

Online Appendices to Mortgage Prepayment and Path-Dependent Effects of Monetary Policy

David Berger, Konstantin Milbradt, Fabrice Tourre and Joseph Vavra

A.1 Data Appendix

A.1.1 Baseline McDash Prepayment Sample

This section describes in more detail our loan-level sample restrictions as well as our identification of prepayment type in CRISM data. See the supplementary replication code for more details (Berger et al. (2021)). Our primary loan-level data set for measuring prepayment and loan gaps is the McDash loan performance and origination data from 1992-2017 produced by Black Knight Financial Services. The origination data provides a number of origination characteristics such as origination date, amount, loan purpose and appraisal value while the performance data provides dynamic info on these loans like current unpaid balance, current interest rate and flags for prepayment. Our prepayment analysis primarily requires information from the dynamic loan performance data: we define loan prepayment in month t as any loan with termination flag "voluntary payoff" in that month and a termination date of month t .

Our primary analysis uses only fixed rate first mortgages, but results are similar when including all mortgages in the McDash data set. In order to maintain a consistent sample when running cross-MSA results, we also drop any loan with missing information on MSA-division. We also drop any loan with missing information on the current interest rate in the McDash loan performance data set, since we cannot measure gaps for these loans. We define the interest rate gap as the current interest rate minus the monthly average 30 year FRM from the Freddie Mac PMMS plus an estimated loan-specific adjustment that is a quadratic function of the borrower's FICO score and the loan's current LTV ratio. We also compute \$ gaps in addition to rate gaps, which we define as the current unpaid balance times the interest rate gap. For that analysis, we drop any loan with missing unpaid balance in the McDash performance data set.

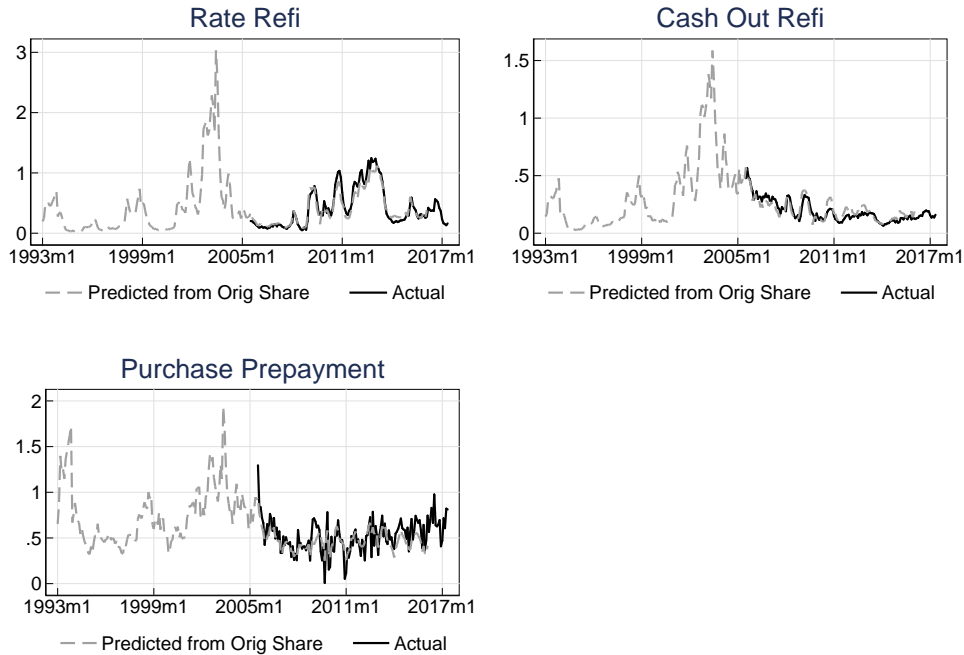
Since most of our analysis is focused on prepayment rates, the majority of our analysis can be performed using only loan performance data, albeit without controls for individual fixed effects or time-varying observables only available in credit records. However, while performance data is required to measure prepayment, it cannot be used to decompose prepayment into that arising from refinancing and moves. This is because loan purpose is collected at origination but not at termination, so the performance data set tells us if a loan prepays but not why. Conversely, origination data can be used to measure the share of *new* originations which are due to refinancing and moves, but it cannot be used to measure the share of *old* mortgages which are repaying. So origination data cannot tell us what share of mortgages prepay and why, since it contains the wrong denominator. This means that measuring the frequency of prepayment by type requires linking information on loan performance for terminating loans with loan origination information for newly originating loans.

After 2005, we are able to use the linked Equifax/CRISM data which we describe in the next subsection to precisely link each individual repaying loan to a newly originating loan so that we can measure exactly why each individual loan is repaying. Prior to 2005, these links are unavailable, so we cannot measure the reason that any individual loan prepays. However, in a stationary environment, performance data and origination data can be combined to proxy for the *share* of prepayment arising from different types. In particular, in an environment with no net flows in and out of the mortgage market, every loan which prepays due to refinancing or due to moving must be matched by a newly originated loan with the same purpose. While we cannot link the new and old loan together, the shares must remain unchanged. This allows us to proxy for rate-refi, cash-out refi and movement frequencies in a month using only loan level data without links to individuals.

In particular, we compute $freq_t^{type} = freq_t^{prepay} \times share_t^{type}$ where $freq_t^{type}$ is the frequency of a given type of prepayment, $freq_t^{prepay}$ is the frequency of prepayment in performance data and $share_t^{type}$ is the share of a given type of loan purpose in origination data. The McDash origination data set only collects loan purpose after 1998, so we measure $share_t^{type}$ using originations data from CoreLogic LLMA. This data set has a structure nearly identical to the McDash data, but it contains reliable loan purpose at origination info as early as 1993. However, we continue to measure $freq_t^{prepay}$ using McDash data

because CoreLogic performance data does not measure prepayment before 1999, and has roughly half the market coverage of the McDash Performance Data set. Combining prepayment frequencies from McDash with loan purposes shares at origination from CoreLogic leverages the comparative advantages of the two data sets. While stationarity is clearly a strong assumption, we can use the CRISM data after 2005 to compare our proxies for frequency by type under the stationarity assumption, with actual frequencies. Figure A-1 shows they are very similar.

Figure A-1: Construction of Prepay Shares by Type



Overall, the McDash Performance data set contains information on approximately 180 million loans. After 2005, the McDash Performance data set covers roughly 50% of total U.S. mortgage debt as measured by the Federal Reserve. Prior to 2005, coverage is somewhat lower, ranging from around 10% market coverage in the early 90s to 20-25% in the late 90s. As a measure of representativeness and external validity, Figure A-2 shows that refinancing in our data closely tracks the refinancing applications index produced by the Mortgage Banker’s Association from 1992-2017.⁷² This suggests that despite the changing sample sizes, the McDash Performance data is broadly representative of the U.S. mortgage market over the entire 1992-2017 period.⁷³

We use the McDash Performance data rather than CoreLogic Performance data because the CoreLogic performance data does not measure prepayment before 1999, and has roughly half the market coverage of the McDash performance data. Using CoreLogic instead of McDash Performance data leads to somewhat noisier refinancing series and eliminates the first 7 years of data. In addition, we cannot link loans to individuals in CoreLogic data like we can in the McDash data using the links we describe

⁷²Note that we measure originations while this index measures applications. According to LendingTree, denials are roughly 8% after the financial crisis due to Dodd-Frank related changes in lending standards. This explains the level difference after the Financial Crisis but the series continue to highly comove.

⁷³Note that even in the months with the fewest observations, we still have more than 5 million mortgages, so only lack of representativeness and not sampling error is a potential concern.

next. Nevertheless, we have repeated our analysis of prepayment using series derived from CoreLogic LLMA Performance data and arrive at similar conclusions.

Figure A-2: Comparison of Refi Measured with McDash Data to Mortgage Bankers’ Association data

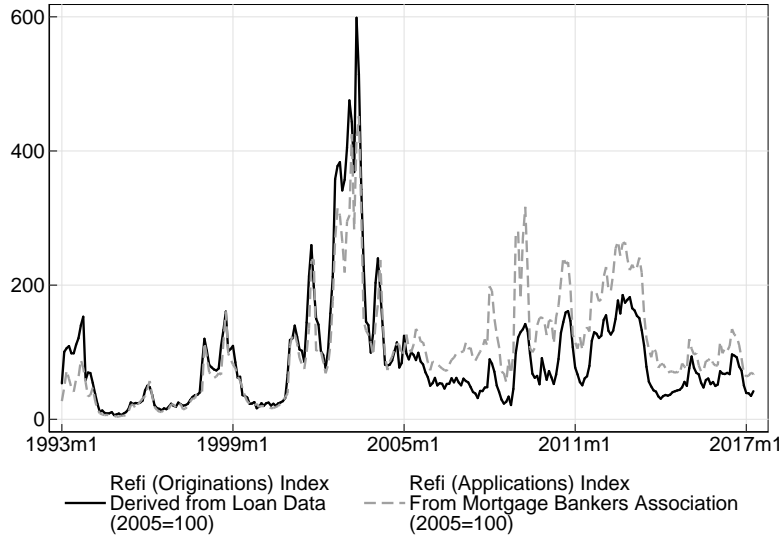


Figure shows an index of refinancing computed using McDash Performance data and CoreLogic Origination Purpose data compared to the Mortgage Banker’s Association Refinancing Application Index. Note that the loan-level index measures originations while the MBA index measures applications. Indices are normalized to 100 in 2005m4.

A.1.2 Linked CRISM Sample: Measuring Refinancing

After 2005, we link loans in the McDash data set to Equifax credit records, which allows us to decompose prepayment into different types and to control for various individual level observables, including individual fixed effects. Our analysis and description of this data closely follows [Beraja et al. \(2019\)](#). The linked Equifax/McDash CRISM data set provides the linked Equifax credit records for each McDash mortgage for the lifetime of the loan, including an additional 6 months before origination and after termination. This link is done directly by Equifax. Credit records provide a consumer’s total outstanding debt amounts in different categories (first-lien mortgages, second-lien mortgages, home equity lines of credit [HELOCs], auto loans, etc.). Additionally, in any month, Equifax provides the origination date, amount, and remaining principal balance of the two largest (in balance terms) first mortgages, closed-end seconds, and HELOCs outstanding for a given consumer.

In order to reduce the computational burden, we begin the analysis of CRISM data by extracting all the loan and individual characteristics from a random 10% sample of all individuals in the Equifax/McDash data at some point between 1992 and 2017. This 10% CRISM sample includes all mortgage loans (approximately 11 million) for approximately 5.9 million individuals⁷⁴ We further restrict our CRISM sample to those consumers who start our sample with two or fewer loans in each category and never have more than three of any of these types of loans outstanding.⁷⁵ These sample restrictions leave

⁷⁴Results are extremely similar when using 5% and 20% samples since the CRISM sample is very large, so sampling error is not important.

⁷⁵Equifax only provides detailed origination information on the two largest loans, so this restriction allows us to infer the origination month, origination balance, and balance of the third largest loan of any loan type even though this information does not appear explicitly in Equifax. We also drop loans that do not have complete consecutive Equifax records.

roughly 96% of the 5.9 million individuals in our analysis sample. In creating this loan-level data set, we assume that the month in which the loan stops appearing in Equifax is the month that it was terminated.

While McDash loans are linked to individual credit records directly by Equifax using social security number, individuals can have multiple loans and the loan information in Equifax does not always exactly match that in McDash, since they come from independent sources.⁷⁶ We thus have to construct a unique match between a loan in McDash and the possible set of linked loans in Equifax. As in [Beraja et al. \(2019\)](#), we consider an Equifax loan/McDash loan pairing a match if the origination date of the Equifax loan is within 1 month and the origination amount is within \$10,000 of the McDash loan. If more than one loan is matched, we use the origination amount, date, termination date, zip code (where available, or 3-digit zip code and MSA-div where not)⁷⁷, and termination balance as tiebreakers. We are able to match roughly 93% of McDash loans to an Equifax loan using these restrictions.

As in our primary analysis, we begin with all remaining outstanding fixed rate first liens in the McDash which are voluntarily paid off. We then look for any loan in the Equifax data set that has an open date within 4 months of the McDash loan's termination date. We classify these new loans as a refinance if either:

- The loan also appears in McDash and is tagged as a refinance in the purpose-type variable.
- The loan also appears in McDash and is tagged as an "Unknown" or "Other" purpose type, and has the same property 5 digit zip code (where available, or 3-digit zip code and MSA-div where not) as the original loan.
- The loan appears only in Equifax but the borrower's Equifax address does not change in the 6 months following the termination of the original loan.

This allows us to compute one of our primary outcomes of interest, the count of first-lien FRM loans which refinance in month t divided by the total number of McDash first-lien FRM loans with Performance data in that month. (We have also considered results which compute balance weighted shares, and they are very similar).

A.1.3 Linked CRISM Sample: Decomposing refinancing into rate and cash-out

To compute the cash-out and rate-refinancing share of loans, we must further break these refinancing loans down by type. In particular, we need to compute how the balance of the new loan compares to the outstanding balance of the loan(s) being prepaid. We begin by labeling any loan in the Equifax data set that terminates between -1 and 4 months from a new McDash loan's close date a "linked" loan, including first mortgages as well as closed-end seconds and HELOCs, and we call the new loan a refinance if:

- The loan is a known refinance in McDash.
- The loan has an "Unknown" or "Other" purpose type in McDash and a linked loan in McDash that has a matching property zip code (5 digit when available or 3-digit + MSA-div when not).
- The loan has an "Unknown" or "Other" purpose type in McDash and a linked loan that appears only in Equifax, but the consumer's Equifax address does not change in the 6 months after the new loan was opened.

If there is more than one linked loan that is a first mortgage in Equifax, we link only the loan that is closest in balance to the origination amount of the new mortgage. We only link those Equifax loans

⁷⁶For example, balances may differ slightly since they may be reported to credit bureaus and servicers at different dates.

⁷⁷To ensure anonymity, McDash reports 5-digit zip code for loans in higher volume locations and 3-digit zip code for loans in lower volume locations.

that exist in the Equifax data for at least three months to prevent the refinanced loan balance from being counted in the old balance of the loan.

For each of these cases, we can then calculate the cash-out amount as the difference between the origination amount on the refinance loan and the balance of the linked loan(s) at termination. In order to capture the correct origination amount on the refinance loan, we want to ensure that we are also including any "piggyback" second liens that are opened with the refinance loan that we find in McDash. Thus, we look for any loan in the Equifax record linked to our refinance loan that has an Equifax open date within three months of our refinance loan and an origination balance of less than 25% of our loan's origination balance if labeled a first mortgage and less than 125% of the refinance loan's origination balance if labeled a HELOC or CES, and add the balance of these piggyback seconds to the refi origination amount when calculating cash-out amounts.⁷⁸ To eliminate outliers, we also drop cash-out and "cash-in" amounts that are greater than \$1,000,000. These amount to dropping less than 0.05% of the refinance loans.

After measuring the change in the balance, we then call a refinancing a cash-out if, after subtracting 2 percent from the new loan to cover closing costs, the new mortgage balance is at least \$5,000 above the old mortgage. Using a more restrictive definition of cash-out reduces the overall share of cash-out and the sensitivity of cash-out to rate gaps while using a less conservative cutoff does the reverse since it reclassifies some rate-refis as cash-outs. By construction, these choices have no effect on the decomposition of prepayment into refi vs. moves.

A.1.4 Leverage controls

In many of our aggregate regressions we control for average leverage in our data set and in our individual level regressions, we control for individual leverage. In order to measure leverage at the loan-level we start with all McDash FRM first mortgages. For each mortgage we estimate its current value as the appraisal value at origination updated using local house price indices from CoreLogic. We use zip code level house price indices to update values when the 5-digit zip is available in McDash and in the CoreLogic indices, and we otherwise use MSA level house price indices. We then compute leverage *LTV* for a given loan as the ratio of the current unpaid balance to this estimate of value. Our aggregate controls then take the average leverage across loans.

This procedure will tend to understate leverage for individuals with multiple loans, but it can be applied over the entire 1992-2017 sample. After 2005, we can construct a more accurate measure of leverage using CRISM data. Following [Beraja et al. \(2019\)](#), we begin with first-lien McDash FRM loans. For each month, we then take the corresponding Equifax record and assign all outstanding second liens to the outstanding first liens in Equifax using the rule that each second lien is assigned to the largest first lien (in balance terms) that was opened on or before the second lien's opening date. We then add the assigned second lien balance(s) to the McDash balance of our original loan as our measure of secured debt on a property, which is the numerator of *CLTV*. We then divide by the value constructed exactly as described above.

⁷⁸We impose these upper bounds because we want to avoid picking up other first lien mortgages (to purchase another property) the borrower might originate at the same time.

A.2 Additional Empirical Results

Figure A-3 is the annualized version of Figure 1 in the main text. The key difference between Figure 1 is that the interest rate gaps are averaged annually (as opposed to monthly), as is the prepayment indicator.⁷⁹ However, a comparison between Figure 1 and Figure A-3 shows that they are very similar. In particular, both show clear evidence of state-dependent prepayment: loans with positive gaps are much more likely to prepay than loans with negative gaps.

Figure A-3: Robustness of Prepayment Hazard to Annual Instead of Monthly Frequency

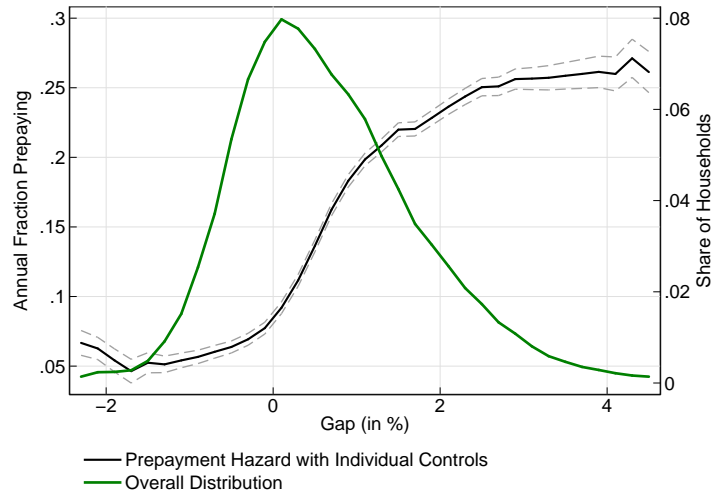


Figure shows the point estimates and 95-percent confidence intervals for the coefficients on the 20 basis point gap bin dummies in regression 1. Gaps are averaged by year, and prepayment is an indicator for any prepayment event over the year. Standard errors are clustered by household (but not year since we only have 12 year observations). In order to include household fixed effects and time-varying characteristics, figure uses CRISM data linked to credit records from 2005m6-2017m4.

Figure A-4 shows that restricting our analysis to households with substantial outstanding mortgage balances delivers results similar to those documented in Figure 1: a step-like behavior of the prepayment hazard, as a function of the rate gap.

Figure A-5 documents the foregone annual savings for households with 300bp+ rate-gaps who do not refinance. As the graph makes clear, the typical household with a large rate gap also has very large potential payment savings. For example, the average (median) annual foregone mortgage payment savings are \$2800 (\$2050). Overall, this supports our conclusion that lack of prepayment despite high individual rate gaps *does not* reflect limited benefits from refinancing.

⁷⁹We also compute standard errors differently. In this regression standard errors are clustered by household, but not year since we only have 12 year observations.

Figure A-4: Prepayment Hazard with Individual Controls: Excluding Mortgage Balances < \$100k

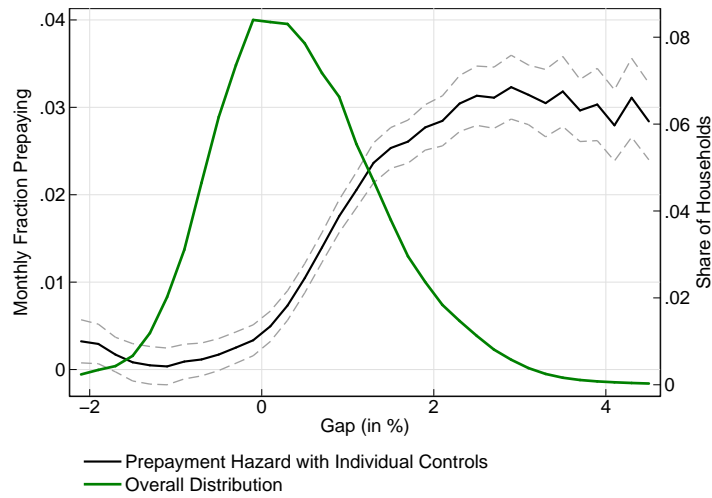


Figure shows point estimates and 95% confidence intervals for coefficients on the 20bp bin dummies in regression 1. We exclude loan-month observations with balances less than \$100,000. Standard errors are two-way clustered by household and month using data from 2005m6-2017m4.

Figure A-5: Annual Payment Savings Lost By Not Refinancing

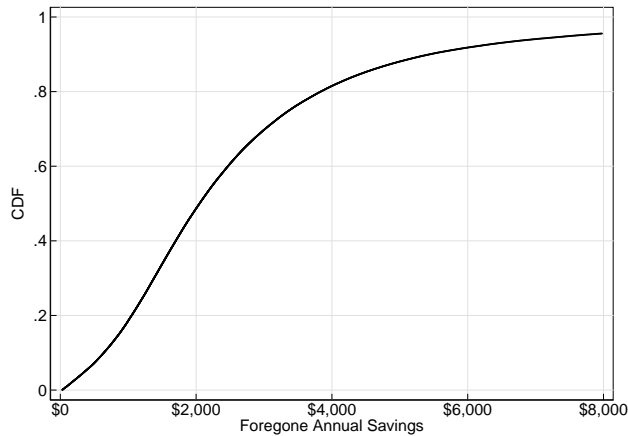
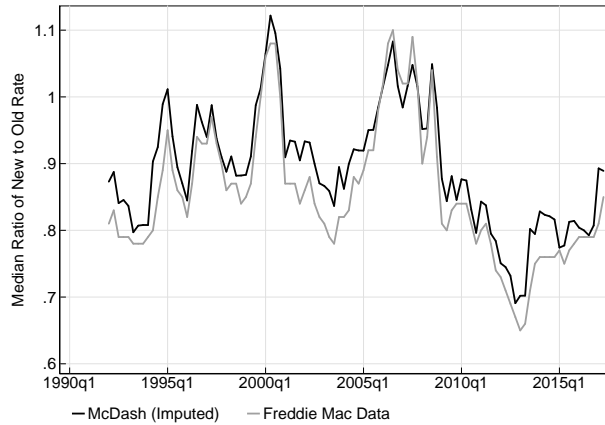


Figure shows CDF of the distribution of foregone annual savings for households with 300bp+ gaps who do not refinance.

In Section 4.2 we argue that rate incentives are a crucial driver of refinancing decisions, even for households taking cash out of their homes. The fact that few loans refinance into higher rates might seem to be at odds with evidence that in many months, most refinancing loans are doing so into higher rates (see [Chen et al. \(2019\)](#) Figure 1, or <https://www.wsj.com/articles/americans-are-taking-cash-out-of-their-homesand-it-is-costing-them-11577529000>). It is not. First, Figure A-6 shows that we can replicate the time-series evidence shown in the publicly available Freddie Mac almost perfectly in our data.

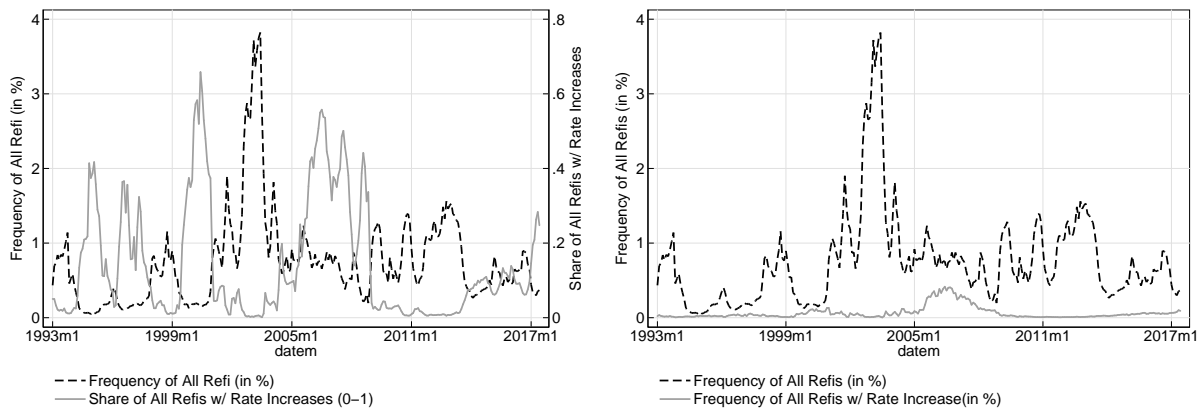
Figure A-6: Median of Ratio of New to Old Rate: McDash vs. Freddie Mac Data



For each month, we estimate the ratio of new rates relative to outstanding old rates for refinancing loans in our data. We then calculate the median of this ratio for each month and compare this to published data for the same object from Freddie Mac.

Figure A-7: Shares and Frequencies of Refinancing into Higher Rates

(a): Share of Refi w/ Increase vs. Refi Freq **(b):** Frequency of All Refi vs. All Refi w/ Increase



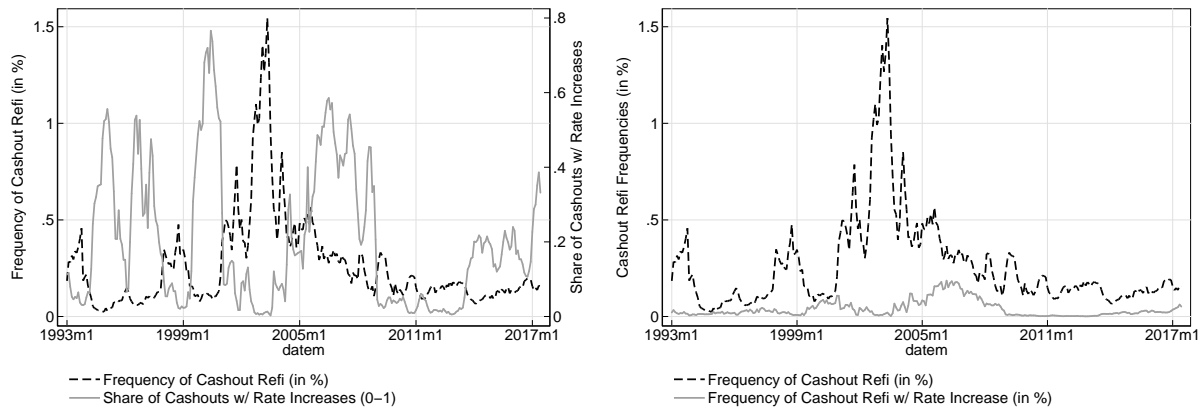
In particular, for each each month and for all prepaid GSE loans in the McDash data, we compute the estimated gap between the outstanding interest rate and the estimated new rate for the loan. We then apply the refinancing hazard computed in CRISM data after 2005 to allocate all of these prepayments between refinancing and moves. For each imputed refi in McDash data, we then compute the ratio of the estimated new rate relative to the outstanding old rate. We then calculate the median of this ratio and Figure A-6 compares it to published data for the same object from Freddie Mac.

Furthermore, Figure A-7 shows that changes in refinancing frequency are key to reconciling this time-series pattern with our cross-sectional fact that refinancing into higher rates is very unusual. Panel (a) shows that the times when most refis result in rate increases are precisely the times when the frequency

Figure A-8: Shares and Frequencies of Cash-out Refinancing into Higher Rates

(a): Share of Refi w/ Increase vs. Refi Freq

(b): Frequency of Refi vs. Refi w/ Increase



of refinancing is extremely low, so the loans refinancing in these months are a small share of overall refinancing activity. Panel (b) more directly shows that the frequency of refinancing into higher rates is nearly zero outside of 2005-2006, and even in those two years it is still low compared to the overall average frequency of refinancing across time. Figure A-8 shows that this is also true when restricting to only cash-out refinancing and excluding rate refis. In sum, these three figures reconcile our results with Chen et al. (2019).

Figure A-9: Distribution of Gaps at Two Dates

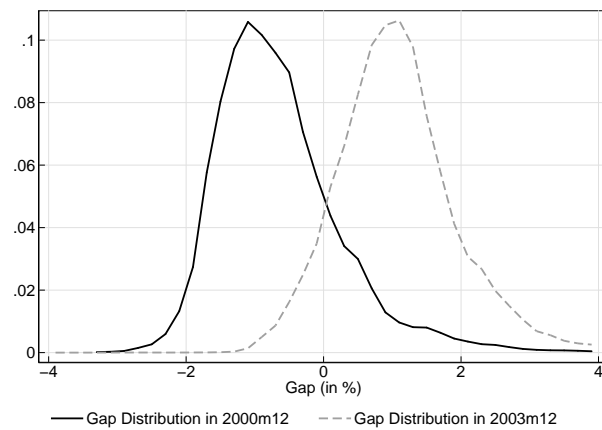
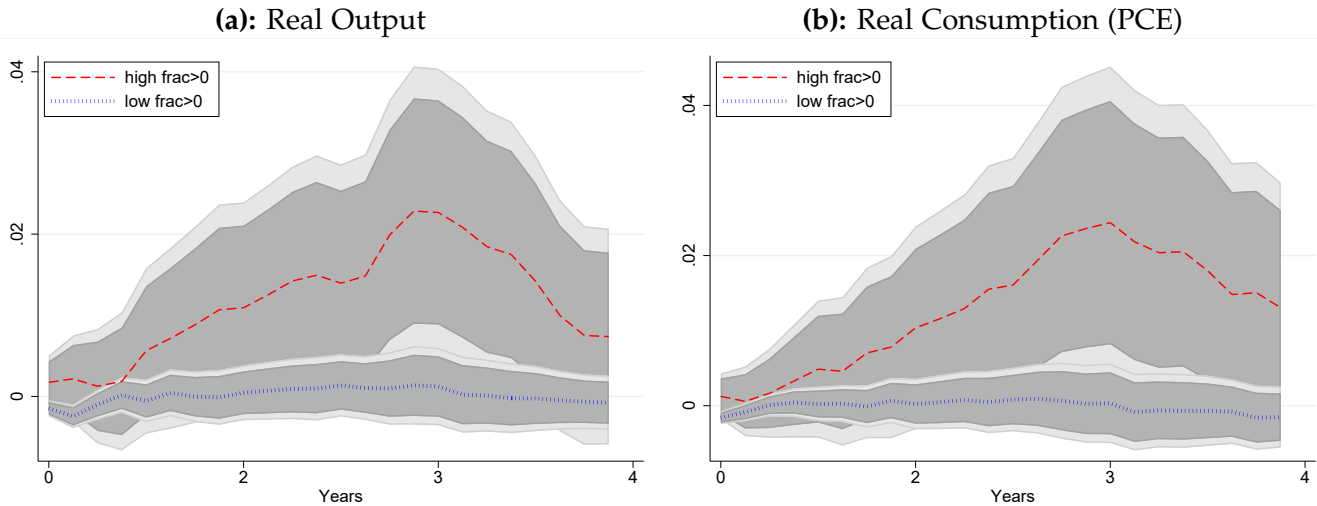


Figure A-9 shows the distribution of interest rate gaps at two points in time. The solid black line shows the distribution of gaps in December of 2000 while the dotted gray line shows the distribution of gaps in December of 2003. The key takeaway is that there is significant variation in these distributions and thus the $frac > 0$ over time. This time-variation along with the state-dependence of the hazard rate is what leads to significant path dependence.

In the main text, we look at time series responses to identified monetary policy shocks by regressing aggregate outcomes directly on monetary policy shocks. That exercise is analogous to our theoretical modeling exercise. In Figure A-10 we instead regress on changes in mortgage rates, instrumented by

these monetary policy shocks. This is conceptually further from our theoretical exercises, but potentially better isolates effects working through prepayment. Point estimates are similar although power is reduced so standard errors are wider.

Figure A-10: Response to Identified Mortgage Rate Shocks



This figure shows the response of log real output and non-durable consumption to the shocks to the 30 year mortgage rate instrumented using identified monetary shocks from [Romer and Romer \(2004\)](#), extended through 2007 by [Wieland and Yang \(2020\)](#). Light (dark) gray shaded areas are 90 (80) percent confidence intervals based on Newey-West standard errors. The red-dashed line indicates the effect of a monetary policy shock when $frac > 0$ is above its median value of 0.596. The blue-dotted line shows the effect when $frac > 0$ is below its median value.

Table A-1 shows that the relationship between $frac > 0$ and prepayment is robust to a number of other potential confounding observables. Columns (2) and (3) add controls for two different measures of the business cycle. Column (4) controls for seasonality with 12 month-of-year dummies. Column (5) provides evidence that adding additional controls to capture non-linearities in rate changes does not alter our conclusions. Specifically we include an indicator for periods of unusually large declines in mortgage rates: for each period, we compute the 3-month change in the mortgage rate $\Delta_{3,t} \equiv M_{t-3} - M_t$ and then add to the baseline regression an indicator for periods in the top ten percent of $\Delta_{3,t}$. Using different cutoffs for the indicator delivers similar conclusions. We have also explored interactions between large rate changes and the relationship between $frac > 0$ and prepayment and find no significant results. Importantly, this does not mean there are not non-linear effects of rate incentives on prepayment: non-linearities are most apparent from the microeconomic hazard in Figure 1. However, the result in Column (5) shows that these non-linearities are already captured by $frac > 0$: when rates decline by large amounts there are large increases in $frac > 0$ and resulting prepayment. Figure 5 shows that there is a large spike in prepayment in 2003. This unusual increase in refinancing coincides exactly with the “Mortgage-Rate Conundrum” documented by [Justiniano et al. \(2017\)](#). We cannot explain this outlier in 2003 based on observables but Column (6) shows that if we introduce a 2003 dummy, leverage and $frac > 0$ explain almost two-thirds of the variation in prepayment. In Column (7), we show that jointly controlling for all of these observables again leads to similar conclusions.

In Column (8), our most stringent empirical specification, we also add controls for one hundred calendar-quarter fixed effects.⁸⁰ In this specification, identification comes only from monthly relationships between prepayment and rate incentives within quarters. These specifications rule out many additional confounding factors such as aging ([Wong \(2019\)](#)) and trends in lender concentration ([Agar-](#)

⁸⁰We also include all of the other controls except the 2003 indicator, which is collinear with quarter fixed effects, but most of these controls are almost constant within quarters and so there is little identifying variation. Redoing the regression with only quarter FE and no additional controls produces similar results.

Table A-1: Relationship Between Rate Incentives and Prepayment with Controls

	(1) None	(2) U	(3) Y	(4) Seas	(5) Large Δm	(6) 2003	(7) All	(8) Quart FE
frac > 0	2.82*** (0.50)	2.82*** (0.50)	2.82*** (0.50)	2.82*** (0.50)	2.83*** (0.50)	2.48*** (0.36)	2.53*** (0.38)	2.14*** (0.39)
LTV	-6.49*** (1.66)	-6.49*** (1.66)	-6.49*** (1.66)	-6.49*** (1.66)	-6.64*** (1.70)	-4.96*** (1.15)	-4.96*** (1.26)	0.53 (5.30)
Constant	3.58*** (0.80)	3.58*** (0.80)	3.58*** (0.80)	3.58*** (0.80)	3.64*** (0.81)	2.79*** (0.57)	2.80*** (0.61)	-0.97 (3.52)
Adj. R^2	0.47	0.47	0.47	0.47	0.48	0.62	0.63	0.93
N	304	304	304	304	304	304	304	304
Date	92-17m4	92-17m4	92-17m4	92-17m4	92-17m4	92-17m4	92-17m4	92-17m4

Newey-West standard errors in parantheses. *=10%, **=5%, ***=1% significance. LTV is average leverage. Prepayment fractions are measured in month $t + 1$ while rate incentives and LTV are measured in month t , since McDash data measures origination not application and there is a 1-2 month lag from application to origination. U is the current unemployment rate, Y is monthly log industrial production detrended with an HP(129600) filter. Seas is 12 month of year indicators. Large Δm includes an indicator for the top ten percent of months by the 3-month decline in mortgage rates. 2003 includes and indicator for the year 2003. All includes all controls in Columns 2-6 simultaneously. Quart includes calendar-quarter fixed effects.

wal et al. (2017) and Scharfstein and Sunderam (2016)) which might influence refinancing incentives and prepayment rates but are unlikely to matter at these high frequencies.

In Appendix Table A-2 we show that the relationship between $frac > 0$ and prepayment also holds at the MSA level, even after including both calendar-month and MSA \times calendar-quarter fixed effects. As the table shows, there is a strong positive relationship between total prepayment (column 1), rate-refi (column 2), cash-out refi (column 3), home purchases (column 4) and $frac > 0$.

Since $frac > 0$ depends on past endogenous interest rates, it is possible that some unobserved confounding factor affects both $frac > 0$ and prepayment propensities even at high frequencies. To address this concern, Table A-3 re-estimates our baseline regressions using the cumulative value of the Gertler and Karadi (2015) high-frequency monetary policy shock series over the past six months as an instrument for $frac > 0$. Unsurprisingly, this reduces power and increases standard errors, but point estimates are nearly identical and $frac > 0$ remains a significant predictor of prepayment activity.⁸¹

In Table A-4 we decompose the positive time-series relationship between total prepayment and $frac > 0$ into its constituent types (rate-refi, cashout-refi and purchase). As suggested by the overall loan-level relationship in Figure 1, $frac > 0$ is most important for explaining rate-refinancing.⁸² $frac > 0$ alone explains roughly 40% of the time-series variance in rate-refinancing. Since leverage directly affects incentives to cash-out and move, we also explore the relationship between leverage and the different prepayment types. Leverage has no effect on rate refinancing, but unsurprisingly, it has a strong negative effect on cash-out and moves. Leverage has stronger independent predictive content (as measured by R^2) for cash-out and moves than does $frac > 0$. However columns (6) and (9) show that including both $frac > 0$ and leverage gives much stronger predictions than either alone. That is, after controlling for leverage, $frac > 0$ has strong additional predictive content for cash-out and moves.

Many of our empirical results focus on prepayment (and its constituent components) as the outcome of interest. However, changes in the average outstanding mortgage rate m^* are arguably more important than prepayment rates since mortgage payments are what enter the household budget con-

⁸¹First-stage F-stats shown in the table exceed the 15% Stock-Yogo critical values for weak instrument bias.

⁸²After 2005 we decompose prepayment using CRISM data; prior to 2005 we assume stationarity and decompose using origination shares. See Section 3. This decomposition requires origination shares data from CoreLogic LLMA data, which has poor coverage prior to 1993. For this reason, regressions start from 1993 rather than 1992 as in other tables.

Table A-2: Effects of Rate Gaps on Prepayment Propensities by MSA

	(1) Tot Prepay	(2) Rate-Refi	(3) Cashout	(4) Purchase
frac > 0	3.05*** (0.33)	1.80*** (0.22)	0.54*** (0.11)	0.69*** (0.24)
Quarter X MSA FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Adj. R^2	0.98	0.93	0.94	0.80
N	112885	112885	112885	112885
Date Range	92-17m4	92-17m4	92-17m4	92-17m4

Standard errors two-way clustered by MSA and month. *=10%, **=5%, ***=1% significance. Prepayment is measured using loan level data from McDash Performance data. After 2005, we decompose prepayment by type using CRISM data which links new and old loans. Prior to 2005, we decompose prepayment by type using origination shares by type from CoreLogic LLMA data. See Appendix for additional discussion. Prepayment fractions are measured in month $t + 1$ while rate incentives are measured in month t , since McDash data measures origination not application and there is a 1-2 month lag from application to origination.

Table A-3: Instrumenting for Rate Gaps with High Frequency Monetary Policy Shocks

	(1)	(2)	(3)
frac > 0	2.36*** (0.83)	3.01*** (0.72)	3.34*** (0.74)
LTV		-7.83*** (1.57)	-6.32*** (1.33)
Constant	-0.073 (0.55)	4.43*** (0.73)	3.50*** (0.56)
Additional Controls	None	None	All
F-Stat	17.2	16.3	21.7
N	247	247	247
Date	92-12m7	92-12m7	92-17m4

Newey-West standard errors in parantheses. *=10%, **=5%, ***=1% significance. This table instruments for $frac > 0$ using the sum over the prior 6 months of the high frequency monetary policy shock series from Gertler and Karadi (2015), available through 2012m7. LTV is average leverage. We calculate leverage for each loan as the ratio of its outstanding balance to value estimated using appraisal values at origination updated using local house price indices from CoreLogic. Loan level data from McDash Performance data+appraisal values from McDash origination data is used to calculate LTV. Prepayment fractions are measured in month $t + 1$ while rate incentives and LTV are measured in month t , since McDash data measures origination not application and there is a 1-2 month lag from application to origination. Additional controls include all those in Column (7) of Table A-1

Table A-4: Effects of Rate Gaps on Prepayment Propensities by Type

	Rate Refi			Cash-out			Purchase		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
frac > 0	1.261*** (0.187)		1.465*** (0.255)	0.296** (0.128)		0.599*** (0.118)	0.400*** (0.132)		0.670*** (0.124)
LTV		0.343 (0.820)	-1.787** (0.844)		-1.786*** (0.303)	-2.657*** (0.399)		-1.385*** (0.442)	-2.358*** (0.401)
Constant	-0.411*** (0.0948)	0.212 (0.533)	0.566 (0.413)	0.0670 (0.0672)	1.375*** (0.206)	1.520*** (0.189)	0.380*** (0.0690)	1.507*** (0.281)	1.669*** (0.229)
Adj. R^2	0.418	0.00244	0.472	0.0713	0.211	0.466	0.101	0.0947	0.335
N	292	292	292	292	292	292	292	292	292
Date Range	93-17m4	93-17m4	93-17m4	93-17m4	93-17m4	93-17m4	93-17m4	93-17m4	93-17m4

Newey-West standard errors in parantheses. *=10%, **=5%, ***=1% significance. Prepayment is measured using loan level data from McDash Performance data. After 2005, we decompose prepayment by type using CRISM data which links new and old loans. Prior to 2005, we decompose prepayment by type using origination shares by type from CoreLogic LLMA data. Regressions begin in 1993 rather than 1992 since reliable CoreLogic origination data on prepayment shares does not begin until 1993. See Appendix for additional discussion. Prepayment fractions are measured in month $t + 1$ while rate incentives and LTV are measured in month t , since McDash data measures origination not application and there is a 1-2 month lag from application to origination.

straint and prepayment matters more if households secure large payment reductions. Prepayment rates and changes in m^* are of course related: in each month the change in average rates is $\Delta \bar{m}^* = \int gap \times f(gap) \times h(gap) dgap$, where f is the density of gaps and h is the prepayment hazard in that month. If gaps are typically positive, then increases in prepayment will lead to declines in \bar{m}^* . However, it is also clear that they need not move perfectly together since average rates will decline by more if the households prepaying have larger gaps.

Table A-5: Effects of FRM Changes and Gaps on Average Coupon Changes

	(1)	(2)	(3)	(4)
$\text{frac} > 0$	-0.0486*** (0.00579)		-0.0470*** (0.00531)	-0.0531*** (0.00654)
ΔFRM		0.0188*** (0.00516)	-0.0286** (0.0124)	0.0369 (0.0356)
$\Delta \text{FRM} \times (\text{frac} > 0)$			0.0542** (0.0231)	0.0679*** (0.0244)
$\Delta \text{FRM} \times \text{mean LTV}$				-0.118* (0.0673)
mean LTV				0.0521** (0.0217)
Constant	0.0180*** (0.00284)	-0.0142*** (0.00178)	0.0175*** (0.00273)	-0.0109 (0.0111)
Adj. R^2	0.497	0.0497	0.516	0.560
N	303	303	303	303
Date Range	92-17m3	92-17m3	92-17m3	92-17m3

Newey-West standard errors in parantheses. *=10%, **=5%, ***=1% significance. Average LTV is average across loans of the ratio of a loan's outstanding balance to value estimated using appraisal values at origination updated using local house price indices from CoreLogic. Loan level data from McDash Performance data+appraisal values from McDash origination data is used to calculate LTV. ΔFRM is the change in the current 30 year FRM is the monthly average of the Freddie Mac weekly PMMS survey 30 year fixed rate mortgage average: <https://fred.stlouisfed.org/series/MORTGAGE30US>. To account for a lag between application and origination, in all specifications, $\text{frac} > 0$ and LTV is measured as of month t , Δm^* is measured between month t and month $t + 1$ and ΔFRM is measured between month $t - 1$ and month t .

Column (1) of Table A-5 documents that there is an extremely strong negative time-series relationship between $\text{frac} > 0$ and Δm^* . Unsurprisingly, Column (2) shows that when the current market interest rate rises, so does the resulting average outstanding rate. More interestingly, Column (3) shows that there is a strong interaction effect between $\text{frac} > 0$ and ΔFRM : interest rate pass-through into average coupons is much stronger when $\text{frac} > 0$ is large. As we discuss below, this increase in rate pass-through with $\text{frac} > 0$ is a central implication of our theoretical model and is a key indicator of path-dependence. Given the importance of leverage for prepayment discussed above, Column (5) also includes interactions of interest rate changes in month t with average leverage in this same month. While we indeed find a negative interaction effect between leverage and pass-through, the interaction between $\text{frac} > 0$ and ΔFRM is if anything mildly amplified.

Table A-6 presents robustness to our main regression results when we only include conforming loans in our sample. Other types of loans can sometimes have different refinancing processes and institutional constraints (e.g. streamlined FHA refi) that could lead to different interactions with rate incentives. Reassuringly the results in Table A-6 are very similar to the results from using our baseline sample as shown in Table 2.

Table A-6: Robustness to Including only Conforming Loans

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Date Range	92-17m4	92-17m4	92-17m4	92-00	01-10	11-17m4	92-00	01-10	11-17m4
frac > 0	2.23*** (0.37)	2.90*** (0.50)		2.48*** (0.39)	2.86*** (0.88)	2.45*** (0.40)	2.55*** (0.37)	5.06*** (0.92)	2.48*** (0.44)
LTV		-5.82*** (1.82)	-1.07 (1.61)				-1.58 (1.70)	-12.7*** (2.89)	-0.11 (0.91)
Constant	0.073 (0.16)	3.29*** (0.92)	2.15** (1.07)	-0.086 (0.19)	0.024 (0.35)	-0.59* (0.30)	0.90 (1.14)	6.18*** (1.28)	-0.54 (0.53)
Adj. R^2	0.33	0.45	0.0020	0.61	0.33	0.67	0.62	0.67	0.66
N	304	304	304	108	120	76	108	120	76

Newey-West standard errors in parantheses. *=10%, **=5%, ***=1% significance. This table redoes the baseline analysis in Table 2, but restricting to only conforming loans.

Conversely, Table A-7 presents robustness to our main regression results when we *exclude* conforming loans from our sample. Housing agencies play a role in credit supply and thus potentially affect transmission of monetary policy into conforming loan mortgage rates, which is outside of the scope of our model. Around 70% of loans over our whole sample period are conforming, so this is a major sample restriction. However, results in Table A-7 are again quite similar to the results from using our baseline sample as shown in Table 2.

Table A-7: Robustness to Excluding Conforming Loans

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Date Range	92-17m4	92-17m4	92-17m4	92-00	01-10	11-17m4	92-00	01-10	11-17m4
frac > 0	2.02*** (0.43)	2.63*** (0.45)		1.77*** (0.33)	2.69*** (0.92)	1.40*** (0.39)	1.85*** (0.30)	4.55*** (0.82)	1.98*** (0.38)
LTV		-6.65*** (1.30)	-4.24*** (1.36)				-2.23** (1.08)	-11.3*** (1.91)	-2.15*** (0.71)
Constant	-0.12 (0.26)	3.60*** (0.62)	4.01*** (0.91)	-0.024 (0.19)	-0.29 (0.56)	-0.039 (0.30)	1.38* (0.76)	4.95*** (0.74)	0.88* (0.44)
Adj. R^2	0.25	0.49	0.11	0.49	0.24	0.41	0.50	0.66	0.54
N	304	304	304	108	120	76	108	120	76
Date Range	92-17m4	92-17m4	92-17m4	92-00	01-10	11-17m4	92-00	01-10	11-17m4

Newey-West standard errors in parantheses. *=10%, **=5%, ***=1% significance. This table redoes the baseline analysis in Table 2, but restricting to only non-conforming loans.

Table A-8 presents robustness to our main regression results when we only include loans which were never delinquent. The main empirical concern here is that if a loan was ever delinquent it may be more difficult to refinance. Thus, if we see a household with a large gap not refinancing, it might be because this household was previously delinquent, not because they are inattentive. Results in Table A-8 are very similar to the results from our baseline sample shown in Table 2.

Table A-8: Robustness to Excluding Loans Which are Ever Delinquent

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Date Range	92-17m4	92-17m4	92-17m4	92-00	01-10	11-17m4	92-00	01-10	11-17m4
frac > 0	2.44*** (0.40)	3.22*** (0.54)		2.41*** (0.39)	3.18*** (0.90)	2.64*** (0.39)	2.50*** (0.37)	5.19*** (0.95)	2.59*** (0.43)
LTV		-6.96*** (1.95)	-1.85 (1.75)				-2.23 (1.51)	-12.6*** (2.93)	0.20 (1.06)
Constant	0.091 (0.19)	3.94*** (0.97)	2.80** (1.16)	-0.037 (0.20)	0.084 (0.37)	-0.56** (0.28)	1.37 (1.02)	6.31*** (1.29)	-0.66 (0.60)
Adj. R^2	0.33	0.48	0.010	0.58	0.36	0.71	0.59	0.66	0.70
N	304	304	304	108	120	76	108	120	76

Newey-West standard errors in parantheses. *=10%, **=5%, ***=1% significance. This table redoes the baseline analysis in Table 2, but restricting to only loans which are never delinquent.

Table A-9 presents robustness using a broader set of loans including those with adjustable rates rather than just the fixed rate loans which are the focus of our main analysis. The main empirical concern here is that if the loan is an adjusted rate mortgage (ARM) then there is no incentive to refinance as the interest rate adjusts automatically (at least once it is past the initial period of fixed rates typical under the hybrid ARMs common in the US). The results in Table A-9 are very similar to the results using our baseline sample. This is because the FRM share is large: overall, just over 80% of mortgage balances are in fixed rate loans rather than adjustable rate loans. In addition, even when focusing on ARMs, most are actually still in the the initial fixed rate period when gaps and refi incentives can emerge. Furthermore, the FRM share varies across time but is generally between 75-85%, except for a moderate dip during the housing boom, and there is little relationship between the FRM share and $frac > 0$.

Table A-9: Robustness to Including all Loans Instead of Only Fixed Rate Loans

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Date Range	92-17m4	92-17m4	92-17m4	92-00	01-10	11-17m4	92-00	01-10	11-17m4
frac > 0	1.91*** (0.40)	2.79*** (0.50)		2.01*** (0.30)	2.40*** (0.91)	1.90*** (0.39)	2.09*** (0.27)	5.09*** (0.88)	2.24*** (0.40)
LTV		-6.89*** (1.58)	-3.04** (1.38)				-2.32** (1.08)	-13.6*** (2.39)	-1.24 (0.77)
Constant	0.22 (0.20)	3.97*** (0.76)	3.34*** (0.93)	0.069 (0.16)	0.27 (0.41)	-0.20 (0.27)	1.53** (0.75)	6.60*** (0.97)	0.35 (0.46)
Adj. R^2	0.23	0.47	0.052	0.57	0.22	0.56	0.58	0.69	0.60
N	304	304	304	108	120	76	108	120	76

Newey-West standard errors in parantheses. *=10%, **=5%, ***=1% significance. This table redoes the baseline analysis in Table 2, but extending the analysis to all loans including those with adjustable rates instead of only looking at fixed rate mortgages.

A.3 Model Appendix

A.3.1 Model Calibration

Table A-10: Model Parameter Values

Panel A: Exogenous Processes				
Parameter	Value			Description
$\ln 2 / \eta_r$	5.3 years			half life of interest-rate shock
\bar{r}	3.5% (p.a.)			(unconditional) interest rate mean
σ_r	6% (p.a.)			interest rate volatility
$\ln 2 / \eta_y$	7.3 years			half-life of (log) income shock
$\mathbb{E}[Y_t]$	\$58,000			(unconditional) income mean
σ_y	21% (p.a.)			log-income volatility
γ	2			inverse IES
F	\$150,000			mortgage debt outstanding

Panel B: Refinancing Frictions				
Parameter	Inattention Baseline	Fixed-Cost	Hybrid	Description
ν	4.1% (p.a.)	4.1% (p.a.)	4.1% (p.a.)	arrival rate of moving shocks
χ_c	22.8% (p.a.)	0	12.5% (p.a.)	arrival rate of zero cost refi opportunity
χ_f	0	2400% (p.a.)	14.5% (p.a.)	arrival rate of positive cost refi opportunity
κ	0	\$2500	\$8250	fixed cost of refinancing for χ_f

A.3.2 Mortgage Rates

In this section, we explain how to map short term rates into mortgage rates when risk-neutral financial intermediaries lend to inattentive households. In the particular case of our pure Calvo model, the value of a mortgage, from the financial intermediary standpoint, is only a function of (a) the current level of short rates r , and (b) the mortgage coupon $m^* = m(r^*)$, since we know the household will refinance whenever he has the opportunity to do so and whenever the short term rate r is below the short term rate r^* that was prevalent at the time of the previous refinancing. Thus, the price of a mortgage (with face value of \$1) can be encoded via the function $P(r, r^*)$:

$$P(r, r^*) = \mathbb{E} \left[\int_0^\tau e^{-\int_0^t r_s ds} m(r^*) dt + e^{-\int_0^\tau r_s ds} \Big| r_0 = r \right]$$

In the above, τ is the prepayment time, a stopping time that is the minimum of (a) an exponentially-distributed time τ_ν representing a move, and (b) the first exponentially distributed attention time τ_χ for which the mortgage rate $m(r_{\tau_\chi})$ is below m^* . If we note \mathcal{L}_r the infinitesimal operator associated with the stochastic process r_t , P satisfies

$$(r + \nu + \chi \mathbb{I}_{\{r < r^*\}}) P = m(r^*) + \nu + \chi \mathbb{I}_{\{r < r^*\}} + \mathcal{L}_r P$$

Assuming the mortgage function $m(r^*)$ is known, this is a standard ordinary differential equation, which can be solved numerically with standard methods. In a risk-neutral environment, it must be the case that the price of the mortgage, at time of origination, is equal to its notional value. In other words, we must have the mortgage market equilibrium condition $P(r, r) = 1$. This latter equation allows us to pin down the implicit function $m(\cdot)$.

A.3.3 Solution to Household Problem

Let $S := (W, m^*, Y)$ be the household idiosyncratic state and $V(r, S)$ be the value function when the short rate is r for a household with income Y , liquid wealth W and a fixed mortgage rate m^* . Let \mathcal{L}_y be the infinitesimal operator associated with the stochastic process Y_t . The household Hamilton-Jacobi-Bellman (HJB) equation can be written:

$$\delta V = \sup_C u(C) + \mathcal{L}_r V + \mathcal{L}_y V + \left(v + \chi \mathbb{I}_{m(r) < m^*} \right) [V(r, S_{-m^*}, m(r)) - V(r, S_{-m^*}, m^*)] + (rW + Y - C - m^*F) \partial_W V \quad (\text{A1})$$

The optimal consumption function $C(r, S)$ solves the first order condition $u'(C(r, S)) = \partial_W V(r, S)$, which can be written:

$$C(r, S) = (\partial_W V(r, S))^{-1/\gamma} \quad (\text{A2})$$

We can then reinject the optimal consumption policy into the HJB equation satisfied by V to obtain a non-linear partial differential equation satisfied by V . The non-linearity stems from the fact that consumption is controlled – its value depends on the first partial derivative of V w.r.t. W . The endogenous savings rate can then be written

$$\mu_W(r, S) := rW + Y - C(r, S) - m^*F$$

Section A.3.6 discusses our numerical method to solve this non-linear PDE.

A.3.4 Fokker Planck Equation

The joint density g_t over (1) the aggregate short rate state r , and (2) the idiosyncratic state vector S , consisting of (a) savings W , (b) coupons m^* and (c) income Y , satisfies the following Fokker Planck equation (for $m^* \neq m(r)$):

$$\partial_t g_t = -\partial_W [\mu_W(r_t, S) g_t(S)] + \mathcal{L}_y^* g_t + \mathcal{L}_r^* g_t - \left(v + \chi \mathbb{1}_{\{m(r_t) < m^*\}} \right) g_t \quad (\text{A3})$$

\mathcal{L}_y^* (resp. \mathcal{L}_r^*) is the adjoint operator of \mathcal{L}_y (resp. \mathcal{L}_r), associated with the stochastic process for Y_t (resp. r_t). This equation describes the inflows and outflows of "particles" in and out of the state (r, S) ; it accounts for changes in short rate r_t , in income Y_t , in savings W_t , as well as refinancings that reset the mortgage coupon of a household. A slightly different equation holds for $m(r) = m^*$, since in this case we must take into account the inflow of households who are refinancing, re-striking their long-term fixed rate mortgage at the rate $m(r)$:

$$\lim_{r \searrow m^{-1}(m^*)} \left[\eta_r(r - \bar{r}) g_t(r, S) + \partial_r \left[\frac{\sigma_r^2 r}{2} g_t(r, S) \right] \right] - \lim_{r \nearrow m^{-1}(m^*)} \left[\eta_r(r - \bar{r}) g_t(r, S) + \partial_r \left[\frac{\sigma_r^2 r}{2} g_t(r, S) \right] \right] + v \int_0^{+\infty} g_t(m^{-1}(m^*), W, x, Y) dx + \chi \int_{m^*}^{+\infty} g_t(m^{-1}(m^*), W, x, Y) dx = 0$$

These equations will be leveraged in our numerical scheme when computing impulse response functions.

A.3.5 Impulse Response Functions

Our impulse response function ("IRF") calculations focus on the following outcome variables: average prepayment rates, average mortgage coupons, and aggregate (per household-annum) consumption. The initial state of the economy is given by a distribution over (a) short rates and (b) liquid savings, coupons

and income, $g_0(r, W, m^*, Y)$ that is degenerate, since the short rate is assumed to be known at time zero. In other words, given our knowledge of r_0 , the initial distribution g_0 satisfies $g_0(r, W, m^*, Y) = 0$ for any $r \neq r_0$, and there exists a density \hat{g}_0 over (W, m^*, Y) that satisfies $g_0(r, W, m^*, Y) = \mathbb{I}_{r=r_0} \hat{g}_0(W, m^*, Y)$. To compute the consumption IRF (for example), we first have to define expected aggregate consumption at time t , $\bar{C}(t; \hat{g}_0, r_0)$, as a function of the initial state of the economy:

$$\bar{C}(t; \hat{g}_0, r_0) := \iiint \mathbb{E} [C(r_t, W_t, m_t^*, Y_t) | r_0 = r, W_0 = W, m_0^* = m^*, Y_0 = Y] \hat{g}_0(W, m^*, Y) dW dm^* dY$$

In the above, $C(r_t, W_t, m_t^*, Y_t)$ is the consumption function for an optimizing household with liquid savings W_t , mortgage coupon m_t^* , income level Y_t , when the current short term rate is r_t (and the corresponding market mortgage rate is $m(r_t)$). The consumption IRF to a 100 bps decline in rates is then simply defined as:

$$IRF_C^{1\%}(t; \hat{g}_0, r_0) := \frac{\bar{C}(t; \hat{g}_0, r_0 - 1\%)}{\bar{C}(t; \hat{g}_0, r_0)} - 1$$

We express consumption IRFs as semi-elasticities, and compute average mortgage coupon and average prepayment IRFs in absolute terms:

$$IRF_{m^*}^{1\%}(t; \hat{g}_0, r_0) := \bar{m}^*(t; \hat{g}_0, r_0 - 1\%) - \bar{m}^*(t; \hat{g}_0, r_0).$$

A.3.6 Numerical Implementation

We compute the equilibrium of the model numerically by determining the value function V at $N := n_w \times n_r \times n_r \times n_y$ discrete points of the state space. We use a standard finite difference scheme with upwinding for solving our PDE – in other words, we use a forward difference for approximating the first partial derivative of V in a given direction whenever the drift in this direction is positive, and a backward difference otherwise. The upwinding strategy ensures that our finite difference scheme is monotone. Since the HJB includes an optimal control, we solve the value function iteratively using a false transient (aka an artificial time-derivative), and at each iteration update the consumption policy using the value function and its derivatives according to equation (A2). Defining $\vec{V}^{(i)}$ as the vector of values of the value function V at each point of our discretization grid at iteration i , our numerical scheme leads us to solve successive linear equation systems of the form

$$\left[(1 + \delta \Delta_t) I - \Delta_t M^{(i)} \right] \vec{V}^{(i+1)} = \vec{V}^{(i)} + \Delta_t \vec{\Phi}^{(i)},$$

where $\Delta_t > 0$ is the time-step of our false transient algorithm, I is the identity matrix (dimension N), $M^{(i)}$ is an $N \times N$ square matrix, $\vec{\Phi}^{(i)}$ is an N dimensional vector with elements $\{u(C_k)\}_{k \leq N}$, and $\vec{V}^{(i+1)}$ is the unknown value vector. The $N \times N$ matrix $M^{(i)}$ is the discrete state counterpart to the infinitesimal operator for the dynamic system (r, S) . It has the interpretation of an "intensity" matrix: its diagonal elements are all negative, its off-diagonal elements are all positive, and its row-sums are all equal to zero.

Our algorithm iterates until the point where the artificial time derivative of our false transient is close to zero. Note that the matrix $M^{(i)}$ then converges to a matrix M at that point. The ergodic distribution of our economic model is then computed by focusing on the implied transition intensity matrix M , and by finding the column vector π that solves $\pi' M = 0$ — in other words, the left-eigenvector of M , associated with the eigen-value 0, that verifies $\sum_{k=1}^N \pi_k = 1$.

Finally, in order to compute impulse response functions, we use the discrete state counterpart of equation (A3) in order to compute the density of our economic system at time t , given initial conditions. Starting from a discretized density of our economic system \vec{g}_0 at time zero, we compute \vec{g}_t iteratively

solving the linear system

$$\frac{\vec{g}_{t+1} - \vec{g}_t}{\Delta_t} = M^T \vec{g}_{t+1},$$

where M^T represents the transpose of M . This allows us to compute statistics of the economic system at time t without relying on Monte-Carlo simulations, but instead by leveraging our value function solution's method.

A.3.7 Caballero-Engel in Continuous Time

Let m_{it}^* be the mortgage coupon of household i at time t , and let m_t be the market mortgage rate. Let $f_t(m^*)$ be the density of mortgage coupons in the economy (and $F_t(m^*)$ its CDF). Let us assume that prepayments are purely driven by a hazard function $h(m^* - m_t)$. In our model, this function is equal to $h(m^* - m) = \nu + \chi \mathbb{I}_{\{m^* - m > 0\}}$. The time- t average (instantaneous) refinancing intensity is equal to

$$\mathbb{E}_i[\rho_{it}] := \int h(m^* - m_t) f_t(m^*) dm^* = \nu + \chi (1 - F_t(m_t))$$

$1 - F_t(m_t)$ is exactly $(frac > 0)_t$, meaning that our benchmark model-implied prepayment rates are purely driven by this moment of the cross-sectional distribution of mortgage coupons. The average coupon rate \bar{m}_t^* can be computed as follows:

$$\bar{m}_t^* := \mathbb{E}_i[m_{it}^*] = \int m^* f_t(m^*) dm^*$$

We also know that the density f_t evolves as follows, between t and $t + dt$, for any $m^* \neq m_t$:

$$f_{t+dt}(m^*) \approx (1 - h(m^* - m_t) dt) f_t(m^*)$$

Thus, we have:

$$\bar{m}_{t+dt}^* = \int m^* [1 - h(m^* - m_t) dt] f_t(m^*) dm^* + m_t \int h(m^* - m_t) f_t(m^*) dm^* dt$$

The first term stems from mortgages that have not been refinanced between t and $t + dt$, whereas the second term stems from the new mortgages being refinanced, and which are contractually setting their coupon at m_t . In other words, we have:

$$d\bar{m}_t^* := \bar{m}_{t+dt}^* - \bar{m}_t^* = \int (m_t - m^*) h(m^* - m_t) f_t(m^*) dm^* dt$$

Using our specialized hazard function, we obtain:

$$\frac{d\bar{m}_t^*}{dt} = \nu (m_t - \bar{m}_t^*) + \chi (m_t - \mathbb{E}_i[m_{it}^* | m_{it}^* > m_t]) (1 - F_t(m_t)) \quad (\text{A4})$$

Let us consider a small change in r_t and its impact on $\frac{d\bar{m}_t^*}{dt}$. This is essentially the (time-slope) of the IRF w.r.t. to a small change in the interest rate:

$$\begin{aligned} \frac{\partial}{\partial r_t} \left(\frac{d\bar{m}_t^*}{dt} \right) &= m'(r_t) [\nu + \chi [1 - F_t(m(r_t))]] - \chi \left(\frac{\partial}{\partial r_t} \mathbb{E}_i[m_{it}^* | m_{it}^* > m(r_t)] \right) [1 - F_t(m(r_t))] \\ &\quad + \chi (\mathbb{E}_i[m_{it}^* | m_{it}^* > m(r_t)] + m(r_t)) f_t(m(r_t)) m'(r_t) \end{aligned}$$

Note that

$$\mathbb{E}_i [m_{it}^* | m_{it}^* > m(r_t)] = \frac{\int_{m(r_t)}^{\infty} m^* f_t(m^*) dm^*}{1 - F_t(m(r_t))}$$

so that

$$\begin{aligned} \frac{\partial \mathbb{E}_i [m_{it}^* | m_{it}^* > m(r_t)]}{\partial r_t} &= \frac{-m(r_t) f_t(m(r_t)) m'(r_t)}{1 - F_t(m(r_t))} - \frac{\int_{m(r_t)}^{\infty} m^* f_t(m^*) dm^*}{[1 - F_t(m(r_t))]^2} (-) f_t(m(r_t)) m'(r_t) \\ &= (\mathbb{E}_i [m_{it}^* | m_{it}^* > m(r_t)] - m(r_t)) \frac{f_t(m(r_t)) m'(r_t)}{1 - F_t(m(r_t))} \end{aligned}$$

Plugging in, we are left with

$$\frac{\partial}{\partial r_t} \left(\frac{d\bar{m}^*}{dt} \right) = m'(r_t) [v + \chi [1 - F_t(m(r_t))]] \quad (\text{A5})$$

A.3.8 Analytic Characterization of Consumption Semi-Elasticity

We focus on a representative household with preferences

$$U := \int_0^{+\infty} e^{-\delta t} \frac{c_t^{1-\gamma}}{1-\gamma} dt$$

We are in a complete markets partial equilibrium environment, in which a household is endowed with constant income \bar{y} , and at time zero, savings w . The household budget constraint is

$$\int_0^{+\infty} e^{-\int_0^t r_s ds} c_t dt = \int_0^{+\infty} e^{-\int_0^t r_s ds} \bar{y} dt + w$$

We are interested in computing the household's consumption response to a shock to interest rates, in the neighborhood of a steady-state. In this steady-state, we assume that short rates are constant and equal to δ . The household' savings are invested in a bank account that earns the risk-free rate r_t .

A.3.8.1 No Pre-payable Mortgage Debt

We first analyze the consumption semi-elasticity to rates in the absence of mortgage debt. We focus on a small mean-reverting shock to the nominal rate, such that

$$r_t = \delta + \epsilon_t \quad d\epsilon_t = -\eta_r \epsilon_t dt$$

We interpret this as a shock to real rates, under the assumption that prices in this economy are perfectly sticky. Asymptotically, the short rate converges back to the steady state δ . The persistence of the monetary shock is parametrized via η_r . Define the value function $V(w, \epsilon; \bar{y})$ as:

$$\begin{aligned} V(w, \epsilon; \bar{y}) &:= \max_c \int_0^{+\infty} e^{-\delta t} \frac{c_t^{1-\gamma}}{1-\gamma} dt \\ \text{s.t. } dw_t &= [(\delta + \epsilon_t)w_t + \bar{y} - c_t] dt \\ d\epsilon_t &= -\eta_r \epsilon_t dt \\ (w_0, \epsilon_0) &= (w, \epsilon) \end{aligned}$$

The value function V satisfies the following HJB equation

$$\delta V = \max_c \frac{c^{1-\gamma}}{1-\gamma} + [(\delta + \epsilon)w + \bar{y} - c] V_w - \eta_r \epsilon V_\epsilon \quad (\text{A6})$$

Optimal consumption satisfies $c^{-\gamma} = V_w$. We posit the following Taylor expansion for V and optimal consumption c

$$\begin{aligned} V(w, \epsilon; \bar{y}) &= V_0(w; \bar{y}) + \epsilon V_1(w; \bar{y}) + o(\epsilon) \\ c(w, \epsilon; \bar{y}) &= c_0(w; \bar{y}) + \epsilon c_1(w; \bar{y}) + o(\epsilon) \end{aligned}$$

The asymptotic expansion of the household's consumption optimality equation leads to

$$\begin{aligned} c_0^{-\gamma} &= V'_0 \\ -\gamma \frac{c_1}{c_0} &= \frac{V'_1}{V'_0} \end{aligned}$$

The terms of order zero in our asymptotic expansion must satisfy

$$\begin{aligned} c_0(w; \bar{y}) &= \delta w + \bar{y} \\ V_0(w; \bar{y}) &= \frac{1}{\delta} \frac{(\delta w + \bar{y})^{1-\gamma}}{1-\gamma} \end{aligned}$$

In other words, consumption must be equal to permanent income, and the value function is simply equal to the net present value of the constant flow utilities over consumption. The HJB equation (A6) allows us to derive a differential equation for the first order term V_1

$$\delta V_1 = c_0^{-\gamma} c_1 + (w - c_1) V'_0 + [\delta w + \bar{y} - c_0] V'_1 - \eta_r V_1$$

We use the optimality condition $c_0^{-\gamma} = V'_0$ and $c_0 = \delta w + \bar{y}$ to simplify and obtain

$$\begin{aligned} V_1(w; \bar{y}) &= \frac{w V'_0(w; \bar{y})}{\eta_r + \delta} = \frac{w (\delta w + \bar{y})^{-\gamma}}{\eta_r + \delta} \\ \frac{c_1(w; \bar{y})}{c_0(w; \bar{y})} &= \frac{-1}{\eta_r + \delta} \left[\frac{1}{\gamma} - \frac{\delta w}{c_0(w; \bar{y})} \right] \end{aligned}$$

The fraction c_1/c_0 is the semi-elasticity of consumption to a small shock to interest rates. Our result is consistent with [Kaplan et al. \(2018\)](#): the direct, partial equilibrium, effect of a small interest rate shock onto consumption is higher if (a) the rate of time preference is small, (b) the persistence of the monetary policy shock is high, and (c) the inter-temporal elasticity of substitution is high. The consumption response is slightly muted by the presence of positive savings (for reasonable values of the asset-to-income ratio).

A.3.8.2 Pre-payable Mortgage Debt

Now imagine that parts of the household's financial position is a fixed-rate mortgage liability, with a coupon that can be refinanced at Poisson arrival times (intensity χ_c). We denote m_t the mortgage market rate at time t , and r_t the short term rate at time t . We assume that

$$r_t = \delta + \epsilon_t \quad m_t = \delta + \pi \epsilon_t \quad d\epsilon_t = -\eta_r \epsilon_t dt$$

The parameter $\pi > 0$ is the pass-through, from the short rate to the mortgage rate. We are assuming that the time-zero mortgage coupon of the household is δ , equal to the long run short (and mortgage) rate. We continue to consider small shocks to the interest rate. If the initial interest rate shock is positive, the mortgage market interest rate jumps up on impact, the household will never use his option to refinance (when given the chance to do so), and thus the consumption response of households to a positive interest rate shock is identical to the response computed in [Section A.3.8.1](#), where the permanent income $\bar{y} + \delta(w - F)$ is adjusted to reflect the presence of mortgage debt.

If the initial interest rate shock is negative, the household will want to refinance whenever he has the opportunity to do so, and will do so only once, since mortgage rates are monotone increasing, converging back to their steady state δ . Let J be the household value function before refinancing, which satisfies

$$(\delta + \chi_c) J(w, \epsilon) = \max_c \frac{c^{1-\gamma}}{1-\gamma} + [(\delta + \epsilon)w + \bar{y} - \delta F - c] J_w(w, \epsilon) - \eta_r \epsilon J_\epsilon(w, \epsilon) + \chi_c V(w, \epsilon; \bar{y} - (\delta + \pi \epsilon)F)$$

The last term of this HJB equation relates to refinancing, following which the household value function is equal to the function V computed in [Section A.3.8.1](#). Optimal consumption satisfies once again $c^{-\gamma} = J_w$. The zero order term of our asymptotic expansion satisfies

$$(\delta + \chi_c) J_0 = \frac{(J'_0)^{1-1/\gamma}}{1-\gamma} + \left[\delta(w - F) + \bar{y} - (J'_0)^{-1/\gamma} \right] J'_0 + \chi_c V_0(w; \bar{y} - \delta F)$$

The solution to this equation is $J_0 = V_0(w; \bar{y} - \delta F)$, and $c_0(w) = \delta(w - F) + \bar{y}$. The first order correction term J_1 satisfies

$$(\delta + \chi_c + \eta_r) J_1 = c_0^{-\gamma} c_1 + (w - c_1) J'_0 + [\delta w + \bar{y} - \kappa F - c_0] J'_1 + \chi_c \left(V_1(w; \bar{y} - \delta F) - \pi F \frac{\partial V_0}{\partial \bar{y}}(w; \bar{y} - \delta F) \right)$$

This can be simplified further since $c_0 = \delta(w - F) + \bar{y}$, and we obtain

$$(\delta + \chi_c + \eta_r) J_1 = w J'_0 + \chi_c (\delta(w - F) + \bar{y})^{-\gamma} \left(\frac{w}{\eta_r + \delta} - \frac{\pi F}{\delta} \right)$$

Since J_0 is known, the above equation allows us to pin J_1 . We then compute the first order correction term for consumption c_1 via

$$-\gamma \frac{c_1}{c_0} = \frac{V'_1}{V'_0}$$

Plugging in and summarizing our result, denoting $\mathbb{I}_{\{\epsilon_0 < 0\}}$ the indicator for whether the initial rate shock is negative or not, we have

$$\begin{aligned} J_0 &= \frac{1}{\delta} \frac{(\bar{y} + \delta(w - F))^{1-\gamma}}{1-\gamma} \\ c_0 &= \bar{y} + \delta(w - F) \\ J_1 &= \frac{(\bar{y} + \delta(w - F))^{-\gamma}}{\eta_r + \delta + \chi_c} \left[w + \chi_c \left(\frac{w}{\eta_r + \delta} - \frac{\pi F \mathbb{I}_{\{\epsilon_0 < 0\}}}{\delta} \right) \right] \\ \frac{c_1}{c_0} &= \frac{-1}{\eta_r + \delta} \left[\frac{1}{\gamma} - \frac{\delta w}{c_0} + \left(\frac{\eta_r + \delta}{\eta_r + \delta + \chi_c} \right) \frac{\chi_c \pi F \mathbb{I}_{\{\epsilon_0 < 0\}}}{c_0} \right] \end{aligned}$$

Those expressions allow us to explicitly see the impact of the mortgage refinancing option on the elasticity of consumption to a small shock to interest rates.

We can also characterize analytically the full impulse response of expected prepayment rates $\mathbb{E}_0 [\rho_t]$ and expected coupon rates $\mathbb{E}_0 [m_t^*] - \delta$ to the time-zero shock to short rates. Expected prepayment rates upon a shock are simply equal to

$$\mathbb{E}_0 [\rho_t] = \mathbb{I}_{\{\epsilon_0 < 0\}} \chi_c e^{-\chi_c t}$$

$\mathbb{I}_{\{\epsilon_0 < 0\}}$ is the indicator function for whether the time-zero rate shock is positive or negative. Let us then compute the expected change in mortgage coupon, following the rate shock. Remember that m_t^* is the mortgage coupon of the household at time t . Notice that the current mortgage rate m_t satisfies

$$m_t = \delta + \pi \epsilon_0 e^{-\eta_r t}$$

At time zero, the initial coupon is $m_0^* = \delta$. Our simple model then allows us to compute the expected coupon change of the household:

$$\mathbb{E}_0 [m_t^*] - \delta = \int_0^t \chi_c e^{-\chi_c s} \pi \epsilon_0 e^{-\eta_r s} ds = \frac{\chi_c \pi \epsilon_0}{\chi_c + \eta_r} \left(1 - e^{-(\eta_r + \chi_c)t} \right)$$

Finally, in the case where the rate shock is negative ($\epsilon_0 < 0$), we can compute the change in consumption rate occurring at the time of an actual refinancing (which can in turn be mapped to our empirical and quantitative event studies of individual refinancing). Just before the refinancing event (assumed to occur at time τ), the consumption rate is equal to (at the first order, i.e. excluding terms that are of order $o(\epsilon)$):

$$c_{\tau-} = \bar{y} + \delta (w - F) - \frac{\epsilon}{\eta_r + \delta} \left[\frac{\bar{y} + \delta (w - F)}{\gamma} - \delta w + \left(\frac{\eta_r + \delta}{\eta_r + \delta + \chi_c} \right) \chi_c \pi F \right]$$

Just after the refinancing event, the consumption rate is equal to (at the first order, i.e. excluding terms that are of order $o(\epsilon)$):

$$c_\tau = \bar{y} + \delta (w - F) - \epsilon \pi F - \frac{\epsilon}{\eta_r + \delta} \left[\frac{\bar{y} + \delta (w - F)}{\gamma} - \delta w \right]$$

In other words,

$$c_\tau - c_{\tau-} = -\epsilon \pi F \left(\frac{\eta_r + \delta}{\eta_r + \delta + \chi_c} \right)$$

Since $\epsilon < 0$, consumption jumps upwards following the actual refinancing event.

A.4 Life Cycle Model Details

The extension of our benchmark model that includes life-cycle elements is structured as follows. First, all our households are home-owner households, just as in our benchmark model. Households transition stochastically between "young", "middle-age" and "old" at Poisson arrival times. We define young households in the data as having 25-40 years of age, "middle-age" as having 40-62 years of age, and "old" as those 62-75 years of age. Transition intensities from one age category to another are parametrized accordingly, such that households end up staying (i) young on average 15 years, (ii) middle-age on average 22 years, and old on average 13 years. Whenever a household dies, a young household is born with mortgage debt and assets as described below.

Thus, the (continuous time) generator matrix for the aging process is as specified in table (A-11).

We calibrate various data moments to mortgage holders in the 2001 SCF (see discussion of Figure A-13 for details of data definitions). We calibrate labor income and retirement income as well as mortgage debt and assets using data from the 2001 SCF. In particular, we compute that (i) young households

	“young”	“middle-age”	“old”
“young”	-1/15	1/15	0
“middle-age”	0	-1/22	1/22
“old”	1/13	0	-1/13

Table A-11: Generator matrix for life cycle model

earn 0.946 times average income, (ii) middle-age households earn 1.117 times average income, and (iii) old households earn 0.760 times average income. We choose average income so that ergodic average household income is equal to \$58,000, which corresponds to the ergodic average household income in our benchmark model. We assume that young and middle-age households face labor income risk, with log income that follows the continuous-time counterpart to an AR(1) process, where the speed of mean reversion and the local volatility are calibrated to match those estimated by [Floden and Linde \(2001\)](#). Old households are assumed to have constant retirement income. Young households are borne with assets of \$25,000 to match mean assets for the young age group in SCF and we also calibrate household’s relative mortgage balance again using data from the 2001 SCF. We estimate that (i) young households have 1.115 times average mortgage debt, (ii) middle-age households have 1.005 times average mortgage debt, and (iii) old households have 0.652 times average mortgage debt. We pick the average level of mortgage debt around which these relatives values are chosen so that the ergodic average mortgage debt outstanding is equal to \$150,000 (corresponding to the fixed mortgage balance in our benchmark model). When a household transitions from one age category to another, that household’s mortgage balance declines by the appropriate quantity; in each age category, households pay a per-annum principal amount on their mortgage such that the expected principal repayments made in that age category equals the expected notional reduction at the time of transition into an older age group. Our model thus matches broad life-cycle patterns of mortgage debt and income patterns, by construction. The age-specific relative income and mortgage balances are summarized in table [\(A-12\)](#).

	relative income	relative mortgage balance
“young”	0.946	1.115
“middle-age”	1.117	1.005
“old”	0.760	0.652

Table A-12: Relative income and mortgage debt by age group

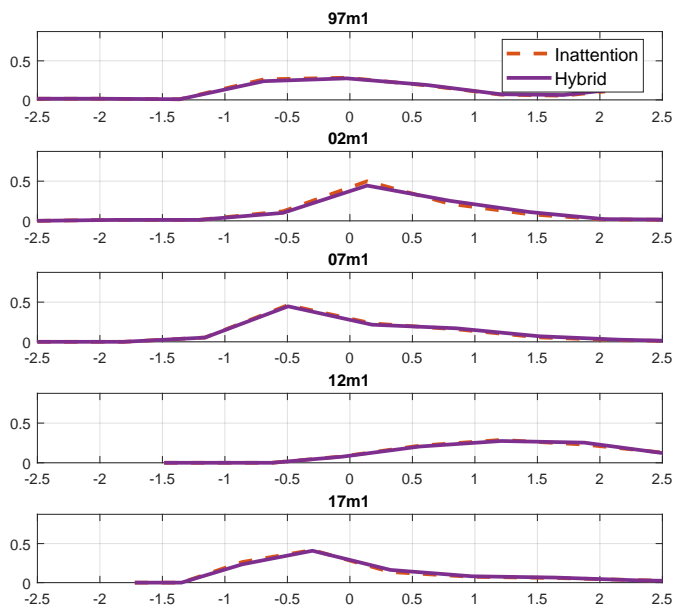
Old households also have a standard bequest motive $bW^{1-\gamma}/(1-\gamma)$, with parameter $b = 1$, as in [Cocco, Gomes and Maenhout \(2005\)](#). In [Figure A-13](#) we show that in addition to matching targeted moments for income and mortgage debt, this model also does a good job of matching untargeted life-cycle profiles of consumption and savings.

A.5 Additional Figures and Tables

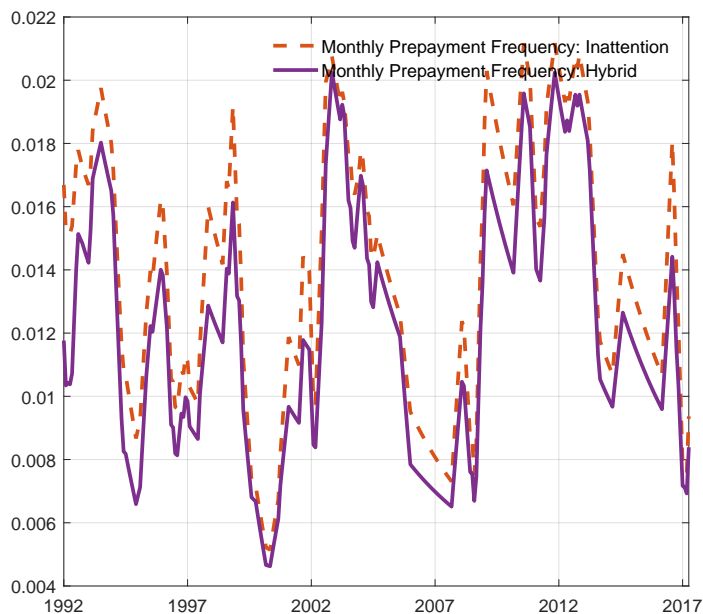
[Figure A-11](#) shows that the time-series predictions of the pure Calvo model are nearly identical to the hybrid model fit to more closely match the micro prepayment hazard. This is because the prepayment hazard for the Calvo model mostly misses the data for small gaps, and refinancing from a small gap to zero has little effect on a household’s mortgage coupon.

Figure A-11: Hybrid Menu Cost+Inattention Model vs. Inattention Model: Time-Series Fit

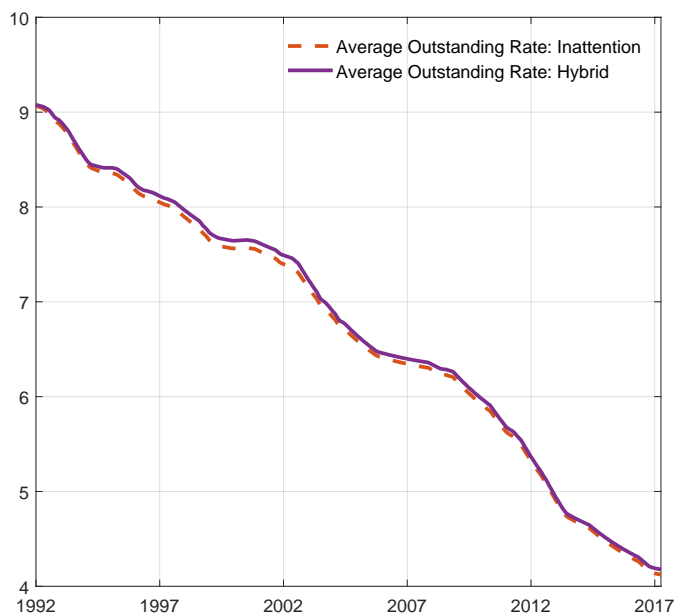
(a): Distribution of Gaps



(b): Frequency



(c): Average Rates



(d): Std Dev of Rates

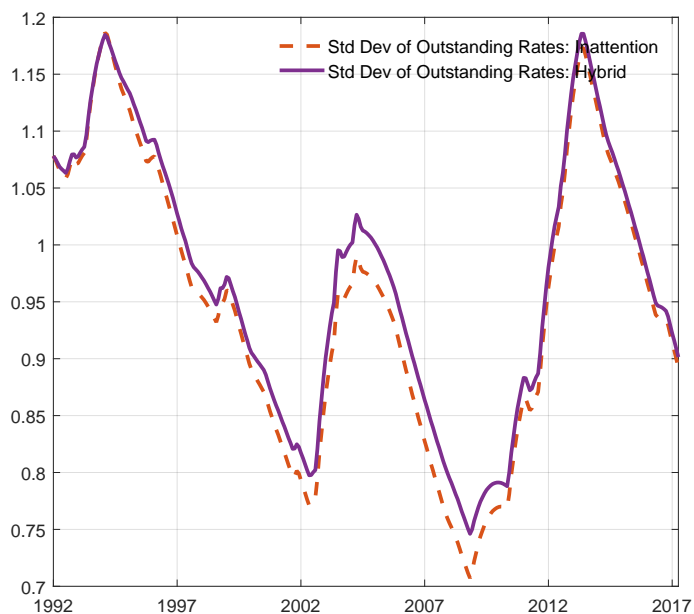


Figure A-12 shows the fit of our baseline Calvo model calibration in Section 6.5.2 to various observable consumption and wealth moments. Panel (a) compares our model implications for the distribution of consumption and wealth, as measured by Lorenz curves. We measure consumption using Consumer Expenditure Survey data from 1988-2018 for homeowners with complete earnings records, weighting spending in the calculation of Lorenz curves by household sampling weights. We measure wealth using 2001 SCF data, and we define liquid wealth in the data as cash, checking accounts, savings accounts, money market mutual funds and directly held stocks and bonds. Panel (b) compares our MPCs to estimates in Lewis et al. (2020). We normalize by the mean MPC and compute the 10th, 25th, 50th, 75th and 90th percentile and compare this to the same object from Lewis et al. (2020) Figure 1.

Figure A-12: Baseline Model Fit to Consumption and Wealth Data

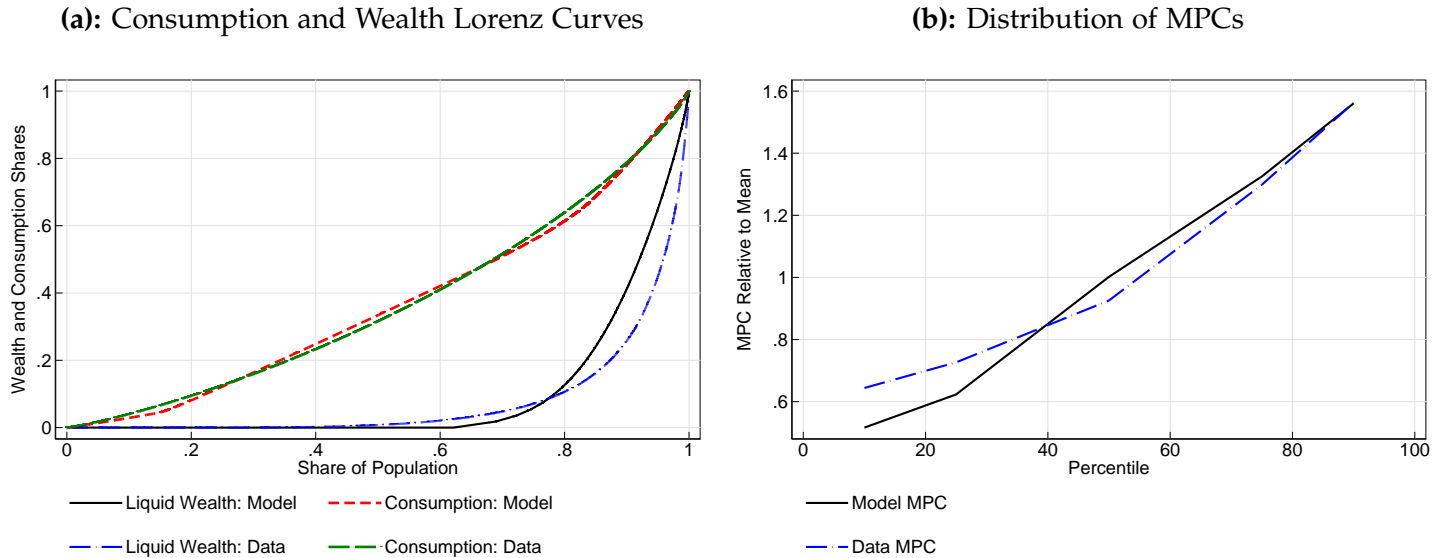


Figure A-13 shows the fit of our lifecycle model in Section 9 to various data moments. We define young homeowners as those 25-40, middle aged as those 40-62, and old households as those 62-75. We compute income, savings and mortgage debt in the 2001 SCF data. Since our model has no homeownership decision and focuses on refinancing, we restrict our analysis to homeowners with non-zero mortgage debt. We define income as labor earnings for the young and middle aged and additionally include broader non-labor income for old households. We define savings as total non-housing wealth since these lifecycle patterns are more straightforward to interpret as lifecycle savings, since they do not confound savings decisions with portfolio effects driven by e.g. the tax treatment and withdrawal rules on retirement accounts. Consumption is measured as total consumption in the consumer expenditure survey using households from 1988-2018. We keep only young and middle aged households with non-missing earnings records but do not impose this restriction on old households. The model fits income and mortgage debt exactly, by construction, while consumption and savings are untargeted.

Figure A-13: Life-cycle Model Fit

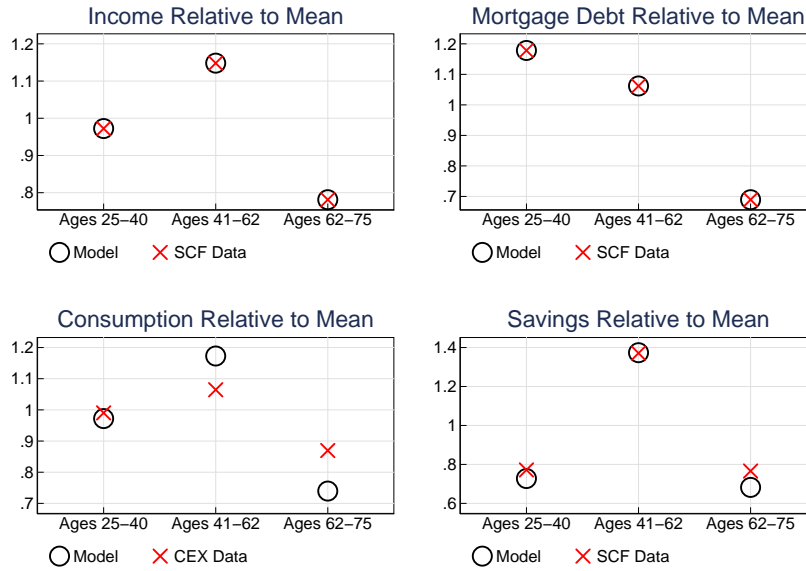


Figure A-14 shows the dynamics of the full IRFs for the regime-shift experiments in Section 7.5. We consider both a 100bps and “max” rate decline. We illustrated the IRF to both the 100bps and “max” shock for the baseline economy in Figure 11. Analogously, Figure A-14 shows the impulse response to both the 100bps and the “max” shock in the regime shift economies, noting that the regime shift occurs in period 0. The horizontal dashed red lines in Figure A-14 (which are identical to those in Figure 11) show the peak coupon response to the 100bps and the “max” shock in the “baseline” economy respectively, in order to help comparing coupon responses in the regime shift economies to the “baseline” economy without regime shift. The date at which the impulse starts indicates how far after the regime shift the monetary shock occurs. For example, the blue line in panel (a) shows the effect of cutting short rates by 100bps one year after the interest rate regime shifts up, while the blue line in panel (b) shows the effect of cutting short rates by 100bps one year after the interest rate regime shifts down.

There are a number of takeaways from Figure A-14. First, looking at 100bps responses on impact in black as compared to the baseline peak responses in red, we can see that 100bps short rate cuts initially have smaller effects on coupons in the “Rate-Shift-Up” economy than in the “Rate-Shift-Down” economy. This arises exactly from the effects emphasized in the previous experiments: when rates rise, ($frac > 0$) decreases and this reduces the effects of a given change in short rates. Moreover, as time passes since the regime shift, 100bps rate cuts become more powerful in the “Rate-Shift-Up” economy and less powerful in the “Rate-Shift-Down” economy.

Second, the maximum stimulus power (given by cutting short rates to zero) increases in the “Rate-Shift-Up” economy and decreases in the “Rate-Shift-Down” economy. A higher average level of interest rates naturally results in a larger max rate cut and resulting response. More interestingly, the difference between maximum stimulus power in the “Rate-Shift-Down” and “Rate-Shift-Up” economy grows with the time since the regime shift (e.g. the green lines in Panel (c) and (d) differ more than the black lines).

Figure A-14: Regime Shift: Average Coupon m^* to 100bps & Max decline in r

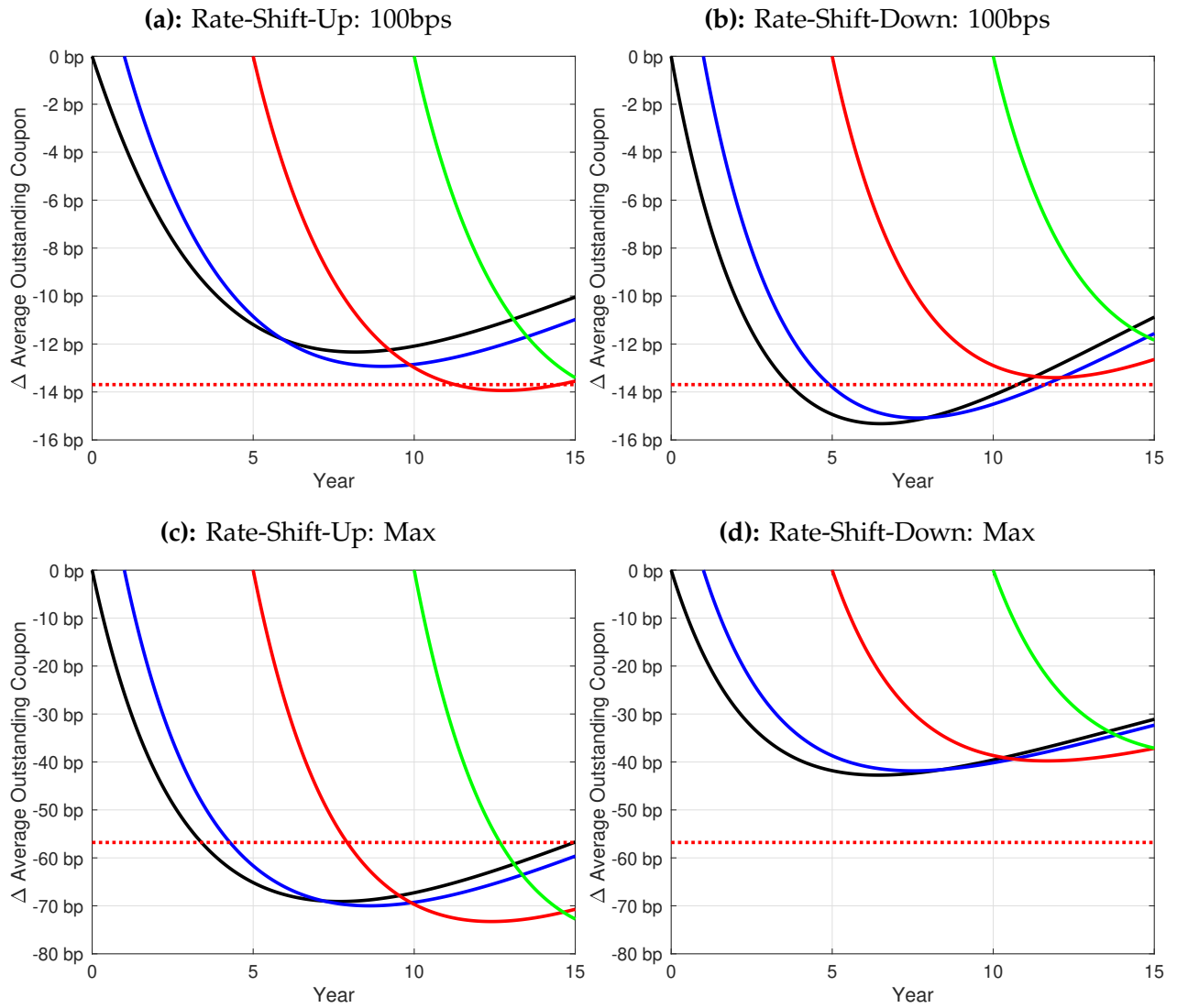


Figure A-15: Robustness of Consumption IRFs to Including Cashout Refi

(a): baseline vs. secular decline

(b): past high vs. past low rates

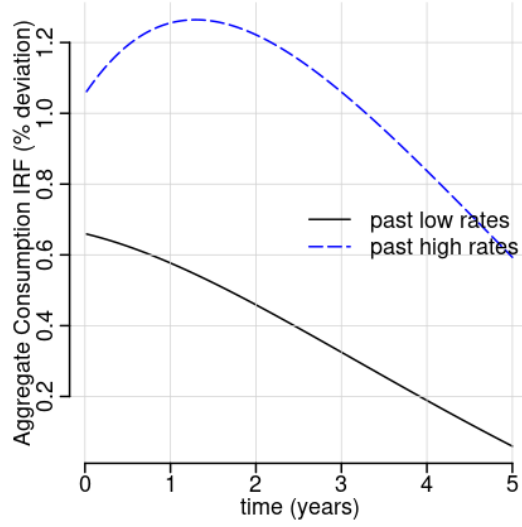
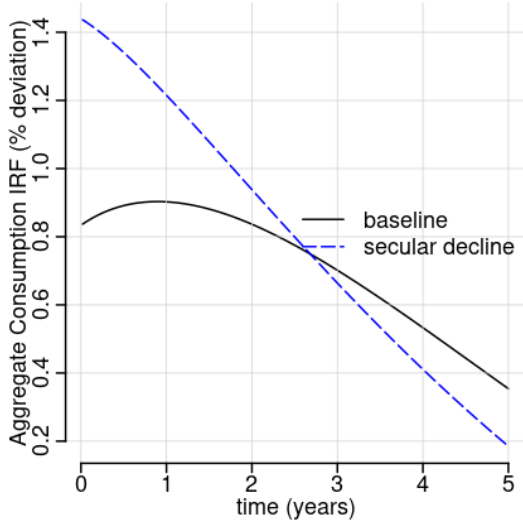


Figure A-16: Robustness of Consumption IRFs to Life-Cycle

(a): baseline vs. secular decline

(b): past high vs. past low rates

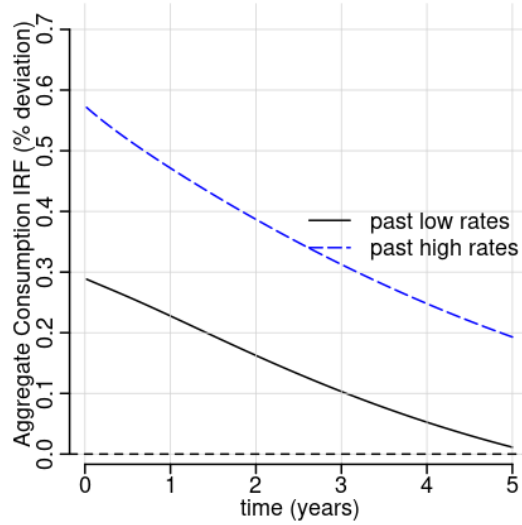
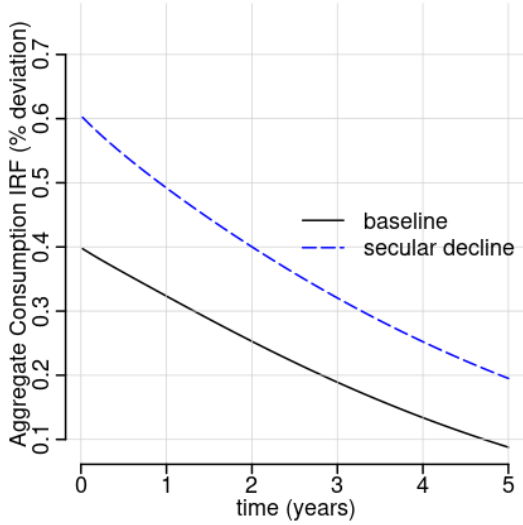


Figure A-17: Robustness of Consumption IRFs to Hybrid Frictions

(a): baseline vs. secular decline

(b): past high vs. past low rates

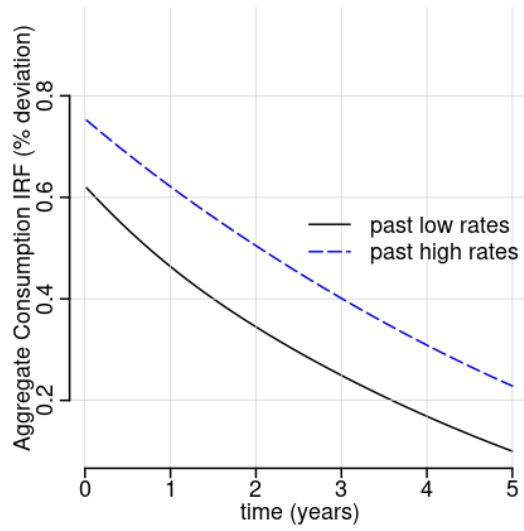
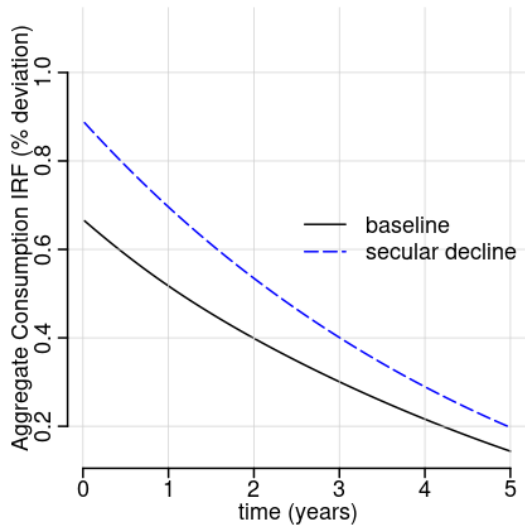
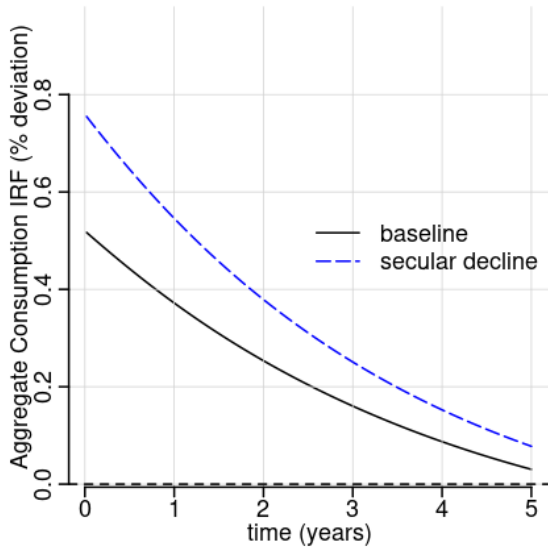
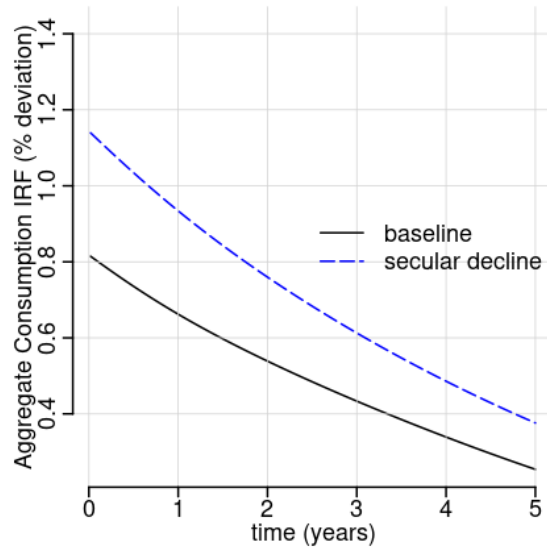


Figure A-18: Robustness of Consumption IRFs to Interest Rate Persistence

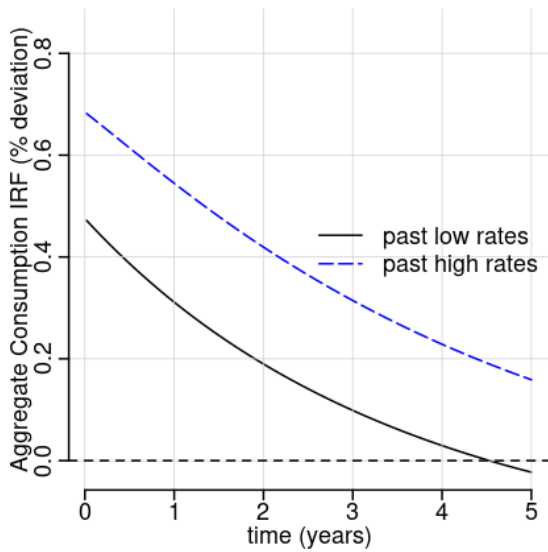
(a): Half-life $0.66 \times h$



(b): Half-life $1.33 \times h$



(c): Half-life $0.66 \times h$



(d): Half-life $1.33 \times h$

