether households move in t (e.g., to capture hassle costs). Thus $h_{it} = h(d_{it}, k_{it}, \epsilon_{idt})$, with k_{it} the market where i lives in t; d_{it} the market chosen by i in period t, where $d_{it} \in \{0, \ldots, J\}$ indicates a period-t move and $d_{it} = s$ the choice to stay put; and ϵ_{idt} the household-specific preference for d. Households choose a sequence of c_{it} and d_{it} to maximize:

$$E_\tau \sum_{t=\tau}^{T} \delta^t \cdot u_\tau(c_{it}, h(d_{it}, k_{it}, \epsilon_{idt})) \tag{A.13}$$

Households can access two types of assets. First, they can buy houses in markets with owner-occupied units financed with fixed-rate mortgages that have constant payments for the life of the loan. Mortgages are due-on-sale, meaning remaining balances must be repaid at face value when owners sell. Payments for new mortgages r_t at origination. Households can't borrow against housing wealth except through purchase mortgages. Second, households can participate in the riskfree credit market, where savings between t and $t+1$ yield r_{t+1}.

In $t = 0$, households enter as occupants of a unit in j with a fixed-rate mortgage (if an owner-occupant). Households also are endowed with financial wealth W_{i0} and income y_{it}.

New construction. In each market, a mass of developers can build new units to maximize static profits. Developers either purchase non-residential properties at price $L_j(\cdot)$ to build residential units, or purchase residential units to convert into non-residential property. Let PV_j^S denote the transaction

price of a residential unit in j. The price of a marginal non-residential property depends on its end-of-period stock, equal to the initial stock S less net construction Q_j^S.[55] Developers must be indifferent between these options in equilibrium, so $PV_j^S = L_j(S - Q_j^S)$.

Equilibrium. Aggregating across households, the choice of d_{it} produces moves $\pi_{jkt}(\cdot)$ between markets j and k. Equilibrium prices $\mathcal{P}_t$ equate the number of available units with total demand in each market:

$$\underbrace{\left(\sum_k \pi_{jkt}(\cdot)\right)}_{:=F_{oj},\text{outflows from } j} + \underbrace{Q_{jt}^S(\cdot)}_{\text{net supply}} = \underbrace{\sum_k \pi_{kjt}(\cdot)}_{:=F_{ij},\text{inflows to } j} \quad \forall j \tag{A.14}$$

This model's complexity depends on (i) how households' dynamic demand is derived and (ii) the distribution of household characteristics. We now consider a special case that admits closed-form solutions for housing demand.

A.3.2 Special case: User-cost formulas with fixed-rate prepayment cost.

First, we eliminate uncertainty. Interest rates are constant and weakly below the discount factor: $r_t = r \leq \delta \ \forall t$; rents grow at rate g_j, so that $\ell_{j,t+1} = \ell_{j0} \cdot (1+g_j)^t$; and the path of income y_{it} is known. Second, on preferences and constraints, we assume households are infinitely-lived ($T = \infty$); can freely buy and sell bonds in credit markets, so long as they remain solvent; have time-invariant, iid idiosyncratic preferences for the value of housing services ($\epsilon_{idt} = \epsilon_{id} \overset{\text{iid}}{\sim} H \ \forall t$); and have dollar-denominated flow utility:

$$\frac{u_\tau(c, h(d,k,\epsilon))}{1-\delta} = c + \alpha^{-1}\left(h_d - \mathbb{1}_{d\neq s}\cdot\lambda + \epsilon\right) \tag{A.15}$$

where α^{-1} is the dollar-value of a unit of housing services and λ represents moving costs.

To represent market segmentation, we model two types of households. Fraction γ consider units in all markets. Fraction $1-\gamma$ only consider units in their initial market k_{i0}: $d \in \{s, k_{i0}\}$.

Finally, we model mortgages as fixed-rate perpetuities. All owners in market j have mortgages with interest rate r_{Fj} and balance M_{Fj} requiring periodic payments $r_{Fj}M_{Fj}$. New mortgages are originated at rate $r > r_{Fj}$ so that existing owners have no incentive to refinance.

Households choose a sequence of c_{it} and d_{it} to maximize period-0 utility, given in equation (A.13), subject to the constraint that the present-value of non-housing consumption, rental payments, and new home purchases equals the present-value of income and assets less liabilities.

Deriving housing demand flows. We conjecture that households make once-and-for-all moving choices, with $d_{it} = s$ for $t > 0$, and verify in Appendix A.4 that this is optimal. We then calculate utility conditional on each initial choice d_i for residents in market j, denoted by v_{ijd}. The household chooses d_i to maximize v_{ijd}.

Since $\delta \geq r$, an optimal consumption path has $c_t = 0$ for $t > 0$. Define PV_j as the present-value cost of a unit in j, where $PV_j = \sum_{t>0} \frac{\ell_{jt}}{(1+r)^t} = \frac{\ell_{j0}}{r-g}$ if j is rented and $PV_j = P_j$ if owner-occupied. For owner-occupants in market j in $t = 0$, the present-value solvency constraint is (suppressing i subscripts):

$$c_0 + \mathbb{1}_{d\neq s}PV_d = \tilde{W}_0 + \mathbb{1}_{d\neq s}(PV_j - M_{Fj}) - \mathbb{1}_{d=s}\frac{r_{Fj}M_{Fj}}{r} \tag{A.16}$$

[55] As a microfoundation, suppose that non-residential property prices are based on spot rents determined in equilibrium. Fewer non-residential properties raise spot rents and hence equilibrium property prices.

where $\tilde{W}_0 := W_0 + \sum_t r^{-t} y_t$. The left-hand side gives the present-value of non-housing consumption and housing costs. Movers to rental units pay rental costs $\frac{\ell_{d0}}{r-g}$; movers to owner-occupied units pay P_{d0} independent of capital structure (because new fixed-rate mortgages have present-value cost equal to the mortgage balance).

In this framework, "prepayment costs" naturally emerge. If households move, they must repay the mortgage at face value. If they stay, they pay $\frac{r_{Fj} M_{Fj}}{r} < M_{Fj}$. Define prepayment costs as $PC_j := M_{Fj} - \frac{r_{Fj} M_{Fj}}{r} = \frac{(r - r_{Fj}) M_{Fj}}{r}$: the wealth that households give up when they move. Prepayment cost reflects the opportunity cost of losing access to a mortgage at rate r_{Fj}. For each dollar borrowed, the mortgage pays an implicit per-period dividend $(r_{Fj} - r)$; scaling by amount borrowed M_{Fj} and discounting by r delivers the expression for PC_j. An increase in r increases PC_j by increasing the flow dividend from borrowing at r_{Fj}.

For renters in $t = 0$, the solvency constraint is $c_0 + \mathbb{1}_{d=s} PV_k + \mathbb{1}_{d \neq s} PV_d = \tilde{W}_0$. Relative to owners, renters have no housing wealth and pay rent in their current market if they stay put. Crucially, renters have no prepayment cost: $PC_j = 0$.

To calculate the payoff from choice d, solve for c_0 in the solvency constraint and substitute into equation (A.13) along with equation (A.15). The optimal choice is:

$$d_i = \arg\max_d \; v_{ijd} - v_{ijs} = \mathbb{1}_{d \neq s} \left[\left(PV_j - PV_d - PC_j \right) + \alpha^{-1} \left(h_d - h_j - \lambda + \epsilon_{id} - \epsilon_{is} \right) \right] \tag{A.17}$$

The flow of households moving from j to $k \neq s$ is therefore:

$$\pi_{jk}(PV_j, PV_k, PC_j) = N_j \cdot \gamma \cdot \Pr\left(k_i = \arg\max_d v_{ijd} - v_{ijs} \right) \tag{A.18}$$

Housing demand has intuitive properties. Flows from j to $d \neq s$ are higher when the relative present-value cost of a move is low or the relative value of housing services are high. Higher α makes household preferences more correlated: different households become more likely to have the same ranking across housing choices d, determined by relative present-value costs.

Equations (A.17) and (A.18) show how fixed-rate mortgages affect housing demand. First, if interest rates were flexible ($r_{Fj} = r$) or mortgages could be repaid at face value, then $PC_j = 0$ and Modigliani-Miller would hold: capital structure would have no influence on equilibrium flows or prices. Second, higher prepayment cost reduces outflows to *all markets*, meaning $\frac{\partial \pi_{jk}}{\partial PC_j} \leq 0$. Third, without consumer heterogeneity, fixed-rate mortgages shut down trade, because at the price that new buyers will pay, existing owners are unwilling to sell. Households in j will buy in d only if d is the best market: $\alpha^{-1} h_d - PV_d \geq \max_{k \neq s} \alpha^{-1} h_k - PV_k$. But then households in d will not sell if $PC_d > 0$, because $0 \geq \max_k PV_d - PV_k + \alpha^{-1}(h_d - h_k) > \max_{k \neq s} v_{idk} - v_{ids}$.

Housing supply. We parameterize housing supply by assuming the price of a marginal non-residential property L_j is locally linear. Imposing market clearing:

$$PV_j = \phi_{0j} - \phi_{Pj}^{-1} \cdot (S - Q_j^S) \implies Q_j^S = S - \phi_{0j} + \phi_{Pj} PV_j \tag{A.19}$$

Collect ϕ_{Pj} in the matrix $\boldsymbol{\Phi} := \text{diag}((\phi_{Pj})_j)$.

A.3.3 Equilibrium properties.

First, we extend the logic of user-cost formulas to the case with fixed-rate mortgages in multiple markets. Fix ℓ_0 as the numeraire, and define $\mathcal{D}_x \boldsymbol{g}$ as the Jacobian matrix with jk^{th} entry $\frac{\partial g_j}{\partial x_k}$. A

first-order Taylor expansion of the system of equations in (A.14) yields:

$$\Delta \boldsymbol{PV} = \frac{\partial P_0}{\partial r}\Delta r + (\mathcal{D}_{PV}(\boldsymbol{F}_{\text{o}} - \boldsymbol{F}_{\text{i}}) + \boldsymbol{\Phi}_P)^{-1} \cdot (-\mathcal{D}_{PC}(\boldsymbol{F}_{\text{o}} - \boldsymbol{F}_{\text{i}})) \cdot \Delta \boldsymbol{PC} \tag{A.20}$$

For owner-occupied markets, $\Delta PV_j = \Delta P_j$. For rental markets, $\Delta PV_j \approx \frac{\partial PV_j}{\partial \ell_j}\Delta \ell_j + \frac{\partial PV_j}{\partial r}\Delta r$, so implied changes in ℓ_j depend on how higher rates impact present-value costs.

The first term on the right-hand side gives the negative effect of higher interest rates on present-value costs in all markets. A higher cost of capital makes renting in the numeraire market more attractive than owning, inducing a common negative price level shift.

The second term shows the impact of fixed-rate mortgages. Higher prepayment cost has two effects: first, it reduces own-market *outflows*, and second, it reduces *inflows* from other markets. Effects on net outflows are captured by $\mathcal{D}_{PC}(\boldsymbol{F}_{\text{o}} - \boldsymbol{F}_{\text{i}})$. Where an increase in prepayment cost *reduces* net outflows, elements of $-\mathcal{D}_{PC}(\boldsymbol{F}_{\text{o}} - \boldsymbol{F}_{\text{i}})$ are positive, indicating upward price pressure.

Housing market flows re-equilibrate by changing prices. The matrix $\mathcal{D}_{PV}\Delta \boldsymbol{F}$ captures impacts of price changes on net outflows from existing units, and $\boldsymbol{\Phi}_P$ captures the impact of price changes on physical supply. Inverting these combined effects maps from net outflow effects due to fixed-rate mortgages to equilibrium price effects.

We next describe equation (A.20)'s empirical predictions. Define, as a benchmark, *steady-state prices* as a price vector that capitalizes amenity value differences across markets, with $PV_k - PV_j = h_k - h_j$ for all k, j. Together with constant interest rates $r_{Fj} = r$ and equal populations $N_j = N$, these prices constitute a *steady state equilibrium* with constant prices, constant population, and no construction, as shown in Appendix A.6.

Proposition 2 (Cross-sectional predictions.)**.** *Consider the regression $\Delta PV_j = a + \beta \cdot \Delta PC_j + \epsilon_j$, where changes are first-order approximations around a steady-state equilibrium.*

1. *$\beta > 0$ if draws of $(\Delta PC_j)_j$ are mean-independent of $(PC_j, h_j, \phi_{Pj})_j$, with constant $E[\Delta PC_j]$ across j, and otherwise has ambiguous sign.*
2. *$\frac{\partial \beta}{\partial \alpha} \geq 0$, so cross-market substitutability does not reduce the price effect of ΔPC.*
3. *β is smaller in cities with more elastic supply ϕ_{Pj}.*
4. *β is larger in cities with higher local outflows: $\frac{\partial \beta}{\partial \gamma} \geq 0$.*

The first result is a fundamental identification challenge: higher PC_j may not correlate with higher prices in j. Although higher own-market prepayment cost reduces local *outflows*, higher prepayment cost in *other markets* might reduce local *inflows* by more.

An exogenous shock to prepayment cost resolves this challenge by ensuring cross-market confounds average out to zero. Then, β depends on the price impact of lower own-market exit to the numeraire rental market. Higher own-market prepayment cost means lower own-market exit to the numeraire rental market, raising own-market prices. As shown below, this is the relevant effect for understanding how much prepayment cost offsets negative price effects of rate increases in the aggregate.

The second result shows an identification opportunity: it is possible to estimate β using a cross section of arbitrarily similar markets. A concern might be that more similar markets have more cross-market arbitrage to eliminate price differences arising from prepayment cost, attenuating β. Proposition 2 shows this concern is wrong, and households can be arbitrarily elastic to present-value cost differences relative to the value of housing consumption without reducing β. For intuition, consider a prepayment cost increase in j in a market with fixed supply. For households in $k \neq j$, higher α increases price arbitrage between j and $k \neq j$. But higher α also makes households in j *more sensitive*

to prepayment cost, *reducing* arbitrage between j and $k \neq j$ for households in j. Given that house prices and prepayment cost are denominated in the same present-value units, these two effects exactly offset. When housing supply is endogenous, β weakly grows with higher α, which also makes household flows more important for prices.[56]

Results (3) and (4) show heterogeneity in the average cross-sectional effect of prepayment cost on house prices. Intuitively, lower local market exit due to rate lock has a bigger impact on prices where existing owners are a larger source of available units for purchase. Existing owners matter less for total available units where supply is elastic (result 3) or typical outflows are lower (result 4).

Causal measurement in the cross-section of owner-occupied housing markets is closely connected to the aggregate effect of prepayment cost on house prices during tightening. Define $\partial_{\Delta\theta}\, g$ as the path derivative of a function g in the direction of a change in parameters $\Delta\boldsymbol{\theta}$.

Proposition 3 (Aggregate implications.)**.** *Consider β_o from the specification described in Proposition 2.1, estimated only on the sample of owner-occupied markets (where $j > J_\ell$).*

The effect of a change in interest rates on average prices is:

$$\Delta\overline{P} = \frac{\partial PV_0}{\partial r}\Delta r + \omega \cdot \Delta\overline{PC} \tag{A.21}$$

where $\omega > 0$, $\overline{P} := J_o^{-1}\sum_{j>J_\ell} P_j$, $\overline{PC} := J_o^{-1}\sum_{j>J_\ell} PC_j$, and changes are first-order approximations around a steady-state equilibrium.

Furthermore, $\frac{\partial_{\Delta\theta}\,\omega}{\partial_{\Delta\theta}\,\beta} > 0$ for parameter changes $\Delta\boldsymbol{\theta}$ with a bigger impact on the outflow effect of prepayment costs than the outflow effects of price changes as given by Condition 1.

The first term in equation (A.21) is negative: a higher cost of capital reduces average prices by making owning less attractive than renting. The second term is positive: prepayment cost unambiguously dampens negative price effects of rate hikes. A higher cross-sectional effect of prepayment cost on house prices implies a greater dampening effect, so long as the cross-sectional effect is estimated with exogenous variation in ΔPC and the higher cross-sectional effect primarily reflects a greater effect of prepayment cost on outflows, in the sense formalized by Condition 1.

A.4 Once-and-for-all location choice is optimal.

Proof. Suppose $d_i = k$. Then:

$$v_{ijk} \geq \max_d\ v_{ijd} - v_{ijs} \tag{A.22}$$

We seek to show that staying in k is optimal in future periods: $0 \geq \max_{d\neq k}\ v_{ikd} - v_{iks}$. Equation (A.22) implies:

$$\begin{aligned}
&(PV_j - PV_k - PC_j) + \alpha^{-1}(h_k - h_j - \lambda + \epsilon_{ik} - \epsilon_{is}) \\
&\qquad \geq \max_{d\neq k}\,(PV_j - PV_d - PC_j) + \alpha^{-1}(h_d - h_j - \lambda + \epsilon_{id} - \epsilon_{is}) && \text{(A.23)}\\
&0 \geq \max_{d\neq k}\,(PV_k - PV_d) + \alpha^{-1}(h_d - h_k - \lambda + \epsilon_{id} - \epsilon_{ik}) && \text{(A.24)}\\
&0 \geq \max_{d\neq k}\ v_{ikd} - v_{iks} && \text{(A.25)}
\end{aligned}$$

where the second line subtracts $\lambda \geq 0$ from the RHS and applies the fact that the idiosyncratic value of a new unit in k after the first period is given by ϵ_{ik} rather than the old ϵ_{is}. □

[56] If higher α is due to increasingly fine-grained market sub-division, then ϕ_{P_j}, which is in units / present-value dollar, may decrease as α increases. If market sub-division changes the mass of households and developers proportionally, then higher α leads to no change in β.

A.5 Formulas for equilibrium flows.

This subsection derives formulas for equilibrium flows from Section 1.

Claim 1. *Equilibrium flows from j to k for $k \neq s$, among people who consider all markets, are:*

$$\pi_{jk} = N_j \gamma \int_z \left[H\left(z - (\lambda + \alpha PC_j)\right) \prod_{d \neq k,s} H\left(z + h_k - h_d - \alpha(PV_k - PV_d)\right) \right] dH(z) \tag{A.26}$$

Furthermore:

$$\frac{\partial \pi_{jk}}{\partial PC_j} = -\alpha N_j \gamma \int_z \left[h\left(z - (\lambda + \alpha PC_j)\right) \prod_{d \neq k,s} H\left(z + h_k - h_d - \alpha(PV_k - PV_d)\right) \right] dH(z) \tag{A.27}$$

$$\frac{\partial \pi_{jk}}{\partial PV_q} = \alpha N_j \gamma \int_z \Bigg[H\left(z - (\lambda + \alpha PC_j)\right) \cdot h\left(z + h_k - h_q - \alpha(PV_k - PV_q)\right) \cdot \prod_{d \neq k,s,q} H\left(z + h_k - h_d - \alpha(PV_k - PV_d)\right) \Bigg] dH(z) \quad \forall q \neq k \tag{A.28}$$

$$\frac{\partial \pi_{jk}}{\partial PV_k} = -\alpha N_j \gamma \int_z \Bigg[H\left(z - (\lambda + \alpha PC_j)\right) \cdot \Bigg(\sum_{d \neq k,s} h\left(z + h_k - h_d - \alpha(PV_k - PV_d)\right) \cdot \prod_{r \neq k,s,d} H\left(z + h_k - h_r - \alpha(PV_k - PV_r)\right) \Bigg) \Bigg] dH(z) \tag{A.29}$$

where $h(\cdot) := H'(\cdot)$.

Proof. Flows for $k \neq s$ are:

$$\pi_{jk} = N_j \gamma \Pr\Big(\epsilon_{ik} + h_k - h_j - \lambda + \alpha(PV_j - PV_k - PC_j) > \max_{d \neq k} \epsilon_{id} + h_d - h_j + \alpha(P_j - P_d) - \mathbf{1}_{d \neq s} \cdot (\lambda + \alpha PC_j) \Big) \tag{A.30}$$

Adding h_j and subtracting αPV_j from both sides:

$$\pi_{jk} = N_j \gamma \Pr\left(\epsilon_{ik} + h_k - \lambda - \alpha(PV_k + PC_j) > \max_{d \neq k} \epsilon_{id} + h_d - \alpha P_d - \mathbf{1}_{d \neq s} \cdot (\lambda + \alpha PC_j) \right) \tag{A.31}$$

$$= N_j \gamma \int_z \Pr\left(\max_{d \neq k} \epsilon_{id} + h_d - \alpha P_d - \mathbf{1}_{d \neq s} \cdot (\lambda + \alpha PC_j) \leq z + h_k - \lambda - \alpha(PV_k + PC_j) \right) dH(z) \tag{A.32}$$

$$= N_j \gamma \int_z \Pr\left(\max_{d \neq k} \epsilon_{id} + h_d - \alpha P_d + \mathbf{1}_{d = s} \cdot (\lambda + \alpha PC_j) \leq z + h_k - \alpha PV_k \right) dH(z) \tag{A.33}$$

$$= N_j \gamma \int_z \left[H\left(z - (\lambda + \alpha PC_j)\right) \prod_{d \neq k,s} H\left(z + h_k - h_d - \alpha(PV_k - PV_d)\right) \right] dH(z) \tag{A.34}$$

Differentiating with respect to PC_j, PV_q (for $q \neq k$), or PV_k delivers equations (A.27), (A.28), and (A.28). □

A.6 Steady-state equilibrium details.

This subsection describes properties of the "steady-state equilibrium" defined in Section 1.

Definition 1 (Steady-state prices.)**.** Steady-state prices $\mathcal{P}^S$ satisfy the following:

- Price differences capitalize amenity differences: $PV_k - PV_j = \alpha^{-1}(h_k - h_j)$ for all j, k.
- Market populations are constant: $\Delta N_j = 0$, implying that $Q_j^S = 0$.

Claim 2. *Steady-state prices are an equilibrium where $r_{Fj} = r$ and $N_j = N$.*

Proof. Since $r_{Fj} = r$ for all j, $\boldsymbol{PC} = \mathbf{0}$.

Substituting $N_j = N$ and steady-state prices into equation (A.26):

$$\pi_{jk} = N \cdot \gamma \cdot \int_z H(z-\lambda)\Pi_{d\neq k,s} H(z) dH(z) \tag{A.35}$$

$$= N \cdot \gamma \cdot \int_z H(z-\lambda) \cdot H(z)^{J-1} dH(z) := N\gamma\overline{\pi} \tag{A.36}$$

Therefore, $\pi_{jk} = N \cdot \gamma \cdot \overline{\pi}$ for all j and k. Substituting into equation (A.14):

$$(J+1) \cdot N \cdot \gamma \cdot \overline{\pi} = (J+1) \cdot N \cdot \gamma \cdot \overline{\pi} + Q_j^S(\cdot) \quad \forall j \tag{A.37}$$

Since steady-state prices maintain constant population, $Q_j^S = 0$, and cross-market flows form an equilibrium. □

Claim 3. *Let $\mathcal{P}^S$ be steady-state prices in a steady-state.*

1. *For all j, k, r and $q \neq 0$:*

$$\frac{\partial \pi_{jk}(\mathcal{P}^S;\cdot)}{\partial PC_j} = \frac{\partial \pi_{kq}(\mathcal{P}^S;\cdot)}{\partial PC_k}; \quad \frac{\partial \pi_{j0}(\mathcal{P}^S;\cdot)}{\partial PC_j} = \frac{\partial \pi_{k0}(\mathcal{P}^S;\cdot)}{\partial PC_k} \tag{A.38}$$

2. *For all j, k, where $q \neq k$ and $r \neq j$:*

$$\frac{\partial \pi_{jq}(\mathcal{P}^S;\cdot)}{\partial PV_k} = \frac{\partial \pi_{qj}(\mathcal{P}^S;\cdot)}{\partial PV_r} \tag{A.39}$$

3. *For all j, k:*

$$\frac{\partial \pi_{jk}(\mathcal{P}^S;\cdot)}{\partial PV_k} = \frac{\partial \pi_{kj}(\mathcal{P}^S;\cdot)}{\partial PV_j} \tag{A.40}$$

Proof. To show part (1), substitute steady-state prices and $N_j = N$ into equation (A.27):

$$\frac{\partial \pi_{jk}}{\partial PC_j} = -\alpha N\gamma \int_z \left[h\left(z - \lambda - \alpha PC_j\right) \cdot \prod_{d\neq k,s} H(z) \right] dH(z) \tag{A.41}$$

where $h(\cdot)$ is the pdf of $H(\cdot)$. Since in the steady-state equilibrium $\boldsymbol{PC} = \mathbf{0}$:

$$\frac{\partial \pi_{jk}(\mathcal{P}^S;\cdot)}{\partial PC_j} = -\alpha N\gamma \int_z \left[h\left(z-\lambda-\alpha PC_j\right) \cdot \prod_{d\neq k,s} H(z) \right] dH(z) \tag{A.42}$$

$$= -\alpha N\gamma \int_z \left[h\left(z-\lambda-\alpha PC_r\right) \cdot \prod_{d\neq q,s} H(z) \right] dH(z) = \frac{\partial \pi_{rq}(\mathcal{P}^S;\cdot)}{\partial PC_k} \tag{A.43}$$

To show part (2), substitute steady-state prices and $N_j = N$ into equation (A.28), and again apply that $PC_j = PC_q = 0$ in steady-state:

$$\frac{\partial \pi_{jq}(\mathcal{P}^S;\cdot)}{\partial PV_k} = \alpha N\gamma \int_z \left[H\left(z-(\lambda+\alpha PC_j)\right) \prod_{d\neq q,s,k} H(z) \right] dH(z) \tag{A.44}$$

$$= \alpha N\gamma \int_z \left[H\left(z-(\lambda+\alpha PC_q)\right) \prod_{d\neq j,s,r} H(z) \right] dH(z) = \frac{\partial \pi_{qj}(\mathcal{P}^S;\cdot)}{\partial PV_r} \tag{A.45}$$

To show part (3), substitute steady-state prices and $N_j = N$ into equation (A.29):

$$\frac{\partial \pi_{jk}(\boldsymbol{P}^S;\cdot)}{\partial PV_k} = \alpha N\gamma \int_z \left[H\left(z-(\lambda+\alpha PC_j)\right) \left(\sum_{d\neq k,s} h(z) \cdot \prod_{r\neq k,s,d} H(z) \right) \right] dH(z) \tag{A.46}$$

$$= \alpha N\gamma \int_z \left[H\left(z-(\lambda+\alpha PC_k)\right) \left(\sum_{d\neq j,s} h(z) \cdot \prod_{r\neq j,s,d} H(z) \right) \right] dH(z) \tag{A.47}$$

$$= \frac{\partial \pi_{kj}(\boldsymbol{P}^S;\cdot)}{\partial PV_j} \tag{A.48}$$

□

A.7 Derivation of equation (A.20).

Proof. In matrix form, the equilibrium defined in equation (A.14) is:

$$\boldsymbol{\Pi} \cdot \mathbf{1}_{J+1} = \boldsymbol{\Pi}' \cdot \mathbf{1}_{J+1} - \boldsymbol{Q}^S \tag{A.49}$$

We will first derive an expression that sets $PV_0 = 1$ as the numeraire in both the initial and perturbed equilibrium. Then, we will show how this expression changes when the numeraire is instead $\ell_0 = 1$.

Note first that with an integrated nonresidential property market, setting PV_0 as the numeraire implicitly defines the numeraire in the nonresidential market as well. Thus, the numeraire price in the nonresidential property market is $PV_0 = L_0 = 1$ as well. The marginal price in other nonresidential property markets is therefore $L_j + L_0$, so that an increase in PV_0 increases the equilibrium nonresidential property price as well.

Assume that $\phi_{P_0} = 0$, and invoke Walras' Law to drop the first row of the system (where $j = 0$).[57] Define J-element vectors $\boldsymbol{F}_\text{o}$ and $\boldsymbol{F}_\text{i}$ as $(\boldsymbol{\Pi} \cdot \mathbf{1}_{J+1})_{j>1}$ and $(\boldsymbol{\Pi}' \cdot \mathbf{1}_{J+1})_{j>1}$, respectively. Further define

[57] If market-clearing holds for $j > 0$, it holds for $j = 0$ as well. A constant population means that $\sum_j Q_j^S = 0$. Substituting using equation (A.14) implies $\sum_j \left(\sum_k \pi_{kj} - \pi_{jk} \right) = 0$. Since $\sum_j Q_j^S - 0$, this also means $\sum_j \left(\sum_k \pi_{kj} - \pi_{jk} + Q_j^S \right) = 0$. If equilibrium holds for $j > 0$, then this becomes $\sum_k \pi_{k0} - \pi_{0k} + Q_0^S = 0$, which is exactly the market-clearing condition for $j = 0$.

$\Delta \boldsymbol{F} := \boldsymbol{F}_{\text{o}} - \boldsymbol{F}_{\text{i}}$. We can write equation (A.49) as:

$$\Delta \boldsymbol{F} = -\boldsymbol{Q}^S \tag{A.50}$$

Take a first-order Taylor expansion of the left-hand side:

$$\Delta \boldsymbol{F} \approx \Delta \boldsymbol{F}^0 + \mathcal{D}_{PV} \Delta \boldsymbol{F} \cdot \Delta \boldsymbol{PV} + \mathcal{D}_{PC} \Delta \boldsymbol{F} \cdot \Delta \boldsymbol{PC} \tag{A.51}$$

For the right-hand side, recognize that, in the initial equilibrium, housing supply satisfies $\boldsymbol{Q}_0^S = \boldsymbol{S} - \boldsymbol{\phi}_0 + \boldsymbol{\Phi}_P \boldsymbol{PV}_0$, and in the new equilibrium, housing supply satisfies $\boldsymbol{Q}_0^S = \boldsymbol{S} - \boldsymbol{\phi}_0 + \boldsymbol{\Phi}_P \boldsymbol{PV}$. Thus, $\boldsymbol{Q}^S = \boldsymbol{Q}_0^S + \boldsymbol{\Phi}_P \Delta \boldsymbol{PV}$.

The equilibrium system becomes:

$$\Delta \boldsymbol{F}^0 + \mathcal{D}_{PV}(\Delta \boldsymbol{F}) \cdot \Delta \boldsymbol{PV} + \mathcal{D}_{PC}(\Delta \boldsymbol{F}) \cdot \Delta \boldsymbol{PC} = -\boldsymbol{Q}_0^S - \boldsymbol{\Phi}_P \cdot \Delta \boldsymbol{PV} \tag{A.52}$$

$$\mathcal{D}_{PV}(\Delta \boldsymbol{F}) \cdot \Delta \boldsymbol{PV} + \mathcal{D}_{PC}(\Delta \boldsymbol{F}) \cdot \Delta \boldsymbol{PC} = -\boldsymbol{\Phi}_P \cdot \Delta \boldsymbol{PV} \tag{A.53}$$

where the second line follows because at the original equilibrium, $\Delta \boldsymbol{F}^0 = \boldsymbol{Q}_0^S$. Rearranging:

$$\Delta \boldsymbol{PV} = -(\mathcal{D}_{PV}(\Delta \boldsymbol{F}) + \boldsymbol{\Phi}_P)^{-1} \cdot \mathcal{D}_{PC}(\Delta \boldsymbol{F}) \cdot \Delta \boldsymbol{PC} \tag{A.54}$$

First, we manipulate equation (A.54) so that the numeraire is ℓ_0 rather than PV_0. Equation (A.14) only defines $\boldsymbol{PV} - PV_0 \cdot \mathbf{1}_{J+1}$, i.e. only fixes $\boldsymbol{PV}$ up to level shifts. This is clearly true for elements of $\boldsymbol{\Pi}$, since all expressions for flows only depend on relative price. This is also true for housing supply $\boldsymbol{Q}^S$; a change in PV_0 that induces a common shift in PV_j shifts L_0, and hence $L_j + L_0$, by an equal amount, leaving market-clearing Q_j^S unchanged.

Therefore, we can write:

$$\Delta \boldsymbol{PV} = \boldsymbol{PV}' - \boldsymbol{PV} - (PV_0' - PV_0) \cdot \mathbf{1}_{J+1} \tag{A.55}$$

where $\boldsymbol{PV}', PV_0'$ are values in the equilibrium where $\boldsymbol{PC}$ and Δr change. When $PV_0 = PV_0' = 1$, equation (A.55) equals $\boldsymbol{PV}' - \boldsymbol{PV}$. However, if the numeraire is $\ell_0 = \ell_0' = 1$, then $(PV_0' - PV_0) \approx \frac{\partial P_0}{\partial r} \Delta r$. Therefore:

$$\Delta \boldsymbol{PV} = -(\mathcal{D}_{PV}(\Delta \boldsymbol{F}) + \boldsymbol{\Phi}_P)^{-1} \cdot \mathcal{D}_{PC}(\Delta \boldsymbol{F}) \cdot \Delta \boldsymbol{PC} + \frac{\partial P_0}{\partial r} \Delta r \tag{A.56}$$

This delivers equation (A.20) in the text. □

A.8 Proof of Proposition 2.

We start with several properties of the objects in equation (A.20).

Lemma 1. *The matrix $B := \mathcal{D}_{PV} \Delta \boldsymbol{F} + \boldsymbol{\Phi}_P$, evaluated at steady-state prices, has these properties:*

1. *$B = D - C$, where D is diagonal and C is symmetric with all elements ≥ 0.*
2. *B is an invertible, positive definite M-matrix, using the definition in Plemmons (1977).*
3. *B^{-1} can be written as:*

$$B^{-1} = D^{-1} \sum_{k=0}^{\infty} (C D^{-1})^k \tag{A.57}$$

Proof. To show part (1), define:

$$D := \boldsymbol{\Phi}_P + \text{diag}\left(\left(\sum_q \frac{\partial(\pi_{jq} - \pi_{qj})}{\partial PV_j}\right)_j\right), \quad C := -\mathcal{D}_{PV}\Delta\boldsymbol{F} + \text{diag}\left(\left(\sum_q \frac{\partial(\pi_{jq} - \pi_{qj})}{\partial PV_j}\right)_j\right) \tag{A.58}$$

The term $\text{diag}\left(\left(\sum_q \frac{\partial(\pi_{jq}-\pi_{qj})}{\partial PV_j}\right)_j\right)$ gives diagonal entries in $\mathcal{D}_{PV}\Delta\boldsymbol{F}$. By construction, $D - C = B$. D is diagonal and positive, since entries are:

$$D_{jj} = \Phi_{Pj} + \sum_q \frac{\partial(\pi_{jq} - \pi_{qj})}{\partial PV_j} > 0 \tag{A.59}$$

which is positive because $\frac{\partial \pi_{jq}}{\partial P_j} > 0$ and $\frac{\partial \pi_{qj}}{\partial P_j} < 0$.

Diagonal entries of C are zero by construction. Off-diagonal entries of C are:

$$-(\mathcal{D}_{PV}\Delta\boldsymbol{F})_{jk} = \sum_q \frac{\partial(\pi_{qj} - \pi_{jq})}{\partial PV_k} \tag{A.60}$$

$$= \frac{\partial(\pi_{kj} - \pi_{jk})}{\partial PV_k} + \sum_{q\neq k} \frac{\partial(\pi_{qj} - \pi_{jq})}{\partial PV_k} = \frac{\partial(\pi_{kj} - \pi_{jk})}{\partial PV_k} \tag{A.61}$$

where the last line applies Claim 3.2. All entries in C are non-negative, as $\frac{\partial \pi_{kj}}{\partial P_k} \geq 0$ and $\frac{\partial \pi_{jk}}{\partial P_k} \leq 0$.

C is also symmetric, since $C_{jk} = \frac{\partial(\pi_{kj}-\pi_{jk})}{\partial PV_k} = \frac{\partial(\pi_{jk}-\pi_{kj})}{\partial PV_j} = C_{kj}$, applying $\frac{\partial \pi_{kj}}{\partial PV_k} = \frac{\partial \pi_{jk}}{\partial PV_j}$ from Claim 3.2 and $\frac{\partial \pi_{jk}}{\partial PV_k} = \frac{\partial \pi_{kj}}{\partial PV_j}$ from Claim 3.3.

To show part (2), note that:

$$B_{kk} = \sum_j \frac{\partial(\pi_{kj} - \pi_{jk})}{\partial PV_k} + \Phi_{Pk} \tag{A.62}$$

$$\sum_{j\neq k,0} \left|B_{jk}\right| = \sum_{j\neq k,0} \frac{\partial(\pi_{kj} - \pi_{jk})}{\partial PV_k} \implies B_{kk} - \sum_{j\neq k} \left|B_{jk}\right| > 0 \tag{A.63}$$

where we make explicit the fact that summation of $\left|B_{jk}\right|$ across rows necessarily omits the numeraire rental market. By the Gershgorin circle theorem, all eigenvalues of B lie in the disk centered at B_{kk} with radius $\sum_{j\neq k}\left|B_{jk}\right|$.[58] Given equation (A.63), all eigenvalues are positive.

Part (1) (off-diagonal entries in B negative, diagonal entries positive), together with the fact that all eigenvalues are positive, implies B is an M-matrix, according to Theorem 1.C.9 in Plemmons (1977).

Finally, B is symmetric because D is symmetric and C is symmetric from part (1). B is therefore positive definite, since all eigenvalues are positive.

As a comment, the fact that B is an M-matrix means that $B^{-1}_{iq} \geq 0$ due to Plemmons (1977) Theorem 1.F.15. Furthermore, since B^{-1} is invertible, it cannot have a column with all zeros.

To show part (3), first recognize that $B = D - C = (I - CD^{-1})D$, so: $B^{-1} = D^{-1}(I - CD^{-1})^{-1}$.

Next, we establish that CD^{-1} has spectral radius $\rho(CD^{-1}) < 1$. The matrix $(CD^{-1})_{jk} = C_{jk}/D_k$. By

[58] This the column version of the Gershrogin circle theorem, which follows immediately from the row version, recognizing that the eigenvalues of matrix A are the same as the eigenvalues of its transpose A'.

definition:

$$D_k = \Phi_{Pj} + \sum_q \frac{\partial(\pi_{kq} - \pi_{qk})}{\partial P_k} > \sum_{j \neq k,0} \frac{\partial(\pi_{kj} - \pi_{jk})}{\partial P_k} = \sum_{j \neq k} \left| C_{jk} \right| \tag{A.64}$$

Therefore: $\frac{\sum_{j \neq k} C_{jk}}{D_k} < 1 \ \forall k$. By the Gershgorin circle theorem, all eigenvalues of CD^{-1} lie in the disk centered at $(CD^{-1})_{kk}$ with radius $\sum_{j \neq k} \left| CD^{-1} \right| = \sum_{j \neq k} \frac{C_{jk}}{D_k} < 1$. Since $C_{kk} = 0$, the disks are centered at zero, meaning the absolute value of all eigenvalues of CD^{-1} are less than one, so $\rho(CD^{-1}) < 1$. Taking the infinite series expansion of $(I - CD^{-1})^{-1}$ and substituting into the definition of B^{-1} yields equation (A.57). □

We first prove part (1): that β is positive in expectation if ΔM_j is mean-independent of $(PC_j, h_j, \phi_{Pj})_j$, and otherwise has ambiguous sign.

Proof. Without loss of generality, assume that $\mathcal{D}_{\mathcal{P}}(\text{diag}(\boldsymbol{PV})) = I$.[59]

Define the annihilator matrix $M := I - J^{-1}\mathbf{1}_J \cdot \mathbf{1}_J'$. M demeans matrix entries by columnwise averages. Further define $\Delta\widetilde{\boldsymbol{PC}} := M \cdot \Delta\boldsymbol{PC}$, $\Delta\widetilde{\mathcal{P}} := M \cdot \Delta\mathcal{P}$.

By the Frisch-Waugh-Lovell theorem:

$$\hat{\beta} = \frac{(M \cdot \Delta\boldsymbol{PC})' \cdot (M \cdot \Delta\mathcal{P})}{(M \cdot \Delta\boldsymbol{PC})'(M \cdot \Delta\boldsymbol{PC})} = \frac{\Delta\boldsymbol{PC}' \cdot (M' \cdot M) \cdot (\Delta\mathcal{P})}{(M \cdot \Delta\boldsymbol{PC})'(M \cdot \Delta\boldsymbol{PC})} \tag{A.65}$$

$$= \frac{\Delta\widetilde{\boldsymbol{PC}}' \Delta\mathcal{P}}{\Delta\widetilde{\boldsymbol{PC}}' \Delta\widetilde{\boldsymbol{PC}}} \tag{A.66}$$

where the last line uses the fact that the annihilator matrix is symmetric ($M = M'$) and idempotent ($M = M^2$).

The denominator of $\hat{\beta}$ is positive, so:

$$\hat{\beta} \propto \Delta\widetilde{\boldsymbol{PC}}' \Delta\mathcal{P} = \sum_j \Delta\tilde{PC}_j \cdot \left(B^{-1} \cdot (-\mathcal{D}_{PC}\Delta\boldsymbol{F}) \cdot \Delta PC \right)_j \tag{A.67}$$

By definition:

$$\left(B^{-1} \cdot (-\mathcal{D}_{PC}\Delta\boldsymbol{F}) \cdot \Delta PC \right)_j = \sum_r B^{-1}_{jr} \left(\sum_q (-\mathcal{D}_{PC}\Delta\boldsymbol{F})_{rq} \Delta PC_q \right) \tag{A.68}$$

Furthermore:

$$\sum_q (\mathcal{D}_{PC}\Delta\boldsymbol{F})_{rq} \Delta PC_q = \sum_q \left[\left(\sum_k \frac{\partial \pi_{rk}}{\partial PC_q} \right) - \frac{\partial \pi_{qr}}{\partial PC_q} \right] \Delta PC_q \tag{A.69}$$

$$= \frac{\partial \pi_{r0}}{\partial PC_r} \Delta PC_r + \left(\sum_{q \neq 0} \frac{\partial \pi_{rq}}{\partial PC_r} \Delta PC_r - \frac{\partial \pi_{qr}}{\partial PC_q} \Delta PC_q \right) \tag{A.70}$$

$$= \frac{\partial \pi_{r0}}{\partial PC_r} \Delta PC_r + J \cdot \overline{D_{PC}\boldsymbol{\pi}_J} \left(\Delta PC_r - J^{-1} \sum_{q \neq 0} \Delta PC_q \right) \tag{A.71}$$

$$= \frac{\partial \pi_{r0}}{\partial PC_r} J^{-1} \mathbf{1}_J \Delta\boldsymbol{PC} + \left(\frac{\partial \pi_{r0}}{\partial PC_r} + J \cdot \overline{D_{PC}\boldsymbol{\pi}_J} \right) \Delta\tilde{PC}_r \tag{A.72}$$

[59] This is without loss. The diagonal of $\mathcal{D}_{\mathcal{P}}(\text{diag}(\boldsymbol{PV}))$ takes positive values for rental markets and equals one for owner-occupied markets, and so just rescales $\Delta\mathcal{P}$ by a series of positive constants. This simplification makes notation easier.

where the second line makes explicit that summing over the q index omits the numeraire rental market; the third line applies that $\frac{\partial \pi_{rq}}{\partial PC_r} = \frac{\partial \pi_{qr}}{\partial PC_q} = \frac{\partial \pi_{rq}}{\partial PC_r}$ from Claim 3.1; and the fourth line adds and subtracts $\frac{\partial \pi_{r0}}{\partial PC_r} J^{-1} \mathbf{1}_J \Delta \boldsymbol{PC}$.

Substituting equation (A.72) into equation (A.67):

$$\hat{\beta} \propto \sum_{j,r} \Delta\widetilde{PC}_j \cdot B^{-1}_{jr} \left(-\frac{\partial \pi_{r0}}{\partial PC_r} J^{-1} \mathbf{1}_J \Delta \boldsymbol{PC} - \left(\frac{\partial \pi_{r0}}{\partial PC_r} + J \cdot \overline{D_{PC} \boldsymbol{\pi}_J} \right) \Delta\tilde{PC}_r \right) \tag{A.73}$$

From Claim 3.1, $\frac{\partial \pi_{r0}}{\partial PC_r} = \overline{D_{PC} \boldsymbol{\pi}_0}$ for all r, and

$$\begin{aligned} \hat{\beta} \propto & -\left(\overline{D_{PC} \boldsymbol{\pi}_0} \cdot J^{-1} \mathbf{1}_J \Delta \boldsymbol{PC} \right) \cdot \left(\sum_{j,r} \Delta\widetilde{PC}_j \cdot B^{-1}_{jr} \right) \\ & \underbrace{- \left(\overline{D_{PC} \boldsymbol{\pi}_0} + J \cdot \overline{D_{PC} \boldsymbol{\pi}_J} \right) \cdot \left(\sum_{j,r} \Delta\widetilde{PC}_j \cdot B^{-1}_{jr} \cdot \Delta\widetilde{PC}_r \right)}_{>0} \end{aligned} \tag{A.74}$$

The second term in equation (A.74) is positive, because (i) $\overline{D_{PC} \boldsymbol{\pi}_0}, \overline{D_{PC} \boldsymbol{\pi}_J} < 0$, and (ii) B symmetric positive-definite implies B^{-1} symmetric positive definite, so $\sum_{j,r} \Delta\widetilde{PC}_j \cdot B^{-1}_{jr} \cdot \Delta\widetilde{PC}_r = \Delta\widetilde{\boldsymbol{PC}}' B^{-1} \Delta\widetilde{\boldsymbol{PC}}$.

The first term in equation (A.74) has ambiguous sign. Although there is much symmetry in the outflow effect of price and prepayment cost changes across markets in steady-state, we still allow markets to differ in their supply elasticity. Hence, the price effect of outflows varies by place, and is encoded in B^{-1}_{jr}. If places with especially high increases in prepayment cost are especially supply-elastic, then there will be negative correlation between $\Delta\widetilde{PC}_j$ and B^{-1}_{jr}, meaning the first term may be negative, and the overall sign of $\hat{\beta}$ ambiguous.

We now apply the fact that ΔPC_j is mean-independent of $(PC_j, h_j, \phi_{Pj})_j$ and $E[PC_j]$ is constant across j. Formally, where the expectation is taken over the j index:

$$E[(\Delta PC_j - \overline{\Delta PC}) \,|\boldsymbol{PC}, \boldsymbol{h}, \boldsymbol{\phi}_P] = E[(\Delta PC_j - \overline{\Delta PC})] = 0 \tag{A.75}$$

where $E[\Delta PC_j] = \overline{\Delta PC}$ for all j. Thus, for any function g: $E[(\Delta PC_j - \overline{\Delta PC}) \cdot g(\boldsymbol{PC}, \boldsymbol{h}, \boldsymbol{\phi}_P)] = 0$. B^{-1} is a function of $(\boldsymbol{PC}, \boldsymbol{h}, \boldsymbol{\phi}_P)$, since all underlying elements are functions of either $\boldsymbol{\phi}_P$ directly, or $(\boldsymbol{PC}, \boldsymbol{h}, \boldsymbol{PV})$, which too are functions of $(\boldsymbol{PC}, \boldsymbol{h}, \boldsymbol{\phi}_P)$ because $\boldsymbol{PV}$ is a function of $(\boldsymbol{PC}, \boldsymbol{h}, \boldsymbol{\phi}_P)$. Taking $g(\boldsymbol{PC}, \boldsymbol{h}, \boldsymbol{\phi}_P) = \sum_r B^{-1}_{jr}$:

$$E\left[\Delta\widetilde{PC}_j \cdot \sum_r B_{jr} \right] = 0 \tag{A.76}$$

where an application of the law of iterated expectations ensures $\Delta\widetilde{PC}_j$ includes $\overline{PC}$.

Because:

$$\begin{aligned} E\left[-\left(\overline{D_{PC} \boldsymbol{\pi}_0} \cdot J^{-1} \mathbf{1}_J \Delta \boldsymbol{PC} \right) \cdot \left(\sum_{j,r} \Delta\widetilde{PC}_j \cdot B^{-1}_{jr} \right) \right] = \\ - \left(\overline{D_{PC} \boldsymbol{\pi}_0} \cdot J^{-1} \mathbf{1}_J \Delta \boldsymbol{PC} \right) \cdot \left(\sum_j E\left[\Delta\widetilde{PC}_j \cdot \sum_r B^{-1}_{jr} \right] \right) = 0 \end{aligned} \tag{A.77}$$

we know that in expectation:

$$E\left[\hat{\beta}\right] = -\left(\overline{D_{PC}\pi_0} + J \cdot \overline{D_{PC}\pi_J}\right) \cdot E\left[\frac{\sum_{j,r} \Delta\widetilde{PC}_j \cdot B^{-1}_{jr} \cdot \Delta\widetilde{PC}_r}{\Delta\widetilde{\boldsymbol{PC}}' \Delta\widetilde{\boldsymbol{PC}}}\right] > 0 \tag{A.78}$$

□

We now prove parts (2)-(4). We maintain that $\Delta\boldsymbol{PV}, \Delta\boldsymbol{PC}$ are first-order expansions around a steady-state equilibrium.

Proof. Formally, we wish to show: (i) $\frac{\partial\beta}{\partial\alpha} \geq 0$ (part 2); (ii) $\frac{\partial\beta}{\partial\varphi} \geq 0$, where $\phi_{Pj} = \varphi \cdot \phi^0_{Pj}$ (part 3); and (iii) $\frac{\partial\beta}{\partial\gamma} \geq 0$ (part 4). From equation (A.78):

$$E\left[\hat{\beta}\right] = -\left(\overline{D_{PC}\pi_0} + J \cdot \overline{D_{PC}\pi_J}\right) \cdot E\left[\frac{\sum_{j,r} \Delta\widetilde{PC}_j \cdot B^{-1}_{jr} \cdot \Delta\widetilde{PC}_r}{\Delta\widetilde{\boldsymbol{PC}}' \Delta\widetilde{\boldsymbol{PC}}}\right] \tag{A.79}$$

$$= -\left(\overline{D_{PC}\pi_0} + J \cdot \overline{D_{PC}\pi_J}\right) \cdot \frac{1}{J} \cdot \sum_j B^{-1}_{jj} \tag{A.80}$$

where the second line follows since the expectation of the Rayleigh quotient of B with respect to a random vector is the average of the trace (e.g., Lemma before Prop 1 in Hutchinson 1989).

The rest of the derivation follows in three steps.

1. Get an expression for B^{-1}_{jr} that we can differentiate with respect to $x \in \{\alpha, \varphi, \gamma\}$.
2. Show that, when evaluated at steady-state prices and prepayment cost, we can express the following partial derivatives as scalar multiples of $x \in \{\alpha, \gamma\}$:

$$\frac{\partial \pi_{jk}}{\partial PC_j} = x \cdot U_x, \quad \frac{\partial \pi_{jk}}{\partial PV_j} = x \cdot Q^o_x, \quad \frac{\partial \pi_{jk}}{\partial PV_k} = x \cdot Q^d_x \tag{A.81}$$

   where U_x, Q^o_x, Q^d_x do not depend on x.
3. Differentiate the right-hand side of equation (A.80) with respect to x.

Step 1. From Lemma 1.3:

$$\sum_i B^{-1}_{iq} = \sum_i \left(D^{-1} \sum_{k=0}^{\infty} (CD^{-1})^k\right)_{iq} = \sum_i \sum_m D^{-1}_{im} \left(\sum_{k=0}^{\infty} (CD^{-1})^k\right)_{mq} = \sum_i D^{-1}_i \left(\sum_{k=0}^{\infty} (CD^{-1})^k\right)_{iq} \tag{A.82}$$

Furthermore, substituting the form of CD^{-1}_{jk} from equation (A.60) and (A.59):

$$\left(\sum_{k=0}^{\infty} (CD^{-1})^k\right)_{iq} = I_{iq} + \frac{C_{iq}}{D_q} + \sum_{m_1} \frac{C_{i,m_1}}{D_{m_1}} \cdot \frac{C_{m_1,q}}{D_q} + \sum_{m_2,m_1} \frac{C_{i,m_1}}{D_{m_1}} \cdot \frac{C_{m_1,m_2}}{D_{m_2}} \cdot \frac{C_{m_2,q}}{D_q} + \ldots \tag{A.83}$$

$$= I_{iq} + \sum_{k=1}^{\infty} S_{kiq}, \quad \text{where } S_{1iq} = \frac{C_{iq}}{D_q}, \quad S_{kiq} = \sum_{m_{k-1}} S_{k-1,i,m_{k-1}} \cdot \frac{C_{m_{k-1},q}}{D_q} \tag{A.84}$$

Hence:

$$B_{iq}^{-1} = D_i^{-1}\left(I_{iq} + \sum_{k=1}^{\infty} S_{kiq}\right) \tag{A.85}$$

Step 2. Define, for $x \in \{\alpha, \gamma\}$:

$$U_x := -\frac{\alpha N \gamma}{x}\int_z \left[h(z-\lambda)\cdot(H(z))^{J-1}\right]dH(z) \tag{A.86}$$

$$Q_x^o := \frac{\alpha N \gamma}{x}\int_z \left[H(z-\lambda)\cdot h(z)\cdot(H(z))^{J-2}\right]dH(z) \tag{A.87}$$

$$Q_x^d := -\frac{\alpha N \gamma}{x}(J-1)\cdot\int_z \left[H(z-\lambda)\cdot h(z)\cdot(H(z))^{J-2}\right] \tag{A.88}$$

It is clear that U_x, Q_x^o, and Q_x^d do not depend on $x \in \{\alpha, \gamma\}$. From equation (A.42), $\frac{\partial \pi_{jk}(\mathcal{P}^S, \Delta \boldsymbol{PC}^S)}{\partial PC_j} = x \cdot U_x$. Using Definition 1 for $\mathcal{P}^S$ and the fact that $\Delta \boldsymbol{PC} = \boldsymbol{0}$ in a steady-state equilibrium, equations (A.28) and (A.29) imply that $\frac{\partial \pi_{jk}(\mathcal{P}^S, \Delta \boldsymbol{PC}^S)}{\partial PV_j} = x \cdot Q_x^o$ and $\frac{\partial \pi_{jk}(\mathcal{P}^S, \Delta \boldsymbol{PC}^S)}{\partial PV_k} = x \cdot Q_x^d$.

Step 3. Let $x \in \{\alpha, \gamma, \varphi\}$. From (A.80):

$$\beta \propto -\frac{x}{x}\cdot(J+1)\cdot x \cdot U_x \cdot \sum_j B_{jj}^{-1} \tag{A.89}$$

$$= -(J+1)\cdot U_x \cdot \sum_j x \cdot B_{jj}^{-1} \tag{A.90}$$

We will use equation (A.90) for $x \in \{\alpha, \gamma\}$ (labeled "Case A" below), and equation (A.80) for $x = \varphi$ (labeled "Case B" below).

Case A. Let $x \in \{\alpha, \gamma\}$. Differentiating equation (A.90) with respect to x gives:

$$\frac{\partial \beta}{\partial x} \propto -(J+1)\cdot U_x \cdot \sum_j \cdot \frac{\partial (x \cdot B_{jj}^{-1})}{\partial x} \tag{A.91}$$

From equation (A.85):

$$\frac{\partial}{\partial x}\left(x \cdot B_{jr}^{-1}\right) = \frac{\partial}{\partial x}\left(\frac{x}{D_j}\left(1 + \sum_{k=1}^{\infty} S_{kjr}\right)\right) \tag{A.92}$$

$$= \sum_i \left(\frac{\partial x / D_j}{\partial x}\right)\cdot\left(1 + \sum_{k=1}^{\infty} S_{kjr}\right) + \frac{x}{D_j}\cdot\left(\sum_{k=1}^{\infty}\frac{\partial S_{kjr}}{\partial x}\right) \tag{A.93}$$

Start with the first term in equation (A.93). By definition: $\frac{x}{D_j} = \frac{1}{x^{-1}\phi_{P_j} + x^{-1}\sum_q \frac{\partial(\pi_{jq} - \pi_{qj})}{\partial P_j}}$. The second term in the denominator of this expression equals $\frac{1}{x}\sum_q x \cdot Q_x^o - x \cdot Q_x^d = (J+1)(Q_x^o - Q_x^d)$, which does not depend on x. Hence: $\frac{\partial x / D_j}{\partial x} = \left(\frac{\phi_{P_j}}{D_j}\right)^2 \geq 0$ where the sign holds with equality when $\phi_{P_j} = 0$.

Move to the second term in equation (A.93). We show by induction that $\frac{\partial S_{kjr}}{\partial x} \geq 0$ for all k, j, r, with equality if $\phi_{P_r} = 0$.

- For $k = 1$:

$$S_{1jr} = \frac{C_{jr}}{D_r} = \frac{\frac{\partial(\pi_{rj}-\pi_{jr})}{\partial PV_r}}{\phi_{P_r} + \sum_p \frac{\partial(\pi_{rp}-\pi_{pr})}{\partial P_r}} = \frac{x \cdot (Q_x^o - Q_x^d)}{\phi_{P_r} + x \cdot (J+1) \cdot (Q_x^o - Q_x^d)} \tag{A.94}$$

$$= \frac{(Q_x^o - Q_x^d)}{x^{-1} \cdot \phi_{P_r} + (J+1)(Q_x^o - Q_x^d)} \tag{A.95}$$

Hence:

$$\frac{\partial S_{1jr}}{\partial x} = \phi_{P_r}^2 \cdot \frac{(Q_x^o - Q_x^d)}{\left(x^{-1} \cdot \phi_{P_r} + (J+1) \cdot (Q_x^o - Q_x^d)\right)^2} = \phi_{P_r}^2 \cdot \frac{C_{jr}}{D_r^2} \geq 0 \tag{A.96}$$

with equality when $\phi_{P_r} = 0$.

- Assume $\frac{\partial S_{k-1,j,r}}{\partial x} \geq 0$. For $k > 1$:

$$S_{kjr} = \sum_{m_{k-1}} S_{k-1,j,m_{k-1}} \cdot \frac{C_{m_{k-1},r}}{D_r} \tag{A.97}$$

$$\implies \frac{\partial S_{kjr}}{\partial x} = \sum_{m_{k-1}} \frac{\partial S_{k-1,j,m_{k-1}}}{\partial x} \cdot \frac{C_{m_{k-1},r}}{D_r} + S_{k-1,j,m_{k-1}} \cdot \frac{\partial C_{m_{k-1},r}/D_r}{\partial x} \tag{A.98}$$

By the induction hypothesis, the first term in equation (A.98) is ≥ 0. From the $k = 1$ case, we know that $\frac{\partial C_{kr}/D_r}{\partial x} \geq 0$ for all k, including $k = m_{k-1}$. Hence, the second term in equation (A.98) ≥ 0. Therefore: $\frac{\partial S_{kjr}}{\partial x} \geq 0$, as required.

Putting everything together, $\frac{\partial}{\partial x}\left(x \cdot B_{jr}^{-1}\right) \geq 0$, and hence $\frac{\partial \beta}{\partial x} \geq 0$ per equation (A.91).

Case B. Let $x = \varphi$. Differentiating the RHS of equation (A.80) with respect to x:

$$\frac{\partial \beta}{\partial \varphi} = -\left(\overline{D_{PC}\pi_0} + J \cdot \overline{D_{PC}\pi_J}\right) \cdot 1J \cdot \sum_j \frac{\partial B_{jj}^{-1}}{\partial \varphi} - \underbrace{\left(\frac{1}{J}\sum_j B_{jj}^{-1}\right)\left(\frac{\partial}{\partial \varphi}\left(\overline{D_{PC}\pi_0} + J \cdot \overline{D_{PC}\pi_J}\right)\right)}_{=0} \tag{A.99}$$

where the second term equals zero because $\frac{\partial \pi_{qk}}{\partial PC_q}$ does not depend on φ. Moreover:

$$\frac{\partial B_{jr}^{-1}}{\partial \varphi} = \frac{\partial}{\partial \varphi}\left(\frac{1}{D_j} \cdot \left(1 + \sum_{k=1}^{\infty} S_{kjr}\right)\right) \tag{A.100}$$

$$= -\left(1 + \sum_{k=1}^{\infty} S_{kjr}\right) \cdot \frac{1}{D_j^2}\frac{\partial D_j}{\partial \varphi} + \frac{1}{D_j} \cdot \sum_{k=1}^{\infty} \frac{\partial S_{kjr}}{\partial \varphi} \tag{A.101}$$

The first term in equation (A.101) is ≤ 0, since $\frac{\partial D_j}{\partial \varphi} = \phi_{P_r}^0 \geq 0$.

We show by induction that the second term is weakly negative as well, since $\frac{\partial S_{kjr}}{\partial \varphi} \leq 0$, with equality when $\phi_{P_r} = 0$.

- For $k = 1$:

$$S_{1jr} = \frac{C_{jr}}{D_q r} = \frac{\frac{\partial(\pi_{rj}-\pi_{jr})}{\partial PV_r}}{\varphi \cdot \phi^0_{P_r} + \sum_p \frac{\partial(\pi_{rp}-\pi_{pr})}{\partial P_r}}; \quad \frac{\partial S_{1jr}}{\partial \varphi} = -\phi^0_{P_r} \cdot \frac{C_{jr}}{D_r^2} \leq 0. \tag{A.102}$$

- Assume $\frac{\partial S_{k-1,jr}}{\partial \varphi} \leq 0$. For $k > 1$:

$$S_{kjr} = \sum_{m_{k-1}} S_{k-1,j,m_{k-1}} \cdot \frac{C_{m_{k-1},r}}{D_r} \tag{A.103}$$

$$\implies \frac{\partial S_{kjr}}{\partial \varphi} = \sum_{m_{k-1}} \frac{\partial S_{k-1,j,m_{k-1}}}{\partial \varphi} \cdot \frac{C_{m_{k-1},r}}{D_r} + S_{k-1,j,m_{k-1}} \cdot \frac{\partial C_{m_{k-1},r}/D_r}{\partial \varphi} \leq 0 \tag{A.104}$$

  The first term is negative by the induction hypothesis. The second term is negative because from the $k = 1$ case, we know that $\frac{\partial C_{kr}/D_r}{\partial \varphi} \leq 0$ for all k. Hence, the second term is negative as well, and so $\frac{\partial S_{kjr}}{\partial \varphi} \leq 0$ for all k.

Putting everything together, $\frac{\partial B^{-1}_{jr}}{\partial \varphi} \leq 0$, implying $\frac{\partial \beta}{\partial \varphi} \leq 0$ from equation (A.99) as required. □

A.9 Proof of Proposition 3.

We first state Condition 1.

Condition 1 (Aggregation). Define the average outflow effect of prepayment cost V, the average net outflow effects of present-value costs Q, and ϕ_j as follows:

$$V := \frac{1}{J_o} \sum_{j>J_\ell} \frac{1}{J_o+1} \sum_k \frac{\partial \pi_{jk}}{\partial PC_j}, \; Q := \frac{1}{J_o} \sum_{j>J_\ell} \frac{1}{J_o+1} \sum_{k \neq j} \frac{\partial(\pi_{jk}-\pi_{kj})}{\partial PV_j}, \; \phi_j := \frac{(J+2)\cdot \Phi_{Pj}}{(J_o+1)Q + \Phi_{Pj}} \tag{A.105}$$

The impact of $\Delta\boldsymbol{\theta}$ on the outflow effect of prepayment cost and new construction is greater than the outflow effect of price changes:

$$\left| \partial_{\Delta\theta} \ln V + \sum_j \phi_j \cdot \partial_{\Delta\theta} \ln \Phi_{Pj} \right| \geq |\partial_{\Delta\theta} \ln Q| \tag{A.106}$$

The LHS gives the effect of a parameter change in the direction of $\Delta\boldsymbol{\theta}$ on the outflow effects of prepayment cost, combined with the housing supply effects of price. The RHS gives the effect on the net outflow effects of present-value costs.

We now prove Proposition 3.

Proof. Restrict to owner-occupied housing markets by right-multiplying (A.20) by $\text{diag}((\boldsymbol{O}'_{J_\ell}, \mathbf{1}'_{J_o})')$:

$$\boldsymbol{P} = -\left((\mathcal{D}_{PV} \Delta\boldsymbol{F} + \boldsymbol{\Phi})^{-1} (\mathcal{D}_{PC} \Delta\boldsymbol{F}) \Delta\boldsymbol{PC} \right)_{j>J_\ell} + \frac{\partial PV_0}{\partial r} \cdot \mathbf{1}_{J_o} \Delta r \tag{A.107}$$

Left-multiply by $J_o^{-1} \cdot \mathbf{1}'_{J_o}$ and substitute notation from Lemma 1 to obtain:

$$\overline{P} = -\frac{1}{J_o} \sum_{j>J_\ell} (B^{-1} \mathcal{D}_{PC} \Delta\boldsymbol{F} \cdot \Delta\boldsymbol{PC})_j + \frac{\partial PV_0}{\partial r} \Delta r \tag{A.108}$$

Following similar steps as in the proof of Prop 2:

$$
\begin{aligned}
&-\frac{1}{J_o}\sum_{j>J_\ell}(B^{-1}\mathcal{D}_{PC}\Delta\boldsymbol{F}\cdot\Delta\boldsymbol{PC})_j \\
&\quad=-\frac{1}{J_o}\sum_{j>J_\ell}\sum_{r}B^{-1}_{jr}\left(\sum_{q}(D_{PC}\Delta\boldsymbol{F})_{rq}\Delta PC_q\right) && \text{(A.109)}\\
&\quad=-\frac{1}{J_o}\sum_{j>J_\ell}\sum_{r}B^{-1}_{jr}\left(\overline{D_{PC}\pi_0}\Delta PC_r+\sum_{q\neq 0}\overline{D_{PC}\pi_J}(\Delta PC_r-\Delta PC_q)\right) && \text{(A.110)}\\
&\quad=-\frac{1}{J_o}\sum_{j>J_\ell,r}B^{-1}_{jr}\left(\overline{D_{PC}\pi_0}\cdot\Delta PC_r+J\cdot\overline{D_{PC}\pi_J}\cdot\Delta\widetilde{PC}_r\right) && \text{(A.111)}\\
&\quad=-\frac{1}{J_o}\overline{D_{PC}\pi_0}\cdot\Delta\overline{PC}\sum_{j>J_\ell,r}B^{-1}_{jr}-\frac{1}{J_o}\left(\overline{D_{PC}\pi_0}+J\cdot\overline{D_{PC}\pi_J}\right)\sum_{j>J_\ell,r}B^{-1}_{jr}\cdot\Delta\widetilde{PC}_r && \text{(A.112)}
\end{aligned}
$$

Since $E\left[B^{-1}_{jr}\cdot\Delta\widetilde{PC}_r\right]=0$ for ΔPC_j mean-independent of fundamentals, in expectation:

$$
\overline{P}=\underbrace{\left(-\frac{1}{J_o}\overline{D_{PC}\pi_0}\sum_{j>J_\ell,r}B^{-1}_{jr}\right)}_{:=\omega>0}\cdot\Delta\overline{PC}+\frac{\partial PV_0}{\partial r}\Delta r \tag{A.113}
$$

where the sign of ω follows from $-\overline{D_{PC}\pi_0}>0$ and $B^{-1}_{jr}>0$ (from Lemma 1.2).

All that remains is to show that $\omega\propto\beta_o$, the coefficient on ΔPC_j a regression of owner-occupied house prices on scaled prepayment cost, $P_j=a_o+\beta_o\cdot\Delta PC_j$. Following steps in the proof of Prop 2, and maintaining the assumption that ΔPC_j are mean-independent of fundamentals, the expected value of $\hat{\beta}_o$ is:

$$
E\left[\hat{\beta}_o\right]=-\left(\overline{D_{PC}\pi_0}+J\cdot\overline{D_{PC}\pi_J}\right)\cdot\frac{1}{J_o}\cdot\sum_{j>J_\ell}B^{-1}_{jj} \tag{A.114}
$$

where we apply the Rayleigh quotient property as in equation (A.80).

We seek to show that:

$$
\partial_{\Delta\boldsymbol{\theta}}\left[E\left[\hat{\beta}_o\right]\right]>0\iff\partial_{\Delta\boldsymbol{\theta}}\left[\omega\right]>0 \tag{A.115}
$$

We apply the steady-state assumptions to simplify expressions for $E\left[\hat{\beta}_o\right]$ and ω. The proof of Claim 3.1 shows that $\overline{D_{PC}\pi_0}=\overline{D_{PC}\pi_J}$. Define $V:=\overline{D_{PC}\pi_0}=\overline{D_{PC}\pi_J}$. Thus,

$$
\partial_{\Delta\boldsymbol{\theta}}\left[E\left[\hat{\beta}_o\right]\right]=-\frac{J_o+1}{J_o}\cdot\left(\sum_{j>J_\ell}B^{-1}_{jj}\right)\cdot V\cdot\left[\frac{\partial_{\Delta\boldsymbol{\theta}}V}{V}+\frac{\sum_{j>J_\ell}\partial_{\Delta\boldsymbol{\theta}}B^{-1}_{jj}}{\left(\sum_{j>J_\ell}B^{-1}_{jj}\right)}\right] \tag{A.116}
$$

$$
\partial_{\Delta\boldsymbol{\theta}}[\omega]=-\frac{1}{J_o}\cdot\left(\sum_{j>J_\ell,r}B^{-1}_{jr}\right)\cdot V\cdot\left[\frac{\partial_{\Delta\boldsymbol{\theta}}V}{V}+\frac{\sum_{j>J_\ell,r}\partial_{\Delta\boldsymbol{\theta}}B^{-1}_{jr}}{\left(\sum_{j>J_\ell,r}B^{-1}_{jr}\right)}\right] \tag{A.117}
$$

For condition (A.115) to hold, the terms in brackets must have the same sign in steady state.

We now analyze $\partial_{\Delta\boldsymbol{\theta}}B^{-1}_{jj}$ and $\partial_{\Delta\boldsymbol{\theta}}B^{-1}_{jr}$. Based on analysis in the proof of Lemma 1, we know that in steady state, $\frac{\partial(\pi_{jq}-\pi_{qj})}{\partial PV_j}=\frac{\partial(\pi_{kr}-\pi_{rk})}{\partial k}$ for all j,q,r,k. Define $Q:=\frac{\partial(\pi_{jq}-\pi_{qj})}{\partial PV_j}$. The diagonal entries of B can

be written as $B_{jj} = (J_o + 1) \cdot Q + \Phi_{Pj}$, and off-diagonal entries are $B_{jr} = Q$, $r \neq j$. Hence:

$$B = (J_o - 1) \cdot Q \cdot I + \boldsymbol{\Phi}_P + Q \cdot (I - \mathbf{1}_{J_o} \mathbf{1}'_{J_o}) \tag{A.118}$$

$$= J_0 \cdot Q \cdot I + \boldsymbol{\Phi}_P - \mathbf{1}_{J_o} \mathbf{1}'_{J_o} \tag{A.119}$$

Applying the Sherman-Morrison formula for B^{-1}:

$$B^{-1}_{jk} = \begin{cases} \frac{1}{d_j} \cdot \left(1 - \frac{1}{\zeta \cdot d_j}\right) = \frac{\zeta d_j - 1}{\zeta d_j^2} & \text{if j=k} \\ \frac{1}{d_j} \cdot \left(-\frac{1}{\zeta \cdot d_k}\right) = -\frac{1}{\zeta d_j d_k} & \text{otherwise} \end{cases} \tag{A.120}$$

$$\text{where } d_j := Q \cdot (J_0 + 1) + \Phi_{Pj}, \quad \zeta := -Q^{-1} + \sum_q d_q^{-1} \tag{A.121}$$

Differentiating $\sum_j B^{-1}_{jj}$:

$$\partial_{\Delta\boldsymbol{\theta}} \sum_j B^{-1}_{jj} = \sum_j -B^{-1}_{jj} \left(\partial_{\Delta\boldsymbol{\theta}} \ln \zeta + 2\partial_{\Delta\boldsymbol{\theta}} \ln d_j\right) + \frac{1}{d_j} \left(\partial_{\Delta\boldsymbol{\theta}} \ln \zeta + \partial_{\Delta\boldsymbol{\theta}} \ln d_j\right) \tag{A.122}$$

Moreover:

$$\partial_{\Delta\boldsymbol{\theta}} \zeta = Q^{-1} \partial_{\Delta\boldsymbol{\theta}} \ln Q - \sum_q d_q^{-1} \partial_{\Delta\boldsymbol{\theta}} \ln d_q \tag{A.123}$$

$$\partial_{\Delta\boldsymbol{\theta}} \ln d_q = \frac{Q(J_o + 1) \partial_{\Delta\boldsymbol{\theta}} \ln Q + \Phi_{Pj} \partial_{\Delta\boldsymbol{\theta}} \ln \Phi_{Pj}}{Q(J_o + 1) + \Phi_{Pj}} \tag{A.124}$$

For exposition, suppose $\partial_{\Delta\boldsymbol{\theta}} \Phi_{Pj} = 0$. We have:

$$\partial_{\Delta\boldsymbol{\theta}} \sum_j B^{-1}_{jj} = \sum_j \left(\frac{1}{d_j} - B^{-1}_{jj}\right) \cdot \left(\partial_{\Delta\boldsymbol{\theta}} \ln \zeta + \partial_{\Delta\boldsymbol{\theta}} \ln d_j\right) - B^{-1}_{jj} \partial_{\Delta\boldsymbol{\theta}} \ln d_j \tag{A.125}$$

$$= \sum_j \underbrace{\frac{1}{\zeta^2 d_j^2}}_{>0} \cdot \underbrace{\left(Q^{-1} + \frac{Q(J_o + 1)}{Q(J_o + 1) + \Phi_{Pj}} \left[\zeta - \sum_q d_q^{-1}\right]\right)}_{\geq 0} \partial_{\Delta\boldsymbol{\theta}} \ln Q - \underbrace{B^{-1}_{jj} \frac{Q(J_o + 1)}{Q(J_o + 1) + \Phi_{Pj}}}_{<0} \partial_{\Delta\boldsymbol{\theta}} \ln Q \tag{A.126}$$

where the sign on the first term follows because $\zeta - \sum_q d_q^{-1} = -Q^{-1}$, and $Q^{-1}\left(1 - \frac{Q(J_o+1)}{Q(J_o+1)+\Phi_{Pj}}\right) > 0$ since $\frac{Q(J_o+1)}{Q(J_o+1)+\Phi_{Pj}} \leq 1$. Further note that:

$$\sum_j -B^{-1}_{jj} \left(\frac{Q(J_o + 1)}{Q(J_o + 1) + \Phi_{Pj}}\right) + B^{-1}_{jj} > 0 \tag{A.127}$$

Define $K_j^\beta := \frac{1}{\zeta^2 d_j^2} \cdot Q^{-1} \cdot \left(1 - \frac{Q(J_o+1)}{Q(J_o+1)+\Phi_{Pj}}\right) - B^{-1}_{jj} \frac{Q(J_o+1)}{Q(J_o+1)+\Phi_{Pj}}$. We can therefore write:

$$\frac{\sum_{j>J_\ell} \partial_{\Delta\boldsymbol{\theta}} B^{-1}_{jj}}{\left(\sum_{j>J_\ell} B^{-1}_{jj}\right)} = \frac{\sum_j K_j^\beta}{\sum_j B_{jj^{-1}}} \partial_{\Delta\boldsymbol{\theta}} \ln Q \tag{A.128}$$

Applying equation (A.127), $\frac{\sum_j K_j^\beta}{\sum_j B^{-1}_{jj}} + 1 = \frac{\sum_j K_j^\beta + B^{-1}_{jj}}{\sum_j B^{-1}_{jj}} > 0$. Thus $\frac{\sum_j K_j^\beta}{\sum_j B^{-1}_{jj}} \geq -1$, and if $|\partial_{\Delta\boldsymbol{\theta}} \ln V| \geq |\partial_{\Delta\boldsymbol{\theta}} \ln Q|$, as implied by Condition 1, the sign of $\partial_{\Delta\boldsymbol{\theta}} \left[E\left[\hat{\beta}_o\right]\right]$ is given by the sign of $\partial_{\Delta\boldsymbol{\theta}} \ln V$.

Next, note that:

$$\partial_{\Delta\boldsymbol{\theta}} \sum_{j,r} B_{jr}^{-1} = \sum_{j,r} -B_{jr}^{-1} \cdot (\partial_{\Delta\boldsymbol{\theta}} \ln \zeta + \partial_{\Delta\boldsymbol{\theta}} \ln d_j + \partial_{\Delta\boldsymbol{\theta}} \ln d_r) \tag{A.129}$$

$$= \frac{-B_{jr}^{-1}}{\zeta} \left(Q^{-1} \left(1 - \frac{Q(J_o+1)}{Q(J_o+1)+\Phi_{Pj}} \right) \right) \partial_{\Delta\boldsymbol{\theta}} \ln Q - B_{jr}^{-1} \frac{Q(J_o+1)}{Q(J_o+1)+\Phi_{Pj}} \tag{A.130}$$

Similar steps as above show that for $K_{jr}^{\omega} := -B_{jr}^{-1} \cdot Q^{-1} \cdot \left(1 - \frac{Q(J_o+1)}{Q(J_o+1)+\Phi_{Pj}}\right) - B_{jj}^{-1} \frac{Q(J_o+1)}{Q(J_o+1)+\Phi_{Pj}}$:

$$\frac{\sum_{j>J_\ell,r} \partial_{\Delta\boldsymbol{\theta}} B_{jr}^{-1}}{\left(\sum_{j>J_\ell,r} B_{jr}^{-1}\right)} = \underbrace{\frac{\sum_{j>J_\ell,r} K_{jr}^{\omega}}{\left(\sum_{j>J_\ell,r} B_{jr}^{-1}\right)}}_{\geq -1} \partial_{\Delta\boldsymbol{\theta}} \ln Q \tag{A.131}$$

so as above, the sign of $\partial_{\Delta\boldsymbol{\theta}}[\omega]$ is given by the sign of $\partial_{\Delta\boldsymbol{\theta}} \ln V$.

Since the signs of both $\partial_{\Delta\boldsymbol{\theta}}[\omega]$ and $\partial_{\Delta\boldsymbol{\theta}}\left[E\left[\hat{\beta}_o\right]\right]$ a are given by the sign of $\partial_{\Delta\boldsymbol{\theta}} \ln V$, condition (A.115) holds, as required.

Now allow that $\partial_{\Delta\boldsymbol{\theta}} \Phi_{Pj} \neq 0$. Setting $\partial_{\Delta\boldsymbol{\theta}} \ln Q = 0$:

$$\partial_{\Delta\boldsymbol{\theta}} \sum_{j} B_{jj}^{-1} = \sum_{j} \left(\frac{1}{d_j} - B_{jj}^{-1} \right) \cdot \frac{-Q^{-1}}{\zeta} \frac{\Phi_{Pj}}{Q(J_o+1)+\Phi_{Pj}} \partial_{\Delta\boldsymbol{\theta}} \ln \Phi_{Pj}$$

$$- B_{jj}^{-1} \frac{\Phi_{Pj}}{Q(J_o+1)+\Phi_{Pj}} \partial_{\Delta\boldsymbol{\theta}} \ln \Phi_{Pj} \tag{A.132}$$

$$> -B_{jj}^{-1} \cdot (J+2) \cdot \frac{\Phi_{Pj}}{Q(J_o+1)+\Phi_{Pj}} \partial_{\Delta\boldsymbol{\theta}} \ln \Phi_{Pj} < 0 \tag{A.133}$$

where we add $1/d_j > 0$ and take the max of $1 - Q^{-1}/\zeta$, which is obtained for $\Phi_{Pj} = 0$ for all j. Following similar steps:

$$\partial_{\Delta\boldsymbol{\theta}} \sum_{j,r} B_{jr}^{-1} > -B_{jr}^{-1} \cdot (J+2) \cdot \frac{\Phi_{Pj}}{Q(J_o+1)+\Phi_{Pj}} \partial_{\Delta\boldsymbol{\theta}} \ln \Phi_{Pj} < 0 \tag{A.134}$$

Applying the expanded version of Condition 1 ensures that the sign of $\partial_{\Delta\boldsymbol{\theta}} \ln V$, reduced by a maximally-pathological supply-elasticity change for both β and ω, still determines the sign of both β and ω as required.

Intuitively, the aggregate and cross-sectional effects vary for two reasons. First, the cross-sectional effect is increasing in relative price effects that may have zero aggregate effect. This implies β may exceed ω. Second, the aggregate effect considers cross-market spillovers due to average increases in prepayment cost everywhere. This implies ω may exceed β.

If there is a change in parameters $\Delta\theta$ that simultaneously *increases* relative price effects and *decreases* the price effects of spillovers, then an increase in β may correspond to a decrease in ω. Condition 1 rules out such pathological perturbations, and focuses attention on parameter changes primarily increase the outflow effects of prepayment costs.

□

B Data construction details

B.1 CRISM and ICE, McDash®details.

We use an individual-level panel for analysis in Section 4, and a local market-level dataset for analysis in Section 3. We also use data from ICE, McDash®on mortgage duration for ancillary analysis and instrument construction.

B.1.1 Individual-level panel.

We randomly draw a 10% sample of individuals from the CRISM monthly panel from June 2005 to December 2024. We restrict to households for whom we observe the purchase of an owner-occupied home with a fixed-rate, 30-year, first-lien mortgage during this period. To ensure an observed mortgage origination corresponds to a home purchase, we require that the mortgage origination date occurs within 180 days of a persistent Equifax address change, where a persistent change is one that does not revert within 6 months, and that the property's zip code of origination aligns with the destination zip code of the persistent address change.

We do not directly observe home sales in CRISM. We infer them based on a combination of mortgage prepayment and moving activity. Specifically, we infer an existing-home sale if three conditions are met: (i) a borrower terminates a purchase mortgage; (ii) a persistent address change occurs within 180 days of mortgage termination; and (iii) the zip code of the property that secured the terminated mortgage aligns with the origin zip code of the persistent address change. To screen out real estate investors and especially distressed sales, we drop sales inferred to be within 1 year of purchase.

We classify an existing-home sale as an own-to-rent move if it meets two additional conditions: (i) households have no outstanding first-lien mortgage balances in the credit records six months after the move; and (ii) a borrower had at least 50% of the origination balance still outstanding on the original purchase loan when the move occurred. The first condition indicates that the borrower did not take out a new mortgage after a move, and the second condition screens out borrowers with enough home equity to purchase a new property in cash without massively downsizing.

The data contains complete Equifax credit bureau information for the full June 2005-Dec 2024 period. However, Equifax waits a year to match loans with ICE, McDash®, which means the most recent matches are from Dec 2025 originations. Since our algorithm for identifying existing-home sales and own-to-rent moves relies on observing activity potentially 6 months after a move, we track activity for up to five years after move-in for our main sample for the 2016 event study. This ensures that originations from December 2017 have a round number of years without truncation issues.

B.1.2 Market-level dataset.

We randomly draw a 50% sample of December observations for individuals in the CRISM monthly panel from 2005-2023. We use this sample to calculate, at the zip code level: (i) churn; (ii) the age distribution; (iii) credit scores; (iv) average appraisal amount, average LTV, average DTI, and ARM share; and (v) scaled prepayment cost.

We calculate (i)-(iii) using averages for local residents, regardless of mortgage characteristics. we restrict the sample to borrowers living in residential addresses to screen out borrowers with zips linked to non-residential business mailing addresses. We estimate churn based on the fraction with a year-over-year change in zip code. We estimate the age distribution by calculating borrower-level age as the difference between year and borrower year of birth, and taking the fraction within each single-year age bucket. Finally, we estimate credit scores by taking the average of borrower Risk Score 3.0.

We calculate (iv) and (v) using averages for local mortgages. We restrict to first-lien mortgages to owner-occupants, excluding mortgages with exotic repayment terms (e.g. weekly or balloon payments) and graduated payment schedules. We assume that the scaled prepayment cost of an adjustable-

rate mortgage is zero, a simplification both because adjustable-rate mortgages have fixation periods which confer present-value prepayment costs, and because our current prepayment cost calculations omit a credit spread relative to the 30-year rate.

The above steps deliver two zip code-by-year panels, which we merge together for our cross-sectional regressions. We then make two sample restrictions. First, we require that a zip code has at least 500 sampled borrowers, both to reduce measurement error in zip-level averages and for data privacy reasons. Second, we drop zip codes with average appraised property values below the 5th or above the 95th percentile. This is because we use FHFA repeat sales indices to estimate price growth. These indices are based on data from mortgages secured by Fannie Mae and Freddie Mac. The indices do not include properties purchased with non-agency jumbo loans or loans backed by the government directly (e.g. FHA, VA loans). Restricting to zip codes with average property values in the middle 90% ensures that price indices are broadly representative of typical transactions in the market.

B.1.3 ICE, McDash® details.

We take a 0.5% sample of loans from 1992Q1-2024Q1 from the unmatched ICE, McDash® dataset. The data has information on loan origination date, loan termination date, contract characteristics (mortgage balance, LTV ratios, DTI ratios, appraised value, interest rates, fixed-vs-variable rate, monthly payments). We use this information to calculate mortgage duration. Our calculations restrict to first-lien, fixed-rate purchase mortgages originated to owner-occupants.

B.2 CoreLogic details.

For analysis in Section 4 of the resale price effect of mortgage rate at origination, we use the parcel-level panel with construction described in Section B.2.1. For analysis on residential investment in Section 3, we use the permitting dataset described in section B.2.2.

B.2.1 Parcel-level panel.

We pull transaction, mortgage, and property characteristic records for all residential properties in CoreLogic. We then identify purchases and sales by owner-occupants with a mortgage where an initial purchase occurred between Jan 2012 and Jan 2018.

We follow several sample selection criteria to focus on purchases by owner-occupants with a mortgage. First, we merge mortgage origination records to property transactions. To match a mortgage record with a property transaction, we require that (i) the transaction record has a flag indicating it was recorded with a mortgage; (ii) the buyer name in the transaction file matches the borrower name in the mortgage file; and (iii) the deed is recorded no more than 10 days before and no more than 21 days after the mortgage is recorded. The fuzzy match in (ii) requires that the Jaro-Winkler edit distance between the two strings is less than 0.2. Second, we drop transactions where the buyer is a corporation or trust, the buyer has a sufficiently different name (to eliminate non-arms length transfers to spouses, children, and other family members), or the buyer mailing address has a different zip code than the property, indicating an investment property or non-primary residence. Third, we restrict to transactions where the matched mortgage is a fixed-rate purchase loan.

After identifying valid purchases, we search for the next valid sale. Our main criteria for a valid sale is that it is arms-length, not an incorporation as a trust or a transfer to a spouse or family member. To screen out the former, we screen out sales to trusts. To screen out the latter, we require that (i) all sellers and buyers on each transaction have sufficiently different names; and (ii) the name of the current seller is different than the name of the next seller.

The CoreLogic mortgage data mostly lack information on interest rates. We impute mortgage rates using the average rate on Freddie Mac purchase mortgages with first payment dates within the same month as the month of first payment for CoreLogic mortgages. We calculate monthly mortgage rates

using the Freddie Mac Single Family Loan-Level dataset. Note this is different than the rates from the Primary Market Mortgage Survey (PMMS), because the PMMS gives posted rates, which may differ from origination rates due to borrowers locking in different rates, and because the PMMS holds fixed borrower characteristics (credit score, LTV) over time.

B.2.2 Permit data.

CoreLogic pulls information on local building permits in some jurisdictions. The building permits dataset records the permitted parcel, the permit issue date, up to three standardized fields describing the project type, and a text field with narrative details from the permit applicant. We use this dataset to identify the number of single-family renovations over 2022-23, as detailed in Section D.13.

First, we restrict to permits for residential investment. We require that the parcel associated with a given unit is zoned for either single-family or mutli-family residential, ensuring that a permit is either for a modification of an existing residential unit or for new residential construction. We use land use codes to classify parcels as either single or multi-family.

Next, we classify permits as new construction or renovation. New construction permits must meet three criteria: (i) "New Construction" is mentioned as one of the project type fields; (ii) no other project types are mentioned; and (iii) if a permit value is listed, the total value exceeds $100,000. The first condition ensures some new construction occurs. The second and third screen out projects with "new construction" of structures that do not add additional units, such as the "new construction" of a shed or garage. Renovation permits mention one of the following: "Home Addition", "Kitchen Remodel", "Bathroom Remodel", "Patios", "Multi-Room Remodel", "Garage Construction", "Decks and Porches", "Pool and Spa Construction".

Finally, we count the total number of permits each year for single-family construction, the number of permits for multi-family construction, and the number of permits for renovations on existing single-family units. To avoid double-counting multiple permits issued for with the same activity (e.g., an HVAC and bathroom permit both for the same new single-family unit property), we allow each parcel to have at most one permit of each type per year-month.

B.3 Supplemental data sources.

American Community Survey. We use annual estimates from the 5-year American Community Survey for zip-code level demographics. Estimates are available from 2011-2022, with the exception of 2020, when ACS did not publish due to the impacts of the Covid-19 pandemic (Census 2021). Where relevant, we impute 2020 values using 2021 estimates. We use the following tables:

- total population (`B01001_001E`);
- total number of households (`B25003_001E`);
- total number of owner-occupied households (`B25003_002E`);
- median family income (`B19113_001E`);
- median owner-assessed housing unit value (`B25077_001E`);
- median gross rent (`B25111_001E`);
- total white population (`B02009_001E`).

Population density. We construct population density as the ratio of ACS population and zip-code land area in meters. Land area in meters comes from ESRI and are sourced from ArcGIS Data and Maps (ESRI 2023).

House price indices. We use FHFA zip-level all-transactions annual house price indices to estimate average market-wide house price growth.

Twin and first births. For data on twin and first births, we use natality information published by the Centers for Disease Control. Our main dataset is drawn from the CDC's WONDER portal, which covers 1995-2023. For our analysis of first births, we supplement the main dataset with information from 1985 and 1987-1994 provided by the NBER.

The CDC's natality data is derived from the National Center for Health Statistics' (NCHS) Vital Statistics Cooperative Program. Through this program, all 50 states and DC send 100% of birth records to the NCHS for processing. The CDC publishes aggregated tables based on this information for each year. The CDC only publishes public-use natality data at the county level for counties with a population of at least 100,000 at the most recent decennial census. When forming instruments that use natality data, we restrict to counties with continuous natality coverage from 1995-present. This reduces the number of zip codes in our sample by 21%.

From the natality dataset, we pull: (i) the annual number of total live births; (ii) the total number of multiple births, including twins and higher (higher-order births account for only 0.1-0.2% of total births); (iii) we use the number of births that are first in maternal birth order (i.e. are "first births."); and (iv) total population, drawn from intercensal population estimates from the Census Bureau. We calculate each of these fields by mother's county of residence, which we map to zip codes using the methodology described below.

Census Building Permit Survey (BPS). The BPS provides information on permits for residential construction in "permit-issuing places". There are about 20,000 permit-issuing places in the survey universe. All permit-issuing places are surveyed annually. Places with regular construction activity (about half of permit-issuing places, covering 99% of total permitting activity) are surveyed monthly. The BPS reports county-level counts of the total number of permits issued for residential construction with 1, 2-4, and 5+ units each month. We pull the county-level file form the Census website as of December 2024, and map from counties to zip codes as described below.

Local housing supply elasticity. We measure the long-term housing supply elasticity by aggregating census tract level estimates from Baum-Snow and Han (2024) (accessed September 2023). We use estimates from the paper's linear IV-finite mixture model. In the replication packet, these estimates are called `gamma01a_newunits_FMM`. To aggregate these estimates, we use Housing and Urban Development crosswalk files mapping from 2000 census tract to zip code. We assign each tract to a unique zip code based on the zip that contains the majority of tract housing units.

Mapping counties to zipcodes. We use 2010 county definitions. We map 2010 counties to zip codes using Housing and Urban Development crosswalk files. For years between 2010-2023, we load Q4 county-to-zip crosswalks. Each zip code is assigned to the unique county where the highest fraction of housing units are located. HUD does not provide data for years before 2010. For those years, we use the 2010 crosswalk files.

Mapping zipcodes to core-based statistical areas. We map zipcodes to CBSAs using (i) the zip-to-county mapping described above and (ii) the county-to-CBSA crosswalk provided on the NBER website, which is sourced from Census.

We manually map counties in Connecticut. Connecticut does not have counties delineated by state or local government. Before July 2022, Census divided Connecticut into 8 county-equivalents for statistical purposes. In July 2022, Census converted from the 8 county-equivalents to 9 (approximately equivalent) planning regions. The planning regions were assigned new FIPS codes. The NBER crosswalk uses the updated planning regions, while CRISM uses legacy county equivalents. We manually map legacy county-equivalents to CBSAs as follows:

- CBSA Bridgeport-Stamford-Norwalk (14860) – Fairfield County (001)
- CBSA Hartford-West Hartford-East Hartford (25540) – Hartford (003), Middlesex (007), and Tolland County (013)
- CBSA New Haven (35300) – New Haven County (009)
- CBSA Norwich-New London (35980) – New London County (011)
- CBSA Torrington (45860) – Litchfield County (005)
- CBSA Worchester MA-CT (49350) – Windham County (015)

B.4 Alternative prepayment cost pricing models.

B.4.1 Agarwal *et al.* (2013) formula.

Define the value of the option to refinance in future periods, assuming no refinancing in t, as $R_t(\cdot)$, and the value of the option to refinance in future periods, assuming refinancing in t, as $R'_t(\cdot)$. Then:

$$R_t(\cdot) := E_t^Q \sum_{0\le q\le n} \frac{m}{1+r_{ft,q}} - E_t^Q \sum_{0\le q\le n} \frac{k^*_{t+q}(\cdot)}{1+r_{ft,q}}, \quad R'_t(\cdot) := E_t^Q \sum_{0\le q\le n} \frac{u(\cdot)}{1+r_{ft,q}} - E_t^Q \sum_{0\le q\le n} \frac{u^*_{t+q}(\cdot)}{1+r_{ft,q}} \tag{B.1}$$

where k^*_{t+q} is the (stochastic) period $t+q$ payment assuming no refinancing in the current period, and u^*_{t+q} is the (stochastic) period $t+q$ payment assuming refinancing in the current period. In the main text, we assume that $k^*_{t+q}(\cdot) = m$ and $u^*_{t+q}(\cdot) = u$, and so $R_t(\cdot) = R'_t(\cdot) = 0$.

We estimate $R_t(\cdot)$ and $R'_t(\cdot)$ each period using updated parameters from the Agarwal *et al.* (2013) paper, detailed below. The main text shows how to estimate hold-to-maturity costs for current and refinanced mortgages, delivering estimates of $E_t^Q \sum_{0\le q\le n} \frac{m}{1+r_{ft,q}}$ and $E_t^Q \sum_{0\le q\le n} \frac{u(\cdot)}{1+r_{ft,q}}$. Taking the difference between hold-to-maturity costs and $R_t(\cdot)$ and $R'_t(\cdot)$ delivers $E_t^Q \sum_{0\le q\le n} \frac{k^*_{t+q}(\cdot)}{1+r_{ft,q}}$ and $E_t^Q \sum_{0\le q\le n} \frac{u^*_{t+q}(\cdot)}{1+r_{ft,q}}$, respectively, We then estimate the market value as:

$$E_t^Q \sum_{0\le q\le n} \frac{m^*_{t+q}(\cdot)}{1+r_{ft,q}} := \min\left[E_t^Q \sum_{0\le q\le n} \frac{k^*_{t+q}(\cdot)}{1+r_{ft,q}}, c_R(M) + E_t^Q \sum_{0\le q\le n} \frac{u^*_{t+q}(\cdot)}{1+r_{ft,q}} \right] \tag{B.2}$$

Appendix Figure B.2 reproduces Figure 1 using our alternative measure of SPC.

Details. In ADL, $R_t = R(x_t)$ is only a function of $x_t := r_t - r_{0t}$, where r_t is the nominal interest rate on new mortgages at date t and r_{0t} is a household's outstanding rate at date t. Define the rate of time preference ρ, the marginal tax rate τ, mortgage balance M, refi closing costs $\kappa(M)$, the "forced move" / exogenous repayment rate λ, and σ is the standard deviation of the new mortgage interest rate. From the proof of Theorem 2:

$$R(x) = K \cdot \exp(-\psi \cdot x), \quad \psi := \frac{\sqrt{2(\rho+\lambda)}}{\sigma}, \tag{B.3}$$

$$K := \frac{M \exp(\psi x^*)}{\psi(\rho+\lambda)}, \quad x^* := \frac{1}{\psi}[\phi + W(-\exp(-\phi))], \quad \phi := 1 + \psi(\rho+\lambda)\frac{\kappa(M)/M}{1-\tau} \tag{B.4}$$

for the optimal refinancing threshold x^*, where $W(\cdot)$ is the Lambert W-function.

We follow the main ADL calibrations of $\rho = 0.05$ and $\kappa(M) := 0.01 \cdot M + 2000$. We update the remaining parameters to reflect developments since 2013. ADL calibrates $\tau = 0.28$, the intermediate tax bracket at the time of the paper, to reflect the value of the mortgage interest rate deduction. We set $\tau = 0.24 \cdot 0.2$, reflecting two changes in the 2017 Tax Cuts and Jobs Act: (i) the middle tax bracket falling to 24%, and (ii) the fact that only about 20% of homeowners with a mortgage deduct mortgage

interest from their taxes.[60] We set $\lambda = 0.077$, which differs from the ADL calibration of 0.147 because we assume (i) move-out rates of 6%, rather than 10%, and (ii) do not consider "exogenous repayment" due to inflation, as our measure of SPC is nominal, not real. Finally, we update σ by computing the annualized standard deviation of monthly observations of the average rate on new 30-year fixed-rate mortgages from the Freddie Mac Primary Mortgage Market Survey over 1971-2019 (equal to $\sqrt{12} \times \sigma_m$, where σ_m is the standard deviation of monthly observations).

B.4.2 Using MBS zero-volatility spreads.

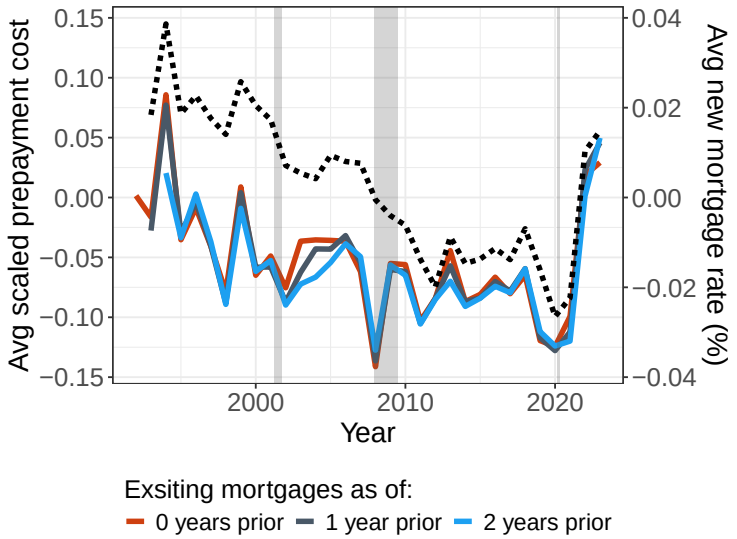
Assume that $E_t^Q \sum_{0 \le q \le n} \frac{m^*_{t+q}(\cdot)}{1+r_{ft,q}} = V_{it}$, the market price of mortgage i. By definition, the zero-volatility spread zv_{it} satisfies $V_{it} = \sum_{0 \le q \le n} \exp(-zv_{it} \cdot q) \cdot m \cdot \rho_t(q)$, where $\rho_t(q)$ is the period-t price of a zero-coupon bond with maturity q. We do not observe traded prices for all outstanding mortgages, but we do observe zero-volatility spreads given daily TBA MBS prices, obtained through JP Morgan. Call these spreads zv_t. We assume:

$$E_t^Q \sum_{0 \le q \le n} \frac{m^*_{t+q}(\cdot)}{1+r_{ft,q}} := \sum_{0 \le q \le n} \exp(-zv_t \cdot q) \cdot m \cdot \rho_t(q) \tag{B.5}$$

$$E_t^Q \sum_{0 \le q \le n} \frac{u^*(\cdot)_{t+q}}{1+r_{ft,q}} := \sum_{0 \le q \le n} \exp(-zv_t \cdot q) \cdot u(\cdot) \cdot \rho_t(q) \tag{B.6}$$

Using zv_t to discount newly-issued mortgages is appropriate if prepayment risk for an individual mortgage approximately equals market expectations for prepayment risk on a basket of mortgages. However, the zero-volatility spread used to discount *existing* mortgages should depend on the original mortgages' maturity and mortgage rate. This assumption is inaccurate if prepayment risk depends on maturity, calendar date, and mortgage issuance date. We reproduce Figure 1 using our alternative measure of SPC in Appendix Figure B.3.

Figure B.1: Scaled prepayment cost by year: different mortgage distribution lags.

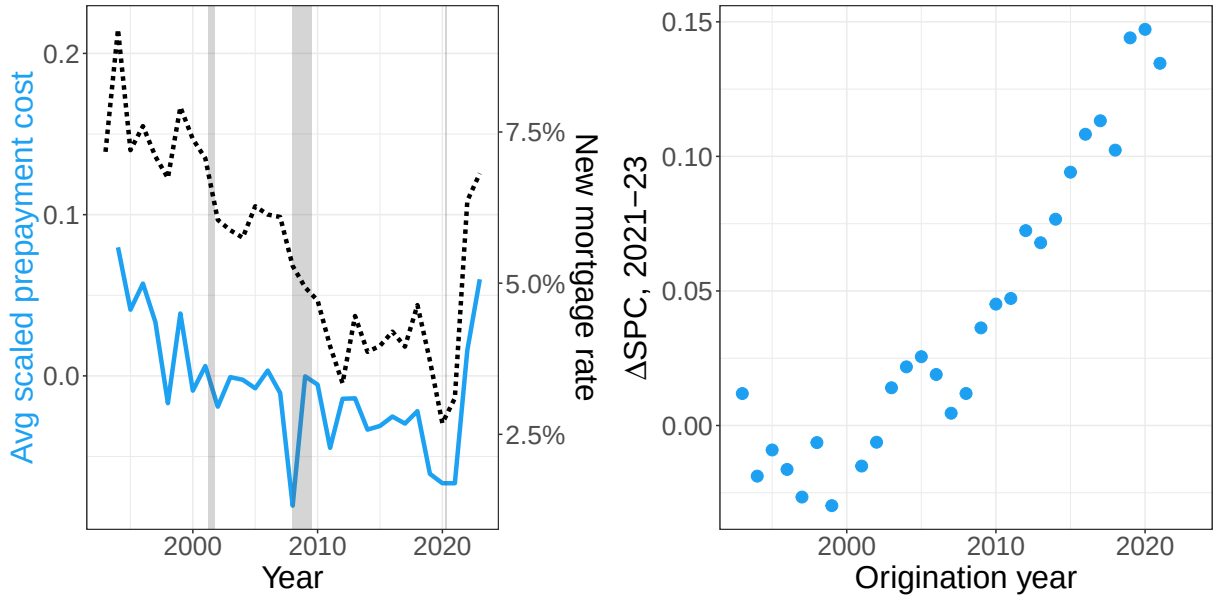


Source: ICE, McDash® and PMMS. The figure plots $SPC_t(I_{t-L})$ for $L = 0, 1, 2$. The blue series on the left scale plots results for $L = 2$ as in the main text, averaging across scaled prepayment costs in t for mortgages outstanding in $t-2$. Red and grey series plot results for $L = 0$ and $L = 1$, respectively. The black dashed series on the right scale plots the Freddie Mac PMMS. Grey regions indicate NBER recessions.

[60] The 2017 Tax Cuts and Jobs Act increased the standard deduction and so decreased itemization. Center (2024) estimates that 8% of households took the mortgage interest rate deduction in 2024; since about 40% of households are owner-occupants with a mortgage, 20% of owner occupants with a mortgage paid their mortgage interest pre-tax.

Figure B.2: Average scaled prepayment costs over time: Measurement with ADL formula.

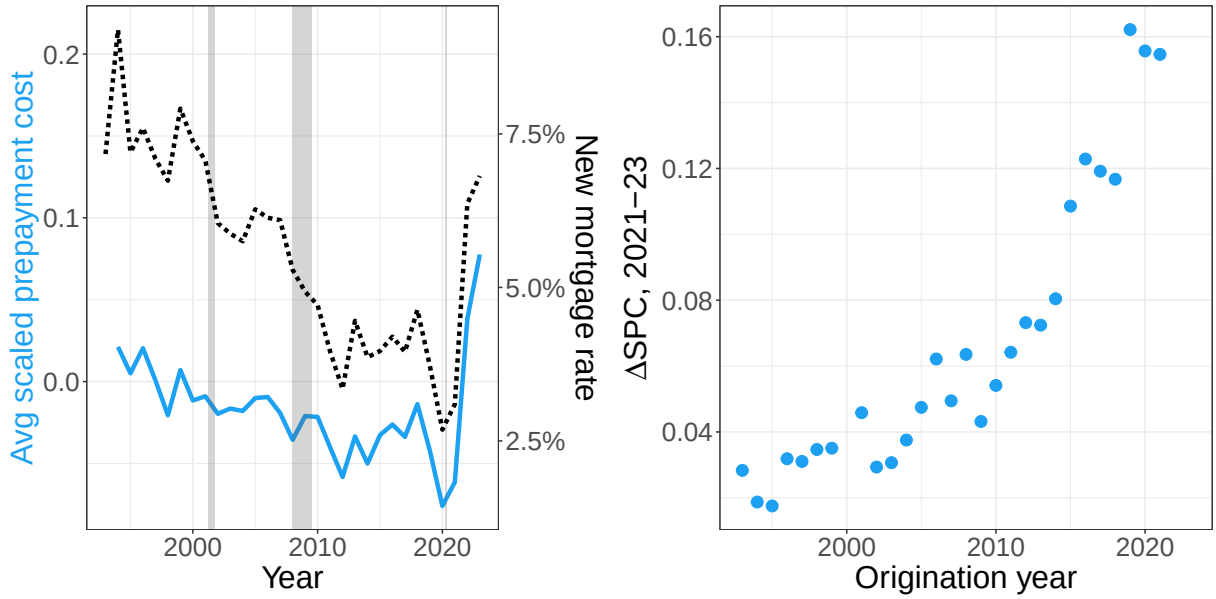
(a) By year. (b) By origination year.



Source: ICE, McDash® and PMMS. Panel (a). The blue solid series on the left scale plots $SPC_t(I_{t-2})$ by t, where SPC_{it} is calculated as described in Section B.4.1. The black dashed series on the right scale plots the Freddie Mac PMMS. Grey regions indicate NBER recessions. Panel (b). The figure plots $SPC_{2023}(I_y) - SPC_{2021}(I_y)$ on the vertical axis against origination year y on the horizontal axis, where I_y is the set of mortgages originated in y.

Figure B.3: Average scaled prepayment costs over time: Measurement with zero-vol spreads.

(a) By year. (b) By origination year.



Source: ICE, McDash® and PMMS. Panel (a). The blue solid series on the left scale plots $SPC_t(I_{t-2})$ by t, where SPC_{it} is calculated as described in Section B.4.1. The black dashed series on the right scale plots the Freddie Mac PMMS. Grey regions indicate NBER recessions. Panel (b). The figure plots $SPC_{2023}(I_y) - SPC_{2021}(I_y)$ on the vertical axis against origination year y on the horizontal axis, where I_y is the set of mortgages originated in y.

C Section 4 details.

C.1 Facts about existing-home sales across the lifecycle.

Data. We supplement the household-level CRISM panel with two surveys. First, we track housing tenure over time using the National Longitudinal Survey of Youth (NLSY79), a high-quality, nationally-representative panel of workers born 1957-1964. Second, we use American Community Survey (ACS) microdata cross-sections from 2001-2023 to connect family size with housing unit characteristics.

1. Moves between owner-occupied and rental housing are common. Figure C.4a plots the probability that owners become renters (red) and renters become owners (blue) by year using NLSY79. The rate of renter entry into owner-occupancy peaks from 1990-2000 at 8-10% per year. The rate of owner exit into renting declines monotonically, from around 14% per year in 1985 to about 2.5% in 2020. From 2000-present, the overall move rate ranged from a low of 8% in 2016 to a high of 13.5% in 2004. Therefore, between 25-39% of total moves among owner-occupants are own-to-rent.

2. Existing-home sales decline with age. In CRISM, we identify households who move into owner-occupied units and track subsequent move-out behavior. We control for residency-duration effects by tracking all households over the four years following move-in, with four years chosen as the time at which move-out rates stabilize (see Appendix Figure C.6a). We restrict to owner-occupants who purchase a home before December 2019 to avoid truncation. Figure C.4b plots in black dots the relationship between age at move-in and existing-home sales over the next four years, controlling for calendar year of origination and property zip code. Existing-home sales over four years decline almost monotonically with age, from 15% at age 20 to 8% at age 70.

3. Own-to-rent moves are U-shaped in age. Figure C.4b plots in red the rate of existing-home sales motivated by own-to-rent moves. Unlike overall moves, the own-to-rent move rate is U-shaped in age.

4. Having kids drives existing-home sales. We estimate the relationship between number of children and unit size through cross-sectional regressions in ACS. Figure C.4c plots the relationship between total number of kids and the number of bedrooms, relative to households with no kids, holding fixed income, county-year, and county-age of household head.

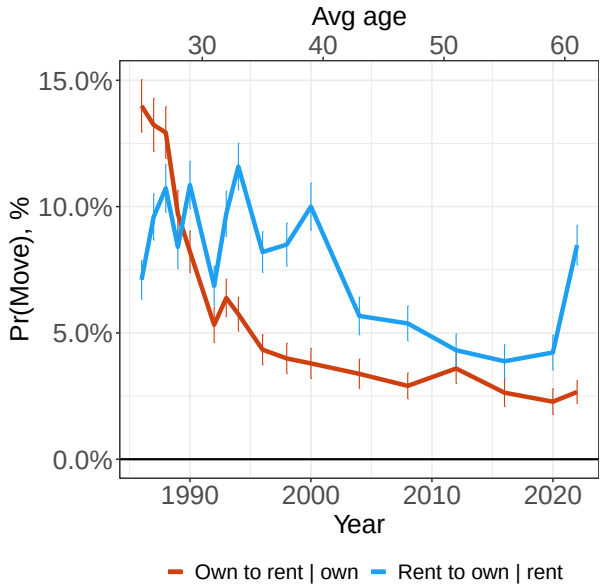
Overall bar heights plot total effects. Households with more children have units with more bedrooms, with an additional child increasing the average number of bedrooms by between 0.25 and 0.3. This implies families purchase new homes as they grow.

Blue bars show the relationship for households with youngest child under age three, while red bars add effects for households where the youngest child is over three. For higher birth orders, a larger share of the incremental effect of an additional child shows up when the youngest is relatively older. Households appear to move around the time they have their first kid – a household with one child has about 0.25 more bedrooms than a household with one child, regardless of the child's age. Higher birth orders lead to additional moves, but with a lag. For example, a household with four kids has only 0.125 more bedrooms than one with three while the youngest is under age three, with another 0.125 more bedrooms once the youngest gets older.

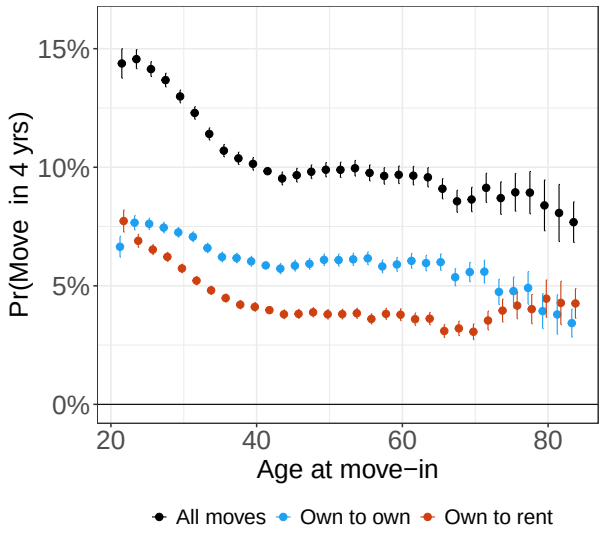
5. Most moves are within-city. Figure C.4d plots existing-home sales motivated by moves across CBSAs. 23% of existing-home sales result in cross-CBSA moves. About 50% of cross-CBSA moves are own-to-rent, possibly reflecting that households enter new markets as renters before buying.

Figure C.4: Lifecycle drivers of existing-home sales.

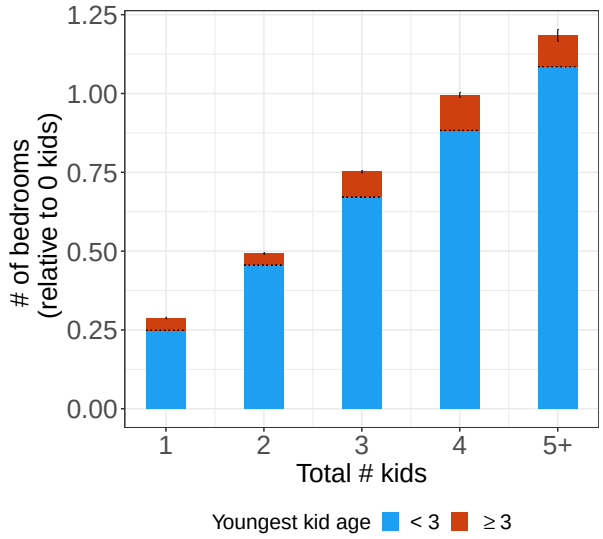
(a) Moves conditional on initial tenure.



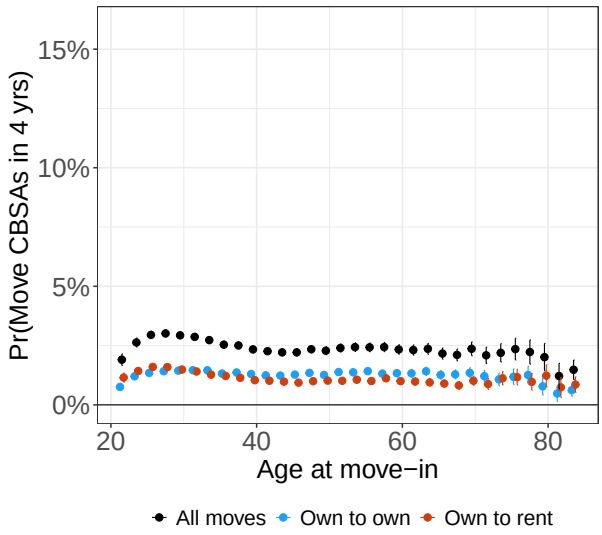
(b) Existing-home sales by age.



(c) Unit size by number of kids.



(d) Cross-city moves by age.



Source: National Longitudinal Survey of Youth, 1979 (a); Equifax Credit Risks Insight Servicing (b, c); and 2001-2023 American Community Survey via IPUMS USA (d). Spikes plot robust 95% confidence intervals. Panel (a). The sample restricts to respondents over age 21 living apart from parents. Respondents are coded as owner-occupants if either they or their spouse owns their current residence. Panel (b). The sample restricts to households who purchase an owner-occupied home with a mortgage between June 2005 and December 2019, and restricts to moves that coincide with an existing-home sale. Own-to-rent moves have no new home purchase within 6 months of initial home sale; own-to-own moves do. Dots plot β_a, plus the average group-wise four-year move rate at age 40, for the regression: $\Pr(Move)_{iq} = \sum_{a \neq 40} \beta_a \cdot I(age_i = a) + \gamma_{z(i)} + \delta_q + \epsilon_{iq}$, where i and q index household and home purchase quarter; $\gamma_{z(i)}$ are zip code fixed effects, and δ_q are origination quarter fixed effects. Spikes ignore the standard error of the mean for the age-40 move rate. Panel (c). The same as panel (b), with outcome variables interacted with an indicator for whether moves crossed CBSA boundaries. Panel (d). The sample restricts to single-family, owner-occupied households with a mortgage. Blue bars plot β_{1n} and red bars plot β_{2n} from the following regression, where i indexes households and t survey year: $Bedrooms_{it} = \sum_n \beta_{1n} \cdot I(Kids_{it} = n) + \beta_{2n} \cdot I(Kids_{it} = n) \times I(MinKidAge_{it} > 3) + \gamma_{g(i),a(i)} + \delta_{g(i),t} + \ln(Income_{it}) + \epsilon_{it}$, where $g(i)$ and $a(i)$ gives county of residence and age of household head, respectively; $I(Kids_{it} = n)$ is an indicator for whether the household n kids under 18;$I(MinKidAge_{it} > 3)$ is an indicator for whether the youngest kid is aged 3-18; and $Income_{it}$ is total family income.

C.2 Confounds from generic origination timing instruments.

On average, the origination timing variation that produces rate lock is confounded. For each person in our panel from 2006-present, we calculate the difference between rate at move-in and the mortgage

rate until the end of the sample, and label deciles of this variable `delta_fwd_q`. Households with higher levels of the variable are more "locked" over the sample period. We then interact the mortrgage rate at origination with `delta_fwd_q`, and project demographics on the interacted variables. Table C.1 shows results. People with relatively high expected rates after move-in are negatively selected on Equifax Risk Score 3.0, LTV, and income.

Table C.1: General timing confounds.
(Source: Equifax Credit Risks Insight Servicing)

Dependent Variables: Model:	Risk score 3.0 (std) (1)	Age (std) (2)	LTV (%, std) (3)	std_move_in_dti_ratio (4)	std_move_in_cur_int_rate (5)	Income (USD '000s, std) (6)	Appraised amt (USD '000s, std) (7)
Variables							
delta_fwd_q = (-2.24,-1.83]-0.139,0.0733]")	0.0149	-0.0329	0.0420	0.1085**	0.4347***	-0.0363	-0.0070
	(0.0209)	(0.0315)	(0.0324)	(0.0468)	(0.0768)	(0.0244)	(0.0165)
delta_fwd_q = (-1.83,-1.53]-0.139,0.0733]")	-0.0173	-0.0317	0.0477	0.0972**	0.3065***	-0.0386*	-0.0164
	(0.0191)	(0.0302)	(0.0319)	(0.0460)	(0.0842)	(0.0230)	(0.0153)
delta_fwd_q = (-1.53,-0.821]-0.139,0.0733]")	0.0104	-0.0182	-0.0032	0.0459	0.1313**	-0.0176	0.0014
	(0.0168)	(0.0181)	(0.0259)	(0.0415)	(0.0499)	(0.0216)	(0.0128)
delta_fwd_q = (-0.821,-0.547]-0.139,0.0733]")	-0.0019	0.0064	0.0094	0.0345	0.0911*	-0.0057	0.0091
	(0.0144)	(0.0123)	(0.0185)	(0.0445)	(0.0526)	(0.0171)	(0.0113)
delta_fwd_q = (-0.547,-0.139]-0.139,0.0733]")	-0.0045	-0.0064	-0.0018	0.0100	0.0447*	-0.0004	0.0015
	(0.0084)	(0.0089)	(0.0065)	(0.0101)	(0.0238)	(0.0102)	(0.0046)
delta_fwd_q = (0.0733,0.27]-0.139,0.0733]")	-0.0093	-0.0147	0.0207	-0.0217**	-0.1211**	0.0034	-0.0125
	(0.0096)	(0.0132)	(0.0157)	(0.0107)	(0.0507)	(0.0060)	(0.0171)
delta_fwd_q = (0.27,0.52]-0.139,0.0733]")	-0.0799***	-0.0120	0.0338*	-0.0065	-0.1928***	-0.0190*	-0.0222
	(0.0196)	(0.0146)	(0.0179)	(0.0134)	(0.0520)	(0.0108)	(0.0213)
delta_fwd_q = (0.52,0.713]-0.139,0.0733]")	-0.0581***	-0.0084	0.0261	-0.0086	-0.2456***	-0.0229**	-0.0186
	(0.0179)	(0.0143)	(0.0178)	(0.0138)	(0.0550)	(0.0102)	(0.0209)
delta_fwd_q = (0.713,0.92]-0.139,0.0733]")	-0.0513***	-0.0065	0.0368**	-0.0204	-0.2837***	-0.0336**	-0.0318
	(0.0185)	(0.0176)	(0.0184)	(0.0189)	(0.0611)	(0.0132)	(0.0239)
Fixed-effects							
move_in_prop_zip	Yes	Yes	Yes	Yes	Yes	Yes	Yes
county-move_in_qtr	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics							
Observations	1,621,058	1,765,896	1,750,213	1,187,223	1,772,978	1,187,102	1,770,121
R^2	0.21391	0.12520	0.23907	0.13108	0.86814	0.30386	0.63711
Within R^2	0.00017	1.6×10^{-5}	4.41×10^{-5}	5.64×10^{-5}	0.01187	3.82×10^{-5}	3.54×10^{-5}

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

C.3 November 2016 event study: Robustness checks

Placebo test: Patterns around November 2014. We replicate in Figure C.5 panels (a) and (c) of Figure 6 two years prior to the November 2016 event. The left panel shows no sharp increase in mortgage rates for originations just before vs. just after November 2014 (although rates do fall slightly). The right panel shows the probability of an existing-home sale overall, or a sale resulting in an own-to-rent move, does not change for originations around November 2014.

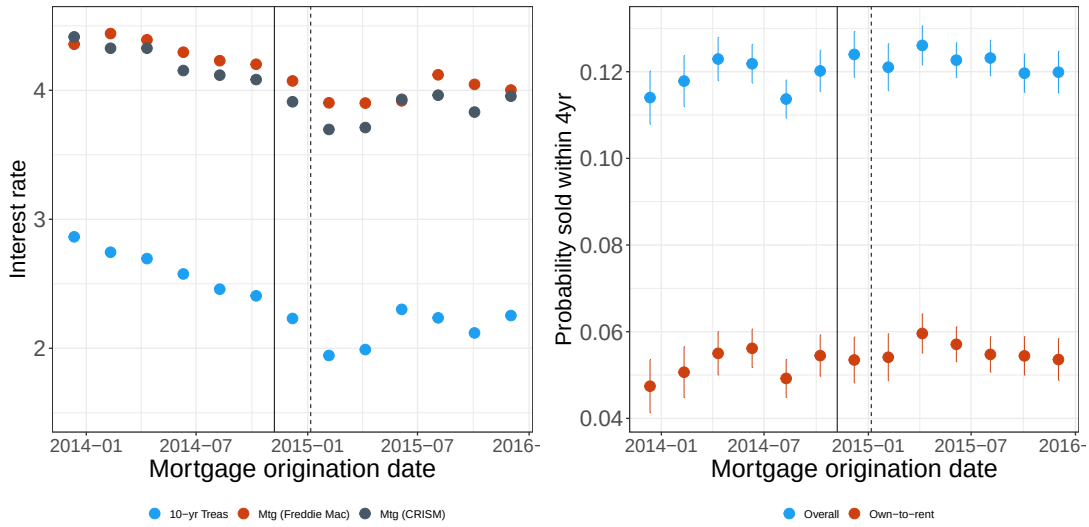
Difference-in-difference design. We use observed seasonal patterns over 2014-2015 to control for any origination-date-of-year effects in the interest rate or four-year existing-home sales rate using an approach similar to difference-in-differences designs. Define $\tau_i \in \{0,1\}$ as an indicator for whether i is in the "treatment group," with origination date between Jan 1, 2016 and Dec 31, 2017, rather than the "control group," with origination date between Jan 1, 2014 and Dec 31, 2015. Define $Post_i \in \{0,1\}$ as an indicator for whether i's origination date is after Nov 7, 2014 (for $\tau_i = 0$) and after Nov 7, 2016 (for $\tau_i = 1$). We estimate the equation:

$$y_i = \beta_0 + \beta_1 \tau_i \times Post_i + \beta_2 \tau_i + \beta_3 Post_i + \eta' X_i + \varepsilon_i \tag{C.1}$$

where the coefficient of interest, β_1, gives the difference in outcome y_i before and after the treatment date for the treatment relative to the control group. We exclude originations in the 60 days after November 7 of the reference year for both the treatment and control groups.

Table C.2 shows estimates of β_1 where the outcome is mortgage rate at origination in Column (1), the unconditional probability of an existing-home sale in Column (2), and the probability of an existing-home sale resulting in an own-to-rent move in Column (3). Estimates are similar to those in

Figure C.5: Placebo: November 2014 change in rates and existing-home sales.



Source: Equifax Credit Risks Insight Servicing and FRED. The figure replicates panels (a) and (c) of Figure 6 for origination dates between Jan 1, 2014 and Dec 31, 2015. Observations are grouped into 60-day bins relative to Nov 7, 2014. See notes to Figure 6 for details.

the main text.

Table C.2: November 2016 event study: Difference-in-differences estimates

Dependent Variables:	Rate (1)	Pr(Moved in 4yr) (2)	Pr(Moved to rent in 4yr) (3)	Rate (4)	Pr(Moved in 4yr) (5)	Pr(Moved to rent in 4yr) (6)
Treated x post	0.9314***	0.0376***	0.0105***	0.6801***	0.0387***	0.0085***
	(0.0102)	(0.0045)	(0.0029)	(0.0123)	(0.0031)	(0.0021)
Treated	-0.5599***	-0.0002	-0.0038**	-0.4391***	0.0026	-0.0014
	(0.0067)	(0.0025)	(0.0019)	(0.0091)	(0.0021)	(0.0015)
Post	-0.3242***	0.0028	0.0018	-0.2878***	0.0024	0.0015
	(0.0069)	(0.0029)	(0.0021)	(0.0077)	(0.0025)	(0.0018)
Window width	4mo	4mo	4mo	1yr	1yr	1yr
Fixed-effects						
Zip-ever refi	Yes	Yes	Yes	Yes	Yes	Yes
Age quintile-Inc quintile-Appraised quintile	Yes	Yes	Yes	Yes	Yes	Yes
Observations	142,011	142,011	142,011	300,375	300,375	300,375
R^2	0.51360	0.16695	0.14499	0.36448	0.11340	0.08883
Within R^2	0.36605	0.01415	0.00519	0.21773	0.01752	0.00583

Clustered (Person & Orig. date) standard-errors in parentheses
****: 0.01, **: 0.05, *: 0.1. Standard errors clustered by person and orig. date in parentheses.*

Source: Equifax Credit Risks Insight Servicing. The table presents coefficient estimates for equation (C.1). Outcomes are fixed mortgage rate in Column (1), the probability of a move resulting in an existing-home sale within 4 years in Column (2), and the probability of an own-to-rent move resulting in an existing-home sale within 4 years in Column (3).

Excluding DC. It is possible that some post-November 2016 reflect churn from political appointees moving into or out of the Washington, DC area due to Trump's election. To investigate whether this affects results, I re-estimate the specification from the main text excluding all moves to and from counties within the Washington-Arlington-Alexandria, DC-VA-MD-WV CBSA. Results are in Table C.3, and are essentially identical to those in the main text.

Effects over time. Figure C.8b shows differences in outstanding mortgage rates for treatment and control by number of years since originatoin. Table C.5 shows dynamic estimates over different horizons, both for reduced-form and IV.

Sticky moves. We re-run our event study considering a subset of households especially inelastic to changes in the move-in mortgage rate: people making cross-CBSA moves. A cross-CBSA move is likely

Table C.3: November 2016 event study: Main estimates, excluding DC

Dependent Variables:	Pr(Moved in 4yr) (1)	Pr(Moved to rent in 4yr) (2)	Pr(Moved in 4yr) (3)	Pr(Moved to rent in 4yr) (4)
Rate	0.052***	0.017**	0.052***	0.018**
	(0.012)	(0.007)	(0.013)	(0.008)
Trend width	4mo	4mo	4mo	4mo
DC?	No	No	Yes	Yes
Fixed-effects				
Zip-ever refi	Yes	Yes	Yes	Yes
Age quintile-Inc quintile-Appraised quintile	Yes	Yes	Yes	Yes
Observations	68,941	68,941	65,403	65,403

Clustered (Person & Orig. date) standard-errors in parentheses
****: 0.01, **: 0.05, *: 0.1. Standard errors clustered by person and orig. date in parentheses.*

Source: Equifax Credit Risks Insight Servicing. The table replicates Table 4, but excludes all moves to or from the Washington-Arlington-Alexandria, DC-VA-MD-WV CBSA. See notes in Table 4 for details.

motivated by a life event like new job, divorce, marriage, or health event, and so is less likely to be canceled due to an increase in mortgage rates. Restricting to cross-CBSA moves reduces our sample size by over 85%, since the vast majority of moves are within-CBSA. We therefore use the difference-in-differences design described above to improve power. Table C.4 shows results, which are similar to those in Table C.2, columns (1)-(2).

Table C.4: November 2016 event study: Sticky moves

Dependent Variables:	Rate (1)	Pr(Moved in 4yr) (2)
Treated x post	0.9634***	0.0386**
	(0.0182)	(0.0196)
Treated	-0.5804***	-0.0099
	(0.0124)	(0.0128)
Post	-0.3471***	0.0148
	(0.0120)	(0.0137)
Window width	4mo	4mo
Fixed-effects		
Zip-ever refi	Yes	Yes
Age quintile-Inc quintile-Appraised val.	Yes	Yes
Observations	18,644	18,644
R^2	0.72522	0.48531
Within R^2	0.41758	0.01963

Clustered (Borrower & Move in date) standard-errors in parentheses
****: 0.01, **: 0.05, *: 0.1. Standard errors clustered by person and orig. date in parentheses.*

Source: Equifax Credit Risks Insight Servicing. The table presents coefficient estimates for equation (C.1), where we restrict to observations where the initial move-in involved a move across CBSAs.

Truncation test. In CRISM, we lose track of households who refinance in our sample period to a servicer not covered in ICE, McDash®. This generally biases downwards our estimate of the effect of rate lock on mobility, since consumers are more likely to refinance (and hence be lost) after the 2016 interest rate increase. We address this by assuming that conditional on choosing to refinance, the fact of refinancing out of the ICE, McDash® sample is uncorrelated with future mobility. We partly test this assumption by regressing an indicator for losing track of a consumer on observable characteristics, controlling for whether a consumer ever refinances. Results are in Table C.9.

Table C.5: November 2016 event study: Dynamic IV effects.
(Source: Equifax Credit Risks Insight Servicing)

Dependent Variables:	Pr(Moved in 3yr)	Pr(Moved in 4yr)	Pr(Moved in 5yr)	Pr(Moved in 6yr)	Pr(Moved in 7yr)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Rate	0.0107	0.0517***	0.0568***	0.0425**	0.0446**
	(0.0088)	(0.0124)	(0.0161)	(0.0175)	(0.0180)
Trend width	4mo	4mo	4mo	4mo	4mo
Fixed-effects					
Zip-ever refi	Yes	Yes	Yes	Yes	Yes
Age quintile-Inc quintile-Appraised quintile	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	68,941	68,941	68,941	68,941	68,941
R^2	0.20911	0.22490	0.25437	0.28070	0.29390
Within Adjusted R^2	0.00304	0.01354	0.03825	0.06539	0.08112
F-test (1st stage), Rate	9,997.1	9,485.7	8,711.1	8,124.4	7,942.1
Wald (1st stage), Rate	2,479.3	2,327.4	2,237.6	2,202.8	2,094.8

Clustered (Person & Orig. date) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

C.4 Implicit borrower discount rate

We calculate the implicit risk-neutral constant discount rate that a representative borrower uses to discount the future using the following formula:

$$\Delta P = \sum_{a=1}^{\tau} \Pr(\text{active in } a) \cdot \frac{\Delta m}{(1+r)^a} \tag{C.2}$$

where ΔP is the price premium, Δm is the difference in annual costs, r is the risk neutral rate the representative borrower uses to discount future costs, τ is the number of years remaining until repayment, $\Pr(\text{active in } a)$ is the probability that a loan outstanding will remain active in a years.

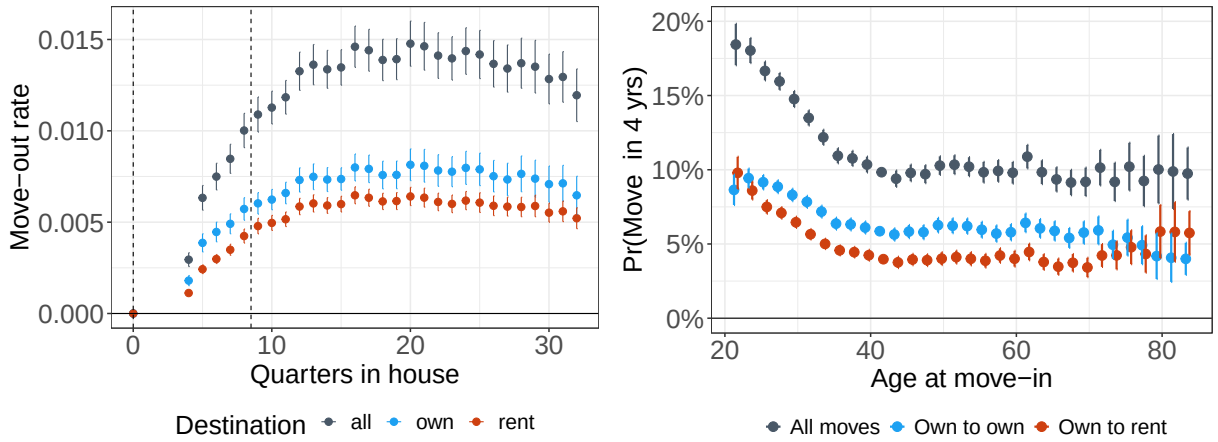
We consider Δm implied on the average loan for a 1pp difference in mortgage origination rate, and use our estimates to predict the implied ΔP. We estimate $K(a)$ as in Appendix Section D.5.3.

C.5 Additional figures

Figure C.6: Existing-home sales summary statistics

(a) Existing-home sales by current home residency duration.

(b) Existing-home sales by age, move-in from 2013-2016.

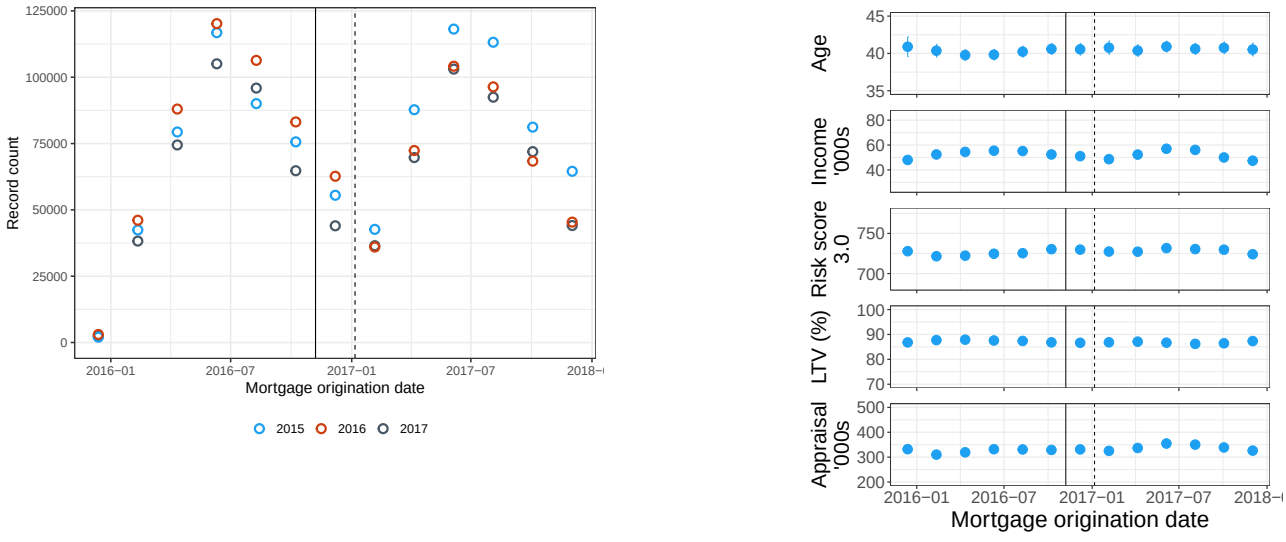


Source: Equifax Credit Risks Insight Servicing. Spikes plot 95% confidence intervals calculated from standard errors clustered by household and calendar date. Panel (a). The sample restricts to households who purchase an owner-occupied home with a mortgage between June 2005 and December 2019, conditions on homeowners who stay for at least a year, and restricts to moves that coincide with an existing-home sale. Own-to-rent moves have no new home purchase within 6 months of initial home sale; own-to-own moves do. Dots plot β_τ for the regression: $\Pr(Move)_{it} = \sum_\tau \beta_\tau \cdot I(\tau = t - q(i)) + \lambda_{a(it)} + \gamma_{c(i)} + \delta_{q(i)} + \omega_t + \epsilon_{it}$, where i and t index households and calendar quarter; $I(\cdot)$ is an indicator function; $q(i)$ is calendar quarter of origination, $\lambda_{a(it)}$ are current age fixed effects, $\gamma_{c(i)}$ are county of residence fixed effects, $\delta_{q(i)}$ are calendar quarter of origination fixed effects, and ω_t are calendar quarter effects. Panel (b). The sample restricts to households who purchase an owner-occupied home with a mortgage between Jan 2013 and December 2016, conditions on homeowners who stay for at least a year, and restricts to moves that coincide with an existing-home sale. Own-to-rent moves have no new home purchase within 6 months of initial home sale; own-to-own moves do. Dots plot β_a, plus the average group-wise four-year move rate at age 40, for the regression: $\Pr(Move)_{iq} = \sum_{a \neq 40} \beta_a \cdot I(age_i = a) + \gamma_{z(i)} + \delta_q + \epsilon_{iq}$, where i and q index household and home purchase quarter; $\gamma_{z(i)}$ are zip code of residence fixed effects, and δ_q are origination quarter fixed effects. Standard errors do not account for sampling error in computing the age 40 mean.

Figure C.7: Balance, November 2016 event study

(a) Origination density.

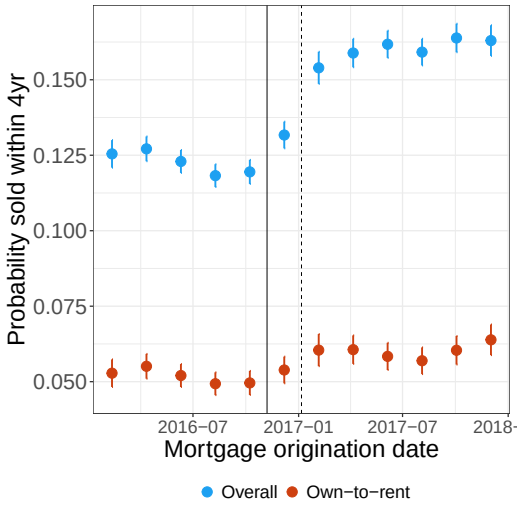
(b) Covariate balance.



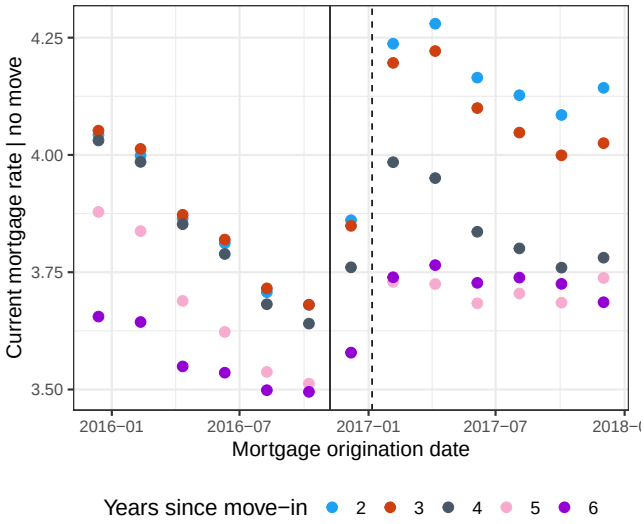
Source: Equifax Credit Risks Insight Servicing. (a) The figure plots raw record counts in the 10% CRISM sample within 60-day buckets. The red dots plot origination counts for 2016-17; the blue dots plot origination counts for 2015-16 at equivalent day-of-year times; and the grey dots plot origination counts for 2017-2018 at equivalent day-of-year times. (b) The figure plots average age, origination LTV, income, Vantage Risk Score 3.0, and appraised value of the property securing the mortgage within 60-day buckets around Nov 7, 2016.

Figure C.8: November 2016 event study: Additional dynamic results.

(a) Existing-home sales by origination date, levels.

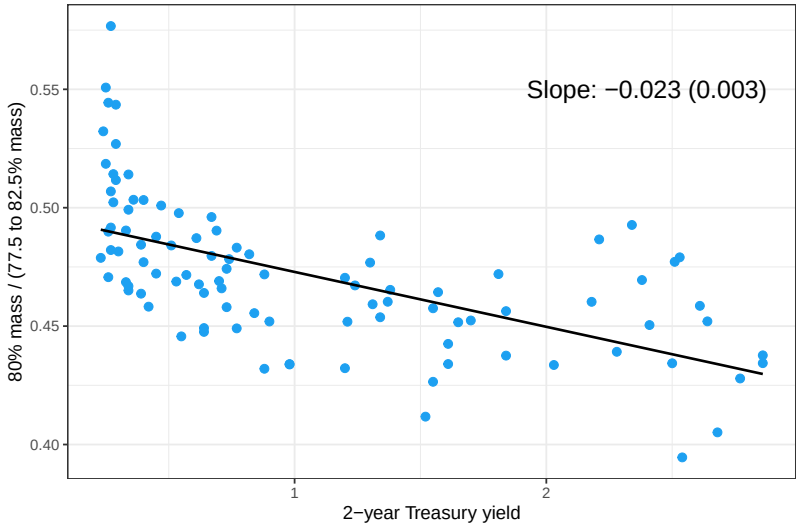


(b) Dynamic effects on outstanding interest rates.



Source: Equifax Credit Risks Insight Servicing. (a) The figure plots the probability of an overall move (blue) or an own-to-rent move (red) within 4-years of move-in for households who initially moved in at the date given on the horizontal axis. Spikes plot 95% confidence intervals. (b) The figure shows the current mortgage rate, conditional on no move before the year given in the legend.

Figure C.9: Bunching at 80% LTV correlates with the risk-free rate.



Source: Equifax Credit Risks Insight Servicing and FRED. The figure restricts to conventional 30-year fixed rate purchase mortgages originated from 2012-2019. Each dot represents a month. The vertical axis gives the fraction of loans with a purchase LTV of exactly 80% relative to the total mass of loans with purchase LTV of between 77.5% and 82.5%. The horizontal axis gives the average monthly 2-year Treasury yield.

C.6 Additional tables

Table C.6: CRISM summary statistics.
(Source: Equifax Credit Risks Insight Servicing)
(a) Full sample.

Statistic	N	Mean	St. Dev.
Age (yrs)	1,946,106	39.896	12.498
Income (USD)	1,251,155	61,639.780	198,187.400
Appraised purchase price (USD)	1,950,683	295,024.300	325,705.900
Equifax Risk Score 3.0	1,719,115	725.289	65.179
Origination interest rate (%)	1,955,627	4.812	1.132
Origination mortgage balance	1,955,627	240,008.600	154,652.400
Monthly payment (USD)	1,953,246	1,518.637	899.262
LTV (%)	1,927,833	86.448	13.156
DTI (%)	1,251,304	36.154	12.091
Pr(Moved in 4 yrs)	1,955,627	0.099	0.299
Pr(Moved to rent in 4 yrs)	1,955,627	0.042	0.200
Pr(Moved outside CBSA in 4 yrs)	1,955,627	0.023	0.150
Pr(Moved outside state in 4 yrs)	1,955,627	0.018	0.133

(b) Nov 2016 event study sample.

Statistic	N	Mean	St. Dev.
Age (yrs)	68,774	40.584	12.609
Income (USD)	42,636	64,540.060	81,109.940
Appraised purchase price (USD)	68,915	333,158.100	267,763.700
Equifax Risk Score 3.0	67,979	730.833	61.078
Origination interest rate (%)	68,941	3.922	0.477
Origination mortgage balance	68,941	272,867.400	170,139.400
Monthly payment (USD)	68,811	1,619.020	909.150
LTV (%)	68,732	87.001	12.339
DTI (%)	42,637	35.300	9.652
Pr(Moved in 4 yrs)	68,941	0.125	0.331
Pr(Moved in 4 yrs), move-in before Nov 2016	40,344	0.114	0.318
Pr(Moved in 4 yrs), move-in after Nov 2016	28,597	0.141	0.348
Pr(Moved to rent in 4 yrs)	68,941	0.050	0.219
Pr(Moved to rent in 4 yrs), move-in before Nov 2016	40,344	0.047	0.213
Pr(Moved to rent in 4 yrs), move-in after Nov 2016	28,597	0.055	0.227
Pr(Moved outside CBSA in 4 yrs)	68,941	0.029	0.169
Pr(Moved outside state in 4 yrs)	68,941	0.023	0.149

Table C.7: CoreLogic summary statistics.
(Source: CoreLogic)
(a) Full sample.

Statistic	N	Mean	St. Dev.
Years since purchase | sale	3,027,288	4.885	1.974
Resale price | sale	3,027,288	326,516.400	216,591.400
Purchase price | sale	3,027,288	240,191.700	2,474,234.000
Purchase interest rate | sale	3,027,288	4.036	0.284
Initial mortgage amount | sale	3,027,256	206,419.400	139,963.800
Origination LTV | sale	3,027,256	0.907	0.641

(b) Nov 2016 event study sample.

Statistic	N	Mean	St. Dev.
Years since purchase | sale	605,057	3.787	1.546
Resale price | sale	605,057	326,224.000	180,873.200
Purchase price | sale	605,057	250,013.900	786,071.700
Purchase interest rate | sale	605,057	4.021	0.265
Initial mortgage amount | sale	605,054	217,005.500	115,000.400
Origination LTV | sale	605,054	0.912	0.796

Table C.8: November 2016 event study: Covariate balance.

Dependent Variables:	Age (std)	Income (USD '000s, std)	Appraised amt (USD '000s, std)	LTV (%, std)	Risk score 3.0 (std), 6 months before move
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Post	0.0165	0.0047	-0.0052	-0.0182	-0.0379
	(0.0210)	(0.0113)	(0.0110)	(0.0177)	(0.0248)
Trend width	4mo	4mo	4mo	4mo	4mo
Pre std dev.	12.30	43.92	201.1	13.48	69.51
Fixed-effects					
Zip-ever refi	Yes	Yes	Yes	Yes	Yes
Age quintile-Inc quintile-Appraised quintile					Yes
Fit statistics					
Observations	68,761	40,691	68,918	68,723	67,543
R^2	0.26499	0.26673	0.68457	0.34480	0.40167
Within R^2	0.00185	0.00052	0.00042	0.00067	0.00208

Clustered (Person & Orig. date) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Source: Equifax Credit Risks Insight Servicing. The table presents estimates of β_1 in equation (12), where outcome variables are demographics as described in each column in standardized units.

Table C.9: November 2016 event study: Truncation test.
(Source: Equifax Credit Risks Insight Servicing)

Dependent Variables:	Risk score 3.0 (std), 6 month lag	Age (std)	LTV (%, std)	DTI (%, std)	Move-in rate (%, std)	Income (USD '000s, std)	Appraised amt (USD '000s, std)
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables							
Truncated	-0.0831***	0.1399***	0.0114	0.0378	0.0331***	0.0047	0.0115
	(0.0290)	(0.0323)	(0.0255)	(0.0410)	(0.0100)	(0.0451)	(0.0210)
Pre std dev.	69.51	12.30	13.48	11.42	1.101	43.92	201.1
Fixed-effects							
Date-ever refi-Zip code	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics							
Observations	40,011	40,934	40,907	20,629	41,070	20,629	41,049
R^2	0.49160	0.45978	0.52089	0.50924	0.70034	0.61914	0.77511
Within R^2	0.00060	0.00138	1.28×10^{-5}	0.00013	0.00082	1.73×10^{-6}	2.12×10^{-5}

Clustered (Borrower & Move in date) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

D Section 3 details.

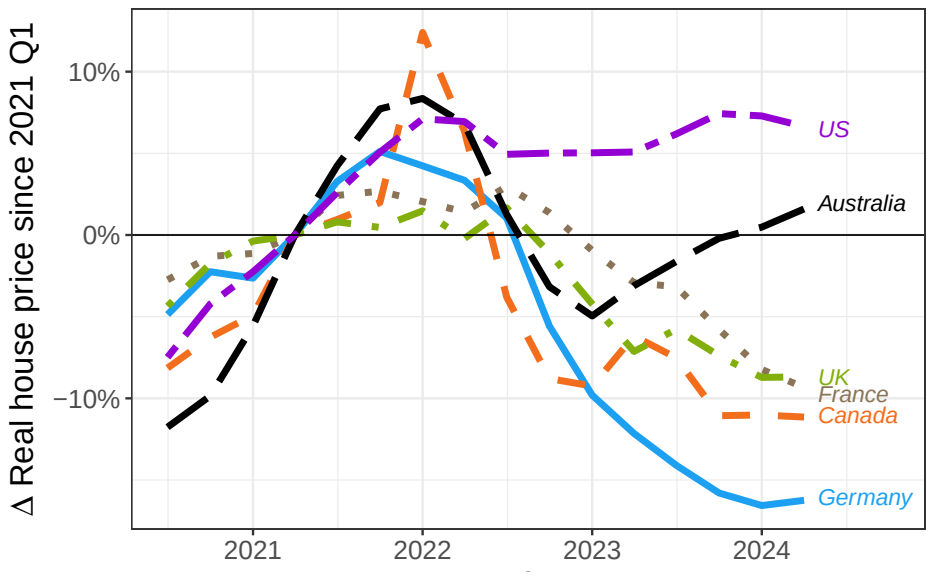
D.1 House price effects of tightening internationally.

The Bureau of International Settlements (BIS) compiles real residential house price series, deflated by national consumer price indices, for many advanced economies. We pull these series from FRED for Germany (`QDER628BIS`), Canada (`QCAR628BIS`), France (`QFRR628BIS`), the UK (`QGBR628BIS`), Australia (`QAUR628BIS`), and the US (`QUSR628BIS`). The BIS compiles series in each country using data available from relevant statistical authorities, and aims to cover both existing and new construction. The BIS uses hedonic regressions, rather than repeat-sales to estimate constant-quality indices, because outside the US existing-home sales are very infrequent. As such, the real series for the US indicates different real growth relative to values quoted in the introduction.

Figure D.10 shows the percentage change in real residential prices in each country since 2021Q1. All countries saw significant increases in nominal long-term yields over this period: nominal 10-year bond yields increased by:

- 2.68% in Germany (FRED series: `IRLTLT01DEM156N`);
- 2.42% in Canada (FRED series: `IRLTLT01CAM156N`);
- 2.96% in France (FRED series: `IRLTLT01FRM156N`);
- 3.49% in the UK (FRED series: `IRLTLT01GBM156N`);
- 3.14% in Australia (FRED series: `IRLTLT01AUM156N`);
- 2.94% in the US (FRED series: `DGS10`).

Figure D.10: Real house price growth since 2021: International comparisons.



Source: Author's calculations and BIS. The figure shows the percentage change in real residential house prices from the BIS Residential Property Price dataset, accessed via FRED. See text for series IDs.

D.2 Calculating $SPC(\cdot)$ with counterfactual mortgage terms.

This appendix gives a formula for calculating SPC_{it} under counter-factual mortgage spreads, origination LTV, and remaining term. We use these formulas to calculate $\Delta SPC_j^{\text{Avg rate x LTV}}$ and $\Delta SPC_j^{\text{Avg rate}}$ in Section 3.2. Define $LTV_{i0} := M_{i0}/P_{ia}$, where M_{i0} is origination balance and LTV_{i0} is origination LTV; and r_{i0}, the outstanding mortgage rate. Recall that remaining term is in years, and interest rates are quoted in APR. We account for fractional years remaining, discounted at the average rate for the full year.

Formulas. We aim to derive a formula for $SPC_{it} := SPC_t(LTV_{i0}, n_i, r_{i0}, \Delta_i, P_{ia})$, assuming fixed-rate amortization with no delinquencies / default and no curtailments. Rearranging equation (4):

$$SPC_{it} := \frac{M_i}{P_{ia}} - \frac{PV_{it}}{P_{ia}} \tag{D.1}$$

where:

$$\frac{PV_{it}}{P_{ia}} := E_t^Q \sum_{0\le q\le n} \frac{m^*_{t+q}(M_i, m_i, n_i, \Delta_i)/P_{ia}}{(1+r_{f\,t,q})} \tag{D.2}$$

$$= \min\left[\frac{m_i}{P_{ia}} \sum_{0\le q\le n_i} \rho_t(q), \frac{2000}{P_{ia}} + 0.01\cdot\frac{M_i}{P_{ia}} + \frac{u(r_{mt}+\Delta_i, M_i, n_i)}{P_{ia}} \sum_{0\le q\le n_i} \rho_t(q)\right] \tag{D.3}$$

It is apparent that all we need are formulas for m_i/P_{ia}, M_i/P_{ia}, and $u(\cdot)/P_{ia}$ as functions of LTV_{i0}, n_i, r_{i0}, and Δ_i. First, we write a formula for m_i/P_{ia}. From the amortization formula, $m_i = 12\cdot M_0 \cdot \frac{r_{i0}/12\cdot(1+r_{i0}/12)^{360}}{(1+r_{i0}/12)^{360}-1}$. Therefore:

$$\frac{m_i}{P_{ia}} = mP(LTV_{i0}, r_{i0}) := 12\cdot LTV_{i0}\cdot\frac{r_{i0}/12\cdot(1+r_{i0}/12)^{360}}{(1+r_{i0}/12)^{360}-1} \tag{D.4}$$

Second, we write a formula for M_i/P_{ia}. After n_i years:

$$\frac{M_i}{P_{ia}} = \frac{M_0}{P_{ia}}\cdot(1+r_{i0}/12)^{12\cdot(30-n_i)} - \frac{m_i}{12\cdot P_{ia}}\cdot\frac{(1+r_{i0}/12)^{12\cdot(30-n_i)}-1}{r_{i0}/12} \tag{D.5}$$

$$= LTV_{i0}\left((1+r_{i0}/12)^{12\cdot(30-n_i)} - \frac{r_{i0}/12\cdot(1+r_{i0}/12)^{360}}{(1+r_{i0}/12)^{360}-1}\cdot\frac{(1+r_{i0}/12)^{12\cdot(30-n_i)}-1}{r_{i0}/12}\right) \tag{D.6}$$

$$= LTV_{i0}\cdot(1+r_{i0}/12)^{360}\cdot\left((1+r_{i0}/12)^{-12n_i} - \frac{(1+r_{i0}/12)^{12\cdot(30-n_i)}-1}{(1+r_{i0}/12)^{360}-1}\right) \tag{D.7}$$

$$:= MP(LTV_{i0}, r_{i0}, n_i) \tag{D.8}$$

The first line writes out the balance assuming interest accrual and no repayments, divided by P_{ia}, minus the cumulative value of payments between origination and year $30-n_i$ (as calculated using the finite geometric sum formula. The second line substitutes in m_i, and the next two lines simplify. Finally, we write a formula for $u(\cdot)/P_{ia}$:

$$\frac{u(\cdot)}{P_{ia}} = 12\cdot\frac{M_i}{P_{ia}}\cdot\frac{(r_{mt}+\Delta_i)/12\cdot(1+(r_{mt}+\Delta_i)/12)^{12\cdot(30-n_i)}}{(1+(r_{mt}+\Delta_i)/12)^{12\cdot(30-n_i)}-1} \tag{D.9}$$

$$= 12\cdot MP(LTV_{i0}, r_{i0}, n_i)\cdot\frac{(r_{mt}+\Delta_i)/12\cdot(1+(r_{mt}+\Delta_i)/12)^{12\cdot(30-n_i)}}{(1+(r_{mt}+\Delta_i)/12)^{12\cdot(30-n_i)}-1} \tag{D.10}$$

$$:= uP_t(LTV_{i0}, r_{i0}, n_i, \Delta_i) \tag{D.11}$$

Define $R_t(n_i) := \sum_{0 \le q \le n_i} \rho_t(q)$. Then:

$$\frac{PV_{it}}{P_{ia}} = \min\Bigg[mP(LTV_{i0}, r_{i0}) R_t(n_i), \\ \frac{2000}{P_{ia}} + 0.01 \cdot MP(LTV_{i0}, r_{i0}, n_i) + uP_t(LTV_{i0}, r_{i0}, n_i, \Delta_i) R_t(n_i) \Bigg] \tag{D.12}$$

$$:= PVP_t(LTV_{i0}, r_{i0}, n_i, \Delta_i, P_{ia}) \tag{D.13}$$

and:

$$SPC_{it} = SPC_t(LTV_{i0}, r_{i0}, n_i, \Delta_i, P_{ia}) \tag{D.14}$$

$$:= MP(LTV_{i0}, r_{i0}, n_i) - PVP_t(LTV_{i0}, r_{i0}, n_i, \Delta_i, P_{ia}) \tag{D.15}$$

Applications. To calculate $\Delta SPC_j^{\text{Avg rate x LTV}}$, we set $\Delta_i = 0$ and calculate:

$$SPC_{it}^{\text{Avg rate x LTV}} := SPC_t(LTV_{i0}, r_{i0} - \Delta_i, n_i, 0, P_{ia}) \tag{D.16}$$

We then take the zip-level average for i outstanding in $t-2$ to calculate $SPC_{jt}^{\text{Avg rate x LTV}}$.

When calculating $\Delta SPC_j^{\text{Avg rate}}$, we first estimate $\overline{n}_t$, $\overline{P}_t$, and $\overline{LTV}_{0t}$ the nationwide average remaining term, appraised value, and origination LTV for mortgages outstanding in year $t-2$. Define $r_{m\tau}$ as the average rate on a new 30-year fixed rate mortgage originated in year-month τ. We then calculate, for each origination year-month τ:

$$SPC_{\tau,t} := SPC_t(\overline{LTV}_{0t}, r_{m\tau}, \overline{n}_t, 0, \overline{P}_t) \tag{D.17}$$

and construct $\Delta SPC_j^{\text{Avg rate}}$ as described in the main text.

D.3 Origination timing fixed effects controls.

This appendix shows that if household-level SPC is correlated with unobserved origination half-year effects, then the controls described in Section 3.4 recover unbiased estimates of β.

Let $h(i)$ indicate the half-year of mortgage origination for property i, where half-years are defined in 6-month buckets (Jan-Jul, Aug-Dec). Suppose at the property level:

$$\%\Delta P_i = \beta \cdot \Delta SPC_i + \eta'_{c(i)} \cdot X_i + \omega_{c(i)} + \zeta_{h(i)} + \varepsilon_i \tag{D.18}$$

where $E[\varepsilon_i \cdot \Delta SPC_i | X_i, c(i), h(i)] = 0$. Taking equal-weighted averages over i:

$$\%\Delta P_j = \beta \cdot \Delta SPC_j + \eta'_{c(j)} \cdot X_j + \omega_{c(j)} + \sum_h \zeta_h \cdot \sum_{\tau : h(\tau) = h} w_{j\tau,2023} + \varepsilon_j \tag{D.19}$$

where ε_j is the equal-weighted average of ε_{ij} and $h(\tau)$ maps from origination month τ to origination half-year h. Adding the additional controls $\left\{\zeta_h \cdot \sum_{\tau : h(\tau)=h} w_{j\tau,2023}\right\}_h$ to the structural equation in (6) thus recovers a consistent estimate of β, since in that case ϵ_j is conditionally uncorrelated with ΔSPC_j.

D.4 Proof of Proposition 1

Let $\chi_j \equiv (c(j), X_j)$ collect observable local housing market characteristics. In this section, unlike in the main text, $\Delta \ln S_{jt}$ refers to the one-period (not two-period) difference in $\ln S_{jt}$.

Formally, Proposition 1 claims that:

$$E\left[\epsilon_j \cdot \Delta SPC_j^S | \chi_j\right] = 0 \tag{D.20}$$

if ϵ_j is mean-independent of $\Delta \ln S_{j\tau'}$:

$$E\left[\epsilon_j | \Delta \ln S_{j\tau'}, \chi_j\right] = E\left[\epsilon_j | \chi_j\right] \quad \forall j\tau' \tag{D.21}$$

Proof. First, that ϵ_j is mean-independent of $\Delta \ln S_{j\tau'}$ implies:

$$E\left[\epsilon_j \cdot G\left(\Delta \ln S_{j\tau'}\right) | \chi_j\right] = 0 \quad \forall G(\cdot),\ j\tau' \tag{D.22}$$

Second, note that where τ'_0 is the first-observed τ':

$$S_{j\tau'} = S_{j\tau'_0} \exp\left(\sum_{q \in (\tau'_0, \tau']} \Delta \ln\left(S_{jq}\right)\right) \tag{D.23}$$

By linearity of $\nu^S_{j\tau}$, we can therefore write, for some function ν'^S_τ:

$$\nu^S_{j\tau} := \nu^S_\tau(\{S_{j\tau'}\}_{\tau'<\tau}) = S_{j\tau'_0}\, \nu'^S_\tau\left(\{\Delta \ln S_{j\tau'}\}_{\tau' \in (\tau'_0, \tau)}\right) \tag{D.24}$$

The predicted origination shares are thus only a function of $\Delta \ln S_{j\tau'}$:

$$w^S_{j\tau,t} = \frac{K(t-\tau)\cdot \nu^S_{j\tau}}{\sum_{\tau' \le t} K(t-\tau)\cdot \nu^S_{j\tau}} = \frac{S_{j\tau'_0} K(t-\tau)\cdot \nu'^S_\tau\left(\{\Delta \ln S_{j\tau'}\}_{\tau' \in (\tau'_0, \tau)}\right)}{S_{j\tau'_0} \sum_{\tau \le t} K(t-\tau)\cdot \nu'^S_\tau\left(\{\Delta \ln S_{j\tau'}\}_{\tau' \in (\tau'_0, \tau)}\right)} \tag{D.25}$$

$$:= w^S_{\tau,t}\left(\{\Delta \ln S_{j\tau'}\}_{\tau' \in (\tau'_0, \tau)}\right) \tag{D.26}$$

Equation (D.26) implies that:

$$SPC^S_{jt} := \sum_{\tau \le t} w^S_{\tau,t}\left(\{\Delta \ln S_{j\tau'}\}_{\tau' \in (\tau'_0, \tau)}\right) \cdot SPC_{\tau,t} \tag{D.27}$$

$$= SPC^S_t\left(\{\Delta \ln S_{j\tau'}\}_{\tau' \in (\tau'_0, \tau)}, \{SPC_{\tau,t}\}_\tau, w^S_{\tau,t}(\cdot)\right) \tag{D.28}$$

The claim that (D.20) holds then follows from defining $G(\cdot) := SPC^S_{2023}(\cdot) - SPC^S_{2021}(\cdot)$ and applying equation (D.22). □

Example 1 (Twin birth rate iid draw from city-by-year distribution.)**.** *Suppose that the twin birth rate $m_{jt} \sim^{iid} \tilde{G}_{c(j),t}$, where $\tilde{G}_{ct}$ is a CBSA-by-year specific distribution. Then $\Delta \ln m_{jt} \sim^{iid} G_{c(j),t}$, a distribution based on $\tilde{G}_{c(j),t}$ and $\tilde{G}_{c(j),t-1}$.*

Conditioning on χ_{jt} includes conditioning on CBSA $c(j)$, which implicitly conditions on the history of $G_{c(j),t}$. We can therefore write $\chi_{jt} \equiv \left(X_{jt}, \left\{G_{c(j),q}\right\}_{q \ge \tau_0}, u_{c(j)}\right)$, where $u_{c(j)}$ include unobserved characteristics of CBSA $c(j)$ besides the twin birth rate distribution history.

For each $j\tau$:

$$E\left[\varepsilon_{jt} | \Delta \ln m_{j\tau}, X_{jt}, \left\{G_{c(j),q}\right\}_{q \ge \tau_0}, u_c(j)\right] = E\left[\varepsilon_{jt} | X_{jt}, \left\{G_{c(j),q}\right\}_{q \ge \tau_0}, u_{c(j)}\right] \tag{D.29}$$

because the realized draw contains no information above the distribution itself. This is equivalent to ε_{jt} being mean-independent of $\Delta \ln m_{j\tau}$ for all $j\tau$ (equation (D.21) being satisfied), which means Proposition 1 holds.

D.5 Twin birth rate instrument construction.

D.5.1 Variation in IVF useage explains little variation in twin birth rate.

This section provides additional analysis demonstrating that variation in IVF usage is not a major driver of variation in the twin birth rate. We focus on data from after 2016, the first year for which the CDC publishes detailed data on both use of assistive reproductive technology (ART, of which IVF is a subset) and multiple births.

First, IVF births are not a significant contributor to the twin birth rate. In 2016, 76,892 infants were born using ART, accounting for 1.8% of all births and 16.4% of all multiple birth infants (Sunderam 2019).

Second, IVF usage rates does not explain much *variation* in the twin birth rate across geography. Table D.10 regresses the twin birth rate on the fraction of births using ART from 2016-2023, the only years for which data are available. Column (1) shows results in levels, and Column (2) shows results in first differences. Higher ART usage does translate into higher twin birth rates. However, most variation in the twin birth rate is not explained by ART. In levels, variation in ART only explains about 2% of variation in the twin birth rate. In differences – using variation relevant for the identifying assumption– changes in ART usage explain only 0.13% of the change in the twin birth rate.

Table D.10: Assistive Reproductive Technology usage explains little variation in twin birth rate.

Dependent Variables:	Pr(Twin)	Δ Pr(Twin)
Model:	(1)	(2)
Variables		
Constant	0.0244***	-0.0004***
	(0.0003)	(7.04×10^{-5})
Pr(ART)	0.0980***	
	(0.0211)	
Δ Pr(ART)		0.0810*
		(0.0426)
Fit statistics		
Observations	4,560	3,990
R^2	0.01989	0.00155
Adjusted R^2	0.01967	0.00129

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

Source: Author's calculations using CDC Natality data from 2016-2023. The table shows estimated coefficients from a regression of the county twin birth rate (twins / total births) on the county ART utilization rate (ART births / total births). The first column estimates in levels, and the second column estimates in first differences. Robust standard errors are clustered by county and calendar year.

D.5.2 Estimating β_a.

As mentioned in the main text, we slightly increase power by parameterizing β_a in equation (7) as a fifth-order polynomial with fixed effects at ages 1, 5, and 9. Specifically, we let:

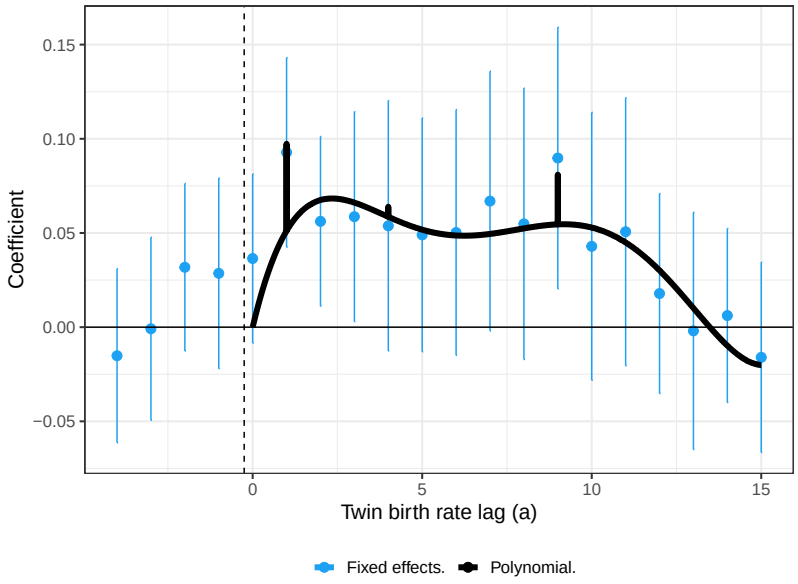
$$\beta_a = \sum_{p=1}^{5} \pi_a \cdot a^p + \sum_{a' \in \{1,5,9\}} \tilde{\beta}_a \mathbb{1}(a = a') \tag{D.30}$$

and estimate π_a and $\tilde{\beta}_a$ via OLS. Figure D.11 shows estimates.

D.5.3 Estimating repayment rates.

We estimate the repayment kernel $K(a)$ using borrower-level records from ICE, McDash®.

Figure D.11: Polynomial estimates: Effect of twin birth rate on housing market churn.



Source: Equifax Credit Risks Insight Servicing. The figure plots fitted values of β_a from equation (D.30). See text for details.

We model repayments, as a fraction of remaining balances, as additively separable in calendar time and tenure effects:

$$\frac{\Delta M_{it}}{M_{i,t-1}} = \gamma_t + \sum_a \lambda_a I(t - T_i = a) + \epsilon_{it} \tag{D.31}$$

where T_i is the origination year of mortgage i. A main source of average ΔM_{it} is borrower full repayment due to moves or refinancing. If the outcome were specified in either levels or as a fraction of initial balances, full repayments would have different effects on estimated coefficients depending on when repayment occurs. This implicitly introduces a remaining balance-by-tenure effect, which the specification in equation (D.31) avoids.

We calculate the fraction of balances remaining a years after origination recursively as:

$$K(0) = 1 \tag{D.32}$$

$$K(a) = K(a-1) \cdot (1 - \hat{\lambda}_a) \tag{D.33}$$

$K(a)$ gives the fraction of balances remaining a years after origination. Assuming a constant inflow of mortgages, the outstanding distribution of origination dates is a weighted average of mortgages originated at each horizon a, with weights proportional to $K(a)$.

D.6 Standard deviation calculations in Figure 3.

This section describes theoretical benchmarks for the standard deviations shown in Figure 3.

Notation. Let j, c, and t index county, CBSA, and year respectively. Let n_{jt}, p_{jt} denote the number of total births and the twin birth rate in county j. The number of twins born is $p_{jt} \cdot n_{jt}$. Let $\mathcal{B}_k$ denote the set of counties in size bin k, with $N_k := |\mathcal{B}|$.

Twin birth rate distributions satisfying Proposition 1. Assume that $p_{jt} \sim^{iid} \tilde{G}_{c(j),t}$, where $\tilde{G}_{ct}$ is a CBSA-by-year specific distribution. Then $\Delta \ln p_{jt} \sim^{iid} G_{c(j),t}$, where G_{ct} is some distribution based on $\tilde{G}_{ct}$ and $\tilde{G}_{c,t-1}$.

Benchmark. Assume twin births are iid binomial draws from a CBSA-by-year specific binomial distribution: $p_{jt} \cdot n_{jt} \sim^{\text{iid}} \text{Binom}(\overline{p}_{c(j),t}, n_{jt})$, where $\overline{p}_{c(j),t}$ is the twin birth rate in city c in year t. By the law of total variance:

$$\text{Var}\left(p_{jt}\right) = E\left[\text{Var}(p_{jt})|c(j)\right] + \text{Var}(E\left[p_{jt}|c(j)\right]) \tag{D.34}$$

$$= E\left[\frac{\overline{p}_{c(j),t} \cdot (1-\overline{p}_{c(j),t})}{n_{jt}}\right] + \text{Var}\left(\overline{p}_{c(j),t}\right) \tag{D.35}$$

We sum the sample counterparts of equation (D.35) within each size bin. We estimate $\widehat{\overline{p}}_{ct} \equiv N_{ct}^{-1} \sum_{j \in c} p_{jt}$, where N_{ct} is the number of county-level observations in year t in CBSA c, and then calculate:

$$\widehat{\text{Var}}(p_{jt}|jt \in \mathcal{B}_k) = N_{ct}^{-1} \sum_{jt \in \mathcal{B}_k} \frac{\hat{\overline{p}}_{ct}(1-\hat{\overline{p}}_{ct})}{n_{jt}} + \widehat{\text{Var}}(\hat{\overline{p}}_{ct}|\mathcal{B}_k) \tag{D.36}$$

where $\widehat{\text{Var}}(\hat{\overline{p}}_{ct}|\mathcal{B}_k)$ is the sample variance of $\hat{\overline{p}}_{ct}$ weighted by observations $jt \in \mathcal{B}_k$.

D.7 Calibrating twin instrument relevance and power.

D.7.1 Twin instrument relevance.

We first analyze coefficient magnitudes for the estimated effect of the twin birth rate on housing market churn in Figure 3a. We focus on the impact of the twin birth rate in year t on moves in $t+1$, which at the family level are moves the year after twins are born rather than one child. The coefficient on the twin birth rate (where the rate runs from 0 to 1) is:

$$\Delta \Pr(\text{Move}) = \Big(\Pr(\text{Has kid}) \times \Delta \Pr(\text{Move} \mid \text{twin parent}) \tag{D.37}$$

$$+ \Pr(\text{Twin-related}) \times \Delta \Pr(\text{Move} \mid \text{twin-related}) \Big) \tag{D.38}$$

$$\times E\left[\text{Vacancy chain length}\right] \tag{D.39}$$

This expression has three important terms. The first term, in line (D.37), shows the *direct* effect of twin births on moves among families that have multiple births rather than singletons. Pr(Has kid), is the birth rate among owner-occupied households with a mortgage. A rough estimate is the equal-weighted average ratio across zip codes of total births to owner-occupied households with a mortgage, equal to 0.090 in the 2015 ACS.

Δ Pr(Move|twins) is the effect of having twins on moves. Twins are unexpected for most families, and increase the number of expected children by a huge amount – by 100% (first births) 50% (second births), or 33% (third births). We calibrate this term as 0.5 based on two data points. First, Figure C.4c shows the cross-sectional relationship in ACS between the number of kids in a house and the number of bedrooms. Each additional child increases the number of bedrooms by about 0.25 on average. Since getting an extra bedroom (typically) requires moving to a new house, this implies than an extra child generates a move for 25% of households. Thus, 50% of households move with two extra kids. Second, in an informal Reddit survey, about half of respondents on the forum `r/parentsofmultiples` said they moved when they found out they were having twins.

The second term in the formula, in line (D.38), reflects *indirect* moves due *directly* to twin births. This reflects parents of multiples asking their parents or siblings to relocate to help with caregiving, or local congestion in childcare and early schooling that causes other families to relocate. Without good data, we set this to zero so our calibration is a lower bound.

The third term in the formula, in line (D.39), reflects *spillover* effects on the local housing market

due to twin-driven moves. When one family moves due to twins, they both leave their previous unit vacant and, if they purchase an existing home, free up their new unit's previous owner to move. Their previous vacancy is filled by another family, who creates their own vacancy. Twin-driven moves thus create broader housing market churn that impacts the timing of moves for a broader set of families unrelated to the one having twins.

It is hard to get precise estimates for how much churn such moves will create. A recent paper by French and Gilbert (2023) studies effective vacancies generated by low-density suburban housing, and tabulates when these vacancies occur by length of "vacancy chain." Estimates from this paper are likely a lower bound in our context, because moves to low-density suburban housing are especially likely to be due to family formation that do not leave vacancies in existing owner-occupied units. Results provided by the authors (not in the working paper) imply an initial move generates 0.92 additional vacancies, implying a lower bound of 1.92 moves. We round this to 2. This is slightly lower than the 2.5 "transactions multiplier effect" estimated due to increased housing demand in "hot" markets in Anenberg and Ringo (2022), which the authors describe as a lower bound on effects on moves.[61]

Our lower bound on the first period effect is thus:

$$\Delta \Pr(\text{Move}|\text{twins}) = 0.09 \times 0.5 \times 2 = 0.09 \in [0.01, 0.16] \tag{D.40}$$

well within the 95% confidence interval for the coefficient estimate of β_1.

D.7.2 Finite sample variation in twin birth rate generates meaningful variation in moves.

We argue empirical housing market churn generated by twin births is due to finite sample variation. This appendix uses simulations to demonstrate this directly and quantify the number of households affected.

We consider the median county in 2015, with 2,811 births / year and 31,233 owner-occupied households (a ratio of births to owner-occupied households with a mortgage of 9.0%). We simulate 10,000 paths for the twin birth rate.

First, there is significant variation in the number of twins born given this county size. Figure D.12a plots the distribution of the annual twin birth rate (left panel, SD = 0.35%) and the distribution of the number of annual twin births (right panel, SD = 10.0).

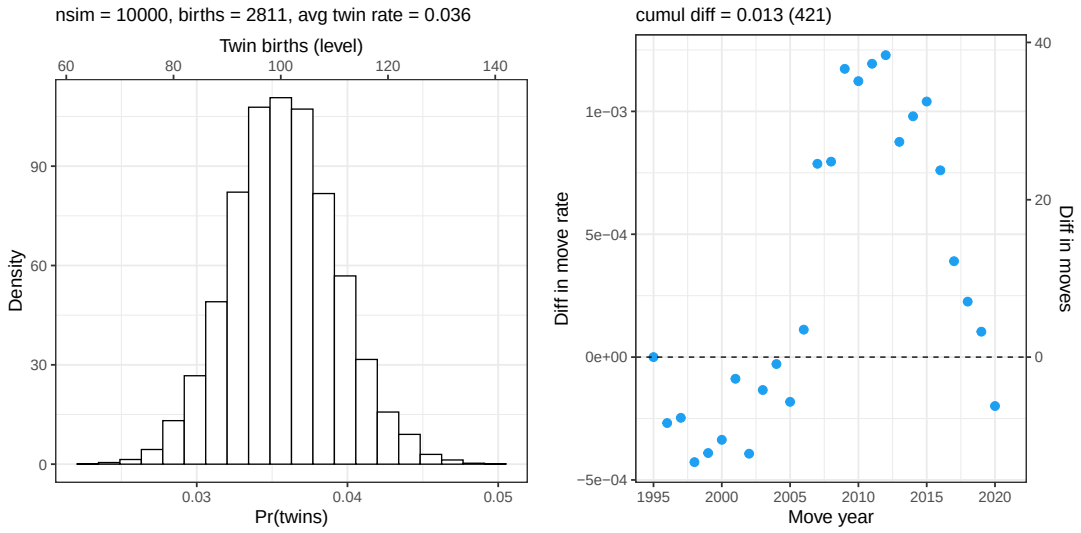
Second, the cumulative effect of variation in the twin birth rate from 1995-2006 creates significant cross-sectional differences in the timing of moves. We simulate twin births from 1995-2006, and run the simulated path through our instrument construction methodology.We then compare annual twin-driven moves that survive until 2021 (formally, $K(2021-\tau)\cdot \nu_{j\tau}^{\text{twin}}$) for a simulation with an instrument value 2 SD above the mean to those of a simulation with an instrument value 2 SD below the mean, similar to D.19. Figure D.12b plots the resulting differences, both as a percentage of the 2021 stock (left panel) and in levels (right panel). Draws with a high value of the instrument have more moves in earlier years and fewer in later years. Even though the difference in moves in any one year is small, the small differences accumulate to sizable ones taken over 1995-2020. In these simulations, twin birth rate variation reshuffles move timing for 611 owner-occupied households in 2021, or 1.3% of the local population. A big change in the existing home sales incentives of this group – due to, for instance, rate lock – will easily have a detectable impact on existing home prices.

[61] The unit of analysis in this paper is housing market transactions, not household moves. In Appendix A.1, the authors find that their estimate for the probability that the previous owner of a sold unit repurchases another one is lower than external estimates of the probability that a household who sells property purchases another property, which itself is different (and likely lower) than the probability that a household who sells their primary residence purchases another primary residence.

Figure D.12: Simulation results

(a) Simulated distribution of twin birth rates.

(b) Simulated twin-predicted origination probability, ± 2 SD of avg instrument.



Source: Author's calculations. Values are calculated across 10,000 simulations for a county with 2,811 annual births and an annual twin birth rate of 3.6%. Panel (a). The figure shows the distribution of annual twin birth rates (bottom) and the annual number of twin births (top) across simulations. Panel (b). The figure shows the difference in annual moves for mortgages that survive to 2021 in shares of owner-occupied households (left) or number of owner-occupied households (right) for each year between simulation draws where simulated instruments are 2 standard deviations above vs. 2 standard deviations below the mean.

D.7.3 Over-identification test.

If the procedure for constructing $\Delta SPC_j^{\text{twin}}$ is valid, then the instrument $\Delta SPC_j^{\text{twin}}$ should load on past twin birth rates with the same coefficients as ΔSPC_j (the endogenous regressor). We investigate with an over-identification test. The first stage for the twins instrument is:

$$\Delta SPC_j = \gamma \cdot \Delta SPC_j^{\text{twin}} + \eta'_{c(j)} \cdot X_j + \omega_{c(j)} + \epsilon_j \tag{D.41}$$

where $\Delta SPC_j^{\text{twin}} := \Delta SPC^{\text{twin}}(\boldsymbol{p}_j; \hat{\theta})$ for twin birth rate history $\boldsymbol{p}_j$, parameters estimated offline $\hat{\theta}$ (e.g., $\hat{\beta}$ in equation (7)), and ΔSPC^{twin} a non-linear function described in the main text.

Taking a first-order approximation:

$$\gamma \cdot \Delta SPC_j^{\text{twin}} \approx \boldsymbol{\delta}' \cdot \boldsymbol{p}_j + \eta'_{c(j)} \cdot X_j + \omega_{c(j)} + \epsilon_j \tag{D.42}$$

Suppose that instead of constructing $\Delta SPC_j^{\text{twin}}$, we ran the fist stage with $\boldsymbol{p}_j$ directly:

$$\Delta SPC_j = \boldsymbol{\pi}' \cdot \boldsymbol{p}_j + \eta'_{c(j)} \cdot X_j + \omega_{c(j)} + \epsilon_j \tag{D.43}$$

If our instrument construction approach is valid, then $\hat{\boldsymbol{\delta}}$ and $\hat{\boldsymbol{\pi}}$ should align.

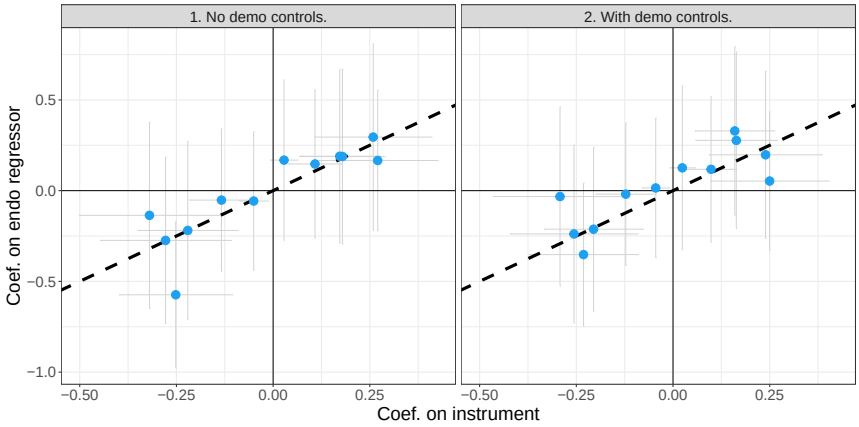
We evaluate whether $\hat{\boldsymbol{\delta}}$ and $\hat{\boldsymbol{\pi}}$ should align with a bootstrapped Sargan-Hansen J test. Under correct specification, the null hypothesis is: $H_0: \ (\boldsymbol{\delta} - \boldsymbol{\pi})' \cdot W \cdot (\boldsymbol{\delta} - \boldsymbol{\pi})$, where $W := \Sigma^{-1}$ for Σ the covariance of $(\boldsymbol{\delta} - \boldsymbol{\pi})$. The test statistic is $J := (\hat{\boldsymbol{\delta}} - \hat{\boldsymbol{\pi}})' \cdot \hat{\Sigma}^{-1} \cdot (\hat{\boldsymbol{\delta}} - \hat{\boldsymbol{\pi}})$. We simulate the distribution of J under the null via 1,000 bootstrap draws. We also use bootstrap draws to calculate 95% confidence intervals for elements of $\hat{\boldsymbol{\delta}}$ and $\hat{\boldsymbol{\pi}}$.

Figure D.13, panel (a) plots coefficient estimates for $\boldsymbol{\delta}$ against coefficient estimates for $\boldsymbol{\pi}$, where coefficients are estimated in specifications with and without demographic controls. Coefficient estimates line up closely. Figure D.13, panel (b) plots the distribution of J under the null. The dashed

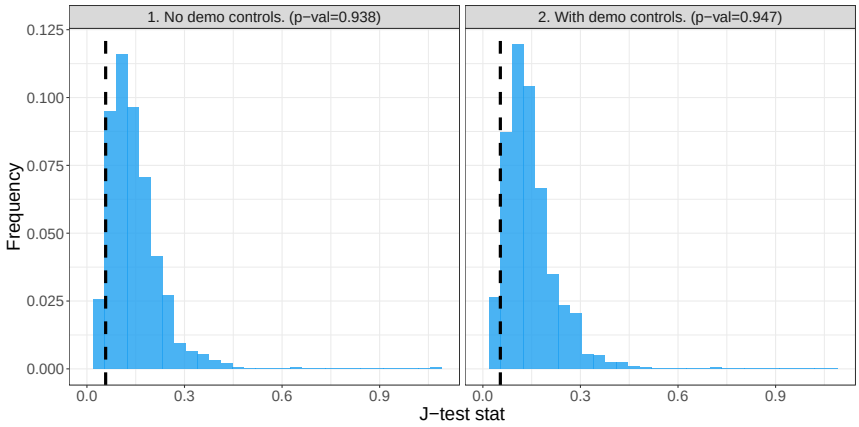
vertical line shows the actual value of the test statistic. We fail to reject the null that coefficient estimates are the same (p-value=0.95 without controls, p-value=0.96 with controls).

Figure D.13: Twins instrument over-identification test.

(a) Coefficient on twin birth rates, 1995-2006



(b) Bootstrapped J-test stats.



Source: Author's calculations using CDC natality data, Equifax Credit Risks Insight Servicing data, and ICE, McDash® data. Panel (a) plots coefficient estimates of $\boldsymbol{\delta}$ from equation (D.42) (x-axis) against $\boldsymbol{\pi}$ from equation (D.43) (y-axis). Spikes plot 95% confidence intervals calculated from 1,000 bootstrapped draws. Panel (b) plots the distribution of J-test statistics under the null H_0 calculated from 1,000 bootstrapped draws. See text for details.

D.8 First births instrument construction details.

D.8.1 Identifying assumption.

This subsection describes how the equation for predicted first-birth driven moves in (**??**) relates to the framework for Proposition 1.

Let $S_{jt} = \left(B_{jt} + \frac{\overline{Moves}_j}{\sum_a \pi_a}\right)$. Then:

$$\nu^{\text{births}}_{j\tau} = \sum_a \pi_a \left(B_{j,t-a} + \frac{\overline{Moves}_j}{\sum_a \pi_a}\right) \tag{D.44}$$

$$= \sum_a \pi_a B_{j,t-a} + \overline{Moves}_j \tag{D.45}$$

S_{jt} defined as such thus can be factored out of $\nu^{\text{births}}_{j\tau}$ as required for equation (D.24) in the proof of Proposition 1.

The identifying assumption behind the first births instrument is therefore that ΔS_{jt} is mean-independent of ε_{jt}. This essentially requires that the change in first births, as a fraction of average moves, is uncorrelated with ε_{jt}. To see this, note that:

$$\Delta S_{jt} = \ln\left(B_{jt} + \frac{\overline{Moves}_j}{\sum_a \pi_a}\right) - \ln\left(B_{j,t-1} + \frac{\overline{Moves}_j}{\sum_a \pi_a}\right) \tag{D.46}$$

$$= \ln\left(\frac{\sum_a \pi_a}{\overline{Moves}_j} B_{jt} + 1\right) - \ln\left(\frac{\sum_a \pi_a}{\overline{Moves}_j} B_{jt} + 1\right) \tag{D.47}$$

$$\approx \sum_a \pi_a \frac{\Delta B_{jt}}{\overline{Moves}_j} \tag{D.48}$$

where the second line follows by factoring out $\frac{\overline{Moves}_j}{\sum_a \pi_a}$ and the third line follows by applying the approximation $\ln(1+x) \approx x$. This approximation should be good because $\frac{\Delta B_{jt}}{\overline{Moves}_j} << 1$ (it is less than the number of births to number of moves ratio, which is the same as the birth rate to move rate ratio, which is about 1.5% / 8%) and $\sum_a \pi_a < 1$.

D.8.2 Estimating π_a.

As mentioned in the main text, we slightly increase power by parameterizing π_a in equation (8) as a tenth-order polynomial with fixed effects at ages -3, 1, 2, 4, and 7. Specifically, we let:

$$\tilde{\pi}_a = \sum_{p=1}^{10} b_a \cdot a^p + \sum_{a' \in \{-3,1,2,4,7\}} c_a \mathbb{1}(a = a') \tag{D.49}$$

and estimate b_a and c_a via OLS. Figure D.14 shows estimates.

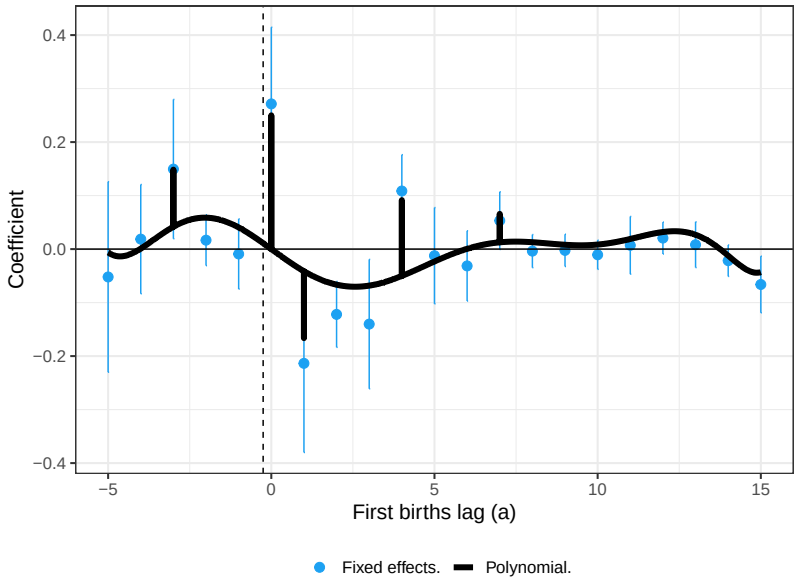
D.9 Twins robustness.

D.9.1 Addressing IVF variation.

Appendix Section D.5.1 shows that cross-sectional variation in IVF usage is not an important driver of variation in in the twin birth rate due to its rarity. This subsection provides evidence that variation in IVF does not drive our twin IV estimates.

We cannot remove twin births due to IVF directly, because the CDC only started reporting IVF-by-plurality birth cross tabs in public datasets in 2016. In 2016, restricting to births to mothers under the

Figure D.14: The effect of the twin birth rate on local housing market churn.



Source: Equifax Credit Risks Insight Servicing. The dark lines in the figure plot fitted values of $\tilde{\pi}_a$ from equation (D.49). The blue dots plot fixed effects estimates of π_a from equation (8), along with 95% confidence intervals calculated with county and year clustered standard errors. All specifications include zip fixed effects, CBSA-by-time fixed effects, time-varying population controls, and time-varying controls for the local age distribution. See text for details.

age of 35 removes about 75% of IVF births. We therefore re-construct our instrument using the twin birth rate for mothers under 35.

Columns (1) and (2) of Table D.12 show results, with and without CBSA-by-demographic controls. Point estimates are nearly identical to those in the main results, although estimates are slightly more precise. This suggests that variation in IVF usage does not drive our estimates.

D.9.2 Control for county-level expected twin birth rate changes.

A sufficient condition for the estimates in the main text to consistently identify β in equation (6) is if changes in the twin birth rate are iid draws from a CBSA-by-time specific distribution.

In this section, we relax this assumption, and allow that the twin birth rate is an iid draw from a local market-specific, time-varying, autocorrelated binomial distribution. As in Borusyak and Hull (2023), we show that including a linear control for the expected value of the instrument given this process consistently identifies β by ensuring that all identifying variation comes exclusively from finite sample variation.

Estimating local twin birth rate process. We model the twin birth rate p_{jt} in location j in period t as a draw from a binomial distribution with success rate $\alpha p_{j,t-1} + \gamma_t + \rho_j$ and number of births n_{jt}, where γ_t is a calendar time trend, ρ_j is a location-specific level, and α is an autoregressive coefficient:

$$p_{jt} \cdot n_{jt} \sim \text{Binom}\left(n_{jt}, \alpha p_{j,t-1} + \gamma_t + \rho_j\right) \tag{D.50}$$

This process implies that:

$$p_{jt} = \alpha p_{j,t-1} + \gamma_t + \rho_j + \epsilon_{jt} \tag{D.51}$$

In this specification, $p_{j,t-1}$ is correlated with $\epsilon_{j,t+k}$ for $k \geq 1$, because $p_{j,t-1}$ impacts p_{jt}, which in turn impacts the distribution from which $\epsilon_{j,t+1}$ is drawn, and so on. To avoid lookahead bias,[62] we follow

[62]The fixed-effect estimate α is lower than the estimate below, in line with negative autocorrelation and lookahead bias.

a two-step procedure. First, we estimate α in a first differences regression:

$$\Delta p_{jt} = \alpha \Delta p_{j,t-1} + \Delta \gamma_t + \Delta \epsilon_{jt} \tag{D.52}$$

Then, we estimate γ_t in a specification that residualizes out $p_{j,t-1}$:

$$p_{jt} - \hat{\alpha} p_{j,t-1} = \gamma_t + \rho_j + \epsilon_{jt} \tag{D.53}$$

Finally, we estimate $\hat{\rho}_j = E[p_{jt} - \hat{\alpha} p_{j,t-1}] - E[\hat{\gamma}_t]$. Table D.11 presents parameter estimates for α and γ_t. The autoregressive coefficient $\hat{\alpha}$ is negative, indicating that the twin birth rate mean reverts. This makes sense. Suppose the local population has a distribution of latent genetic predisposition towards having twins (e.g., due to excess FSH secretion). The more people with twin genetics give birth in year t, the fewer are available to give birth in $t+1$, meaning the twin birth rate should fall.

Expected instrument. Using simulated draws from the estimated distribution, we form many counterfactual instruments $\Delta SPC_j^{\text{twin,k}}$. The estimated distribution fits the observed distribution of predicted twin-driven mortgage rates well. Figure D.15 compares the pdf of CBSA-residualized variation for observed $\Delta SPC_j^{\text{twin}}$ and the simulated distribution, given 1,000 draws. The two distributions broadly align.

Estimation using only finite sample variation. Let $z_j = \Delta SPC_j^{twins}$ denote the twin instrument, and let $z_j^{\perp}$ denote its value residualized on fixed effects and controls in equation (6). Write $z_j^{\perp} = f_j(\omega; \alpha, \{\gamma_t\}, \rho_j, p_{0j})$, where ω is the vector of randomly drawn shocks due to realized finite sample variation. The identifying assumption for a consistent estimate of β is:

$$(\omega \perp \epsilon_j) | \rho_j, \alpha, p_{0j}, \{\gamma_t\} \tag{D.54}$$

This allows ϵ_j to correlate with ρ_j, p_{0j}, or $\{\gamma_t\}$.

Define $\mu_{jt} \equiv E_\omega \left[f_j(\omega; \alpha, \{\gamma_t\}, \rho_j, p_{0j}) \right]$. The expected cross-sectional covariance between the instrument and ϵ_j is:

$$E\left[N^{-1} \sum_j z_j^{\perp} \epsilon_j \right] = N^{-1} \sum_j E\left[z_j^{\perp} \epsilon_j \right] \tag{D.55}$$

$$= N^{-1} \sum_j E\left[E\left[z_j^{\perp} \epsilon_j | \alpha, \{\gamma_t\}, \rho_j, p_{0j} \right] \right] \tag{D.56}$$

$$= N^{-1} \sum_j E\left[\mu_j E\left[\epsilon_j | \alpha, \{\gamma_t\}, \rho_j, p_{0j} \right] \right] = E\left[N^{-1} \sum_j \mu_j \epsilon_j \right] \tag{D.57}$$

where the first line follows from linearity of expectations, the second from the law of iterated expectations, the third from the identifying assumption in (D.54), and the final line from another application of the law of iterated expectations.

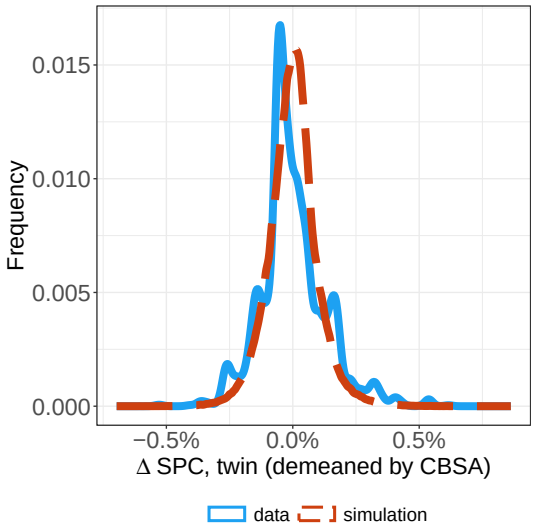
Thus, after projecting $z_{jt}^{\perp}$ on μ_{jt}, the residualized instrument will have no expected covariance with ε_{jt}, making the TSLS estimate of β consistent. This is achieved by including μ_{jt} as a TSLS control.

Columns (3) and (4) of Table D.12 thus control for $K^{-1} \sum_k \Delta SPC_j^{\text{twins, k}}$ across $K = 1,000$ simulations, with and without CBSA-by-demographic controls.

D.9.3 Use first-difference estimates of β.

In the main text, we estimate β_a in equation (7) via a smoothed polynomial to increase power, described in section D.5.2. Columns (5) and (6) of Table D.12 show robustness to using the first differ-

Figure D.15: Simulated $\Delta SPC_j^{\text{twin}}$ distribution, county-level, autocorrelated twin birth rate process.



Source: Equifax Credit Risks Insight Servicing, ICE, McDash®, and CDC Natality. The figure compares the PDF of the CBSA-residualized estimate of $\Delta SPC_j^{\text{twin}}$ for the data (blue) and values of the instrument based on simulated draws (red) from the estimated process in equation (D.50).

Table D.11: Twin autoregression estimates

Dependent Variables:	Δ Pr(Twins) (1)	Pr(twins) - α× (Lag Pr(Twins)) (2)
Lag Δ Pr(Twins)	-0.4971***	
	(0.0140)	
Year = 1997		0.0016***
		(0.0003)
Year = 1998	0.0008**	0.0039***
	(0.0004)	(0.0004)
Year = 1999	0.0003	0.0058***
	(0.0004)	(0.0004)
Year = 2000	-0.0010***	0.0064***
	(0.0004)	(0.0003)
Year = 2001	-0.0004	0.0076***
	(0.0004)	(0.0004)
Year = 2002	-0.0004	0.0088***
	(0.0004)	(0.0004)
Year = 2003	-0.0006	0.0097***
	(0.0004)	(0.0004)
Year = 2004	-0.0005	0.0108***
	(0.0004)	(0.0004)
Year = 2005	-0.0016***	0.0107***
	(0.0004)	(0.0004)
Year = 2006	-0.0015***	0.0108***
	(0.0004)	(0.0004)
Constant	0.0016***	
	(0.0003)	
Fixed-effects		
County		Yes
Sample	1997-2006	1996-2006
Observations	4,611	5,130
R^2	0.25288	0.76880
Within R^2		0.33753

County-clustered standard-errors in parentheses
****: 0.01, **: 0.05, *: 0.1*

Source: Equifax Credit Risks Insight Servicing and CDC natality data. Column (1) shows coefficient estimates for equation (D.52), and Column (2) shows estimates for equation (D.53).

ence estimates displayed in Figure 3a.

D.10 November 2016 event study.

This appendix describes how we use the November 2016 event study in Section 4 to construct a market-level instrument. Intuitively, we use the difference in number of moves just before and just after November 2016 to predict zip-level origination just before and just after the interest rate increase.

Table D.12: Effect of rate lock on local equilibrium prices: Twins robustness

	%Δ Price, 21-23					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	2.239***	2.316**	1.469**	1.865**	1.486**	1.805**
	(0.7892)	(0.9326)	(0.6083)	(0.7237)	(0.7055)	(0.7450)
$E[\Delta SPC]$, 2021-23			-12.68***	-10.12**		
			(3.802)	(4.321)		
Fixed-Effects						
CBSA	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls?	Yes	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls?	No	Yes	No	Yes	No	Yes
Instrument?	Twin < 35	Twin < 35	Twin	Twin	Twin, FD	Twin, FD
Observations	5,975	5,975	5,975	5,975	5,975	5,975
F-test (1st stage), Δ SPC, 2021-23	65.409	59.515	72.827	61.079	61.387	67.579
Wald (1st stage), Δ SPC, 2021-23	19.095	16.457	26.928	22.197	18.069	17.900

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

Source: Author's calculations using Equifax Credit Risks Insight Servicing, ICE, McDash®, and CDC natality data. Each column reports an estimate of β in equation (6) using a variant of the main twins methodology. In columns (1) and (2), we construct $\Delta SPC_j^{\text{twin}}$ using variation in the twin birth rate among mothers under age 35. In columns (3) and (4), we control for the expected value of the instrument under an autocorrelated process for the change in the twin birth rate, as described in Section D.9.2. In columns (5) and (6), we use first-difference estimates of β_a in equation (7), rather than polynomial estimates, to construct $\Delta SPC_j^{\text{twin}}$.

We then use average origination timing shares to predict origination in all other periods.

First, for the 50% sample used to construct our cross-sectional market-level dataset, we follow steps in Section B.1.1 to identify exact move-in dates. Second, we calculate the number of move-ins for each zip code during the 120 days before November 7, 2016 and after January 6, 2017. Define the B_{pre} and B_{post} as the set of months before and after the interest rate increase, with $B := B_{\text{pre}} \cup B_{\text{post}}$. Third, we calculate the difference in the number of moves just before and just after the mortgage rate increase as a fraction of the number of owner-occupied housing units in zip code j in 2017 measured in ACS. Define Δf_j as the difference, and n_j as the sum of originations during the full window. A higher Δf_j corresponds to a greater share of households who moved in prior to the increase.

We then apply the difference in flow to impute the difference in outstanding origination shares as of $t \in \{2019, 2021\}$. We transform flow Δf_j into a difference in the period-t stock as follows. Where $w_{jt}^B := \sum_{\tau' \in B} w_{j\tau',t}$, we estimate $\Delta F_{jt} := \Delta f_j \cdot \frac{\sum_j w_{jt}^B}{\sum_j n_j}$. We then calculate:

$$w_{j\tau',t}^{\text{Timing 16}} := \begin{cases} \frac{J^{-1}\left(\sum_j w_{jt}^{B_{\text{pre}}}\right) + 0.5 \cdot \Delta F_{jt}}{|B_{\text{pre}}|} & \text{if } \tau' \in B_{\text{pre}} \\ \frac{J^{-1}\left(\sum_j w_{jt}^{B_{\text{post}}}\right) - 0.5 \cdot \Delta F_{jt}}{|B_{\text{post}}|} & \text{if } \tau' \in B_{\text{post}} \\ J^{-1} \sum_j w_{j\tau',t} & \text{if } \tau' \notin B \end{cases} \tag{D.58}$$

and use $w_{j\tau,t}^{\text{Timing 16}}$ to construct $\Delta SPC_j^{\text{Timing 16}}$. We truncate $w_{j\tau,t}^{\text{Timing 16}}$ to be non-negative.

Table D.20, column (6) regresses $\Delta SPC_j^{\text{Timing 16}}$ on local area demographics. There is a statistically-significant correlation with the homeowner rate, but the instrument is otherwise balanced.

Table D.13 shows estimates of β in specification (6) where outcomes are 2021-23 price growth and 2022-23 outflows. Columns (1) and (2) show price effects, and columns (3) and (4) show outflow effects. Results are similar to the main text. The effective F-stat below 10 in columns (1) and (3), which calls for caution in interpreting asymptotic standard errors in those columns. However, that both point estimates and standard errors align with both (i) versions with controls and (ii) coefficient

estimates in the main text gives confidence that weak IV issues do not severely distort results.

Table D.13: Effect of rate lock on local equilibrium, November 2016 event study

Dependent Variables:	%Δ Price, 21-23		Pr(move), 2022	
	(1)	(2)	(3)	(4)
Δ SPC, 2021-23	1.52**	1.75***	-0.55**	-0.60***
	(0.61)	(0.59)	(0.25)	(0.20)
Instrument	Timing 2016	Timing 2016	Timing 2016	Timing 2016
CBSA FEs	Yes	Yes	Yes	Yes
Rent/price x CBSA controls	Yes	Yes	Yes	Yes
Demo x CBSA controls	No	Yes	No	Yes
Orig timing controls	None	None	None	None
Observations	6,370	6,370	6,370	6,370
F-test (1st stage), Δ SPC, 2021-23	33.0	46.7	33.0	46.7
Wald (1st stage), Δ SPC, 2021-23	8.2	14.8	8.2	14.8

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

Source: Equifax Credit Risks Insight Servicing. Columns (1) and (2) show TSLS estimates of $\hat{\beta}$ from specification (6) where the outcome is price growth from 2021-23, and columns (3) and (4) show estimates where the outcome is the probability of a move from 2022-23. All estimates use $\Delta SPC_j^{\text{Timing 16}}$ described in Section D.10 as an instrument for ΔSPC_j. Columns (2) and (4) additionally include the CBSA-by-demographic controls included in Tables D.24 and D.25.

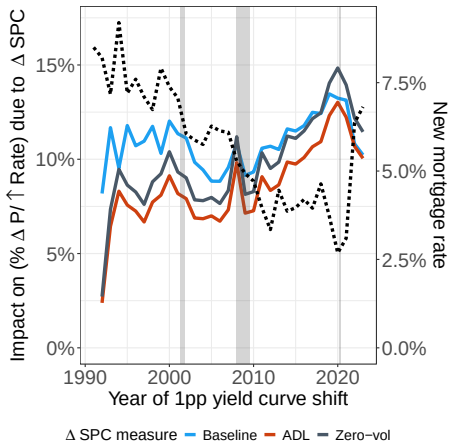
D.11 Alternative prepayment cost measurement.

This appendix replicates price and outflow results from Tables 1, D.24, 2, D.25, and Figure 4 using alternative approaches to measure ΔSPC_j. As outlined in Section B.4, we value the option to refinance in future periods using either the model of optimal refinancing behavior from Agarwal *et al.* (2013) (hereafter, "ADL") or using mortgage-backed securities prices in the TBA market (hereafter, "zero-vol").

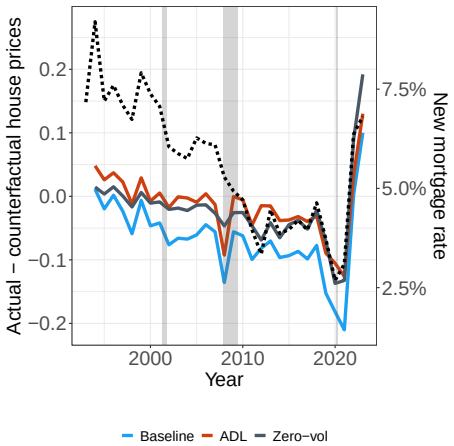
ADL estimates replicating Tables 1 and 2 are in Table D.14. Zero-vol estimates replicating Tables 1 and 2 are in Table D.14. Figure D.16 replicates Figure 4, adding comparisons for estimates that use the ADL or zero-vol approach to estimating ΔSPC_j.

Figure D.16: Effect of rate lock on house prices over time: Alternative measurement.

(a) Offset to price-rate semi-elasticity.



(b) Price effect of removing rate lock.



Source: Equifax Credit Risks Insight Servicing and ICE, McDash®. This figure presents aggregate implications of cross-sectional estimates where ΔSPC_j is measured as in the main text (blue), valuing the future option to refinance using a calibration of the model in Agarwal *et al.* (2013) (red), or valuing the future option to refinance using MBS prices in the TBA market (grey). Panel (a) replicates 4a, and Panel (b) replicates 4b.

Table D.14: Effect of rate lock on prices and outflows: Measurement with ADL formula

(i) Prices.

	Panel (a). Endogenous origination timing. %Δ HPI, 21-23			Panel (b). Origination timing instruments. %Δ HPI, 21-23		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	0.1115	0.1044	0.2582	1.892***	2.517***	1.963**
	(0.0820)	(0.0857)	(0.1736)	(0.6884)	(0.6370)	(0.8311)
Fixed-Effects						
CBSA	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls?	Yes	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls?	No	No	No	No	No	No
Origination timing controls?	None	None	None	None	None	Half-yr
Instrument?		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
Observations	8,000	8,000	8,000	5,975	6,377	6,377
F-test (1st stage), Δ SPC, 2021-23		87,906.7	2,111.3	85.223	114.55	137.87
Wald (1st stage), Δ SPC, 2021-23		28,600.8	271.34	25.372	36.485	66.581

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

(ii) Outflows.

	Panel (a). Endogenous origination timing. Pr(Move), 2022			Panel (b). Origination timing instruments. Pr(Move), 2022		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	-0.1484***	-0.1328***	-0.3069***	-1.162***	-0.9766***	-0.8839***
	(0.0294)	(0.0297)	(0.0738)	(0.2968)	(0.2272)	(0.2541)
Fixed-Effects						
CBSA	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls?	Yes	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls?	No	No	No	No	No	No
Origination timing controls?	None	None	None	None	None	Half-yr
Instrument?		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
Observations	8,000	8,000	8,000	5,975	6,377	6,377
F-test (1st stage), Δ SPC, 2021-23		70,655.6	1,278.2	78.858	77.776	129.23
Wald (1st stage), Δ SPC, 2021-23		16,628.3	139.57	25.067	23.655	62.686

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

Source: Author's calculations using Equifax Credit Risks Insight Servicing, ICE, McDash®, and CDC natality data. Each column reports an estimate of β in equation (6) using a different methodology. Unlike in the main text, the value of the future option to refinance in ΔSPC_j is measured using a calibration of the model in Agarwal *et al.* (2013), as described in Appendix Section B.4. Panel (i) replicates Table 1, and Panel (ii) replicates Table 2.

Table D.15: Effect of rate lock on prices and outflows: Measurement with zero-vol spreads.
(i) Prices.

	Panel (a). Endogenous origination timing. %Δ HPI, 21-23			Panel (b). Origination timing instruments. %Δ HPI, 21-23		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	0.1527*	0.1153	0.3451	2.157***	3.338***	2.151**
	(0.0911)	(0.0946)	(0.2331)	(0.8119)	(0.9354)	(0.9053)
Fixed-Effects						
CBSA	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls?	Yes	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls?	No	No	No	No	No	No
Origination timing controls?	None	None	None	None	None	Half-yr
Instrument?		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
Observations	8,000	8,000	8,000	5,975	6,377	6,377
F-test (1st stage), Δ SPC, 2021-23		70,655.6	1,278.2	78.858	77.776	129.23
Wald (1st stage), Δ SPC, 2021-23		16,628.3	139.57	25.067	23.655	62.686

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

(ii) Outflows.

	Panel (a). Endogenous origination timing. Pr(Move), 2022			Panel (b). Origination timing instruments. Pr(Move), 2022		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	-0.1484***	-0.1328***	-0.3069***	-1.162***	-0.9766***	-0.8839***
	(0.0294)	(0.0297)	(0.0738)	(0.2968)	(0.2272)	(0.2541)
Fixed-Effects						
CBSA	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls?	Yes	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls?	No	No	No	No	No	No
Origination timing controls?	None	None	None	None	None	Half-yr
Instrument?		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
Observations	8,000	8,000	8,000	5,975	6,377	6,377
F-test (1st stage), Δ SPC, 2021-23		70,655.6	1,278.2	78.858	77.776	129.23
Wald (1st stage), Δ SPC, 2021-23		16,628.3	139.57	25.067	23.655	62.686

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

Source: Author's calculations using Equifax Credit Risks Insight Servicing, ICE, McDash®, and CDC natality data. Each column reports an estimate of β in equation (6) using a different methodology. Unlike in the main text, the value of the future option to refinance in ΔSPC_j is measured using MBS pricing in the TBA market – captured by zero-volatility spreads – as described in Appendix Section B.4. Panel (i) replicates Table 1, and Panel (ii) replicates Table 2.

D.12 Interpreting estimates.

From 1992-present, we calculate average ΔSPC caused by a 1pp increase in the short rate. We then use our estimates from Section 3.3 to calculate how much rate lock offsets the negative house price effects of higher rates.

To calculate SPC_{it} after a 1pp increase in the short rate, we need: (i) the post-increase yield curve to discount future payments; (ii) the post-increase new mortgage origination interest rate; and (iii) the post-increase zero-volatility spread, for robustness exercises that price the prepayment option using market MBS yields.

The first step is to estimate the yield curve after a 1pp increase in the short rate. Recall we discount future payments using the Fed's estimates of the nominal zero-coupon yield curve, described in Gürkaynak *et al.* (2007). These estimates model the instantaneous forward rate at maturity n each day d using the functional form from Svensson (1994):

$$f_d(n,0) = \beta_0 + \beta_1 \exp(-n/\tau_1) + \beta_2 \cdot (n/\tau_1) \cdot \exp(-n/\tau_1) + \beta_3 \cdot (n/\tau_2) \cdot \exp(-n/\tau_2) \tag{D.59}$$

where $\beta_0, \beta_1, \beta_2, \beta_3, \tau_1, \tau_2$ are parameters estimated daily by minimizing the sum of squared deviations between predicted and observed bond prices, weighted by inverse duration.

We shift the forward curve by incrementing β_0 by 0.25bp relative to estimates from Gürkaynak *et al.* (2007). To form discount functions, we integrate under the forward curve to compute yields: $y_d(n) = n^{-1} \int_0^n f_d(x,0)dx$ This delivers an estimate of $y_d(n)$ for each day d. For each horizon n, we calculate the average yield over the month of December in each year: $\overline{y}_t(n)$. We then discount at horizon n using the price of the zero-coupon bond that matures n years hence: $1 + r'_{ft,n} = \exp\big(\overline{y}_t(n) \cdot n\big)$.

To calculate the new mortgage origination interest rate after 25bps yield curve shift, we calculate the impact of on modeled 10-year yields: $\Delta\overline{y}_t(10)$. We then assume 30-year fixed mortgage rates change by the same amount, so the post-increase available rate in year t is $r'_{mt} = r_{mt} + \Delta\overline{y}_t(10)$. This is valid if mortgage rates have a constant spread over the 10-year Treasury yield.[63]

To calculate new zero-volatility spreads after a 25bp yield curve shift, we first regress the zero-volatility spread to LIBOR on β_0 and β_0^2 to capture non-nonlinearities: $zv_d = \kappa_0 + \kappa_1 \cdot \beta_{0,d} + \kappa_2 \cdot \beta_{0,d}^2 + e_d$. The coefficient on β_0 is -3.8, the coefficient on β_0^2 is 0.34, and the R-squared is 1.6%. We then predict the post-increase zero-volatility spread as $zv'_d = zv_d + \hat{\kappa}_1 \cdot \beta_{0,d} + \hat{\kappa}_2 \cdot \beta_{0,d}^2$, and take the average across d each year t to calculate $\overline{zv}'_t$. This information is enough to estimate, for each i, SPC'_{it}, which calculates SPC_{it} but replaces $r_{ft,q}, r_{mt}$, and $\overline{zv}_t$ with $r'_{ft,q}, r'_{mt}$, and $\overline{zv}'_t$. We then calculate $SPC'_t := |I_t|^{-1} \sum_{i \in I_t} SPC'_{it}$, the average of SPC'_{it} among borrowers with outstanding mortgages at the end of year t.

Figure 7b plots $\hat{\beta} \cdot (SPC'_t - SPC_t)$: the predicted effect of rate lock on price growth from a 25bp increase in the short rate, by calendar year, where β is the baseline twins estimate in Table 1.

D.13 Effects on residential investment.

This appendix describes our analysis of the residential investment effects of increased prepayment cost. We study effects on the construction of new single-family and multi-family units, as well as the effects on single-family renovations.

Data on new unit construction. To estimate single-family and multi-family residential construction, we use estimates from the Census Building Permit Survey (BPS) – see Appendix Section B.3 for details. Our measure of single-family construction is the annual change in the number of permits for construc-

[63] This approach will under-estimate the increase in mortgage rates if long-run yields are mean-reverting. In that case, an increase in yields should increase the spread to the 10-year Treasury, since a rate increase predicts higher refi risk.

tion with 1 unit. Our measure of multi-family construction is the annual change in the number of permits for construction with 5+ units. We focus on construction over 2022-23. We scale outcomes by the estimated number of single and multi-family units in year-end 2021.

Data on renovations. We estimate renovations in single-family units using CoreLogic permits data. We calculate the zip-level annual count of permits for renovations and divide by the zip-level annual count of local single-family permits. We restrict to zip codes within counties where implied CoreLogic county-level single-family permit counts cover at least 25% of the Census BPS permit count.

Estimates. We estimate γ in the following equation via two-stage least squares:

$$\Delta Q_j/Q_j = \gamma \cdot \hat{\beta} \cdot \Delta SPC_j + \eta'_{c(j)} \cdot X_j + \omega_{c(j)} + \epsilon_j \tag{D.60}$$

where $\Delta Q_j/Q_j$ is the change in units over units in 2021, and $\hat{\beta}$ is the coefficient estimate from equation (6) estimated with a matched instrument and reported in Tables 1 and D.24 as relevant. Multiplying ΔSPC_j by $\hat{\beta}$ implies the right-hand side variable is predicted price growth due to ΔSPC_j, so γ can be interpreted as the price elasticity of supply for new units.

Table D.16 shows baseline estimates with specifications as in Table 1. Panel (a) shows increased single-family construction due to price increases resulting from rate lock. Estimates in Columns (3)-(5), where we address bias due to endogenous origination timing, show that new construction is consistent with a price elasticity of supply ranging from around 0.48 to 0.9. 95% confidence intervals overlap with the average supply elasticity in Baum-Snow and Han (2024) of around 0.4.

Estimates in columns (1)-(3), which do not address bias due to origination timing endogeneity, are much higher – indicating an unrealistic price elasticity of supply of around 2. Upwards bias indicates places with high ΔSPC_j due to especially high origination during low-rate periods have especially elastic supply, consistent with results in Section 3.7.

Panel (b) shows no effect on multi-family construction in columns (3)-(5), but a positive effect in columns (1)-(2).

Panel (c) shows somewhat variable effects on remodeling permits. In Columns (3)-(5), where estimates better avoid biases, the effect appears flat to negative. This makes sense, as we expect impacts on renovations to be ambiguous. On the one hand, prepayment costs to moving means that owners expect to stay in their current home for longer. This could encourage renovation so that units better meet owners' needs. On the other hand, much renovation is driven by new buyers of existing-homes who want to make improvements to update their new space. Since rate lock reduces turnover, it will also have a negative impact on new construction.

D.14 Effects on rents.

This appendix documents the effects of changes in prepayment costs on *rent* growth from 2021 to 2023. We estimate β in equation (6), where the outcome is changes in rents rather than changes in prices. We use the same empirical methodology as in Section 3. To compute rent indices by zip code, we use building-level rent data from RealPage. To smooth out some monthly noise and maximize the number of buildings in the differences we will compute, we first form quarterly averages of reported rents and occupancy rates at the building level. We then compute the change in rent in each building from 2021 Q4 to 2023 Q4. We aggregate these building-level rent changes to the zip code level by taking weighted averages of those rent changes, where the weights are given by the number of occupied units in the building. Analogously to the definition of scaled prepayment costs, we then scale these rent changes at the zip-code level by the average value of a housing unit in that zip code in 2021. Finally, we multiply this value by 12 to annualize the measured change in monthly rent.

Table D.16: Effect of rate lock on residential investment

	Panel A. Δ Single-fam units/units, 2022-23				
	(1)	(2)	(3)	(4)	(5)
Δ SPC, 2021-23	2.165***	2.282***	0.9041***	0.8344***	0.4785***
	(0.3142)	(0.3530)	(0.2465)	(0.1374)	(0.1562)
	Panel B. Δ Multi-fam units/units, 2022-23				
	(1)	(2)	(3)	(4)	(5)
Δ SPC, 2021-23	0.0349**	0.0739***	0.0119	0.0083	0.0052
	(0.0158)	(0.0172)	(0.0121)	(0.0057)	(0.0064)
Fixed-Effects					
CBSA	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls?	No	No	No	No	No
Demo x CBSA controls?	No	No	No	No	No
Origination qtr controls?	No	No	No	Yes	No
Instrument?		Avg rate	Twin	First birth	Lag avg rate
Fit statistics					
Observations	8,000	8,000	5,975	6,377	8,000
F-test (1st stage), Δ SPC, 2021-23		829.15	62.458	145.79	214.88
Wald (1st stage), Δ SPC, 2021-23		145.01	19.855	43.885	106.89
	Panel C. Remodeling permits/units, 2022-23				
	(1)	(2)	(3)	(4)	(5)
Δ SPC, 2021-23	-1.185*	2.948***	-0.3653	-0.2081*	0.3640
	(0.6306)	(0.6182)	(0.2820)	(0.1203)	(0.2979)
Fixed-Effects					
CBSA	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls?	No	No	No	No	No
Demo x CBSA controls?	No	No	No	No	No
Origination qtr controls?	No	No	No	Yes	No
Instrument?		Avg rate	Twin	First birth	Lag avg rate
Fit statistics					
Observations	6,499	6,499	4,972	5,285	6,499
F-test (1st stage), Δ SPC, 2021-23		647.77	56.084	105.23	150.44
Wald (1st stage), Δ SPC, 2021-23		112.99	19.146	30.493	72.339

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

Source: Equifax Credit Risks Insight Servicing, ICE, McDash®, CoreLogic, ACS, and Census. Each column reports estimates of γ in equation (D.60). Column (1) is the fixed effects estimate. Columns (2)-(5) report estimates using $\Delta SPC_j^{\text{Avg}}$, SPC_j^{Twin}, $SPC_j^{\text{First birth}}$, and $\Delta SPC_j^{\text{Lag avg rate}}$. See text and notes to Table 1 for details.

RealPage has somewhat limited geographic coverage, and so our sample size declines by about 30%.

Table D.17 shows main estimates. Point estimates are negative in panel (a), where we use endogenous origination timing variation. Estimates are statistically zero or negative in panel (b), where we address instrument timing endogeneity.

The theory in Section 1 does not provide clear guidance on what should happen to rents *within* a city in market segments with higher or lower prepayment costs. However, it does suggest that *across cities*, cities with higher average prepayment cost should have lower rents due to decreased own-to-rent moves. We test this hypothesis in Table D.18, where we drop CBSA fixed effects. Point estimates broadly remain negative, although some of the origination timing instruments in Panel (b) lose precision.

Table D.17: Effect of rate lock on local equilibrium rents.
(i) No CBSA x demographic controls.

	Panel (a). Endogenous origination timing.			Panel (b). Origination timing instruments.		
	Δ Rent, 2021-23 / House price, 2021					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	-0.0122	-0.0227***	-0.3424***	-0.0029	-0.2185***	-0.1758**
	(0.0080)	(0.0087)	(0.0384)	(0.0609)	(0.0465)	(0.0680)
Instrument		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
CBSA FEs	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls	No	No	No	No	No	No
Demo x CBSA controls	No	No	No	No	No	No
Orig timing controls	None	None	None	None	None	Half-yr
Observations	5,270	5,270	5,270	4,318	4,573	4,573
F-stat (1st stage)		87,231.8	532.9	46.2	138.2	155.0
Wald (1st stage)		11,383.2	103.5	18.1	46.1	66.6

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

(ii) With CBSA x demographic controls.

	Panel (a). Endogenous origination timing.			Panel (b). Origination timing instruments.		
	Δ Rent, 2021-23 / House price, 2021					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	-0.0193*	-0.0189*	-0.0539***	0.0333	-0.0867**	-0.1418*
	(0.0099)	(0.0098)	(0.0384)	(0.0578)	(0.0368)	(0.0831)
Instrument		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
CBSA FEs	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls	No	No	No	No	No	No
Demo x CBSA controls	Yes	Yes	Yes	Yes	Yes	Yes
Orig timing controls	None	None	None	None	None	Half-yr
Observations	5,270	5,270	5,270	4,318	4,573	4,573
F-stat (1st stage)		64,871.5	3,408.3	50.9	162.5	156.1
Wald (1st stage)		20,686.4	455.3	17.0	41.8	68.4

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

Source: Author's calculations using Equifax Credit Risks Insight Servicing, ICE, McDash®, RealPage, and CDC natality data. Each column reports an estimate of β in equation (6) using a different methodology, where the outcome variable is the constant-quality change in rents from 2021-23 divided by 2021 zip-level house prices. See notes to Table 1 for methodological details. Panel (i) has no CBSA x demographic controls. Panel (ii) adds the CBSA x demographic controls described in the notes to Table D.24.

Table D.18: Effect of rate lock on local equilibrium rents: No CBSA fixed effects.
(i) No CBSA x demographic controls.

	Panel (a). Endogenous origination timing.			Panel (b). Origination timing instruments.		
	Δ Rent, 2021-23 / House price, 2021					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	-0.0498***	-0.0480***	-0.2942***	-0.1413	-0.0476	-0.1542**
	(0.0108)	(0.0099)	(0.0268)	(0.1201)	(0.0569)	(0.0652)
Instrument		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
CBSA FEs	No	No	No	No	No	No
Rent/price x CBSA controls	No	No	No	No	No	No
Demo x CBSA controls	No	No	No	No	No	No
Orig timing controls	None	None	None	None	None	Half-yr
Observations	5,270	5,270	5,270	4,318	4,573	4,573
F-stat (1st stage)		77,390.2	936.4	33.2	199.8	135.7
Wald (1st stage)		4,204.2	203.0	6.9	15.7	54.8

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

(ii) With CBSA x demographic controls.

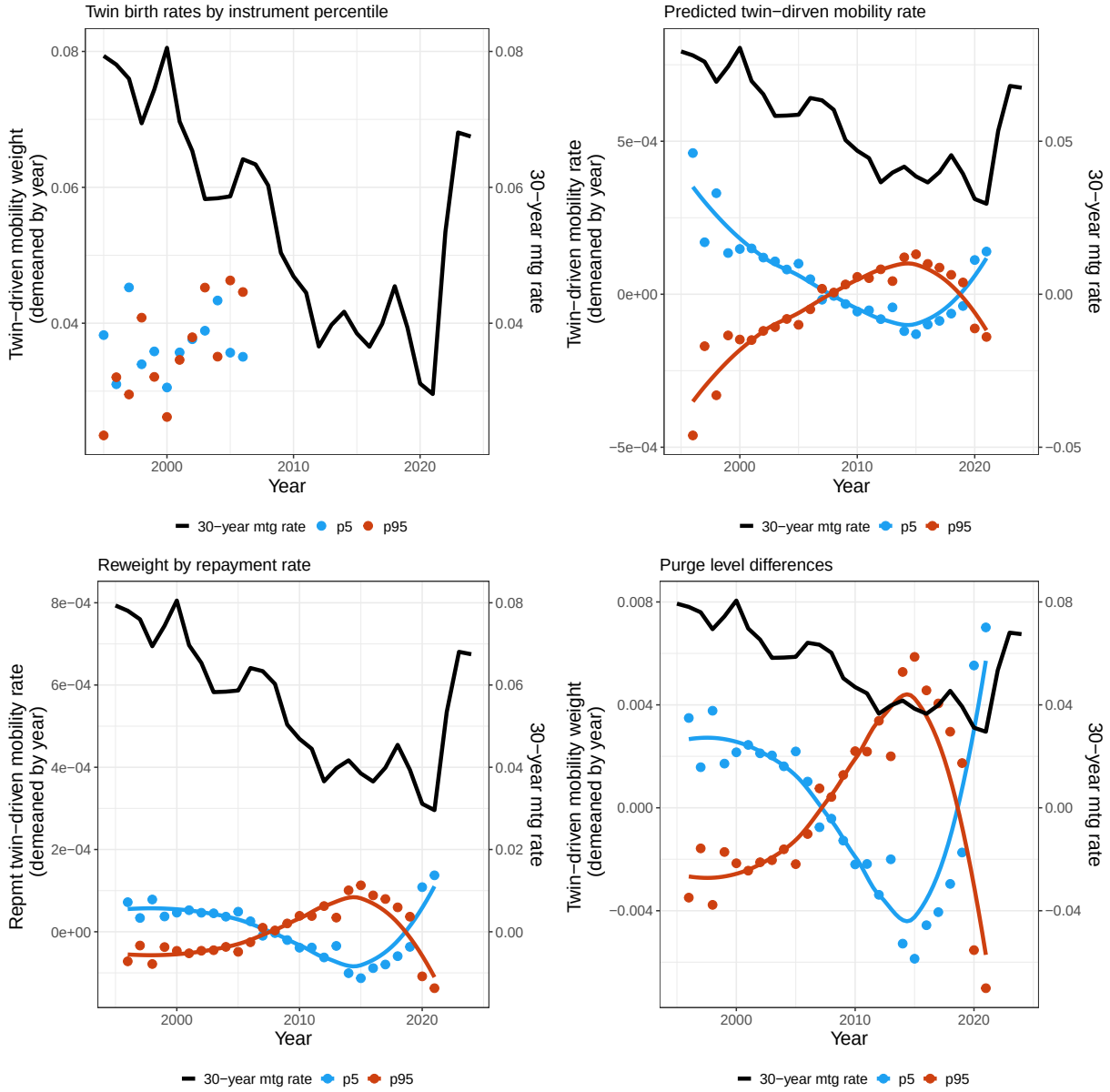
	Panel (a). Endogenous origination timing.			Panel (b). Origination timing instruments.		
	Δ Rent, 2021-23 / House price, 2021					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	-0.0542***	-0.0444***	-0.0833***	-0.2828	0.0045	-0.1523**
	(0.0122)	(0.0110)	(0.0384)	(0.0139)	(0.1833)	(0.0622)
Instrument		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
CBSA FEs	No	No	No	No	No	No
Rent/price x CBSA controls	No	No	No	No	No	No
Demo x CBSA controls	Yes	Yes	Yes	Yes	Yes	Yes
Orig timing controls	None	None	None	None	None	Half-yr
Observations	5,270	5,270	5,270	4,318	4,573	4,573
F-stat (1st stage)		79,206.0	3,939.7	15.5	212.4	172.9
Wald (1st stage)		4,572.6	565.1	3.0	13.0	66.7

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

Source: Author's calculations using Equifax Credit Risks Insight Servicing, ICE, McDash®, RealPage, and CDC natality data. Each column reports an estimate of β in equation (6) using a different methodology, where the outcome variable is the constant-quality change in rents from 2021-23 divided by 2021 zip-level house prices. Unlike in the main text and Table D.17, we omit CBSA fixed effects. See notes to Table 1 for methodological details. Panel (i) has no CBSA x demographic controls. Panel (ii) adds the CBSA x demographic controls described in the notes to Table D.24.

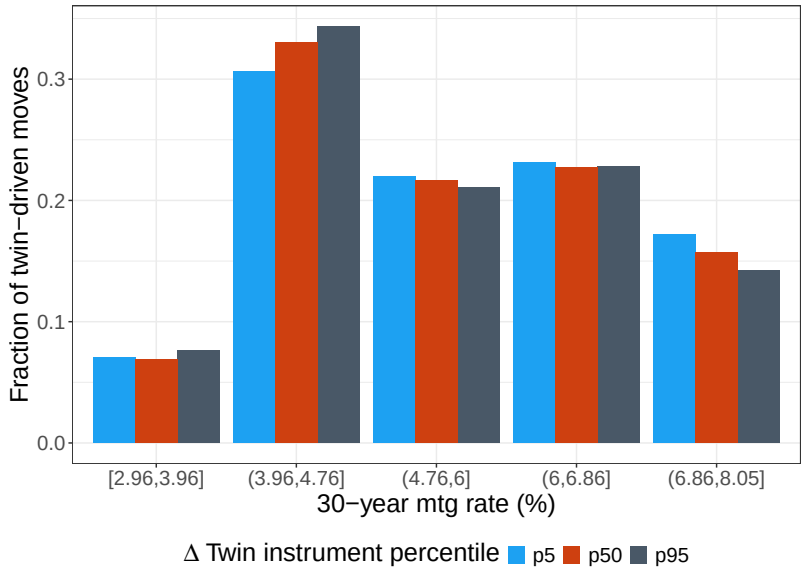
D.15 Additional figures.

Figure D.17: Visual construction of $\Delta SPC^{\text{twin}}_{j,2023}$.



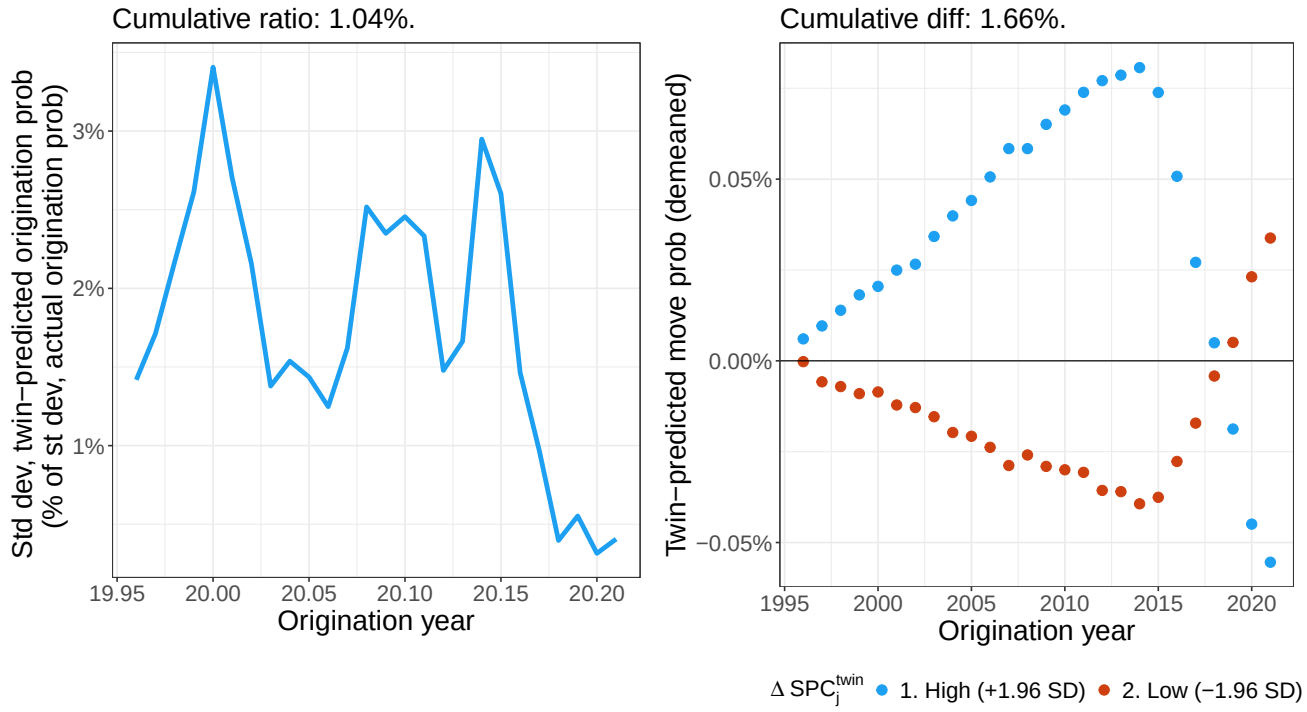
Source: Equifax Credit Risks Insight Servicing, ICE, McDash®, and CDC Natality. The figure plots steps for constructing $w^{twin}_{j\tau,2023}$ for zip codes at the 5th and 95th percentiles of the $\Delta SPC^{\text{twin}}_{j}$ distribution. The top-left panel shows raw twin birth rates for 1995-2006 along with the 30-year mortgage rate. The top-right panel shows $\nu^{\text{twin}}_{j\tau}$, demeaned for each τ to show relative differences. The bottom-left panel shows $K(2021-\tau)\cdot\nu^{\text{twin}}_{j\tau}$. The bottom-right panel shows $w^{twin}_{j\tau,2023}$. See main text for details.

Figure D.18: Distribution of mortgage rate at origination.



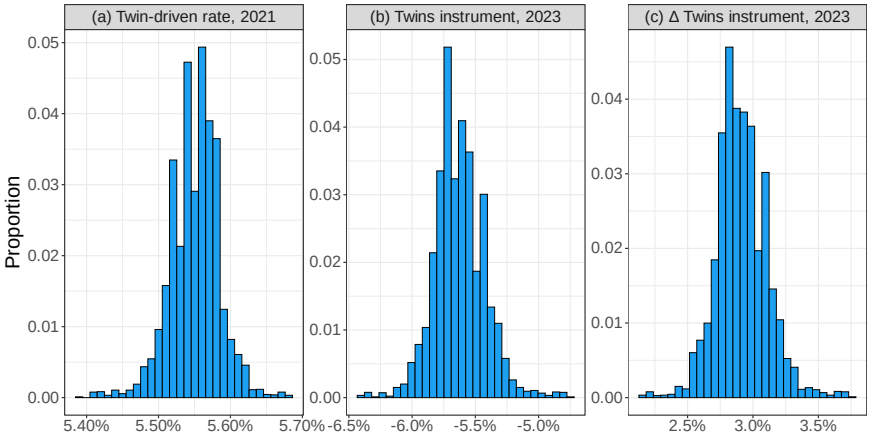
Source: Author's calculations using CDC natality data, Equifax Credit Risks Insight Servicing data, and ICE, McDash® data. For each market j, we predict the fraction of households who had their mortgage originated when the interest rate was in each 30-year Treasury rate quintile: $\Pr(T_{ijt} \in \mathcal{T}_q) = \sum_{\tau \leq 2021} w_{j\tau,t}^{twin} \mathbb{1}(\tau \in \mathcal{T}_q)$, where $\mathcal{T}_q$ is the set of calendar years for which the 30-year Treasury is within the qth quintile. The figure plots $\Pr(T_{ijt} \in \mathcal{T}_q)$, split by percentiles of $\Delta SPC_j^{\text{twin}}$.

Figure D.19: Variation in twin-predicted moves by year.



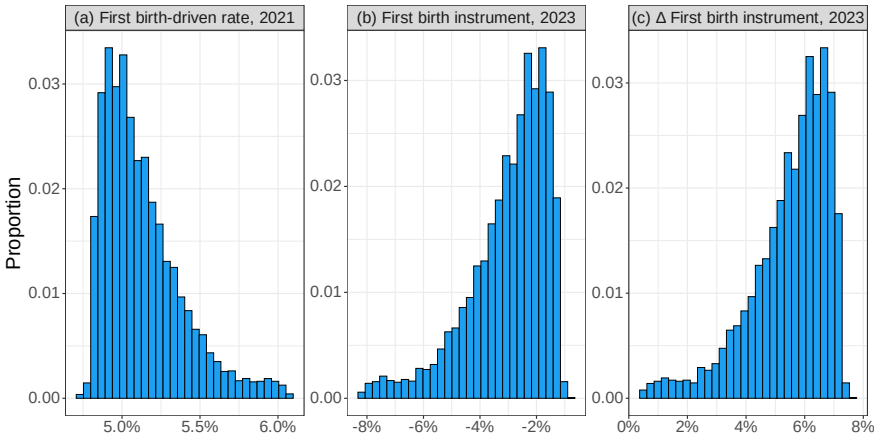
Source: Equifax Credit Risks Insight Servicing, ICE, McDash®, and CDC Natality. The series on the left panel plots, for each year τ, $\sigma(w_{j\tau,2023}^{\text{twin}})/\sigma(w_{j\tau,2023})$, where $\sigma(\cdot)$ is the cross-sectional standard deviation, $w_{j\tau,2023}^{\text{twin}}$) is the twin-driven probability an outstanding mortgage in 2021 in zip j was originated in year τ, and $w_{j\tau,2023}$ is the actual probability an outstanding mortgage in 2021 in zip j was originated in year τ. The right panel plots takes zip codes 1.96 standard deviations above and below the mean of the instrument $\Delta SPC_j^{\text{twin}}$ and plots (demeaned) $w_{j\tau,2023}^{\text{twin}}$ for each year. The title displays the sum of the absolute difference across years. See text for details.

Figure D.20: Distribution of twins instrument.



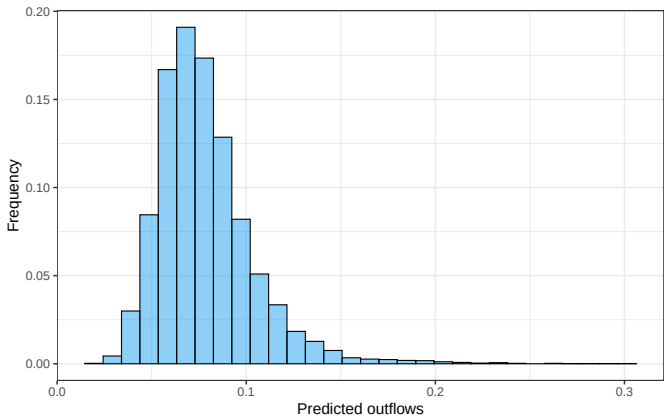
Source: Author's calculations using CDC natality data, Equifax Credit Risks Insight Servicing data, and ICE, McDash® data. Panel (a) plots the pdf of $r^{twin}_{j,2023} := \sum_{\tau \leq 2021} w^{\text{twin}}_{j\tau,2023} \cdot r_{m\tau}$. Panels (b) and (c) plot the pdf of $SPC^{\text{twin}}_{j,2023}$ and $\Delta SPC^{\text{twin}}_{j}$, respectively.

Figure D.21: Distribution of first birth instrument.



Source: Equifax Credit Risks Insight Servicing, ICE, McDash®, and CDC Natality. Panel (a) plots the pdf of $r^{births}_{j,2023} := \sum_{\tau \leq 2021} w^{\text{births}}_{j\tau,2023} \cdot r_{m\tau}$. Panels (b) and (c) plot the pdf of $SPC^{\text{births}}_{j,2023}$ and $\Delta SPC^{\text{births}}_{j}$, respectively.

Figure D.22: Predicted outflows in 2022-23 based on 2019 age distribution.



Source: Equifax Credit Risks Insight Servicing. The figure plots the distribution of predicted outflows based on coefficient estimates in equation (10) and the 2019 age distribution.

D.16 Additional tables.

Table D.19: Local housing market summary statistics.

	N	Mean	St. Dev.	p25	Median	p75
% Δ Price, 21-23	8,000	0.242	0.0867	0.186	0.235	0.294
SPC, 2021	8,000	−0.123	0.0171	−0.134	−0.123	−0.111
SPC, 2023	8,000	0.0447	0.0151	0.0363	0.0459	0.0544
Δ SPC, 2023	8,000	0.168	0.0211	0.154	0.167	0.18
SPC (ADL), 2021	8,000	−0.0702	0.0158	−0.0803	−0.0696	−0.0593
SPC (ADL), 2023	8,000	0.0556	0.0146	0.0476	0.0564	0.0645
Δ SPC (ADL), 2023	8,000	0.126	0.019	0.114	0.125	0.137
SPC (Zero−vol), 2021	8,000	−0.0633	0.012	−0.0712	−0.0628	−0.0546
SPC (Zero−vol), 2023	8,000	0.0756	0.0138	0.0682	0.0757	0.0836
Δ SPC (Zero−vol), 2023	8,000	0.139	0.0168	0.128	0.138	0.149
Pr(Move next year), 2022-23	8,000	0.058	0.0198	0.0448	0.0546	0.0669
Pr(Move next year), 2021-22	8,000	0.0709	0.0242	0.055	0.0673	0.0821
Pr(Move next year), 2020-21	8,000	0.073	0.0236	0.0573	0.0696	0.0846
Pr(Move next year), 2019-20	8,000	0.0782	0.0271	0.0606	0.0737	0.0904
Pr(Move next year), 2006-19 avg	8,000	0.0761	0.0248	0.0593	0.0717	0.0876
N hhlds ('000s), 2021	8,000	10,700	5,680	6,220	9,740	14,000
Median family income, 2021 (USD '000s)	8,000	90,500	28,700	69,200	85,600	107,000
Population density (pop / sq km)	8,000	954	1,500	124	446	1,250
Avg age (yrs), 2021	8,000	51.7	2.83	50.1	51.4	53
Frac white, 2021	8,000	0.777	0.197	0.694	0.841	0.923
Unit value, 2021 (USD)	8,000	268,000	126,000	169,000	242,000	343,000
Rent to price ratio, 2021	8,000	0.0605	0.0202	0.0466	0.0572	0.0702
Δ rent, 2021-21 / Price, 2021	8,000	0.00893	0.00551	0.00575	0.00835	0.0116
Homeowner rate, 2021	8,000	0.68	0.151	0.585	0.698	0.794
Avg LTV ratio, 2021 (%)	8,000	78.3	6.73	73.7	78.7	82.9
Avg DTI ratio, 2021 (%)	8,000	33.3	2.31	31.8	33.3	34.8
Fixed-rate mtg share, 2021	8,000	0.974	0.0143	0.967	0.977	0.984
Avg Risk Score 3.0, 2021	8,000	756	22.7	740	758	774
Housing supply elasticity	6,194	0.208	0.151	0.104	0.207	0.31
Births / population, 2005	6,377	0.0141	0.00254	0.0123	0.014	0.0157
Twin birth rate, 2005	5,975	0.0352	0.00708	0.0306	0.0346	0.0389
Δ Twin birth rate, 2005	5,975	−0.000392	0.00519	−0.00328	−0.000155	0.00246
First birth / population, 2005	6,377	0.00561	0.00091	0.00495	0.0056	0.00622

Source: Equifax Credit Risks Insight Servicing and ACS. The table shows summary statistics at the zip code level for the sample used to estimate coefficients in equation (6). House price growth is calculated from zip code-level FHFA all-transaction repeat sale price indices. *SPC* variables, the annual probability of moving for years 2006-2023, average age, average LTV ratio, average DTI ratio, the fixed-rate mortgage share, and the average Vantage Risk Score 3.0 are calculated from the 50% CRISM sample. Fertility-related variables, such as the birth rate, twin birth rate, and first birth rate, are calculated using CDC WONDER. Remaining variables are calculated from ACS. See Section B for details.

Table D.20: Instrument balance on area demographics

Dependent Variables:	Δ SPC, 2021-23	IV for Δ SPC, 2021-23				
	(1)	(2)	(3)	(4)	(5)	(6)
Ln inc, 2021, sd	0.11***	0.07***	0.03	0.04	0.02**	-0.07
	(0.03)	(0.02)	(0.03)	(0.04)	(0.01)	(0.12)
Homeowner rate, 2021, sd	0.17***	0.09***	0.02	-0.03	0.01	0.18**
	(0.04)	(0.02)	(0.04)	(0.06)	(0.01)	(0.08)
Pop density, 2021, sd	0.90***	0.62***	-0.05	-0.17	0.16***	-0.09
	(0.19)	(0.12)	(0.14)	(0.20)	(0.04)	(0.16)
1-Nonwhite share, 2021, sd	0.10**	0.05	0.07	0.10	-0.02*	-0.01
	(0.05)	(0.04)	(0.05)	(0.07)	(0.01)	(0.10)
Origination LTV ratio, 2021, sd	0.16*	-0.04	0.03	-0.06	0.00	0.03
	(0.09)	(0.03)	(0.04)	(0.06)	(0.01)	(0.05)
Origination DTI ratio, 2021, sd	0.08**	0.07***	-0.03	0.13***	0.04***	0.00
	(0.04)	(0.02)	(0.04)	(0.05)	(0.01)	(0.06)
FRM share, 2021, sd	0.25***	0.24***	-0.01	0.18***	-0.02**	0.04
	(0.03)	(0.02)	(0.02)	(0.03)	(0.01)	(0.04)
Rent / price, 2021, sd	-0.07**	-0.06***	0.02	0.09**	0.01	0.00
	(0.03)	(0.02)	(0.03)	(0.04)	(0.01)	(0.04)
Equifax Risk Score 3.0, 2021, sd	0.37***	0.36***	-0.10	-0.03	-0.06	-0.03
	(0.10)	(0.08)	(0.08)	(0.09)	(0.04)	(0.10)
Age, 2021, sd	-0.63***	-0.35***	0.02	0.26***	0.09***	-0.02
	(0.07)	(0.04)	(0.05)	(0.07)	(0.02)	(0.12)
Ln inc, 2019, sd	0.10***	0.16***	-0.01	0.16***	0.06***	0.17
	(0.04)	(0.02)	(0.03)	(0.04)	(0.01)	(0.14)
Homeowner rate, 2019, sd	-0.18***	-0.08***	-0.01	-0.11*	0.00	-0.18**
	(0.04)	(0.03)	(0.05)	(0.06)	(0.01)	(0.08)
Pop density, 2019, sd	-0.94***	-0.69***	0.01	0.17	-0.15***	0.08
	(0.18)	(0.12)	(0.13)	(0.19)	(0.04)	(0.17)
1-Nonwhite share, 2019, sd	0.00	0.02	0.04	0.01	0.02	0.00
	(0.06)	(0.05)	(0.05)	(0.06)	(0.01)	(0.09)
Origination LTV ratio, 2019, sd	0.34***	0.06***	0.03	0.15***	0.01	0.00
	(0.08)	(0.02)	(0.03)	(0.05)	(0.01)	(0.04)
Origination DTI ratio, 2019, sd	0.15***	0.03	0.03	0.00	0.02**	0.10*
	(0.03)	(0.02)	(0.04)	(0.04)	(0.01)	(0.05)
FRM share, 2019, sd	0.09***	-0.02	-0.01	-0.09***	0.04***	-0.05
	(0.03)	(0.02)	(0.03)	(0.03)	(0.01)	(0.04)
Rent / price, 2019, sd	-0.01	-0.06***	-0.03	0.02	-0.01	0.00
	(0.04)	(0.02)	(0.03)	(0.05)	(0.01)	(0.05)
Equifax Risk Score 3.0, 2019, sd	-0.27***	0.22***	0.08	0.20**	0.11***	0.03
	(0.09)	(0.08)	(0.07)	(0.08)	(0.04)	(0.10)
Age, 2019, sd	0.36***	0.18***	-0.03	-0.22***	-0.10***	0.02
	(0.06)	(0.04)	(0.05)	(0.07)	(0.02)	(0.12)
Instrument		Avg rate	Twin	First birth	Lag avg rate	Timing 2016
CBSA FEs	Yes	Yes	Yes	Yes	Yes	Yes
Orig timing controls	None	None	None	None	Half-yr	None
Observations	7,982	7,982	5,963	6,363	6,363	6,363
Within Adjusted R^2	0.60	0.86	0.03	0.10	0.15	0.00
Wald (joint nullity)	141.29	616.82	2.40	11.39	24.31	2.39

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

Source: Equifax Credit Risks Insight Servicing and ICE, McDash®. The table regresses ΔSPC_j, $\Delta SPC_j^{\text{Avg rate}}$, $\Delta SPC_j^{\text{Twin}}$, $\Delta SPC_j^{\text{First birth}}$, $\Delta SPC_j^{\text{Lag avg rate}}$, and $\Delta SPC_j^{\text{Timing 16}}$ on various 2021 and 2019 zip-level demographics. All specifications include CBSA fixed effects. The specification with $\Delta SPC_j^{\text{Lag avg rate}}$ includes controls for the fraction of 2021 mortgages originated in each pre-2019 half-year bin, as described in Section 3.4. County-clustered standard errors are in parentheses.

Table D.21: Demographics do not explain twin birth rate variation.

	Pr(twins) (%), 2019		Δ Pr(twins) (%), 2020	
	(1)	(2)	(3)	(4)
Pr(twins) (%), 2018	0.1459**			
	(0.0739)			
Δ Pr(twins) (%), 2019			-0.4156***	
			(0.0662)	
Ln inc, 2019, sd		0.0178		0.0146
		(0.0173)		(0.0232)
Ln unit val, 2019, sd		0.0182		-0.0303
		(0.0296)		(0.0392)
Rent / price, 2019, sd		0.0062		-0.0042
		(0.0099)		(0.0123)
Homeowner rate, 2019, sd		-0.0005		-0.0065
		(0.0111)		(0.0136)
Pop density, 2019, sd		-0.0052		-0.0102
		(0.0161)		(0.0160)
1-Nonwhite share, 2019, sd		0.0005		-0.0119
		(0.0121)		(0.0156)
Age, 2019, sd		0.0480		-0.0580
		(0.0299)		(0.0401)
Age, 2018, sd		-0.0407		0.0395
		(0.0275)		(0.0375)
FRM share, 2019, sd		-0.0012		-0.0012
		(0.0125)		(0.0162)
Origination LTV ratio, 2019, sd		0.0157		-0.0127
		(0.0157)		(0.0191)
Origination DTI ratio, 2019, sd		-0.0115		-0.0026
		(0.0128)		(0.0169)
Equifax Risk Score 3.0, 2019, sd		-0.0093		0.0027
		(0.0209)		(0.0264)
Ln population, 2019, sd		-0.0190		0.0456
		(0.0356)		(0.0484)
Pr(Move in 2018), sd		0.0020		0.0046
		(0.0048)		(0.0080)
Fixed-effects				
CBSA	Yes	Yes	Yes	Yes
Observations	5,975	5,963	5,975	5,963
Within Adjusted R^2	0.02655	0.00571	0.18007	0.00672

Clustered (County) standard-errors in parentheses
****: 0.01, **: 0.05, *: 0.1. County-clustered standard erro1rs in parentheses.*

Source: Equifax Credit Risks Insight Servicing, CDC natality data, and ACS. The table shows estimated coefficients from a regression of the levels of and first differences in the twin birth rate on local demographics. The outcome in Columns (1)-(2) is the 2019 twin birth rate in percent (1%=1), and the outcome in Columns (3)-(4) is the 2019-2020 change in the twin birth rate in percent (1%=1).

Table D.22: Twin-predicted moves predict moves and outstanding 2021 rate.

Dependent Variables:	Pr(Originate)	Avg mtg rate at origination, 2021
Model:	(1)	(2)
Variables		
Twin-predicted Pr(Originate)	0.1436***	
	(0.0124)	
Twin-predicted mtg rate, 2021		1.546***
		(0.3502)
Fixed-effects		
CBSA	Yes	Yes
Fit statistics		
Observations	117,800	5,975
Within Adjusted R^2	0.01333	0.01129

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Source: Equifax Credit Risks Insight Servicing, ICE, McDash®, and CDC natality data. Column (1) shows coefficients from a regression of observed origination shares for the distribution of mortgages outstanding in 2021 ($w_{j\tau,2023}$) on twin-predicted origination shares ($w^{twin}_{j\tau,2023}$). Column (2) shows results of a regression of the average origination-timing driven outstanding 30-year fixed mortgage rate in 2021 ($r_{j,2023}$) on twin-driven mortgage rates ($r^{twin}_{j,2023} := \sum_{\tau \leq 2021} w_{j\tau,2023} \cdot r_{m\tau}$).

Table D.23: Predicted vs. actual mortgage rate, first births IV.

Dependent Variables:	Avg mtg rate, 2021	Avg mtg rate at origination, 2021
Model:	(1)	(2)
Variables		
First birth-driven rate, 2021	0.4171***	0.4064***
	(0.0438)	(0.0365)
Fixed-effects		
CBSA	Yes	Yes
Fit statistics		
Observations	6,377	6,377
Within Adjusted R^2	0.06243	0.06686

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

Equifax Credit Risks Insight Servicing, ICE, McDash®, CDC Natality data, and ACS. Column (1) regresses the 2021 average outstanding mortgage rate on the first-birth driven analogue. Column (2) replaces the outcome variable with $r^{\text{Avg}}_{j,2023} := \sum_{\tau \leq 2021} w_{j\tau,2023} r_{m\tau}$, the average outstanding origination-timing driven mortgage rate.

Table D.24: Effect of rate lock on local equilibrium prices, with CBSA-by-demographic controls.

	Panel (a). Endogenous origination timing. %Δ Price, 21-23			Panel (b). Origination timing instruments. %Δ Price, 21-23		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	0.25***	0.27***	0.38***	2.12***	1.88***	1.48***
	(0.09)	(0.09)	(0.13)	(0.69)	(0.54)	(0.53)
Instrument		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
CBSA FEs	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls	Yes	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls	Yes	Yes	Yes	Yes	Yes	Yes
Orig timing controls	None	None	None	None	None	Half-yr
Observations	8,000	8,000	8,000	5,975	6,377	6,377
F-stat (1st stage)		92,528.1	4,849.7	71.0	163.5	294.1
Wald (1st stage)		25,921.7	577.0	19.6	38.5	136.7

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

Source: Equifax Credit Risks Insight Servicing, ICE, McDash®, CDC Natality data, and ACS. We include CBSA fixed effects interacted with controls for 2021 log income, log population, log unit value, the homeownership rate, population density, age fixed effects, average origination LTV ratio, average origination DTI ratio, the fixed-rate mortgage share, the nonwhite share, the rent-price ratio, and the ratio of 2023 rents to 2021 prices.

Table D.25: Effect of rate lock on local outflows, with CBSA-by-demographic controls.

	Panel (a). Endogenous origination timing. Pr(Move), 2022			Panel (b). Origination timing instruments. Pr(Move), 2022		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	-0.26***	-0.25***	-0.52***	-1.14***	-0.55***	-0.64***
	(0.03)	(0.03)	(0.05)	(0.25)	(0.13)	(0.19)
Instrument		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
CBSA FEs	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls	Yes	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls	Yes	Yes	Yes	Yes	Yes	Yes
Orig timing controls	None	None	None	None	None	Half-yr
Observations	8,000	8,000	8,000	5,975	6,377	6,377
F-stat (1st stage)		92,528.1	4,849.7	71.0	163.5	294.1
Wald (1st stage)		25,921.7	577.0	19.6	38.5	136.7

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

Source: Equifax Credit Risks Insight Servicing, ICE, McDash®, CDC Natality data, and ACS. We include CBSA fixed effects interacted with controls for 2021 log income, log population, log unit value, the homeownership rate, population density, age fixed effects, average origination LTV ratio, average origination DTI ratio, the fixed-rate mortgage share, the nonwhite share, the rent-price ratio, and the ratio of 2023 rents to 2021 prices.

Table D.26: Effect of rate lock on moves, by destination.

(a) Cross-CBSA moves.

	Panel (a). Endogenous origination timing.			Panel (b). Origination timing instruments.		
	Pr(move), diff CBSA, 2022					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	0.03**	0.05***	0.15***	-0.09	0.21***	0.02
	(0.01)	(0.01)	(0.02)	(0.06)	(0.05)	(0.05)
Instrument		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
CBSA FEs	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls	Yes	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls	No	No	No	No	No	No
Orig timing controls	None	None	None	None	None	Half-yr
Observations	7,876	7,876	7,876	5,853	6,255	6,255
F-stat (1st stage)		115,491.5	1,585.2	77.5	97.1	186.1
Wald (1st stage)		24,585.2	161.9	24.0	30.1	90.8

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

(b) Own-to-rent moves.

	Panel (a). Endogenous origination timing.			Panel (b). Origination timing instruments.		
	Pr(move) to rent, 2022					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	-0.01	-0.01	-0.10***	-0.28***	-0.19**	-0.24**
	(0.01)	(0.01)	(0.03)	(0.09)	(0.08)	(0.10)
Instrument		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
CBSA FEs	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls	Yes	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls	No	No	No	No	No	No
Orig timing controls	None	None	None	None	None	Half-yr
Observations	7,994	7,994	7,994	5,970	6,372	6,372
F-stat (1st stage)		117,279.1	1,609.6	74.9	97.6	188.3
Wald (1st stage)		24,582.3	166.7	23.7	30.4	92.6

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

Source: Equifax Credit Risks Insight Servicing, ICE, McDash®, CDC Natality data, and ACS. Top panel. The outcome is the fraction of households in zip j moving to a different CBSA from 2022-23. See notes to Table 1 for additional details. Bottom panel. The outcome is the imputed fraction of households in zip j moving from owning to renting from 2022-23. For each zip, we calculate the overall $Pr(Move, 2022)_j$, as well as $Pr(OwnToRent|Move, 2022)_j$ using the methodology described in Section B.1.1. The outcome is $Pr(OwnToRent, 2022)_j := Pr(Move, 2022)_j \times Pr(OwnToRent|Move, 2022)_j$. See notes to Table 1 for additional details.

Table D.27: Cross-sectional price effects: Timing robustness.
(a) Effect of 2021-23 ΔSPC on 2019-2021 house price growth.

	Panel (a). Endogenous origination timing.			Panel (b). Origination timing instruments.		
			pg_2019_2021			
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2021-23	0.2971***	0.2332***	-1.137***	0.0794	0.0285	-0.0429
	(0.0474)	(0.0472)	(0.1755)	(0.4359)	(0.3291)	(0.3929)
Fixed-Effects						
CBSA	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls?	Yes	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls?	No	No	No	No	No	No
Origination timing controls?	None	None	None	None	None	Half-yr
Instrument?		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
Observations	8,000	8,000	8,000	5,975	6,377	6,377
F-test (1st stage), Δ SPC, 2021-23		117,326.6	1,610.0	74.947	97.154	188.62
Wald (1st stage), Δ SPC, 2021-23		24,639.5	166.80	23.740	30.185	92.866

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

(b) Effect of 2019-23 ΔSPC on 2019-23 house price growth.

	Panel (a). Endogenous origination timing.			Panel (b). Origination timing instruments.		
			%Δ HPI, 19-23			
	(1)	(2)	(3)	(4)	(5)	(6)
Δ SPC, 2019-23	0.0828	0.0932	0.2121	1.786***	2.087***	1.128**
	(0.0723)	(0.0766)	(0.1434)	(0.6449)	(0.4888)	(0.4825)
Fixed-Effects						
CBSA	Yes	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls?	Yes	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls?	No	No	No	No	No	No
Origination timing controls?	None	None	None	None	None	Half-yr
Instrument?		Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
Observations	8,000	8,000	8,000	5,975	6,377	6,377
F-test (1st stage), Δ SPC, 2019-23		79,780.1	2,588.3	73.830	140.52	294.92
Wald (1st stage), Δ SPC, 2019-23		22,854.6	331.66	22.701	48.662	142.09

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

Source: Equifax Credit Risks Insight Servicing, ICE, McDash®, CDC Natality data, and ACS. Each panel replicates Table 1, with different timing for ΔP and ΔSPC. In (a), the outcome is price growth from 2019-21, and the regressor is ΔSPC from 2021-23. In (b), the outcome is price growth from 2019-23, and the regressor is ΔSPC from 2019-23. See notes to Table 1 for additional details.

Table D.28: Effect of rate lock on local price growth: first stage and reduced form.

(a) First stage.

	Δ SPC, 2021-23				
	(1)	(2)	(3)	(4)	(5)
IV for Δ SPC, 2021-23	1.06***	0.70***	1.79***	0.23***	0.74***
	(0.01)	(0.05)	(0.37)	(0.04)	(0.08)
Instrument	Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
CBSA FEs	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls	No	No	No	No	No
Orig timing controls	None	None	None	None	Half-yr
Observations	8,000	8,000	5,975	6,377	6,377

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

(b) Reduced form.

	%Δ Price, 21-23				
	(1)	(2)	(3)	(4)	(5)
IV for Δ SPC, 2021-23	0.10	0.18	3.14**	0.54***	1.05**
	(0.08)	(0.12)	(1.26)	(0.12)	(0.43)
Instrument	Avg rate x LTV	Avg rate	Twin	First birth	Lag avg rate
CBSA FEs	Yes	Yes	Yes	Yes	Yes
Rent/price x CBSA controls	Yes	Yes	Yes	Yes	Yes
Demo x CBSA controls	No	No	No	No	No
Orig timing controls	None	None	None	None	Half-yr
Observations	8,000	8,000	5,975	6,377	6,377

****: 0.01, **: 0.05, *: 0.1. County-clustered standard errors in parentheses.*

Source: Equifax Credit Risks Insight Servicing, ICE, McDash®, CDC Natality data, and ACS. The table reports first stage (in panel a) and reduced form (in panel b) estimates for the five IVs used in Table 1. See notes to Table 1 for additional details.

E Section 5 details.

E.1 Household expectations

Expectations. Households expect that aggregate state variables $\theta_t = (r_t, \ell_t, P_t)$ follow a correlated first-order Markov process:

$$E[\log \theta_{t+1} | \theta_t] = \log \theta_t + \boldsymbol{\mu} + \boldsymbol{\rho} \cdot r_t + \boldsymbol{\Sigma}(r_t) \cdot \boldsymbol{\omega}_{t+1} \tag{E.1}$$

where $\boldsymbol{\omega}_{t+1} \equiv (\omega^r_{t+1}, \omega^\ell_{t+1}, \omega^P_{t+1})' \overset{\text{iid}}{\sim} N(0, I)$. In this expression, $\boldsymbol{\mu} \equiv (\mu_r, \mu_\ell, \mu_P)'$ encodes average growth and $\boldsymbol{\rho} \equiv (\rho_r, \rho_\ell, \rho_P)'$ the correlation between growth and the short rate. The matrix $\boldsymbol{\Sigma}(r_t)$ is lower diagonal, and determines households' expectations for how innovations co-vary. It depends on r_t because the variance of innovations to the short rate depend on its current level given the Cox, Ingersoll, Ross parametrization.

Estimation. We estimate parameters for the Cox, Ingersoll, Ross process using monthly data on the 2-year Treasury yield from January 1987 to July 2024 using maximum likelihood. Our estimation routine adapts the Matlab code from Kladívko (2007).

Nominal rental cost growth comes from the CPI-U's rent of primary residence series, and nominal house price growth comes from the FHFA's all-transactions annual house price index. We choose these indices over alternatives because they have the longest time series. We use data from 1976-2019, and deflate nominal values to 2019 dollars using the CPI-U.

To estimate the non-CIR coefficients in (E.1), we first estimate ω^r_{t+1} from the 2-year nominal Treasury series and estimated CIR coefficients. We then estimate $\mu_\ell, \mu_P, \rho_\ell, \rho_P$ using seemingly unrelated regressions, and form estimates of $\boldsymbol{\omega}_{t+1}$. To estimate the non-zero coefficients in $\boldsymbol{\Sigma}(r)$, where the ijth coefficient is σ_{ij}, we use a lower triangular Cholesky decomposition of the estimated covariance matrix of $\boldsymbol{\omega}$. Estimates are presented in equation (E.2) below.

$$\begin{pmatrix} r^s_{t+1} \\ \log \frac{l_{t+1}}{l_t} \\ \log \frac{P_{t+1}}{P_t} \end{pmatrix} = \underbrace{\begin{pmatrix} 0.0041 \\ 0.015 \\ 0.066 \end{pmatrix}}_{\equiv \mu} + \underbrace{\begin{pmatrix} 0.90 \\ -0.15 \\ -0.79 \end{pmatrix}}_{\equiv \rho} r_t + \underbrace{\begin{pmatrix} 0.077\sqrt{r_t} & 0 & 0 \\ -0.0095 & 0.024 & 0 \\ 0.033 & 0.028 & 0.056 \end{pmatrix}}_{\equiv \Sigma(r_t)} \begin{pmatrix} \omega^r_{t+1} \\ \omega^\ell_{t+1} \\ \omega^P_{t+1} \end{pmatrix} \tag{E.2}$$

E.2 Offline calibration.

This subsection provides additional details on calibrations for a subset of parameters described in Table E.29.

Per-period owner-occupied housing costs. Per-period owner-occupied housing costs include maintenance, insurance, and property taxes. These are modeled as ad valorem costs. For maintenance costs, Fannie Mae (2025) recommends budgeting 1-4% of the property value each year for maintenance expenses; we select 2% as a benchmark. For insurance costs, we use the 2019 average premium cited in III (2025), which is 0.3% of the 2019 average home value. For property taxes, we use the average 2019 tax rate of 1.14%. This gives a total flow cost of 3.44%.

Mortgage spread. We calculate the monthly spread between the average 30-year fixed mortgage rate reported in the Freddie Mac Primary Mortgage Market Survey and the 30-year Treasury yield, and take the average over 2005-2019. This gives an average spread of 1.05%.

Household state distribution. We calibrate the household state distribution using the 2013, 2016, 2019, and 2022 Survey of Consumer Finances (SCF), which has information on mortgages and other

forms of financial wealth.

We primarily use the Summary Extract dataset, supplemented with some items from the Full Public dataset. We deflate variables to 2019 dollars and use household-level survey weights throughout.

We restrict to households aged 22-82, to align with 30 time steps in our model. Following Laibson *et al.* (2024), we further restrict to households with meaningful labor force attachment by dropping all households with annual income less than $1,000 \times \$7.25$, equivalent to a half-time salary at the federal minimum wage.

To calculate financial wealth, we start from the `fin` (Total Financial Assets) variable. We subtract unsecured liabilities: credit card debt (`ccbal`), student loans (`edn_inst`), and other unsecured installment debt (`oth_inst`). We further subtract components current or future pensions / retirement savings included in `fin` that respondents say cannot be withdrawn as a lump sum, which requires pulling account-level responses from the Full Public dataset. These components are included in the Social Security and retirement income variable (`ssretinc`) used to calibrate the income process below, and so exluding them avoids double-counting pension resources.

To calculate mortgage characteristics, we first strip out differences in mortgages due to differences in the quality of homes purchased. We calculate current LTVs as the ratio between mortgage debt (excluding HELOCs) and estimated home value. We then compute the normalized mortgage balance as the household LTV multiplied by the median home price of homes sold in the survey year. We further compute normalized mortgage payments as first mortgage payments, as a share of mortgage debt, multiplied by the normalized mortgage balance.[64]

Finally, we compute the state distribution in each survey year on a discretized grid. We use 15 gridpoints for financial wealth, 15 gridpoints for mortgage balances, and 10 gridpoints for mortgage payments. For each survey, we calculate the fraction of households in each ownership status $\times$ age $\times$ financial wealth $\times$ mortgage balance $\times$ mortgage payment bin.

Income by age. We calculate the income profile by age using the sample described above for survey years 2013-2019. We define income as the sum of wages (`wageinc`), business, sole proprietorship, and farm income (`bussefarminc`), social security and other retirement income (`ssretinc`), and income from unemployment, alimony/child support, TANF/food stamps/SSI, and other income sources (`transfothinc`). We regress log income on age dummies and calendar year fixed effects. The income profile is calibrated as the estimated age dummies plus the estimated effect for calendar year 2016 – chosen to match the focal year of the moments used for minimum distance estimation. As the SCF asks about pre-tax income, we apply a uniform 25% income tax to approximate federal, state, and local taxes.

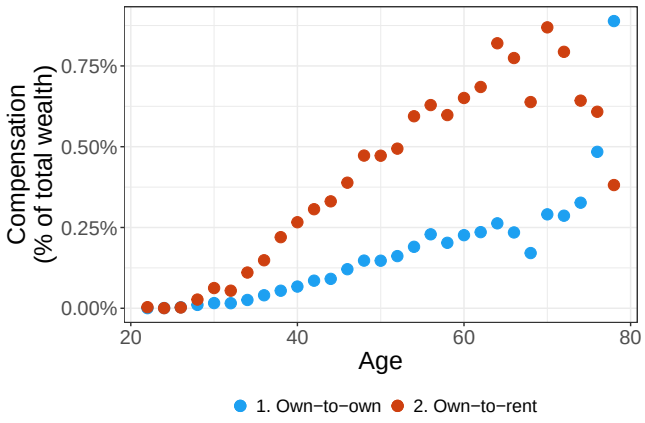
LTV by age. We estimate the age profile of origination LTV using the same SCF dataset described above for survey years 2013-2019. We restrict to owners who moved in the last year. We calculate the age profile of household LTV relative to the average LTV by regressing the LTV, demeaned on the sample average, on a fifth-degree polynomial in household age. We take fitted values from that regression for each age, and add to the mean sample LTV of 20%.

[64] This operation preserves the amortization schedule of the original mortgage, and hence preserves interest rates.

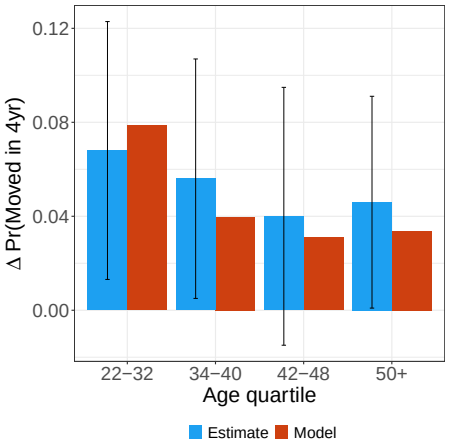
E.3 Additional figures and tables.

Figure E.23: Model mobility effects by age.

(a) Compensation required for forced moves.

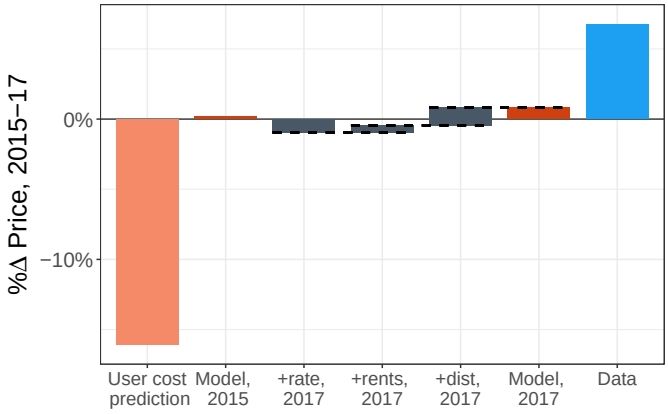


(b) Model predicts heterogeneous impact of rate lock by age.



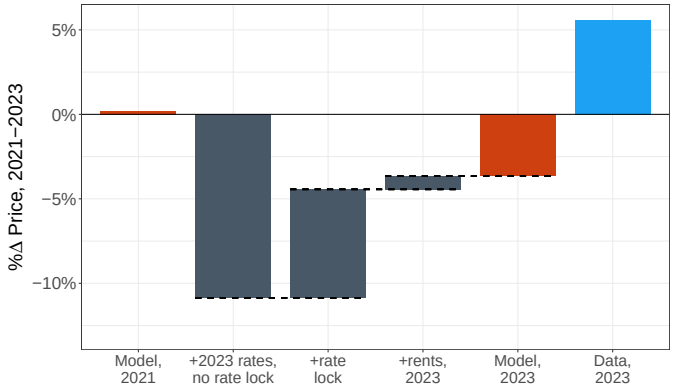
Source: Author's calculations. (a) The figure plots compensation required to make an owner with age given on the x-axis indifferent between staying in their current unit and being forced to move to a different unit. Compensation is a percentage of total wealth (the sum of financial wealth, home equity, and the present-value of future income). Blue dots show compensation required for a forced own-to-own move, and red dots show compensation required for a forced own-to-rent move. (b) The blue bars plot estimates and 95% confidence intervals of the impact of rate lock on subsequent exiting-home sales from the November 2016 event study presented in Table 4. The red bars show predictions of the event study estimates based on simulations of estimated model.

Figure E.24: Model-predicted vs. actual house price growth from 2015-2016 to 2017-2018



Source:Author's calculations. The grey bars show how changing rents, 2-year Treasury rates, and the household state distribution one at a time from 2015-2016 to 2017-2018 levels impacts price growth. The light red bar shows calibrations based on parameters from Himmelberg *et al.* (2005), described in the main text. The dark red bar shows the model-implied total change in price growth. The blue bar plots the empirical change in the real price of the average house sold.

Figure E.25: Modeled impact of 2021-23 tightening, with rent increases.



Source:Author's calculations. The figure reproduces Figure 7b, but adds the impact of increased real rents over 2021-23.

Table E.29: Offline calibrated and estimated parameters

Symbol	Description	Value	Source
δ	Discount factor	0.90	Calibrated
γ	Coef. of relative risk aversion	2	
(A_{min}, A_{max})	Min / max age	$(22, 82)$	
(φ_1, φ_2)	Bequest motive	$(\hat{\alpha}, 0)$	
σ	Nesting parameter	0.5	
τ_s	Transaction cost	0.03	Industry sources:
τ_{ft}	Per-period owner-occupied costs	$0.0344 \cdot P_t$	Fannie Mae (2025), Attom (2020), III (2025)
Δ	Mortgage spread	0.0105	Calibrated / SCF
$\{\phi_a\}$	Origination LTV		
$\{y_a\}$	Income by age		SCF (2013-2019)
G_{kt}, N_{kt}	Household state distribution		SCF (2013-2022)
$\boldsymbol{\mu}, \boldsymbol{\rho}, \boldsymbol{\Sigma}(r)$	Agg. state transitions		See Section E.1

Source: Author's calculations. The table shows parameters calibrated and estimated offline. See Section 5.3 for details.

Table E.30: Minimum distance estimates

Symbol	Description	Estimate	Identifying empirical moment	Empirical moment	Modeled moment
h_o	Flow util of owner-occ housing	5.7	Pr(Moved to own in 4 yr), 2016 (overall)	0.075	0.075
λ_0	Net util benefit of move	-1.8	Pr(Moved to rent in 4 yr), 2016	0.050	0.050
α	Non-housing consumption util coef	10.9	$\frac{\uparrow \Pr(\text{Moved in 4 yr})}{\Pr(\text{Moved in 4 yr})}$, 2016 event study	0.239	0.239
λ_s	Net util benefit of tenure change	11.7	$\frac{\uparrow \Pr(\text{Moved to rent 4 yr})}{\Pr(\text{Moved in 4 yr})}$, 2016 event study	0.079	0.079
χ	Prob of considering move	0.39	Variance in Pr(Move), 2011-2019	0.012	0.012

Source: Author's calculations. The table shows parameters estimated with the minimum distance procedure described in Section 5.3. The first two columns describe parameters to be estimated. The third column present minimum distance estimates. The fourth and fifth column describes the empirical moment primarily used to identify each parameter. The sixth column shows the model-generated moment value