Household Mobility and Mortgage Rate Lock

Online Appendix

Jack Liebersohn & Jesse Rothstein

November 12, 2024

A   Detailed Data Construction

A.1   Main Sample

Our main analysis sample is a panel of households with mortgages. To track mobility for these households, we first identify unique mortgages, defined as mortgages originated by a particular borrower for a specific principal amount on a specific date. We drop mortgages which are duplicated in the credit records and match these mortgages to the panel of households.

We identify the mortgage location as the household's zip code two quarters after origination, and measure moves as households that change ZIP codes (or states) thereafter. We identify new purchase mortgages as those where the household did not live in the mortgage location ZIP code prior to mortgage origination, and refinance mortgages as those where the household did.

For each mortgage, we track the ZIP code of the household for the subsequent forty quarters, even if they have paid off the original mortgage. We also include information on the household's total number of mortgages, debt, and data as of the origination date, such as the loan origination amount. Using data 40 quarters after mortgage origination ensures that we continue to track households who refinance their mortgage or prepay for other reasons. Including these individuals is important for our empirical strategy. The decision to prepay is an endogenous result of ex post mortgage rates, so excluding households that refinance would lead to sample attrition that is correlated with the outcomes of interest.

We count loans taken out to refinance an earlier mortgage as new mortgages. Each new mortgage begins a new spell. This means that a household can have several overlapping spells - one beginning when it originally purchases the home and others beginning each time it refinances. We select one at random for each household to ensure that observations are not dependent and that our sample appropriately represents purchase loans and refinances.

A.2 Credit Registry Variables

We measure moving using an indicator variable if an individual changes ZIP codes. The UC-CCP data also contains census tract and block information for most households starting in 2010. Identifying census tract moves would be an alternative way to measure mobility but we prefer to use ZIP codes because it is available for all households. Also, mobility measurement is more difficult with census block information because census block codes change over time to reflect changing census definitions. About one-third of moves across census tracts do not result in a change in ZIP code, so we will miss these moves. However, aggregate patterns of ZIP code and census block moves are very similar, so we think our findings are likely to generalize to other levels of geography.

To distinguish between renters and homeowners without a mortgage, we rely on a UC-CCP field that identifies known homeowners from public records data. Of people who do not have a mortgage at the beginning of a spell in a ZIP code, we label those who are identified at that point as homeowners as cash buyers, and those who are never identified as homeowners as renters. (Spells that start as non-homeowners but transition within the spell to be homeowners are excluded.)

A.3 Cash Buyer Sample

In some of our analyses, we use a sample of cash buyers - homeowners without mortgages - as a control group. This includes households that are indicated as homeowners in the UC-CCP data, based on public records data, but who do not have mortgages at any time between their arrival into a ZIP code and their departure from that ZIP code. In our primary analyses of this sample, we include only cash buyers who arrive in the ZIP code in 2013 or later, though in a robustness check we extend the sample to include all post-2004 arrivals. (The UC-CCP begins in 2004, so we cannot identify arrival dates prior to that year.)

As discussed in the text, we reweight the cash buyer sample to resemble the mortgage

holder sample on observables. Specifically, we stack our sample of cash buyer and mortgage holder quarterly observations and estimate a logit model where the dependent variable is an indicator for being a mortgage holder. Explanatory variables are origination quarter, time and ZIP code group fixed effects, and householder age, credit score, and outstanding debt. We obtain very similar results when using either a subset of the predictors or when using interaction terms as well. We construct fitted probabilities from this model $\hat{p}$, and reweight the cash buyer sample by $\hat{p}/(1-\hat{p})$ before constructing our cell mean mobility hazards.

B   Additional Tables and Figures

In this appendix we present several additional results. Table B.3 presents the first-stage relationship between the interest rate gap calculated based on the origination rate, g^*, and the gap calculated from the rate that applies to the most recent refinance, g. We implement this as an OLS regression applied to the panel of mortgage-by-quarter observations, with mortgages excluded after the borrower leaves the ZIP code, and we cluster standard errors at the mortgage level. The table also shows a "reduced form" model that uses g^* directly in (5) in place of g.

Figure B.1: Actual and Counterfactual Quarterly ZIP Code Moving Hazards
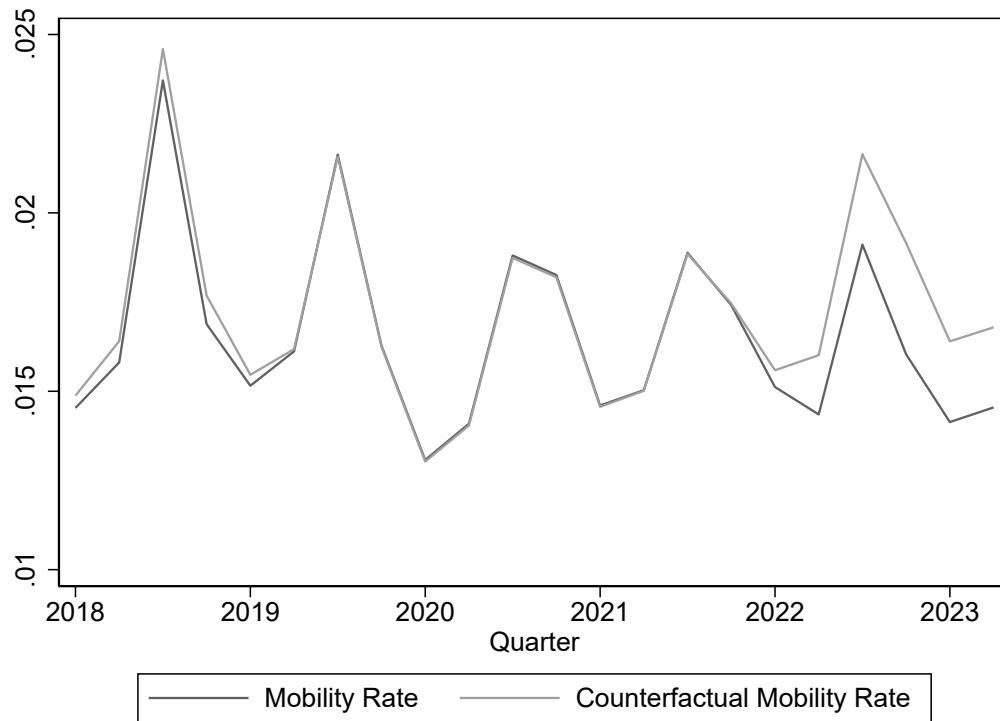
Notes: Counterfactual moving rates are estimated by using estimates from our preferred OLS model (Table 2 row 1, column 3) and setting the counterfactual rates to 0 when there is a positive rate gap as discussed in section 5.3.

Figure B.2: Kaplan-Meier survival curves, by period

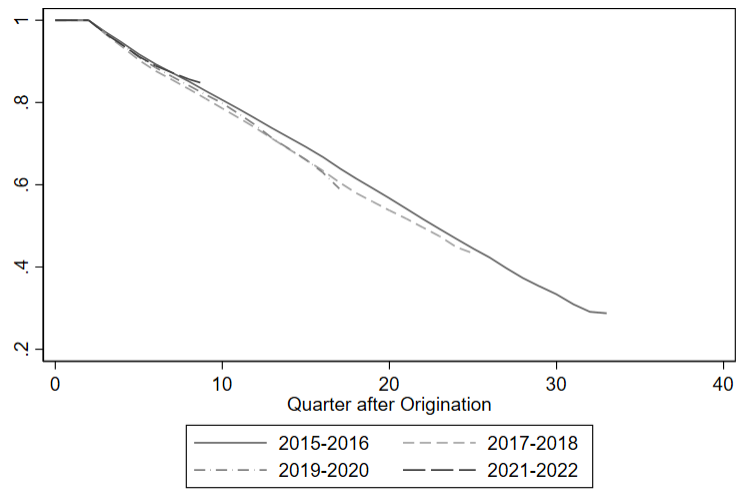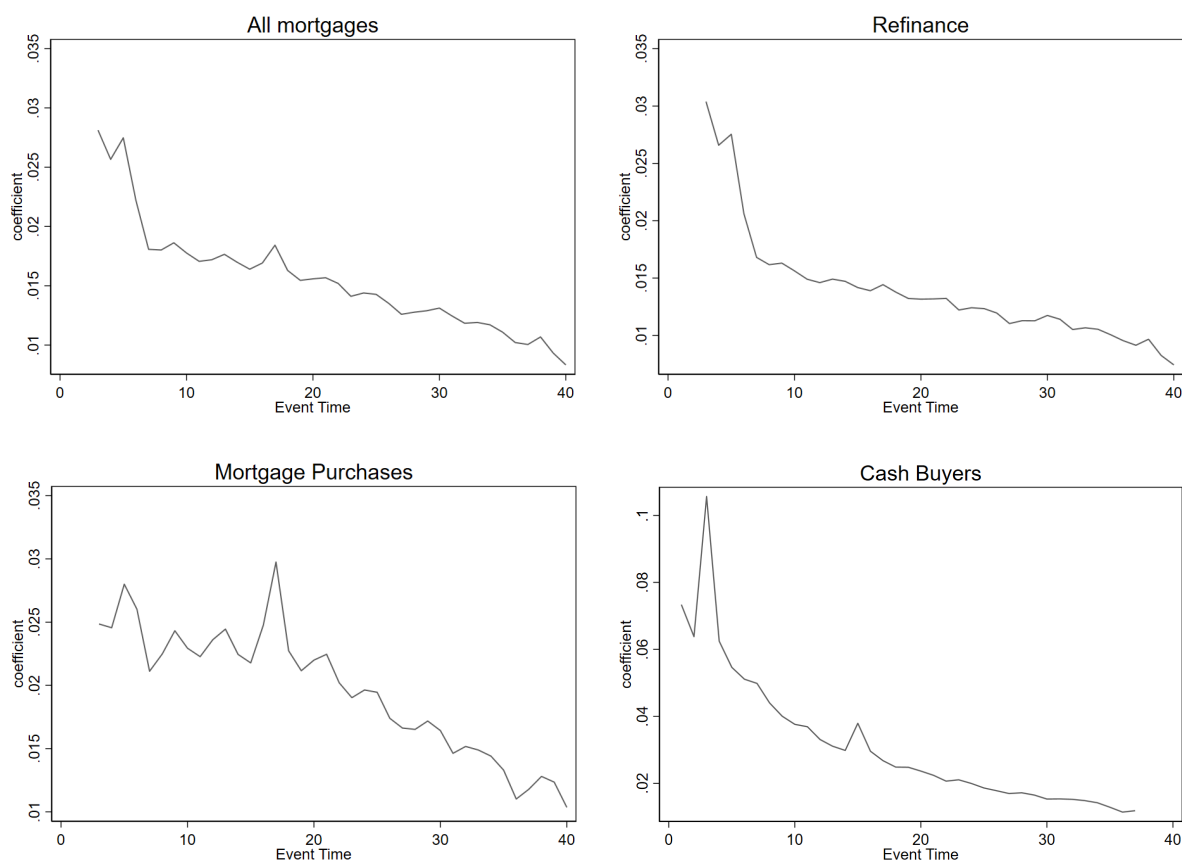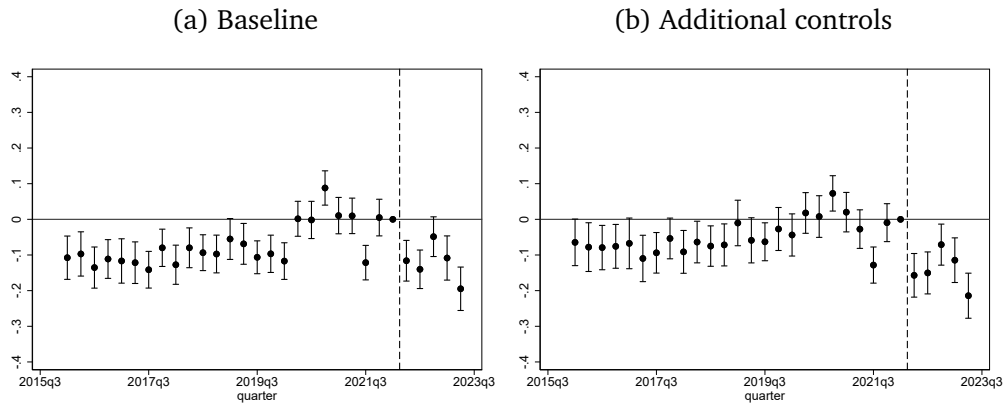Figure B.3: Baseline Hazards

Notes: Figures show estimates of baseline hazards from estimates of equation (5), with calendar quarter indicators as the only controls. Calendar quarter coefficients are reported in the corresponding panels of Figure 3.

Figure B.4: Difference-in-Hazard Event Study Figures

(a) Baseline

(b) Additional controls

Source: University of California (UC-CCP)

Notes: Figures show estimates of baseline hazards from estimates of equation (5) on a sample including mortgagers and cash buyers. Specification for the "Baseline" panel shows quarter-by-cash buyer fixed effects while controlling for quarter fixed effects. Specification for "Additional Controls" panel adds controls for a cubic polynomial in house price growth interacted with a positive house price growth variable and a sample indicator, and origination quarter-by-current quarter fixed effects. Standard errors clustered by origination quarter and sample.

Figure B.5: Rate gap histograms

A. Full sample

B. Pre-2022 subsample
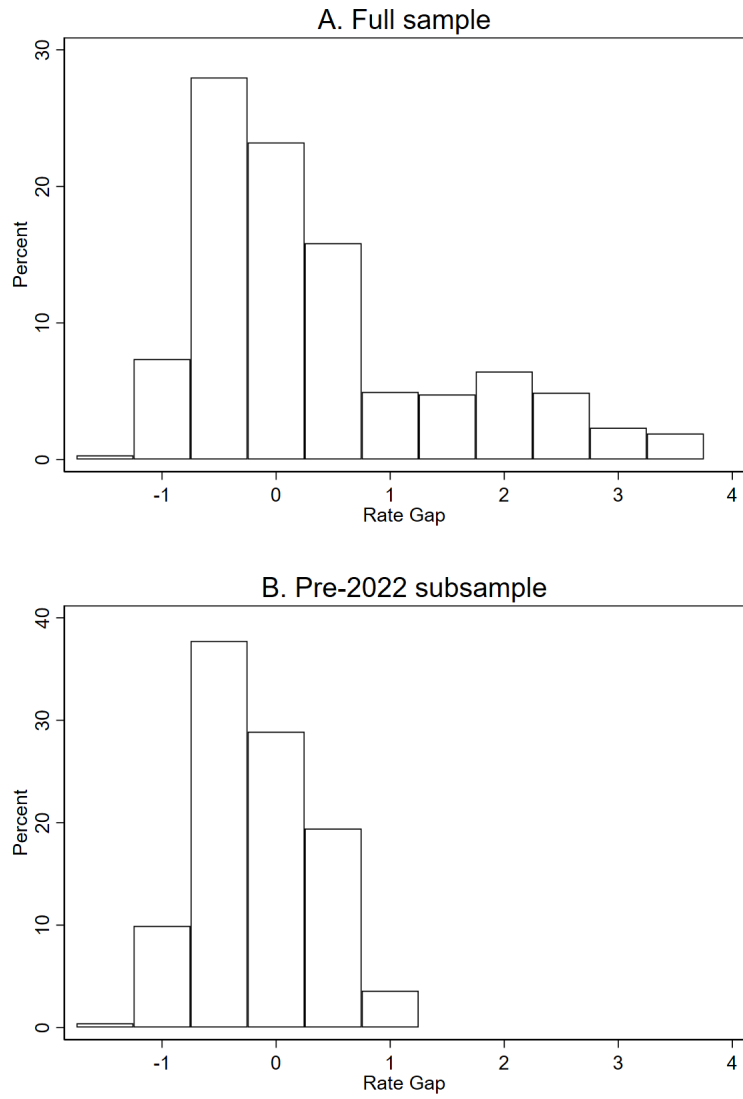
Source: University of California (UC-CCP)

Notes: Figures show histograms of the imputed rate gap for the full sample and for the period prior to quarter 2, 2022.

Figure B.6: Comparing pre-2022 estimates to full sample

A. Full sample

B. Pre-2022 subsample
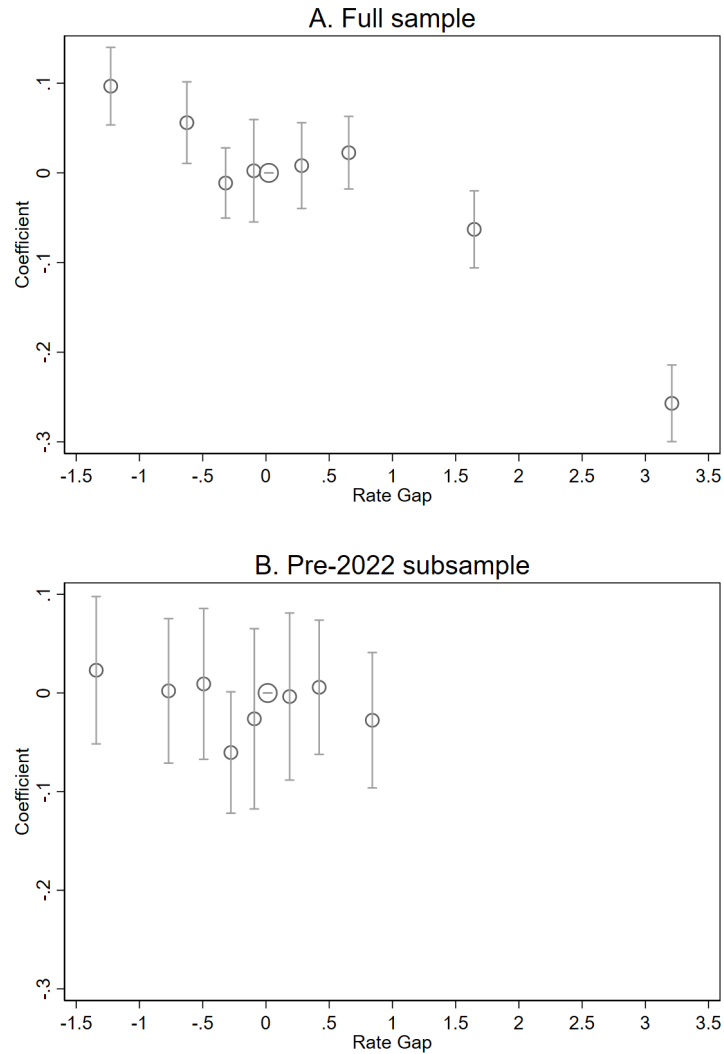
Source: University of California (UC-CCP)

Notes: Figure shows coefficients on indicators for 10 deciles of the rate gap, including both positive and negative values. Specification otherwise matches Table 2, column 3. The excluded category is the bin that includes zero rate gap; due to a substantial share of observations with exactly zero gap, this bin accounts for 21% of the sample. Spikes show 95% confidence intervals. Panel B restricts the sample to the pre-2022 observations.

Table B.1: Summary statistics: Cash buyers

	Mean	Median	Standard deviation	Min	Max	N
Year	2019.327	2020.00	2.789	2013.00	2023.00	6298990
Cash Buyer	1.000	1.00	0.000	1.00	1.00	6298990
Rate Gap (positive)	1.113	0.69	1.063	0.00	4.01	2668808
Rate Gap (negative)	-0.593	-0.49	0.455	-1.95	-0.00	2974564
Rate Gap (unconditional)	0.191	0.00	1.109	-1.95	4.01	6298990
Log ZHVI Change	0.175	0.11	0.193	-2.11	3.51	5677353
Log ZHVI change, conditional on positive [exclude 0/negative]	0.202	0.14	0.193	0.00	3.51	4927981
Household still in ZIP	0.945	1.00	0.228	0.00	1.00	6298990
Household Moved this Quarter	0.055	0.00	0.228	0.00	1.00	6298990

Note: Cash buyers are defined as homeowners that do not have a mortgage at the time that they move into a new ZIP code.

Table B.2: Estimates identified by cohort instead of time variation

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Main sample	-0.077	-0.023	-0.011	-0.079	-0.031	-0.020
	(0.005)	(0.016)	(0.020)	(0.005)	(0.016)	(0.023)
Controls						
ZHVI & negative rate gap	X	X	X	X	X	X
Origination time controls						
Origination quarter FE	X			X		
Origination year FE			X			X
Calendar time controls						
Season FE	X			X		
Linear time	X			X		
Quarter FE		X	X		X	X

Notes: This table presents estimates of Specification 5 similar to the first row of Table 2. Columns 1 and 4 here repeat the estimates in columns 3 and 6 of Table 2. Columns 2 and 5 replace the fixed effects for the quarter of mortgage origination with fixed effects for the quarter in which the outcome is measured. Columns 3 and 6 add back fixed effects for the year of origination, leaving only within-year variation in origination rates. Standard errors are clustered at the origination quarter by zip home price index group level.

Table B.3: First stage and reduced form specifications for IV model

	First stage			Reduced form		
	(1)	(2)	(3)	(4)	(5)	(6)
Main sample	0.898	0.881	0.860	-0.046	-0.068	-0.066
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Mortgage purchases	0.902	0.885	0.857	-0.023	-0.016	-0.023
	(0.003)	(0.003)	(0.003)	(0.005)	(0.004)	(0.005)
Refinances	0.896	0.878	0.862	-0.058	-0.093	-0.085
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.006)
Controls						
Season FE	X	X	X	X	X	X
Linear Time Controls		X	X		X	X
ZHVI & negative rate gap			X			X

Notes: This table presents first stage and reduced form estimates from the IV model. Standard errors are clustered at the origination quarter by zip home price index group level.

Table B.4: Specifications allowing for nonlinear effects of rate gaps

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rate gap	-0.077	-0.089	-0.015	-0.077	-0.058	-0.078	0.014	0.019
	(0.005)	(0.009)	(0.011)	(0.005)	(0.015)	(0.006)	(0.017)	(0.018)
Rate gap squared		0.001			-0.007			
		(0.003)			(0.004)			
Rate gap * (rate gap > 1)			-0.091				-0.098	-0.101
			(0.011)				(0.012)	(0.012)
Rate gap * (rate gap > 2)			0.019				0.010	0.009
			(0.006)				(0.006)	(0.007)
Negative rate gap				-0.029	-0.036	-0.022	-0.042	-0.051
				(0.008)	(0.010)	(0.010)	(0.010)	(0.013)
Negative rate gap × (rate gap < -1)						-0.009		0.011
						(0.008)		(0.008)

Notes: Rate gap is the difference between the current market rate at time t and the market rate at the time the mortgage was last refinanced, g_{it}, and is measured in percentage points. Other than rate gap specification, controls are as in Table 2, column 3, and models are estimated by OLS. Standard errors are clustered at the origination quarter by zip home price index group level.

Table B.5: Robustness

	OLS	IV
	(1)	(2)
Preferred results	-0.077	-0.079
	(0.005)	(0.005)
Alternative specifications		
Alternative clustering	-0.077	-0.079
	(0.013)	(0.012)
Poisson	-0.066	
	(0.005)	
Log hazard	-0.076	-0.078
	(0.005)	(0.005)
Hazard	-0.0012	-0.0012
	(0.0001)	(0.0001)
No entry-quarter FEs	-0.067	-0.078
	(0.005)	(0.005)
Add year-quarter FEs	-0.023	-0.031
	(0.016)	(0.016)
Alternative samples		
Pre-COVID originations	-0.081	-0.082
	(0.006)	(0.005)
Sample starts in 2015	-0.061	-0.065
	(0.005)	(0.005)
Sample ends in 2021	-0.008	-0.003
	(0.016)	(0.015)
Alternative duration measures		
Refinance spells start at move in	-0.062	-0.064
	(0.007)	(0.007)
Cash buyer sample includes all post-2004 entrants	-0.019	
	(0.004)	

Notes: This table presents OLS and IV estimates for a variety of specifications and subsamples. All specifications include controls from Table 2 column 3, unless otherwise noted. "No entry-quarter FEs" includes controls from Table 2 column 2 excluding entry-quarter FEs. "Add year-quarter FEs" include controls from Table 2 column 2 excluding entry-quarter FEs but including current quarter FEs. "Refinance spells start at move in" includes both new originations and refinances, but measures duration for both as time since moving to the ZIP code. "Cash buyer sample includes post-2004 entrants" extends the cash buyer sample to include all who moved to the ZIP after 2004 (rather than 2013 as in the main results) and does not use propensity weights. Standard errors are clustered at the origination quarter by zip home price index group level unless otherwise noted.