Mortgage Prepayment and Path-Dependent Effects of Monetary Policy[*]

David Berger[†]   Konstantin Milbradt[‡]   Fabrice Tourre[§]   Joseph Vavra[¶]

February 2020

Abstract

How much ability does the Fed have to stimulate the economy by cutting interest rates? We argue that the presence of substantial debt in fixed-rate, prepayable mortgages means that the ability to stimulate the economy by cutting interest rates depends not just on their current level but also on their previous path. Using a household model of mortgage prepayment matched to detailed loan-level evidence on the relationship between prepayment and rate incentives, we argue that recent interest rate paths will generate substantial headwinds for future monetary stimulus.

Keywords: Monetary Policy, Path-Dependence, Refinancing, Mortgage Debt

JEL codes: E50, E21, G21

[†]Duke University and NBER; david.berger@duke.edu
[‡]Northwestern University and NBER; milbradt@northwestern.edu
[§]Copenhagen Business School; ft.fi@cbs.dk
[¶]University of Chicago and NBER; joseph.vavra@chicagobooth.edu.

1 Introduction

How much can the Fed stimulate the economy by cutting interest rates? There is growing evidence that mortgage refinancing plays an important role in the transmission of monetary policy to real economic activity.[1] We argue that the current strength of this channel will depend on the past history of interest rates: rate cuts can encourage borrowers to refinance their mortgages, but only if they have not already locked in lower fixed rates before. This means that past Fed decisions affect the sensitivity of the economy to today's actions, and today's actions in turn affect future "policy space".

We demonstrate the importance of this path-dependence using a heterogeneous agent incomplete markets model with prepayable fixed-rate mortgages, which we discipline using empirical patterns obtained from monthly panel data on millions of borrower credit records linked to those borrowers' mortgage loan information. This micro-data consistent model leads to a macro environment with complex non-linear dynamics and path-dependent transmission of monetary policy to the real economy. Despite these complicated dynamics, our model delivers a practical rule-of-thumb to help guide policy making: the fraction of outstanding loans with mortgage rates above the current market rate, a measurable object we refer to as $frac > 0$, summarizes information about past rates relevant for predicting current stimulus power. In addition to this current guidance about the sensitivity of the economy to rate changes, our model also provides simple predictions about how $frac > 0$ evolves under different hypothetical policy paths and thus guidance about how current actions will affect future policy space.

While our model can shed light on many different scenarios, we highlight several implications for policy making in the current macro environment: 1) The secular decline in mortgage rates over the last thirty years has steadily pushed up $frac > 0$ and the effectiveness of monetary policy over this time period. Policy makers should anticipate weaker responses to monetary stimulus in stable or increasing rate environments. 2) Monetary policy is less powerful today because rates were kept low for a long time after the Great Recession, during which time many households refinanced and locked in low fixed rate mortgages. 3) It will take longer for the Fed to reload its "ammunition" as rates return to normal than it took to use up its ammunition when it cut rates. This asymmetry stems from the fact that households avoid prepaying when rates increase but actively refinance when rates decrease. All three forces constrain the Fed's ability to stimulate the economy if it needs to in the near future.

We now discuss the empirical facts that guide our modeling choices and resulting policy conclusions. Using linked borrower-loan panel data, we begin by characterizing the prepayment hazard as a non-parametric function of the "rate gap" (the difference between the contractual mortgage coupon on a loan m^* and the current market interest rate m on similar mortgages). We find that the prepayment hazard exhibits a "step-like" shape: prepayment rates are low and constant for loans with negative rate gaps, increase sharply for rate gaps between 0 and 100bps and then plateau at around 2% per month for rate gaps above 100bps. This illustrates both the well-known state-dependent nature of prepayment rates (cf. Schwartz and Torous (1989)) but also the fact that most households nevertheless do not refinance even with strong rate incentives (cf. Keys, Pope and Pope (2016)). We contribute to this literature by estimating a non-parametric prepayment hazard which exploits linked borrower-loan data to isolate the influence of rate incentives separately from confounding factors such as borrower credit worthiness, loan

---

[1]Cf. Greenwald (2018), Wong (2019), Beraja et al. (2019)

age related "burnout" effects, leverage, as well as any permanent borrower heterogeneity.[2]

In the presence of state-dependent micro behavior like we observe for mortgage prepayment, it is well-known that the micro adjustment hazard plays a crucial role in determining how aggregate shocks transmit to macro outcomes.[3] In our context, this means that a credible model of interest rate transmission through mortgage markets must match the empirical prepayment hazard. After documenting its basic step-like shape, we dive deeper into the micro data to further inform our theoretical modeling. We show that the modest prepayment rates for loans with large positive gaps are not driven by refinancing constraints or by limited benefits from refinancing and instead suggest an important role for inattention and time-dependence, consistent with Andersen et al. (2019).

We then explore how different types of prepayment (rate refi, cashout refi and moves) respond to rate incentives. We focus mostly on total prepayment in both the data and model, since any prepayment resets a household's rate gap to zero and is thus equally relevant for determining the evolution of rate incentives over time. Nevertheless, different types of prepayment could respond differently to rate incentives, with different implications for modeling. With our borrower-loan linked data, we can match prepaying loans to newly originated loans by the same borrower, which allows us to construct separate hazards for each prepayment type.[4] Notably, we find that the probability of *both* rate and cash-out refinancing is very low for loans with negative rate gaps and both hazards exhibits step-like nonlinearities. Thus rate incentives are crucial for all refinancing decisions: even households taking cash out of their homes rarely do so absent a simultaneous rate decrease.[5] Most observed prepayment into higher rates instead occurs from households moving.

We turn next to time-series implications for aggregate mortgage prepayment. We show that the distribution of rate gaps varies substantially across time and predicts aggregate prepayment rates in a way consistent with the average loan-level hazard. In particular, the fraction of loans with positive rate gaps in the data, $frac > 0$, has key predictive power for aggregate prepayment rates. If the empirical hazard was an exact step function at 0, then $frac > 0$ would fully summarize all information about how rate incentives affect aggregate prepayment. In practice, $frac > 0$ predicts 92.5% of the variation in aggregate prepayment that can be explained using the entire distribution of rate gaps. This formal metric shows that the qualitative step-like prepayment hazard is a good quantitative fit for the data, which will in turn be a crucial feature guiding our modeling choice of prepayment frictions.[6] The strong time-series relationship we find between $frac > 0$ and prepayment is stable across time and very robust. It holds before, during and after the housing boom-bust, after controlling for a host of covariates and non-linearities, and when instrumenting for rate incentives to address endogeneity concerns. It is also robust to various measurement and sample selection issues, holds at the regional level and shows up after decomposing total prepayment into its constituent components.

Thus, rate incentives matter crucially for aggregate prepayment rates. We next argue that mortgage prepayment also matters for the transmission of interest rates into spending. Using an event study design similar to Beraja et al. (2019), we show that households are much more likely to buy a car after

---

[2]Our data also covers a much larger loan sample than typical studies, allowing for more flexible non-parametric estimation.

[3]cf. Caballero and Engel (2007)

[4]To our knowledge, there is no prior micro evidence on how different types of prepayment respond to rate incentives, since typical loan data sets can measure which loans prepay but not the reason for prepayment.

[5]We later reconcile this result with prior time-series evidence suggesting refinancing into higher rates is somewhat common.

[6]We also show $frac > 0$ has more predictive power than many alternative summary statistics for the gap distribution.

refinancing their mortgage.[7] To provide direct evidence that rate savings matter for spending, we then study how car buying interacts with the mortgage interest savings obtained when refinancing. We show that amongst refinancing households, those obtaining large rate savings are much more likely to buy a car than those obtaining small savings. Strikingly, this holds *both* for rate and cash-out refinancing. More generally, the increased purchase propensity after rate refinancing is 75-88% as large as after cash-out refinancing, indicating that both types of refinancing are sensitive to interest savings and matter for spending. We then document that these micro results extend more broadly to regional aggregates, using cross-region relationships between rate incentives, prepayment and regional auto purchases.

Finally, we provide empirical evidence that time-varying mortgage prepayment matters for aggregate monetary transmission. First, we use a simple back-of-the-envelope to argue that observed variation in mortgage prepayment can lead to transfers between borrowers and lenders worth hundreds of billions of dollars in present value, and thus plausibly matter for aggregate GDP. Second, we use a local projections approach to show that monetary policy shocks indeed have stronger effects on aggregate economic activity when $frac > 0$ is large. This suggests that the micro patterns we identify matter for monetary policy transmission to the real economy. However, this type of aggregate evidence is merely suggestive, since it does not isolate any one particular transmission mechanism. Furthermore, even if it did reveal precisely how $frac > 0$ matters for monetary transmission through prepayment, it would leave several crucial policy questions unanswered: How does $frac > 0$ evolve when the Fed changes rates? Are there non-linearities so that this evolution depends on the size of rate changes or their particular path?

In the second half of the paper, we build a theoretical framework to explore these questions and characterize how monetary policy affects aggregate spending through its effect on mortgage prepayment. Importantly, the goal is not to quantify all channels of monetary transmission or even all effects working through housing and mortgage markets. Instead, the goal of the model is two-fold: 1) Argue that the microeconomic patterns we document indeed have important implications for how aggregate spending responds to monetary policy. 2) Provide guidance about the potency of this rate incentive channel at a moment in time and how this potency will evolve in the future given current policy choices.

We explore the role of this rate incentive channel in an otherwise standard economic environment. In particular, we start with a continuous time incomplete markets consumption-savings model with labor income risk. To this standard setup, we add interest rate fluctuations and fixed rate mortgage debt which can be refinanced subject to some frictions, which we discipline using the micro evidence from the first half of the paper. Finally, we introduce a risk-neutral financial intermediary that offers competitively priced mortgage contracts, generating an endogenous equilibrium link between short rates and mortgage rates. This leads to an important role for redistribution in equilibrium: rate declines reduce debt payments for borrowers who refinance, but at the same time lower returns for lenders.[8]

Several other elements of our baseline model are kept intentionally simple: all households have identical constant mortgage balances and we abstract from cashout refinancing, lifecycle effects, and house price dynamics. Abstracting from these forces in our baseline model allows us to isolate and more precisely characterize the independent influence of refinancing frictions, which we can directly discipline using our micro evidence. However, we show in robustness results that adding additional

---

[7]While our empirical specification deals with some confounding concerns by controlling for borrower and time fixed effects, the timing of refinancing is clearly endogenous and so these may not be causal effects.

[8]Greenwald (2018) shows that these interactions have important aggregate implications in a representative agent framework.

richness along these dimensions complicates the model but leaves the main insights unchanged.

We allow for refinancing frictions emphasized in the literature by assuming that households get opportunities to refinance without cost at some random times, while at other random times households can refinance only by paying a fixed cost. For different parameterizations, this setup nests a pure state-dependent "menu cost" model, a pure time-dependent "Calvo" model of inattentive refinancing, as well as intermediate mixed frictions. To pin down their importance, we initialize models with different frictions to the actual 1992 loan-level distribution of mortgage rates, expose them to actual monthly mortgage rates from 1992-2017 and calibrate each model to match the average prepayment frequency in the data. We then study how each model fits untargeted time-series moments. We begin by comparing a pure Calvo attention model to a pure menu cost model. The Calvo model is a *much* better fit: it generates mortgage coupon distributions and prepayment patterns which track the data fairly closely across time, while the menu cost model generates many time-series patterns starkly at odds with the data. This poor time-series fit arises because a menu cost model implies a prepayment hazard at odds with the micro data: prepayment rates are too low for moderate positive gaps and too high for large positive gaps, since a large enough rate incentive leads almost everyone to refinance, despite our model including substantial household heterogeneity and scope for heterogeneous refinancing decisions.

We next show that a hybrid model with both frictions best matches the prepayment hazard but that its time-series implications are nearly identical to the pure Calvo model. However, the Calvo model has crucial advantages over this hybrid model. For counterfactual experiments, we must solve the financial intermediary's mortgage pricing problem to endogenize mortgage rates and capture redistributive effects between borrowers and lenders. For general frictions, this requires treating the endogenous joint distribution of households as a state-variable. Under the Calvo model, a key simplification arises: household prepayment decisions are orthogonal to consumption-savings decisions, so new mortgage contracts can be priced without knowing household state-variables. This then allows us to pin down monetary policy transmission to mortgage outcomes and disposable income before specifying features of the model's "consumption block" like preferences, labor income, and wealth.[9] Resulting equilibrium mortgage outcomes are then easily integrated into the more standard consumption block of the model. Given the dramatic computational advantages of the Calvo model over the hybrid model despite similar observable implications, we focus primarily on this specification for most of our results.

We show that our model generates non-linear, path-dependent implications for monetary policy, which cannot be easily captured with reduced form statistical relationships. In addition to its numerical advantages, our baseline Calvo model provides a second key advantage: it allows us to transparently characterize this path-dependence and provide concrete policy guidance. Adapting results in Caballero and Engel (2007) to our continuous time setting, we show that in this model, the current value of $frac > 0$ encodes the information about past rates necessary to predict how average mortgage coupons respond to current shocks. This model also delivers simple solutions for the dynamics of $frac > 0$, making it easy to determine how the Fed's current actions will affect its future ability to stimulate mortgage markets. However, our key outcome of interest remains aggregate demand. We find in our numerical results that the implications for mortgage outcomes extend to broader stimulus power: monetary policy has large

---

[9]This separation between liquidity and rate motives for refinancing might seem like a disadvantage, but it is useful to again highlight the empirical observation that refinancing without rate incentives is rare. Models with strong liquidity motives for refinancing instead typically imply that refinancing into higher rates is quite common (Chen, Michaux and Roussanov (2019)).

and state-dependent effects on aggregate spending which are tightly connected to $frac > 0$.

While these spending effects cannot be characterized analytically in our incomplete markets environment, we solve in closed-form a simplified complete markets version of our model to provide additional intuition for the magnitude of various channels at work. We characterize the semi-elasticity of consumption to short rates and show that these responses can be amplified substantially by the presence of fixed-rate prepayable debt. However, the strength of this prepayment channel depends crucially on the pass-through from short rates to mortgage rates, household asset positions, and how refinancing frictions interact with rate incentives. These are central, endogenous components of our main model.

The strong theoretical relationship between $frac > 0$ and responses to rate changes together with the dynamics of $frac > 0$ implied by our prepayment model naturally deliver all the policy implications discussed earlier: 1) $frac > 0$ is large when rates are trending down, which increases responsiveness to monetary policy. 2) $frac > 0$ is small if previous rates were lower than today, which decreases responsiveness to monetary policy. 3) Households actively refinance when rates fall but avoid prepaying when rates rise, so monetary policy uses its ammunition up more rapidly when lowering rates than it recovers it when raising rates. All of these forces will constrain the Fed in the near-term.

We explore various robustness checks in the Calvo refinancing framework with endogenous mortgage rates and in more complicated environments using exogenous rates. Our basic conclusions are robust to introducing: 1) cash-out refinancing and thus heterogeneous time-varying mortgage balances; 2) life-cycle effects and inflow of new mortgages from population growth; 3) more complex forms of refinancing frictions that improve further model fit; 4) various alternative processes for the dynamics and persistence of short term interest rates and resulting endogenous passthrough to mortgage rates.

2   Related literature

A large empirical literature using loan-level data shows that rate incentives matter for prepayment.[10]. Our linked borrower-loan data lets us extend this prior literature in several important ways. These links allow us to isolate the role of rate effects on prepayment independent from confounding factors emphasized in the literature such as "loan age", "burnout" or permanent heterogeneity. They also let us measure the sensitivity of some durable spending outcomes to refinancing and to decompose prepayment hazards into subcomponents (rate refi, cash-out and moves), in order to show the crucial role of rate incentives. Finally, our sample has both broader loan coverage and longer time dimension than typical studies. This allows us to estimate a non-parametric hazard and quantify its implications for aggregate prepayment over time. We show that the hazard, as a function of interest rate gaps, exhibits a non-linear shape not well-described by simple, commonly used quadratic or cubic relationships. On the micro data front, we relate most closely to Andersen et al. (2019). They use Danish rather than US mortgage data and a different estimation strategy to identify refinancing frictions. While their work studies only mortgage outcomes and not spending, they reach conclusions similar to ours about the important role of time-dependent in addition to more standard state-dependent frictions. We further extend this literature by exploring the macroeconomic spending implications of these microeconomic relationships, showing that they lead to important path-dependent consequences of monetary policy.[11]

---

[10]Cf. Green and Shoven (1986), Schwartz and Torous (1989), Deng, Quigley and Order (2000) for some prominent examples.

[11]Andersen et al. (2019) briefly explores the effects of monetary policy on aggregate refinancing under some counterfactual mortgage systems, but the focus of the paper is on estimating microeconomic frictions.

A large literature argues that mortgage markets matter for monetary policy transmission.[12] Our central argument – the fact that time-varying refinancing incentives lead to time-varying effects of monetary policy – is similar to insights in Beraja et al. (2019). They focus on variation in refinancing incentives which arise from house price movements and resulting home equity, while we focus on interest rate incentives. This distinction is crucial: interest rates and resulting rate incentives respond almost immediately to monetary policy while house prices are indirectly and more slowly affected by monetary policy.[13] This means that the current distribution of rate gaps and the effectiveness of monetary policy is very directly influenced by the past history of interest rates, and it is this intertemporal feedback between today's actions and tomorrow's rate gaps and policy effectiveness that distinguishes our results from prior studies in which time-varying monetary policy effectiveness is driven by exogenous shocks.[14]

Monetary policy transmission in our model relates closely to the interest rate exposure channel in Auclert (2019). In our model, households' maturing liabilities and interest rate exposure depend on mortgage prepayment decisions and thus the distribution of rate gaps. Since this distribution depends on past rates, interest rate exposure and monetary policy effects are path-dependent. We focus on aggregate spending effects arising from these changing mortgage payments but note that monetary policy also has separate, welfare-relevant redistributional consequences from inflation and other channels.[15]

Our paper also relates to concurrent work in Eichenbaum, Rebelo and Wong (2019). They make similar arguments for state-dependent monetary transmission through refinancing. Our paper begins with micro data analysis, with a focus on the prepayment hazard and its implications for the entire distribution of rate incentives over time. Their data work uses regions as the unit of observation, similar to the second half of our empirical analysis. Our richer micro data in turn motivates a focus on different frictions as a source of infrequent refinancing: they model households that face a fixed cost of refinancing, while we focus mostly on inattention. Inattentive refinancing can help explain the empirical evolution of the loan-level rate distribution over time and makes it feasible to calculate equilibrium counterfactuals with endogenous mortgage rates and borrower-lender redistribution. While we include cash-out refinancing and lifecycle elements in some robustness results, their partial equilibrium model of borrower behavior is nevertheless richer than ours: it includes decisions about home ownership and house sizes, movements in aggregate house prices and income with interest rates, and finite duration rather than perpetual mortgage contracts. We thus view their richer quantitative model as complementary to ours and find it reassuring that our simplifications do not appear central for understanding path-dependence.

3 Data description

We briefly describe our primary mortgage-related data here. Appendix A.1.1 provides additional details as well as discussion of other data used in our analysis.

Our main prepayment measures come from BKFS McDash loan origination and mortgage servicing records from approximately 180 million loans over the period 1992-2017. This data set includes detailed

---

[12]See Di Maggio et al. (2017), Agarwal et al. (2017), Greenwald (2018), Wong (2019), and Beraja et al. (2019).

[13]Gertler and Karadi (2015) finds pass-through of current FFR (one-year rates) into mortgage rates of 0.27 (0.54-0.80) using FOMC surprises. This high-frequency identification literature also explores real vs. nominal pass-through, effects of expected rates vs. risk premia and decomposes transmission into rate/information effects (cf. Nakamura and Steinsson (2018)). These distinctions are unimportant for us: we need only the simpler fact that Fed policy moves nominal mortgage rates.

[14]See e.g. Vavra (2014) and Beraja et al. (2019)

[15]See Doepke and Schneider (2006), Doepke, Schneider and Selezneva (2015)

information on loan characteristics such as current interest rate and unpaid balances, appraisal values at origination, type of loan (rate-refi, cash-out, purchase), indicators for prepayment and borrower FICO scores. We measure prepayment shares as the fraction of all fixed rate first liens in the McDash Performance data set in a month with voluntary prepayment indicators.[16] While the data set provides information which distinguishes rate-refi, cash-out and new purchases at the time of loan origination, these are not available at the time a loan is closed due to prepayment. This this data can measure prepayment but it cannot separately measure rate refinancing, cash-out and moves.

To distinguish between different types of prepayment and measure additional individual level outcomes and covariates, we use information from the Equifax Credit Risk Insight Servicing McDash (CRISM) data set. This data set merges McDash mortgage servicing records with Equifax credit bureau data and is available from 2005 onwards. The structure of the data set makes it possible to link multiple loans by the same borrower together, which cannot be done with mortgage servicing data alone. This lets us link the loan being paid off with any potential new loan and thus measure the reason for prepayment to distinguish refinancing from moves. It also lets us measure equity extraction through cash-out refinancing. For time-series analysis prior to 2005 when CRISM starts, we infer the frequency of rate, cash-out and prepayment from moves by multiplying the origination shares of each type by the overall prepayment frequency. Appendix Figure A-1 validates this procedure in the post-2005 data.[17]

The CRISM data set links every loan in the McDash data set to an individual, and covers roughly 50% of outstanding US mortgage balances. Prior to 2005, the McDash data set has somewhat lower coverage, ranging from 10% market coverage in the early 90s to 20-25% in the late 90s. As a measure of representativeness and external validity, Appendix Figure A-2 shows that refinancing in our data closely tracks the refi applications index produced by the Mortgage Banker's Association from 1992-2017.[18]

4   The Prepayment Incentive: Empirical Evidence on Mortgage Outcomes

Our analysis begins with a number of new empirical results relating economic activity to refinancing incentives. In this section, we focus on mortgage related outcomes such as prepayment and payment changes; we begin with household loan-level evidence and then turn to aggregate relationships. Next, in Section 5 we look at implications for spending related outcomes; we begin with household-level event studies, then show results for region-level auto spending, and finally show aggregate time-series evidence that refinancing incentives affect the response of broad GDP to monetary policy shocks.

4.1   Overall Loan Level Prepayment Patterns

We begin our empirical analysis by computing the distribution of loan-level "rate gaps" and its relationship to prepayment, pooling all monthly observations in the CRISM data from 2005-2017. For each

---

[16]Results are very similar with alternative samples, see Appendix A.2. We equally weight mortgages, but redoing all results weighting by balances produces nearly identical results. In line with our model setup, our prepayment indicator does not include default as prepayment. This distinguishes our results from those using MBS pools to estimate prepayment.

[17]We measure origination shares using CoreLogic LLMA data because McDash data has limited loan origination info prior to 1998. The CoreLogic data is very similar to the McDash data set but its performance data does not include prepayment information prior to 1999 and cannot be linked to households. Thus, we focus primarily on CRISM/McDash data and use information from CoreLogic data in only a very limited way.

[18]We measure originations while this index measures applications. According to LendingTree, denials are $\approx$ 8% after Dodd-Frank related changes in lending standards; this explains level difference after the Financial Crisis.

outstanding loan i in month t we define the rate gap as $gap_{i,t} = m^*_{i,t} - m_{i,t}$, where $m^*_{i,t}$ is the current interest rate on the outstanding loan and $m_{i,t}$ is the predicted rate for a new fixed rate loan originated in period t given borrower/loan characteristics for FICO and LTV.[19]

Our main outcome of interest in both the model and the data is total prepayment rather than a particular subset of prepayment (rate refi, cashout refi, or moves). Any prepayment resets the rate gap to zero, so all prepayment types are equally relevant for determining the distribution of rate gaps and its evolution across time. After measuring $gap_{i,t}$, we sort loan-months into 20 bp wide bins and estimate a non-parametric relationship between prepayment and rate gaps using the following regression:

$$prepay_{i,j,t} = \mathbb{1}(gapbin)_{j,t} + X_{i,j,t} + \delta_i + \epsilon_{i,j,t} \tag{1}$$

where $\mathbb{1}(gapbin)_{j,t}$ is a dummy for the gap bin of household i with loan j in month t, $X_{i,j,t}$ is a vector of loan and household-level characteristics and δ_i is a household fixed effect.[20] This specification controls for a number of observables which affect both prepayment and rate incentives in order to isolate the pure rate incentive effect, which can be most directly influenced by interest rate policy.[21]

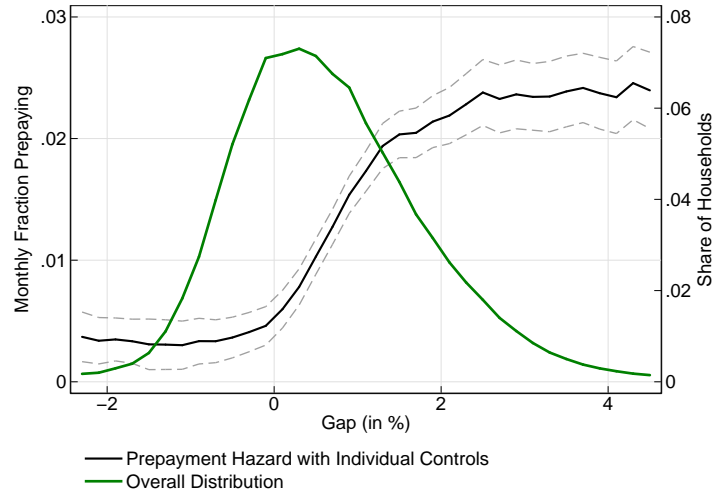**Figure 1: Prepayment Hazard with Individual Controls**

Figure shows point estimates and 95% confidence intervals for coefficients on the 20bp bin dummies in regression 1. Standard errors are two-way clustered by household and month using CRISM data linked to credit records from 2005m6-2017m4.

Figure 1 shows the resulting prepayment hazard and distribution of gaps. The first observation is that there is clear evidence of state-dependent prepayment: loans with positive gaps are much more likely to prepay than loans with negative gaps. While our data set has broader coverage than many prior studies, this result is not new. We instead emphasize the particular shape of the prepayment hazard (which again

---

[19]We assume $m_{i,t}$ is equal to the average market 30-year FRM in month t 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. Instead using a common rate for all new loans in month t delivers very similar results.

[20]$X_{i,j,t}$ controls are: a quadratic in FICO, a quadratic in current LTV, a quadratic in loan age and dummies for whether the current loan was itself a new purchase loan, a cash-out refi or a rate refi, dummies for investor type (GSE, RFC, GNMA, on-balance sheet, private MBS), loan type (FHA, VA, conventional w/ and w/out PMI and HUD).

[21]We estimate a linear probability specification both for numerical efficiency and to reduce concerns about incidental variables bias in non-linear models with fixed effects. However, we have found similar results when using logit specifications; in practice the bias is likely relatively small since most borrowers have a long monthly time series. We also find similar effects when estimating stratified regressions by frequency groups rather than running fixed effect specifications.

is obtained using rich controls to isolate the role of rate incentives from confounding forces): the hazard is low and stable for negative rate gaps, then rises rapidly as gaps become positive before stabilizing again at just over 2% monthly. This implies that prepayment rates are very state-dependent, but most households still do not prepay even when their rate gaps are very large.[22] Furthermore, since Figure 1 controls for both household fixed effects and loan age, this shape is not driven by the burnout effects and permanent heterogeneity often emphasized in the literature.[23]

For a model to deliver credible policy analysis, it must match this "step-like" shape of the prepayment hazard as function of rate gaps. In particular, it must deliver a low and stable hazard when gaps are negative, and stable, higher hazard when rate gaps are large. Two straightforward explanations might rationalize the low sensitivity of prepayment to rate incentives when gaps are large: 1) Lender constraints: perhaps households with large rate gaps would like to refinance but are prevented from doing so by lenders. 2) Small benefits: maybe the benefits of refinancing for households with large gaps are actually small, because they have small balances or small remaining mortgage durations. We now show that neither explanation is supported by the data. Instead (and consistent with empirical evidence in Andersen et al. (2019); Keys, Pope and Pope (2016)), we argue that these patterns are more plausibly explained by inattention to mortgage rates and suboptimal refinancing decisions.

4.1.1 Lender-based Refinancing Constraints?

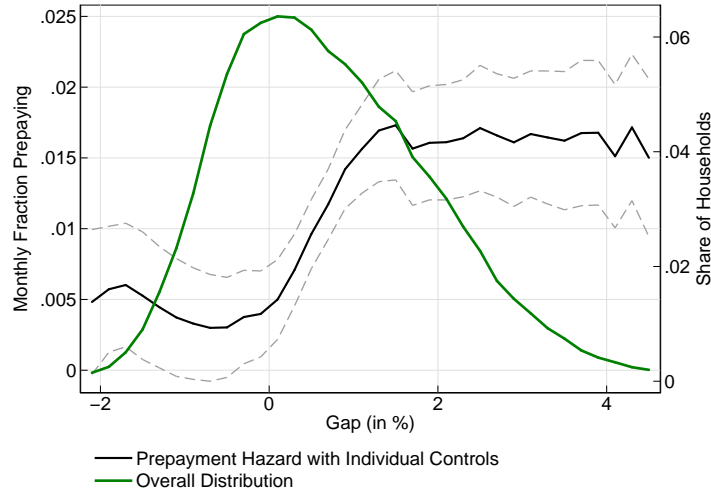**Figure 2: Prepayment Hazard with Individual Controls: Unconstrained Households**

Figure shows point estimates and 95% confidence intervals for coefficients on the 20bp bin dummies in regression 1. Standard errors are two-way clustered by household and month. The figure includes only borrowers with FICO>650 and LTV<0.65 and excludes data from the years 2007-2011.

Maybe households with large rate gaps would like to refinance but are prevented from doing so by lenders due to a high LTV ratio, a low credit score, or low income (DeFusco and Mondragon, 2020). Figure 2 suggests that such lender constraints do not drive our results. In this figure, we recompute

---

[22]While we focus on monthly hazards, the time unit of analysis does not drive our conclusions. See Appendix Figure A-3.

[23]Cf. Deng, Quigley and Order (2000). Importantly, although the regression has household fixed effects, we have a household panel and thus within household variation in gaps across time, so the regression is identified even if individual households only prepay once rather than multiple times. Indeed, separately recomputing Figure 1 only including households with one, two, or three plus prepayment events produces very similar results.

Regression ([1](#)) but we now only include households with high FICO scores (>650), low LTV (<0.65) and we exclude the time-period 2007-2011 when households were more likely to be unemployed. These households can very likely obtain a new loan if they want to refinance. The results are somewhat noisier since this is a more restrictive sample with a shorter time dimension, but they are qualitatively unchanged relative to the baseline. Thus, the shape of the prepayment hazard is not caused by lender-driven constraints.

Barriers to refinancing might also arise from liquidity constrained households who cannot pay various costs like appraisal and title fees incurred when refinancing. We do not observe household balance sheets or income, so we cannot totally rule this out, but the institutional setting makes this explanation unlikely: costs associated with refinancing can typically be rolled into the new loan or paid for using negative points. Thus, households can typically refinance with little out-of-pocket cost. This is especially true for households with substantial home equity as in Figure 2.

4.1.2   Limited Benefits from Refinancing?

It might also be the case that our results are explained by heterogeneity in mortgage balances: if most households with large rate gaps have low mortgage balances or remaining durations, their savings from refinancing may be low despite a large rate gap. We show this is not the case. Focusing just on loans with very large rate gaps (above 300 bps) which do not prepay, the average (median) annual foregone mortgage payment savings are $2800 ($2050) with an average (median) remaining mortgage duration of 22.2 (20.8) years.[24] Furthermore, while Figure 1 controls for mortgage balances, Appendix Figure A-4 further shows that restricting our analysis to households with substantial outstanding mortgage balances delivers very similar results. Thus, the low sensitivity of the prepayment hazard to rate incentives for large gaps cannot be rationalized by a limited refinancing benefit for these loans.

4.2   Prepayment Determinants: Rate Incentives vs. Equity Extraction

While we mainly focus on total prepayment, we also study the three prepayment sub-components: (1) Moves, (2) Cash-out refinancing, and (3) Rate refinancing (all refinancing without cash-out). This empirical exercise allows us to analyze the potentially different role of rate incentives for each prepayment type, and in turn influences our modeling choices. While many papers estimate prepayment hazards using loan-level data, to the best of our knowledge, we are the first paper to estimate relationships between prepayment type and loan-level rate incentives. We can perform these calculations since our CRISM data links loans to borrowers. This lets us match prepaid loans to information on new loans originated by the borrower at the same time, which allows us to determine the type of prepayment. Figure 3 shows that refinancing incentives strongly affect all three components of prepayment.

The first takeaway is that rate incentives are a crucial driver of refinancing decisions, even for households taking cash out of their homes: the probability of *both* rate and cash-out refinancing is very low for loans with negative rate gaps. Most prepayment into higher rates occurs when households move rather than when they refinance. For the 2005-2017 period covered by our CRISM data where we can precisely measure which loans refinance and their realized rates changes, only 6.3% of all refinancing loans do so

---

[24]Appendix Figure A-5 shows the entire distribution of potential payment reductions for these loans. The vast majority could obtain substantial payment reductions by refinancing.

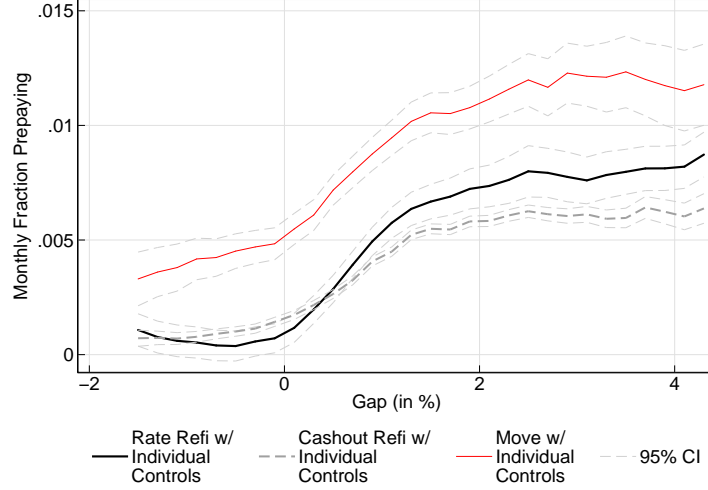**Figure 3: Prepayment Hazard Decomposition with Individual Controls**

Figure shows the point estimates and 95-percent confidence intervals for the coefficients on the 20 basis point gap bin dummies in regression 1 performed separately for each type of prepayment. Standard errors are two-way clustered by household and month.

into a higher rate and only 3.3% raise rates by 50 bps or more. Focusing only on cash-out refis, there are more rate increases but they are still infrequent: 14.7% of cash-out refis lead to any rate increase and only 7.7% lead to a rate increase of 50 bps or more. Extending to the 1992-2017 sample under an additional stability assumption implies that the share of all refis (cash-out refis) into any higher rate is 7.0% (14.0%) and into a 50 bps or higher rate is 3.2% (6.6%).[25] Furthermore, much of the small share of refinancing into higher rates that does occur is concentrated in the unusual 2004-2006 boom period.

In sum, Figure 3 makes clear that rate incentives are a crucial driver of *all* types of refinancing decisions. Our earlier results show that a strong rate incentive is far from sufficient to explain refinancing behavior. The results in this section show that rate incentives are almost a necessary condition for refinancing. This strong interaction between rate incentives and cash-out refinancing is directly in line with the evidence in Bhutta and Keys (2016), but in our case using individual loan-level measures of rate incentives rather than a proxy that relies on current aggregate rates.

In one sense, this result is unsurprising: if a household wants to tap home equity, taking out a HELOC will typically dominate refinancing an entire first lien into a higher rate, unless the HELOC-mortgage rate spread is very large. However, our evidence that few loans refinance into higher rates might nevertheless seem to be at odds with evidence that in many months, most refinancing loans are doing so into higher rates.[26] It is not. Appendix Figure A-6 shows that we replicate this time-series evidence almost perfectly in our data. Appendix Figures A-7 and A-8 show that changes in refinancing frequency are key to explaining these joint cross-section and time-series patterns: the times when most refis result in rate increases are precisely the times when the frequency of refinancing is extremely low, so the loans refinancing in these months are a small share of overall refinancing activity.

Further exploring the determinants of prepayment behavior, Figure 4 shows the prepayment hazard

---

[25]These calculations assume the hazard prepayment type from 1992-2005 is the same as what we estimate in the post 2005 CRISM data. We cannot fully test this assumption, but Appendix Figures A-6 shows that median implied rate changes from this procedure line up extremely closely with published data from Freddie Mac.

[26]See for example Chen, Michaux and Roussanov (2019) Figure 1.

11

as a non-parametric function of a loan's current LTV ratio rather than interest rate gaps.[27] Hazards clearly depend on the LTV ratio, but mostly for underwater households. These are a small share of all loans, and we have already shown in Section 4.1.1 that they do not drive the non-linear relationship between rate incentives and prepayment. For more typical households, there is a modest hump-shaped relationship between home equity and prepayment probabilities, so that prepayment rates are highest for those with some home equity but who are not close to paying off their homes. These effects, however, are small in magnitude compared to the effects of rate incentives we emphasize.

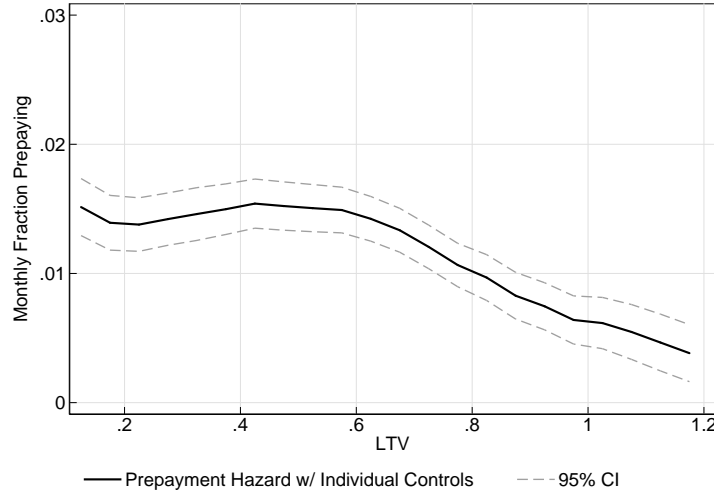**Figure 4: Prepayment Hazard by Leverage with Individual Controls**

Figure shows the point estimates and 95-percent confidence intervals for the coefficients on leverage dummies in regression 1. Standard errors are two-way clustered by household and month. In order to include household fixed effects and time-varying characteristics, figure uses CRISM data linked to credit records from 2005m6-2017m4.

4.3   Time-Series Variation in Rate Incentives and Aggregate Prepayment

The micro evidence thus far shows a strong non-linear relationship between rate gaps and prepayment propensities when pooling the data across all months. We next move from these pooled relationships to time-series analysis in order to show that: 1) The distribution of rate gaps varies substantially across time. 2) This time-series variation strongly predicts time-series variation in prepayment in a way that is easily summarized given our micro patterns. In particular, since our prepayment hazard exhibits a "step-like" pattern, our preferred summary statistic for the distribution of gaps at a point in time is the fraction of loans with positive rate gaps, which we label as $frac > 0$. This statistic is important for the following reason: if the empirical hazard was an exact step function at 0, then $frac > 0$ would capture all relevant information about the gap distribution necessary to predict current aggregate prepayment rates. Figure 5 shows that $frac > 0$ moves substantially across time, ranging from less than 0.2 in early 2000 to nearly 1 in 2003 and 2010.Figure 5 also shows that $frac > 0$ is highly correlated with the fraction of loans prepaying in a given month, with a correlation of 0.53.

Table 1 uses a time-series regression to more explicitly assess the predictive content of $frac > 0$ for aggregate prepayment relative to alternative moments of the rate gap distribution. Column 1 shows that there is a significant positive relationship between $frac > 0$ and prepayment. The R^2 of 0.282 means

---

[27]We also control non-parametrically for rate gaps, but the hazard as a function of LTV is insensitive to these controls.

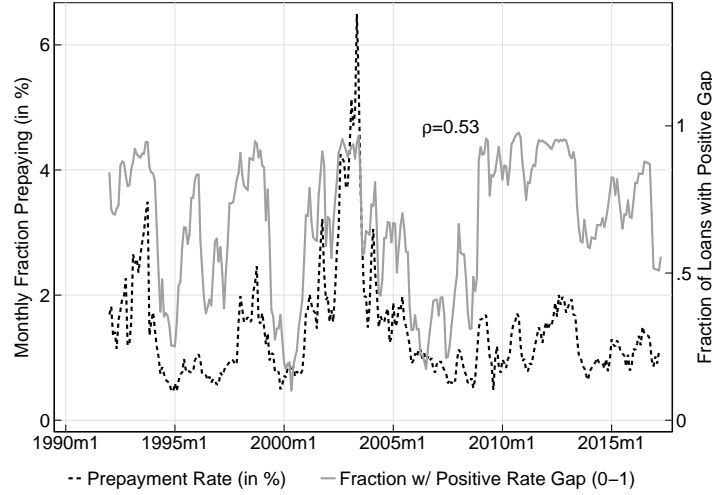**Figure 5: Prepayment vs. Fraction with Positive Rate Gap Time-Series**

Figure shows the fraction of loans in McDash Performance data with positive rate gaps in each month as well as the fraction of loans prepaying in each month. The time sample is 1992m1-2017m4.

that this single variable explains just under thirty percent of the time-series variation in prepayment. Interestingly, if we run a regression of monthly prepayment on the monthly fraction of loans in the full set of bins used in Figure 1, the R^2 rises only from 0.282 to 0.305. This means that $frac > 0$ captures 92.5% (.282/.305) of the relevant information in the full distribution of rate gaps for predicting current prepayment. Since $frac > 0$ would capture 100% of the relevant information if the hazard were exactly a step function at zero, such a hazard is thus a close approximation to the data in terms of predictions for aggregate prepayment rates. The remaining columns of Table 1 explore the predictive power of alternative statistics summarizing the distribution of rate incentives. While these moments are highly correlated, they are all less successful (as measured by R^2) at predicting prepayment than $frac > 0$.

Having established the basic time-series relationship between $frac > 0$ and prepayment, the remainder of this section explores various robustness checks and additional outcomes including: 1) controlling for a host of additional covariates, 2) allowing for time-varying or non-linear relationships 3) looking at relationships across regions, 4) instrumenting for rate incentives to address endogeneity concerns, 5) decomposing total prepayment into its constituent components and analyzing changes in coupons in addition to the frequency of prepayment, and 6) various measurement and sample selection issues.

The upshot of this analysis is that $frac > 0$ has very strong predictive power for prepayment and all of its subcomponents that is remarkably stable across time and little related to other observables. The only other covariate we find with predictive power for prepayment is the LTV ratio, and it matters mostly during the housing boom-bust, and only after controlling for the role of rate incentives, as captured by $frac > 0$. We briefly discuss results in the text but relegate detailed analysis to Appendix A.2.

4.3.1 The Housing Boom-Bust and the Stable Role of Rate Incentives

Motivated by the analysis in Beraja et al. (2019) and the micro effects of LTV ratios shown in Figure 4, in this section we explore the role of the LTV ratio and whether its inclusion in predictive regressions changes the relationship between $frac > 0$ and prepayment. Table 2 Column (2) shows that a greater LTV ratio is indeed associated with lower prepayment rates. Looking at the R^2 implies that the average LTV

ratio and $frac > 0$ together explain roughly half of the time-series variation in prepayment.[28] However, Column (3) shows that the explanatory power of the average LTV ratio by itself is much lower, with an R^2 of only 0.027. This is not surprising in light of our earlier micro evidence that individual prepayment hazards are more closely related to rate gaps than to LTV ratios.

Figure 5 suggests that the relationship between $frac > 0$ and prepayment rates may have shifted over time. To assess this, Table 2 Columns (4)-(6) re-estimate regressions prior to, during, and after the housing boom-bust. From these regressions it is clear that in all sub-periods the response of prepayment to $frac > 0$ is very strong and similar in magnitude: thus, even though average prepayment rates are lower after 2010, the effect of $frac > 0$ on prepayment rates is unchanged. Finally, comparing the R^2 in Columns (4)-(6) with that in (7)-(9) shows that outside of the large boom-bust period, including the LTV ratio along with $frac > 0$ provides little additional predictive power. Furthermore, even during this boom-bust period, LTV only adds predictive power after controlling for rate incentives: re-estimating the regression in Column (8) with controls for LTV but not $frac > 0$ lowers the R^2 from 0.67 to 0.03.

4.3.2 Controls for Other Observables and Endogeneity

In Appendix Table A-1, we show that the relationship between $frac > 0$ and prepayment is robust to controls for the business cycle, seasonality, non-linearities and outliers. Most importantly, we show that our results are very similar when we control for one hundred calendar-quarter fixed effects and so identify off only within-quarter variation. This rules out many confounding factors such as aging (Wong (2019)) or changing lender concentration (Agarwal et al. (2017) and Scharfstein and Sunderam (2016)) that might influence refinancing incentives and prepayment rates but are unlikely to change at high frequencies. In Appendix Table A-2 we show similar relationships hold at the MSA level, even after including both calendar-month and MSA $\times$ calendar-quarter fixed effects.

However, this still may not reflect a causal relationship since $frac > 0$ depends on past endogenous interest rates, and it is possible that some unobserved confounding factor affects both $frac > 0$ and prepayment propensities even at high frequencies. To address this concern, Appendix Table A-3 shows that re-estimating our main regression using high-frequency monetary policy shocks as an instrument for $frac > 0$ delivers almost identical conclusions.

4.3.3 Decompositions, Additional Outcomes and Loan Composition

While we concentrate on total prepayment, in Appendix Table A-4 we replicate our analysis, instead focusing on individual prepayment types (rate refi, cashout refi and moves). As suggested by Figure 3, the independent effect of $frac > 0$ is strongest for rate-refinancing but is also important for cash-out and moves. Unsurprisingly, the LTV ratio has a significant effect on cash-out refinancing, but together $frac > 0$ and LTV have much more predictive power than either LTV or $frac > 0$ alone.

In Appendix Table A-5, we show that greater rate incentives also translate directly into faster reductions in actual mortgage coupons, not just into greater prepayment. While this is not surprising, it need not necessarily hold mechanically. More interestingly, interest rate pass-through into average coupons is

---

[28]We have also explored specifications which interact $frac > 0$ with the LTV ratio. The point estimates in these specifications imply that $frac > 0$ has less effect on refinancing when average leverage in the economy is high, but interaction effects are imprecisely estimated and effects on predictive power for prepayment are negligible.

much stronger when $frac > 0$ is large. This increase in rate pass-through with $frac > 0$ will be a central implication of our theoretical model and is a key indicator of path-dependence.

Finally, Appendix Tables A-6-A-9 show that our conclusions do not depend on the particular loans in our baseline sample: our baseline includes all fixed rate mortgages, but we can use broader samples including all loans or more narrow samples which restrict to only conforming fixed rate mortgages or those fixed rate mortgages which have never been delinquent.

5 Spending and GDP responses

Does this prepayment activity matter for aggregate monetary transmission into spending? To provide an initial assessment of aggregate scale, we begin with a simple back-of-the-envelope calculation of how much redistribution in mortgage payments is induced by interest rate reductions under different scenarios for prepayment rates. The overall level of residential mortgage debt is approximately $12.6 trillion. For simplicity, consider a reduction in mortgage rates of 100 basis points that lasts for one year.[29] This leads to a decline in borrowers' aggregate *annual* mortgage payments of roughly $44.7 billion if borrowers prepay at the 2003 rate of 35.4% p.a. but only a reduction of $11.6 billion if they prepay at the 2000 rate of 9.2% p.a.[30] With fixed rates, these annual payment reductions accrue for the life of the mortgage. A $44.7 ($11.6) billion annual reduction in payments thus translates into a present value of around $340 ($88) billion.[31] Dividing by 128.5 million US households implies an average annual payment reduction per household of roughly $350 ($90) translating to a present value of $2644 ($684). Furthermore, if rate reductions also spur equity extraction, as suggested by Figure 3, then this calculation understates the redistribution of disposable income arising after rate cuts.

These simple calculations suggest that rate changes could lead to mortgage payment changes big enough to matter for aggregate spending. Whether this is the case will depend on how spending responds to changes in mortgage payments. In this section we explore this question. We begin with individual borrower-level event studies to explore how auto purchases are related to refinancing and any resulting rate savings or equity extraction. We then show strong regional relationships between auto purchases and rate incentives. However, these empirical results capture only spending by borrowers and not potential offsetting effects from declines in lender income after rate reductions. In addition, although auto spending is an important component of business cycle fluctuations in aggregate spending, it is clearly somewhat special. We thus finish by providing aggregate evidence that monetary policy shocks indeed have larger effects on aggregate economic activity when $frac > 0$ is large.

5.1 Refinancing Event Studies

Our data does not measure broad spending, but we can proxy for car purchases using new car loans in borrower credit records. We run a borrower-level event study of this spending measure on refinancing, similar to Beraja et al. (2019), but extended to 2005-2017 rather than focusing only on 2009. In particular,

---

[29]This is a large rate decline, but not out of line with reductions seen after QE1 and other easing cycles. These calculations scale linearly with the rate decline, so e.g. a 50 basis point decline leads to payment reductions half as large.

[30]Residential debt in 2019Q3 Fed financial accounts is $12.638 trillion. $.01 \times 12638 \times .354 = 44.8$ and $.01 \times 12638 \times .092 = 11.6$

[31]This discounts payments over 30 years at a 15% rate to account for both prepayment and time discounting.

we regress a new car loan indicator on household fixed effects, calendar-month fixed effects as well as months-from-refinancing indicators, interacted with whether the refinancing involves cash-out.[32]

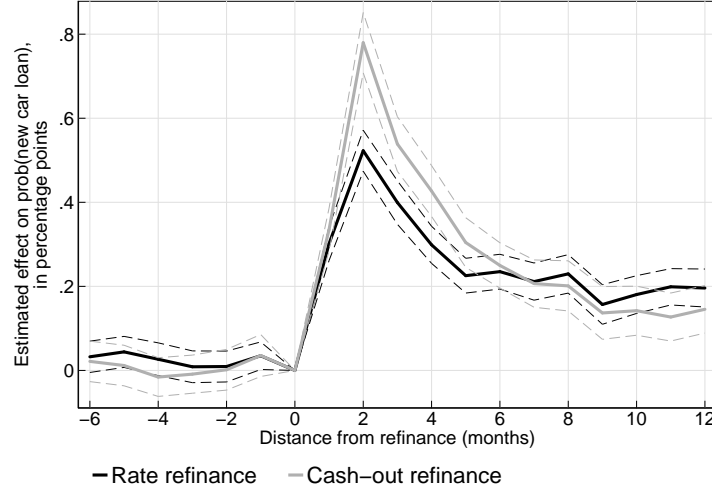**Figure 6: Individual Level response of Car Purchase to Refinancing**

Figure shows coefficients on month indicators (relative to refinancing month) from regressing a new car loan indicator on household fixed effects, calendar-month fixed effects and months-from-refinancing indicators, interacted with whether the refinancing involves cash-out. Dashed lines are 95% confidence intervals with standard errors two-way clustered by household and calendar month. Our indicator for a new car loan is an increase in auto loan balances of $5000 or more.

Figure 6 shows there are large increases in car buying after refinancing. Adding point estimates implies an increase in the probability of purchasing a car in the 12-months after rate (cash-out) refinancing of 3.16 (3.59) percentage points. This a large effect: the baseline annual purchase probability is 17.6%, so borrowers who refinance are 18-20 percent more likely than usual to purchase a car over the following year. The response of car purchases to cash-out refinancing is more front-loaded within the year, but the 12-month effect of rate refinancing is 88% as large as that of cash-out refinancing (0.88=3.16/3.59).[33] Despite the fixed effects in our regressions which address some obvious threats to identification, refinancing decisions are clearly endogenous, so these relationships may not be causal.[34] Nevertheless, they suggest that both cash-out and rate refinancing matter for spending. Furthermore, while cash-out refinancing has a stronger relationship with spending, the type of refinancing matters much less for subsequent spending than does the fact that a refinancing of either type has occurred.

Figure 7 explores the relationship between how much households save on interest payments when they refinance and their subsequent spending behavior. We split the sample into those households with the top quartile (>$3150) and bottom quartile (<$1250) of annual rate savings when refinancing.[35] We then re-estimate our event study but interacting the effects of cash-out and rate refi with indicators for these separate savings groups. We again include calendar-month and household fixed effects.[36] Panel

---

[32]Calendar-month fixed effects deal with concerns about confounding effects from time-varying aggregate conditions. For example, common movements in mortgage and auto loan rates might generate misleading relationships between refinancing and spending. Our fixed effect specification absorbs common interest rate movements by comparing the purchase decisions of borrowers who refinance to those who do not *within* a given month. Household fixed effects deal with concerns that households who refinance more often on average might also have different average auto purchase behavior.

[33]Over a shorter 3-month horizon, the effect of rate-refinancing is still 74% that of cash-out refinancing (.74=1.22/1.65).

[34]Abel and Fuster (2018) provide evidence (albeit with a less representative sample) that exogenous rate refinancing has a *causal* effect on borrower behavior; see also causal of ARM resets on spending in Di Maggio et al. (2017).

[35]We define annual savings as $balance_{pre} \times (m^*_{pre} - m^*_{post})$, so savings are computed using the initial balance if it changes.

[36]This means results are not driven by the obvious potential confounding factor in months when mortgage and auto interest

16

Figure 7: Heterogeneity by Annual Interest Rate Savings

(a): Rate Refi

(b): Cash-out Refi
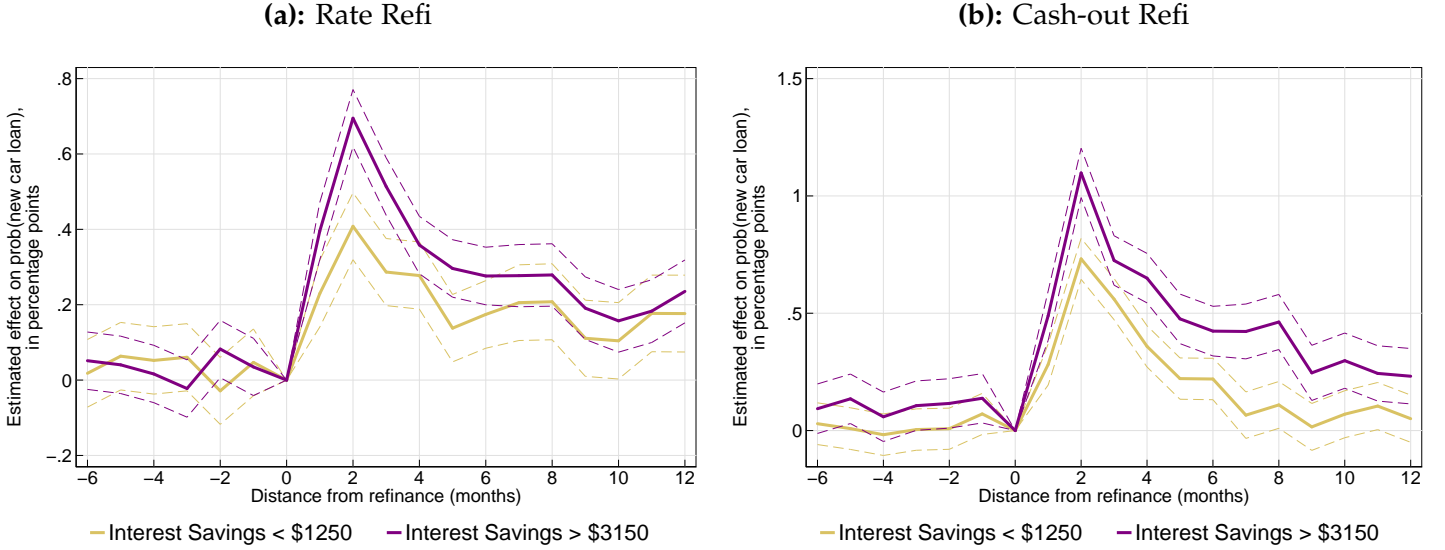
This figure repeats the analysis in Figure 6 but splitting also by the highest and lowest quartile of interest rate savings.

(a) shows effects of interest savings when rate refinancing, and Panel (b) shows effects of interest savings when cash-out refinancing. There are two key takeaways: 1) Those with substantial interest savings are much more likely to buy a car after refinancing. 2) The relationship between car buying and mortgage interest savings holds both for rate refinancing *and* for cash-out refinancing, so rate savings also matter for households who exctract equity from their homes. Indeed borrowers in the high savings-rate refi group are more likely to purchase a car than borrowers in the low savings-cashout refi group.[37]

The borrower-level evidence in this section shows that refinancing is associated with strong effects on spending, and the strength of rate incentives matters for these relationships: the greater the interest savings upon refinancing, the greater the subsequent increase in spending. We now turn to more aggregated evidence that mortgage rate incentives matter for spending and thus for monetary transmission.

5.2 Cross-Region Evidence

We now explore regional relationships between rate incentives, prepayment and car purchases using R.L. Polk zip code registration data. Appendix Table A-2 showed a strong MSA-level relationship between prepayment rates and $frac > 0$. In Table 3, we show there are also strong relationships with MSA car sales growth. In Column (1), we regress month t to $t + 1$ auto sales growth on the change in prepayment between $t - 1$ and t, plus a month fixed effect so we identify off cross-region but not aggregate time-variation. This shows that regions with larger increases in prepayment see larger increases in auto sales.[38] Results are slightly stronger in Column (2), which also includes MSA $\times$ quarter fixed effects.

While this shows a strong relationship between changing prepayment and changing car purchases,

---

rates are low, mortgage savings are high and it is also more attractive to buy a car.

[37]The 12-month purchase probability increases by 3.86 percentage points for rate refinancers with the largest savings and by 2.79 percentage points for cash-out refinancers with the smallest savings. Clearly the strongest effect is for cash-out borrowers with large interest rate savings, whose 12-month purchase propensity increases by 5.77 percentage points (almost 33 percent).

[38]The standard deviation of Δ freq is 0.388, so a 2 SD increase in frequency is associated with a 1.5 percentage point increase in the growth rate of auto sales (.388*2*.019=.0147). This is relative to an average growth rate of 2.2%.

this may not be a causal relationship. For example, increases in expected future income might lead to more car buying and more refinancing to finance that spending, in which case Δfreq is biased upward. Conversely, current income shocks could bias the coefficient down if greater income leads to greater spending but a decrease in the need to refinance to fund that spending. Motivated by our empirical evidence that $frac > 0$ affects prepayment, we thus instrument for Δfreq using $\Delta frac > 0$ in Columns (3) and (4). The identifying assumption is that changes in $frac > 0$ (after controlling for month and MSA $\times$ quarter FE) only affects car sales growth through refinancing. Point estimates in these IV specifications are increased substantially, suggesting that the second type of OLS bias above is more important.

Columns (4)-(8) show show that interest rate changes affect spending more in locations where more households are prepaying their mortgages. Specifically, we regress auto growth from month t to $t + 1$ on the frequency of prepayment in month t (rather than the change in prepayment) but now interacted with the change in mortgage rates between month $t - 1$ and t. This shows that mortgage rate declines correlate with greater spending growth when more households are refinancing; a central prediction of our theoretical model. It is important to note that interest rate endogeneity is not a particular concern for this specification, since any endogenous relationship between rate changes and aggregate conditions is absorbed by time fixed effects. However, prepayment frequencies may again be related to other transitory local conditions which affect auto spending growth and confound causal interpretations. We are primarily interested in the interaction term, which will be unaffected by shocks which just move prepayment and spending together. Nevertheless, we next instrument for the frequency of prepayment using the level (rather than the change) in $frac > 0$. The identifying assumptions are similar to those explained before, and again relationships are strengthened.

5.3 Aggregate Time-Series Evidence

While the borrower-level event studies and regional patterns are highly suggestive, cross-sectional relationships can potentially differ from aggregate relationships. Furthermore, while car purchases are an important share of business cycle fluctuations, they are clearly only one, potentially non-representative component of total spending.[39] To address these concerns, we now provide aggregate time-series evidence that the effect of identified monetary policy shocks on aggregate GDP varies with $frac > 0$.

We estimate the response of GDP and aggregate consumption to identified monetary policy shocks by local projections. We use identified shocks from Romer and Romer (2004) extended through 2008 by Coibion (2012). The key question of interest is whether aggregate responses to an identified monetary policy shock vary with $frac > 0$.

We incorporate non-linearities following Auerbach and Gorodnichenko (2012) and estimate a regime-specific local projection at each horizon $h = 0, ..., H$:

$$y_{t+h} = I_{t-1}\left[\alpha_{h,1} + \sum_{k=1}^{K} \theta_{h,k}^1 X_{t-k} + \beta_h^1 \epsilon_t\right] + (1 - I_{t-1})\left[\alpha_{h,2} + \sum_{k=1}^{K} \theta_{h,k}^2 X_{t-k} + \beta_h^2 \epsilon_t\right] + \delta_{h,1}t + \delta_{h,2}t^2 + \varepsilon_t$$

---

[39] Real chained GDP fell by \$628 billion in the Great Recession from 2007Q4 to 2009Q2, while vehicle spending fell by \$91 billion, implying roughly a 15% contribution. Similar effects hold in earlier recessions.

where y_t is the outcome variable, α_h is a regime specific constant, X_t is a vector of controls including lagged values of y_t, ϵ_t is an identified monetary policy shock, $t + t^2$ is a quadratic time trend and I_{t-1} is an indicator function equal to 1 when $frac > 0$ is greater than its median value and zero otherwise.[40] The IRFs are the sequences $\{\beta_h^l, \beta_h^2\}_{h=0}^H$, representing output and consumption responses at future horizons to shocks today when $frac > 0$ is above or below the median, respectively.

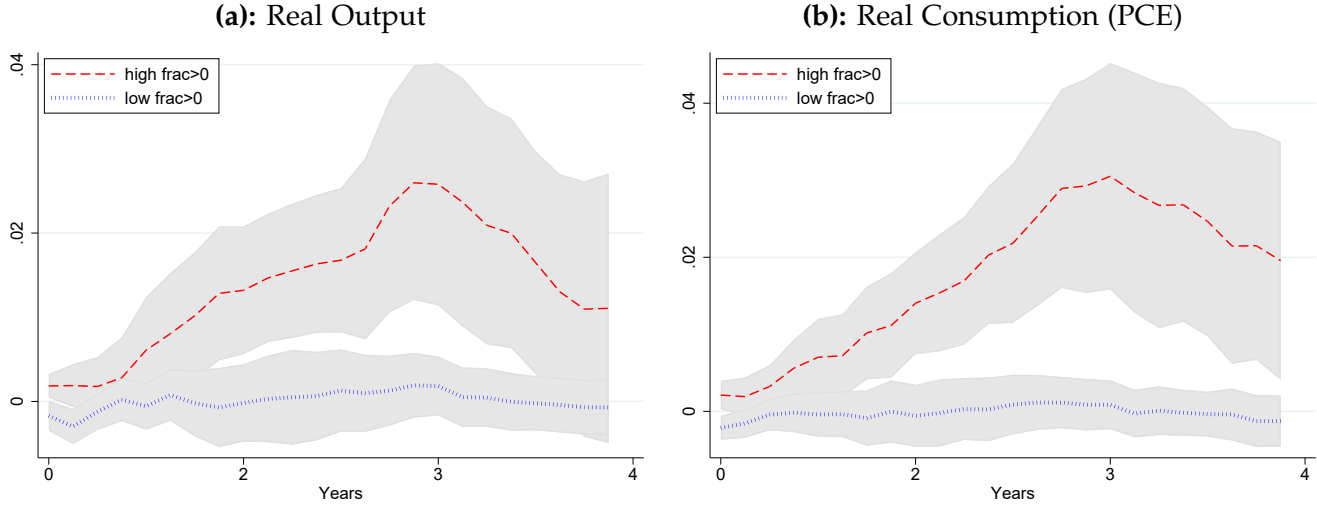**Figure 8:** Response to Identified Monetary Policy Shocks

(a): Real Output   **(b):** Real Consumption (PCE)

Figure shows the response of log real output and non-durable consumption to the identified monetary shocks from Romer and Romer (2004), extended through 2008 by Coibion (2012). Dashed lines are 90 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.

Figure 8 shows results. The left (right) panel shows the response of log real GDP (PCE-Personal Consumption Expenditure) to a 100 bps expansionary monetary policy shock. The red-dashed line shows the effect when $frac > 0$ is above its median value and the blue-dotted line shows the effect when $frac > 0$ is below its median value. Both output and consumption increase significantly more in response to an expansionary monetary policy shock when $frac > 0$ is high.

6   A Model of Mortgage Prepayment

We now turn to our theoretical model which we use to interpret our empirical results and characterize how monetary policy affects aggregate spending through mortgage prepayment. The goal of the model is two-fold: 1) Argue that the prepayment patterns we document have aggregate implications for real spending that are large enough to matter for monetary policy. 2) Provide guidance about the potency of this rate incentive channel at a moment in time and its evolution over time given current policy choices. We show that mortgage prepayment leads to non-linear, path-dependent dynamics. This means that assessing the effect of current actions on future policy space without theory is likely to generate

---

[40]Ramey (2016) highlights the importance in including time-trends when estimating IRFs by local projections. To extend our sample before 1992, we forecast $frac > 0$ within our sample period 1992-2017 using nine lags of the fixed rate mortgage rate and then use this prediction to forecast out of sample. We run these forecast regressions in log space so that predictions for $frac > 0$ are bounded between 0 and 1. The within sample correlation between the actual and predicted $frac > 0$ is over 92%.

incorrect conclusions. Our micro data consistent model has the added benefit to deliver a rule-of-thumb for quantifying these intertemporal effects.

We use a continuous-time open economy model which embeds a household mortgage refinancing problem into an incomplete markets environment with endogenous mortgage pricing. Our model includes (a) a continuum of households making consumption, savings and refinancing decisions, and (b) a competitive, risk-neutral financial intermediary that extends fixed rate prepayable mortgage loans and finances itself via deposits and short term debt from domestic savers and international capital markets.

Households are subject to idiosyncratic labor income risk and choose to consume or save in a liquid asset subject to a borrowing constraint, as in Aiyagari (1994). To this standard environment we add prepayable mortgage debt and interest rate fluctuations. Households take the stochastic process for both short-term interest rates and mortgage rates as given. The mortgage interest rate is pinned down by the financial intermediary's zero profit condition, and any net funding surplus or deficit run by the financial intermediary is filled in international capital markets at the short term interest rate. This endogenous relationship between short rates and mortgage rates leads to endogenous redistributional effects of rate changes: when rates decline, it frees up disposable income for those borrowers who refinance, but it also lowers returns for lenders. This in turn influences the strength of monetary policy.

6.1 Uncertainty

Household i receives uninsurable idiosyncratic labor income Y_{it} per unit of time, with $\ln Y_{it}$ following the continuous time counterpart to an AR(1) process:

$$d \ln Y_{it} = -\eta_y \left(\ln Y_{it} - \ln \bar{Y} \right) dt + \sigma_y dZ_{it}, \tag{2}$$

where Z_{it} is a standard Brownian motion that is both independent across households and independent from any aggregate states of the economy, $\ln \bar{Y}$ is the ergodic mean log income, σ_y^2 is the instantaneous variance (per unit of time) of log income, and η_y measures the persistence of log income.

Agents also face aggregate interest rate risk. We assume the short term interest rate set by the Fed, r_t, follows a one-dimensional Feller square-root process (see Cox, Ingersoll Jr and Ross (2005))

$$dr_t = -\eta_r \left(r_t - \bar{r} \right) dt + \sigma_r \sqrt{r_t} dZ_t, \tag{3}$$

where Z_t is a standard Brownian motion, $\bar{r}$ is the ergodic average short term rate, $r_t \sigma_r^2$ is the instantaneous variance (per unit of time), and η_r measures the persistence of the process.

Mortgage market interest rates m_t also follow a stochastic process, determined in equilibrium via a risk-neutral asset pricing equation that ensures financial intermediaries break-even when extending fixed-rate prepayable mortgage debt (see Section 6.3). We note that fluctuations in $m_t = m(r_t)$ will arise from fluctuations in r_t in equilibrium.

6.2 Household Balance-Sheets and Refinancing Frictions

Each household is born at $t = 0$ with liquid savings W_0 and a house financed with a fixed-rate pre-payable mortgage with balance F and coupon rate m_t^*.[41] In our benchmark model we assume that all households have identical, constant mortgage balances, but in Section 9.1 we show that introducing cash-out refinancing and heterogeneous mortgage balances only amplifies our conclusions.

Our choice to only focus on rate refinancing in the benchmark model warrants additional discussion. Recall our earlier evidence that cash-out refinancing without a simultaneous rate decline is unusual: the overwhelming majority of all refinancing is associated with rate reductions. Our benchmark model captures this crucial role of rate incentives, it can be easily solved numerically, and allows us to transparently illustrate and discipline the key mechanism driving the path-dependence of monetary policy.[42]

We assume that each mortgage can be refinanced at the discretion of the household only at random, exponentially distributed attention times (with arrival intensity $\chi_c + \chi_f$). When these opportunities arise, the household can choose to keep its existing mortgage or to refinance at the prevailing mortgage market rate m_t, either (i) "for free" (with probability $\chi_c / (\chi_c + \chi_f)$) or (ii) by paying a fixed cost $\kappa > 0$ (with probability $\chi_f / (\chi_c + \chi_f)$). Our model thus nests two different types of refinancing frictions: when $\chi_f = 0$, we obtain a pure "Calvo" model in which households obtain opportunities to refinance at no cost at Poisson arrival times, and they exercise their option if and only if the current market interest rate is below their outstanding coupon rate. This model delivers a prepayment hazard that is a step function (at zero) in interest rate gap space. When $\chi_f = +\infty$ and $\chi_c = 0$, we obtain a pure "menu cost" model in which households have an option to refinance at any time but only by paying a fixed cost. This implies a refinancing probability of zero for small rate gaps and $+\infty$ above the (endogenous) individual gap threshold at which it is optimal to pay the fixed refinancing cost.

Finally, we also assume that households face exogenous moving shocks that arrive at Poisson rate ν, forcing them to move to a different (identical) house and reset their mortgage coupon in the process.[43]

6.3 Financial Intermediaries and Mortgage Pricing Equilibrium

Mortgages in our model are priced by risk-neutral competitive financial intermediaries who originate fixed rate prepayable mortgages financed with short term floating rate debt. Intermediaries take the short term interest rate r_t as given and have access to a perfectly elastic supply of capital at that rate in international markets. When the short term rate is r, the value of an existing mortgage loan with coupon m^* and face value \$1 is determined by the risk-neutral asset pricing equation:

$$P(r, m^*, S) = \mathbb{E}\left[\int_0^\tau e^{-\int_0^t r_u du} m^* dt + e^{-\int_0^\tau r_u du} \,\middle|\, r_0 = r, S_0 = S \right] \tag{4}$$

---

[41]We assume the house has constant price and so abstract from interest rate effects on consumption arising through house price effects documented in Mian, Rao and Sufi (2013), Berger et al. (2018) and Guren et al. (2018). Modeling these effects would likely amplify our conclusions: interest rate histories leading to greater house price growth, prepayment activity and cash-out should lead to greater resulting consumption.

[42]The Fed can fairly directly affect rate incentives, and we can in turn directly discipline the effect of rate incentives on prepayment using our micro data. In contrast, the Fed has much less direct control over cash-out incentives via house prices.

[43]Landvoight, Piazzesi and Schneider (2015) provides empirical evidence for these exogenous moving shocks. Krivenko (2018) documents that they are important for matching the dynamics of house prices.

where S_t is a vector of household state variables and τ is the (endogenous) prepayment time, which can potentially depend on both aggregate interest rates and idiosyncratic household states. In our perfectly competitive mortgage market equilibrium, the market value of a mortgage at origination must be equal to its notional balance. In other words, for financial intermediaries to break-even, the "fair-market" coupon $m(r, S)$ of a mortgage extended to a household in state S when the short rate is r must satisfy:

$$P\left(r, m(r, S), S\right) = 1, \tag{5}$$

For the general specification of refinancing frictions described in Section 6.2, solving the mortgage market equilibrium is extremely difficult: financial intermediaries must forecast household prepayment decisions to determine m, but prepayment decisions depend both on idiosyncratic household states and on the equilibrium evolution of m, which we are trying to determine. However, the micro data patterns documented in Section 4 will ultimately lead us to a particular benchmark model with key simplifications that imply an equilibrium function $m(\cdot)$ that depends only on the short rate r.

6.4   Household Problem

Households have identical CRRA preferences with rate of time preference δ and inter-temporal rate of substitution $1/\gamma$. Households can save in a liquid savings account W_t with return r_t to insure against labor income shocks, but they cannot take on unsecured short-term debt, so $W_t \geq 0$. Thus, their only liability is their outstanding mortgage, and their net financial asset position is equal to $W_t - F$. Households in our model do not have any option to default.

The relevant state vector for household's decision problem is $(r, S_i) := (r, W_i, m_i^*, Y_i)$. Households make consumption $\{C_t\}_{t \geq 0}$ and refinancing decisions $\{\rho_t^{(c,f)}\}_{t \geq 0}$, and solve the following problem (where we drop the subscript i for notational convenience):[44]

$$V\left(r, S\right) := \sup_{C, \rho^{(c)}, \rho^{(f)}} \mathbb{E}\left[\int_0^{+\infty} e^{-\delta t} \frac{C_t^{1-\gamma}}{1 - \gamma} dt \ \middle| \ r_0 = r, S_0 = S\right]$$

$$\text{s.t. } dW_t = \left(Y_t - C_t + r_t W_t - m_t^* F\right) dt - \rho_t^{(f)} \kappa F dN_t^{(\tau_f)}, \qquad W_t \geq 0$$

$$dm_t^* = \left(m_t - m_{t-}^*\right) \left[\rho_t^{(c)} dN_t^{(\tau_c)} + \rho_t^{(f)} dN_t^{(\tau_f)} + dN_t^{(\tau_m)}\right],$$

with Y_t following (3), r_t following (2), and $m_t = m\left(r_t, S_t\right)$ solving (5). Here τ_c is the sequence of times when refinancing is costless, τ_f is the sequence of times when refinancing requires a cost $\kappa > 0$, and τ_m is the sequence of times the household is forced to move. $dN_t^{(\tau_c)}, dN_t^{(\tau_f)}, dN_t^{(\tau_f)}$ are changes in the associated counting processes (and thus equal 1 at the relevant arrival times and 0 otherwise).

At any "zero cost" attention time τ_c, it is optimal for the household to refinance whenever the prevailing mortgage rate is below that household's outstanding coupon:

$$\rho_t^{(c)} = \begin{cases} 1, & \text{if } m_t < m_{t-}^* \\ 0, & \text{otherwise} \end{cases}$$

---

[44]See Appendix Section A.3.3 for more details on our Hamilton-Jacobi-Bellman equation corresponding to the household value function V and numerical solutions' method we employ to solve such equation.

At any "fixed-cost" attention time τ_f, it is optimal for the household to refinance whenever the value of paying the refinancing cost and obtaining a lower rate exceeds the value of inaction:

$$\rho_t^{(f)} = \begin{cases} 1, & \text{if } V\left(W_t - \kappa, r_t, m_t, Y_t\right) > V\left(W_t, r_t, m_{t-}^*, Y_t\right) \\ 0, & \text{otherwise} \end{cases}$$

6.5 Calibration

We calibrate our model in two steps. We first describe how we discipline the parameters governing refinancing frictions and then discuss the remaining parameter values.

6.5.1 The Nature of Refinancing Frictions

We argue in this section for a benchmark calibration based on the Calvo inattention friction. This choice dramatically simplifies the calculation of the mortgage market equilibrium, allows us to provide intuition for the dynamics of path-dependence, and most importantly, fits the data much better than a pure menu cost model while giving predictions nearly identical to models with both frictions.

We begin by comparing a pure Calvo inattention model to a pure menu cost model.[45] We initialize both models to the actual 1992 loan-level distribution of mortgage rates and expose them to the actual monthly mortgage rate time series from 1992 to 2017. We pick refinancing parameters so that each model matches the average prepayment frequency in the data from 1992 to 2017, and we study how each model fits the time-series of various (untargeted) moments of the data.[46]

We calibrate the annual moving rate ν to 4.1% to match the empirical prepayment hazard for loans with negative rate gaps. This leaves one free parameter in both models: we set $\chi_c = 22.8\%$ in the inattention model and $\kappa = \$2500$ in the menu cost model, which implies an average monthly prepayment frequency from 1992-2017 in both models of 1.5%. We then see how these models match the time-series behavior of prepayment rates and the loan-level distribution of coupons and rate gaps over time.

Figure 9 shows that the model with inattention-based refinancing frictions is a dramatically better fit than the menu cost model. Panel (a) shows a snapshot of the entire cross-sectional distribution of coupons in the models vs. the data every five years and Panel (b) - Panel (d) show time-series fits for various summary statistics. Considering that only the initial distribution in 1992 and average prepayment rates are targeted, the fit generated by the inattention model is overall quite good (with the notable exception of missing the large outlier in 2003).

In contrast, the menu cost model calibrated to match the same average frequency of adjustment in the data does not fit time-series patterns. Panel (b) shows that the prepayment frequency is much spikier than in the data. Panel (c) shows that such menu cost leads to an average outstanding coupon m_t^* that closely tracks the running minimum of the market rate m_t, in contrast to the more sluggish evolution in

---

[45]For computational reasons, we specify a large but finite arrival rate in the menu cost model. Specifically, we pick $\chi_f = 24$ so that households get refinancing opportunities on average every 2 weeks.

[46]In the Calvo model, refinancing behavior is independent of preferences, income and the short rate process. In the menu cost model, households' refinancing behavior depends on all model parameters, so we jointly calibrate all parameters. Although we feed in rates from the data, we must also specify interest rate and income expectations, which we do using the eventual model calibration. None of our conclusions are sensitive to these calibration choices.

the data. Panel (d) shows that the menu cost model generates far too little cross-sectional dispersion in rates and misses the time-series of dispersion.[47]

Figure 9: Inattention and Menu Cost Models vs. Data: Time-Series Fit

(a): Distribution of Gaps

(b): Frequency

**(c):** Average Rates

(d): Std Dev of Rates
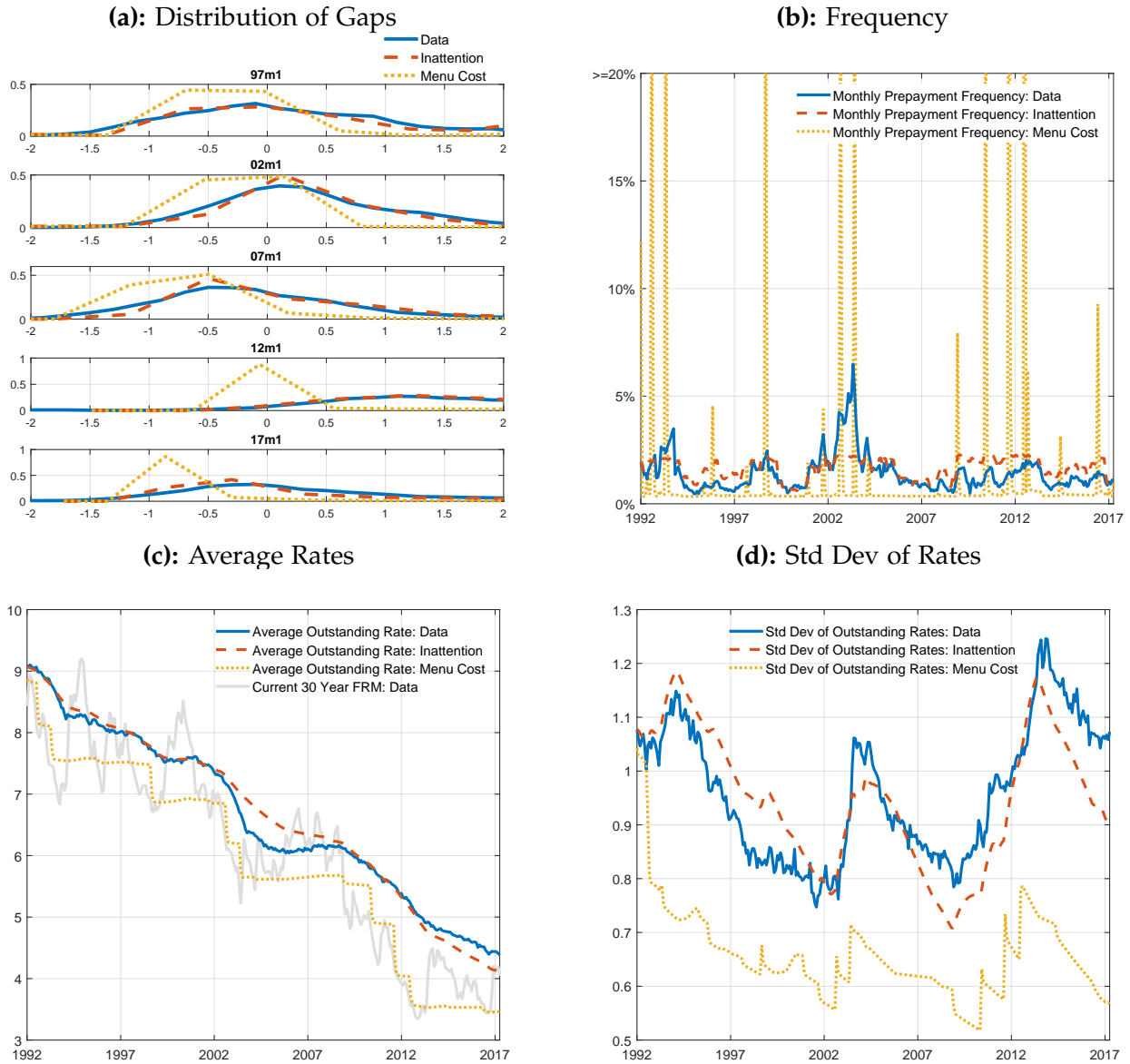
This poor time-series fit arises because the menu cost model generates a prepayment hazard at odds with the micro data, as shown in Panel (a) of Figure 10. The menu cost model implies a hazard which is too low for moderate positive gaps but which then becomes too high for large positive gaps. This occurs despite the fact that our model includes substantial cross-household heterogeneity in income and liquid wealth and thus scope for heterogeneous refinancing decisions: with a fixed cost of adjustment, a large enough rate incentive eventually leads almost all households to refinance.

By construction, the inattention model generates a prepayment hazard with a step exactly at zero,

---

[47]These conclusions are robust to different fixed costs as well as permanent heterogeneity in fixed costs. Of course, a random menu cost model can be made isomorphic to the attention model and so fit the data, but this requires a cost which is typically very high punctuated by brief periods near zero. Such a process arises naturally via inattention.

Figure 10: Prepayment Hazards

(a): Inattention and Menu Cost vs. Data   **(b):** Hybrid vs. Data
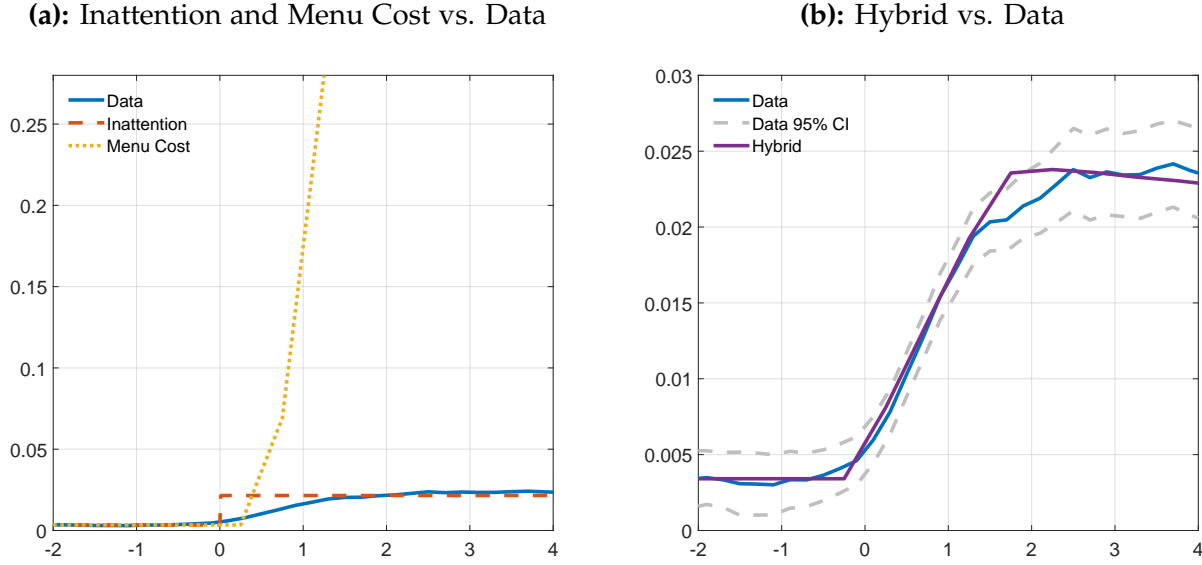
which is much closer to the empirical hazard. However, the inattention model does not fit the data perfectly either: it implies prepayment rates which are too high at rate gaps between 0 and 100bps. We thus explore a "hybrid" model with both inattention and fixed costs of adjustment. In particular, we jointly pick χ_c, χ_f and κ to match the entire empirical hazard. Panel (b) of Figure 10 shows that a model with both frictions indeed better fits the prepayment hazard.[48] However, Appendix Figure A-10 shows that this hybrid model generates time-series implications almost identical to the pure Calvo inattention model. Indeed, for small rate gaps, refinancing makes little difference for a household's mortgage coupon. This means that the fact that the Calvo model overstates prepayment rates for households with small positive gaps is mostly irrelevant for the actual coupons households obtain.

While both models have similar time-series implications, key simplifications arise for calculating interest rate counterfactuals in the pure Calvo inattention model. In this model, household prepayment decisions are orthogonal to consumption-savings decisions, and so household states S can be eliminated from the mortgage pricing equation (4). We can then easily solve the fixed point mortgage equilibrium equation (5), pinning down the mortgage market rate $m_t = m(r_t)$ as a function of the short rate, as shown in Section A.3.2. In addition, this model delivers simple analytical insights, and thus intuitive characterizations of monetary policy path-dependence, that are unavailable in environments with other frictions. Since the time-series implications of the pure inattention and hybrid models are almost identical under the historical time-series of interest rates but the pure inattention model has key computational and pedagogical advantages, we focus primarily on this model. However we show in robustness results that our policy conclusions are nearly identical if we use the hybrid model in a simpler rate environment without solving for the full mortgage market equilibrium.[49]

6.5.2 Calibrating Additional Parameters

We summarize parameters in Table A-10. We set $\gamma = 2$ following standard values in the macro literature. We fix the mortgage balance F to the average in our data of \$150,000. Log income is calibrated following

---

[48]The best fit parameters $\kappa = \$8250, \chi_f = 0.145$ and $\chi_c = 0.125$ imply an average cost when refinancing of \$1934.

[49]That is. we solve the household model with general frictions but equilibrium mortgage pricing from the pure Calvo model.

Floden and Linde (2001), implying mean reversion parameter $\eta_y = 9.3\%$ (corresponding to a half-life of 7.3 years), conditional volatility $\sigma_y = 21\%$, and an ergodic mean log income $\ln \bar{Y}$ that leads to an ergodic average income of $\mathbb{E}[Y_t] = \$58,000$ per year, consistent with average US household income.

We pick the rate of time preference δ so that after interest rate shocks, the present value change in aggregate mortgage payments is the same as the present value change in aggregate capital income.[50] One can interpret this calibration strategy as targeting a closed economy "on average" but with stochastic foreign capital flows. This requires $\delta = 14\%$ per annum, and generates an ergodic average liquid savings $\mathbb{E}[W_t] = \$35,000$ and an average MPC out of liquid wealth of 0.26. More interestingly, we find that households in the lower 50 percentile of the liquid savings distribution have MPCs roughly twice that of households in the upper 50 percentile of the liquid savings distribution. This heterogeneity of MPCs with interest rate exposure is directly in line with recent estimates in Crawley and Kuchler (2018).

The short rate is calibrated to have an ergodic average equal to $\bar{r} = 3.5\%$, a speed of mean reversion $\eta_r = 13\%$ (corresponding to a half life of 5.3 years), and a volatility parameter $\sigma_r = 6\%$ (corresponding to an ergodic standard deviation of short rates of 2.2%). Those parameters are obtained via a maximum-likelihood estimation procedure using daily data for 3-months treasury yields from 1982 to 2018.

The short rate process combined with "Calvo" prepayment frictions leads to an equilibrium mapping $m(\cdot)$ between short term interest rates r_t and mortgage rates m_t. The average slope of the mortgage pricing function $m(\cdot)$ delivers a pass-through of 0.57, which is in line with various estimates from the literature using high-frequency identification strategies.[51] We also explore the robustness of our results to alternative interest rate calibrations and resulting pass-through to mortgage rates in Section 9.

7 Path-Dependent Monetary Transmission: Mortgage Market Outcomes

In this section, we explore the effect of monetary policy on prepayments and average coupons, and thus on disposable income. We begin with an analysis of these outcomes since under our benchmark frictions ($\chi_f = 0$ and $\chi_c > 0$), these outcomes depend solely on χ_c, ν and the process for r_t and not on any of the more complicated model elements and parameters which determine consumption. We then turn to our model's implications for consumption in Section 8.

7.1 $frac > 0$ as a Key Statistic: Cabellero-Engel in Continuous Time

Our analysis focuses on comparing impulse response functions (IRFs) at time $t = 0$ in experiments with different $t < 0$ interest rate histories and thus different time-zero cross-sectional coupon distribution. While each experiment leads to different cross-sectional distributions and resulting IRFs, we preview now that these differences can be explained through a single moment of the time-zero cross-sectional coupon distribution: the proportion of households with a positive mortgage rate gap, $frac > 0$.

---

[50]More precisely, we focus on an economy where (a) the short term rate is equal to its ergodic mean, and (b) the cross-sectional household distribution is equal to its ergodic distribution conditional on the short rate equaling its ergodic mean. We then compute the impulse response of capital income $r_t W_t$ and mortgage payments $m_t^* F_t$ to a 100bp decrease in short rates and pick δ (and thus the ergodic mean of liquid savings W_t) so that these streams are equal in present value.

[51]Gilchrist, Lopez-Salido and Zakrajsek (2015) finds a pass-through of 0.68 (Table 6), Wong (2019) finds a pass-through of 0.392 (Table 5) and Gertler and Karadi (2015) find pass-through of current Fed Funds Rates into mortgage rates of 0.27 and pass-through of one-year rates into mortgage rates ranging from 0.54-0.80. (Table 1 Columns 3+7). Importantly, these relatively high pass-through numbers reflect the fact that even though mortgages have a maturity of 30 years, their pricing is closer to medium-term treasuries since they typically prepay in around seven years.

To see this, let $f_t(m^*)$ be the time t cross-sectional density of mortgage coupons (with cumulative density F_t) and $h(m^* - m)$ be the instantaneous prepayment hazard for a gap of $m^* - m$. Under our Calvo model of refinancing behavior, cross-sectional average prepayment intensities are then

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

$[1 - F_t(m_t)]$ is exactly $(frac > 0)_t$, so model prepayment rates are driven solely by this moment of the cross-sectional coupon distribution. Thus, if two economies with different time-zero cross-sectional coupon distributions are exposed to the same rate shock, any resulting differences in prepayment IRFs can be explained entirely by differential changes in $(frac > 0)_t$ between t and $t + dt$.

Next, consider the average coupon $\bar{m}_t^* := \mathbb{E}_i[m_{it}^*] = \int m^* f_t(m^*) dm^*$. In Appendix A.3.7, we show that the slope of the IRF of $\bar{m}_t^*$ at $t = 0$ to a small $\varepsilon \approx 0$ shock to r at time 0 is:

$$\lim_{t \to 0} \frac{dIRF_{m^*}^\varepsilon(t)}{dt} \approx \varepsilon \frac{\partial}{\partial r_0} \left(\frac{d\bar{m}_0^*}{dt} \right) = \varepsilon \cdot m'(r_0)(\nu + \chi_c[1 - F_0(m_0)]),$$

where $m'(r_0)$ is represents the "local" pass-through from short rates to mortgage rates. This is essentially the continuous-time counterpart to discrete-time results in Caballero and Engel (2007) (see their equation (17)), specialized to a hazard which is a step-function at zero. As a simple numerical example, suppose that $\chi_c = 0.228$, $\nu = .041$, $\varepsilon = -100bps$, $m'(r_0) = .5$, and $frac > 0 = [1 - F_0(m_0)] = .8$. Then the slope of the average coupon IRF is -11.2bps per year. If $frac > 0$ falls to 0.2, this slope is instead -4.3bps per year. This emphasizes the importance of $frac > 0$ for understanding monetary policy transmission from short rates to disposable income.

7.2 "Baseline" Economy: the Ergodic Distribution

Let $t = 0$ be the current time in the economy, from which we begin all policy experiments. To isolate the effects of different rate histories, most of our experiments compare two economies with identical current interest rate r_0 but different histories $\{r_t\}_{t<0}$ and thus different initial cross-sectional coupon and rate gap distributions.[52] Most of our experiments compare IRFs of (1) a "baseline" economy with an initial cross-sectional distribution equal to the economy's ergodic density when the short rate $r_t = r_0$, to (2) alternative cross-sectional distributions generated by specific $\{r_t\}_{t<0}$ paths of interest rates. In most experiments, we set $r_0 = \bar{r}$. We concentrate mostly on IRFs to a 100bps cut in r_0 (which lowers m_0 by roughly 50bps), but we also show effects of lowering r_0 to zero, which we refer to as the "max" rate stimulus. Following the time-zero impulse, r_t follows the dynamics specified in equation (3).

Figure 11 shows the IRF of average coupons to these two stimulus policies occurring at $t = 0$ in our "baseline" economy. We depict the peak coupon responses to each shock, which we use as a reference point in some later scenarios, as horizontal lines in red, but we defer discussion of magnitudes until introducing additional experiments and bringing consumption back into the model

7.3 Effects of the Secular Mortgage Rate Decline

In our first experiment, we investigate how the actual path of interest rates from 1992-2017 affects monetary policy effectiveness. During that time frame, we witnessed a secular decline in interest rates,

---

[52]The distributions of income and wealth also differ, but in our baseline model this is irrelevant for mortgage outcomes.

Figure 11: Baseline: IRF of Average Coupon m^* to 100bps & Max decline in r

(a): Average Coupon m^* **(b):** Monthly Prepayment Flows
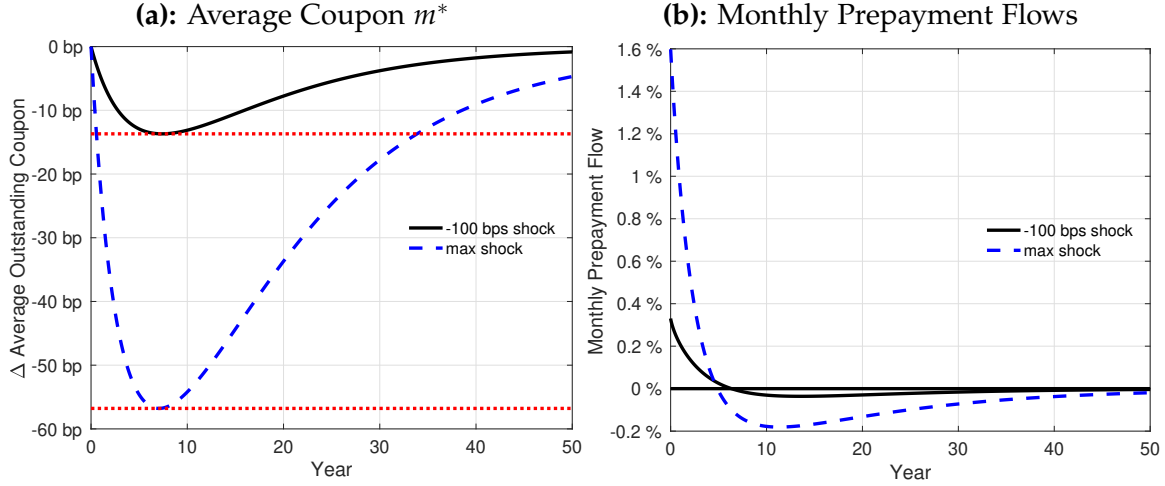
with the 30-year mortgage rate ultimately falling from roughly 9% to 4%. To explore the effects of this trend, we follow the procedure of Section 6.5.1 by initializing our model with the 1992 empirical coupon distribution and exposing it to the observed sequence of interest rates from 1992 to 2017 to obtain a model-implied cross-sectional coupon distribution in 2017. We then compare IRFs computed starting from this distribution to those obtained under the "baseline" ergodic economy.[53] In our comparison, both economies are started at $t = 0$ with the same $r_0 = \bar{r}$, have the same $\{r_t\}_{t \geq 0}$ and differ only in $\{r_t\}_{t < 0}$ and thus their initial cross-sectional gap distributions. Figure 12 shows that monthly prepayment and average coupons respond much more to a 100bps rate cut in the "secular decline" economy exposed to the 1992-2017 rate history than in the "baseline" ergodic economy.[54]

Figure 12: Impulse Response Functions to 100bps decline in r

(a): Average Coupon m^* **(b):** Monthly Prepayment Flows
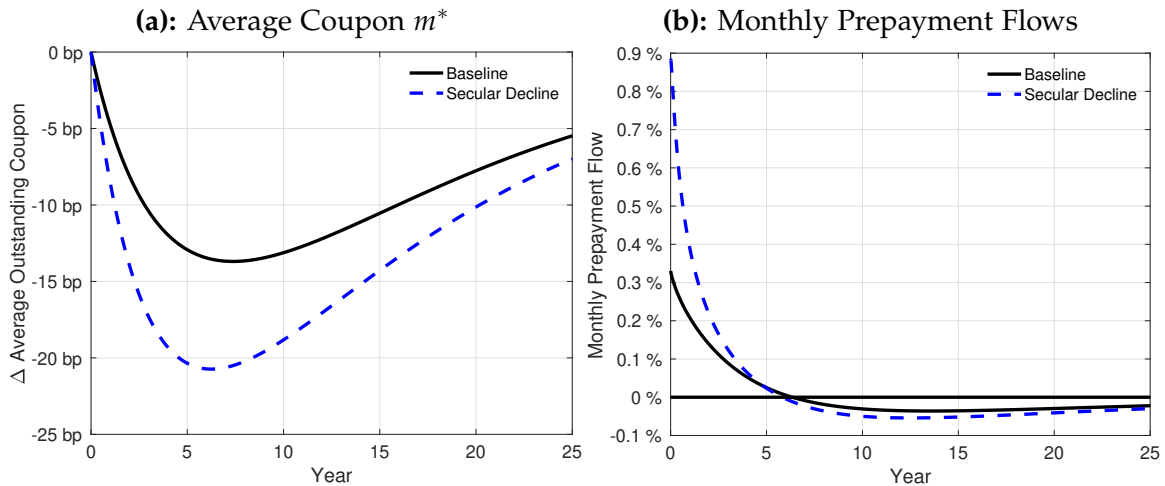
This is because the economy with a history of declining rates starts with larger $frac > 0$ today, as

---

[53]Figure 9 shows that this model implied distribution in 2017 is very similar to the 2017 empirical distribution, so if we use 2017 empirical coupons as our "initial distribution", IRFs are very similar.

[54]IRFs for the "max" shock are roughly scaled up versions of the IRFs to the 100bps shock and yield the same path-dependence conclusions, so we focus mostly on 100bps shocks to simplify exposition.

shown in Figure 13. This in turn amplifies the response of this economy's prepayment rates, average coupons, and ultimately disposable income to current rate cuts, as discussed in Section 7.1

Figure 13: Distribution of rate gap $m^* - m$ in Ergodic and Actual Simulation
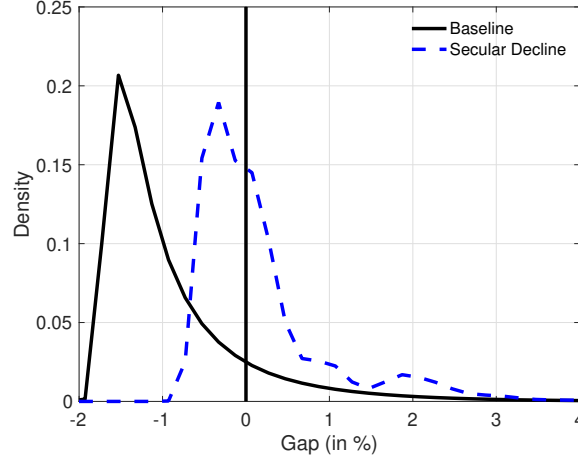
7.4 Effects of Past Easing and Tightening Cycles

We next compare IRFs in the baseline ergodic economy to two alternative scenarios in which previous rates were either (i) high or (ii) low for an extended period of time.[55] In all three economies, time-zero interest rates are set to $r_0 = \bar{r}$, future rates $\{r_t\}_{t \geq 0}$ follow (3), and the economies differ only in $\{r_t\}_{t<0}$. The behavior of the "past high rates" economy illustrates the potency of monetary policy at the end of a tightening cycle while the "past low rates" economy illustrates the potency of monetary policy following a long period of monetary easing – such as that observed in the aftermath of the Great Recession.

Once again, we find substantial path-dependence: Figure 14 shows that average coupons respond to the same 100bps cut in r more strongly in the economy with previously high than previously low rates. The intuition is the same as before: given identical current interest rates, the economy with previously high rates has larger $frac > 0$ than the economy with previously low rates. This means that after exiting a long period of low rates, the Fed will have reduced stimulus power if it needs to reverse course.

7.5 Asymmetries and Reloading

In the prior experiments, r_0 and $\{r_t\}_{t \geq 0}$ were identical, but the time-zero cross-sectional distributions differed due to differences in $\{r_t\}_{t<0}$. In this section, we illustrate the dynamic effects of actions today on future monetary policy potency by comparing two economies which are initially identical but that differ in $\{r_t\}_{t \geq 0}$. In particular, we consider two economies that are initially in the "baseline" environment described in Section 7.2. In one of them (the "Rate-Shift-Up" economy), the ergodic mean short rate permanently increases by 100bps at time zero, while in the other (the "Rate-Shift-Down" economy), it permanently decreases by 100bps. We then compute the effects of rate cuts at different points in time after the regime shift. This exercise is meant to assess the speed with which the Fed regains

---

[55]In the "past high rates" (resp. "past low rates") economy, the time-zero household cross-sectional distribution corresponds to the ergodic household cross-sectional density arising when interest rates follow equation (3), with $\bar{r}_{past}$ 100bps above (100bps below) our baseline. When we analyze consumption outcomes, these cross-sectional distributions are computed assuming that households have rational expectations – they understand that short term rates have an ergodic mean different from the baseline rate $\bar{r}$, thus affecting not only their refinancing but also their savings behavior.

Figure 14: IRF of Average Coupon m^* to 100bps decline in r

(a): Average Coupon m^* **(b):** Monthly Prepayment Flows
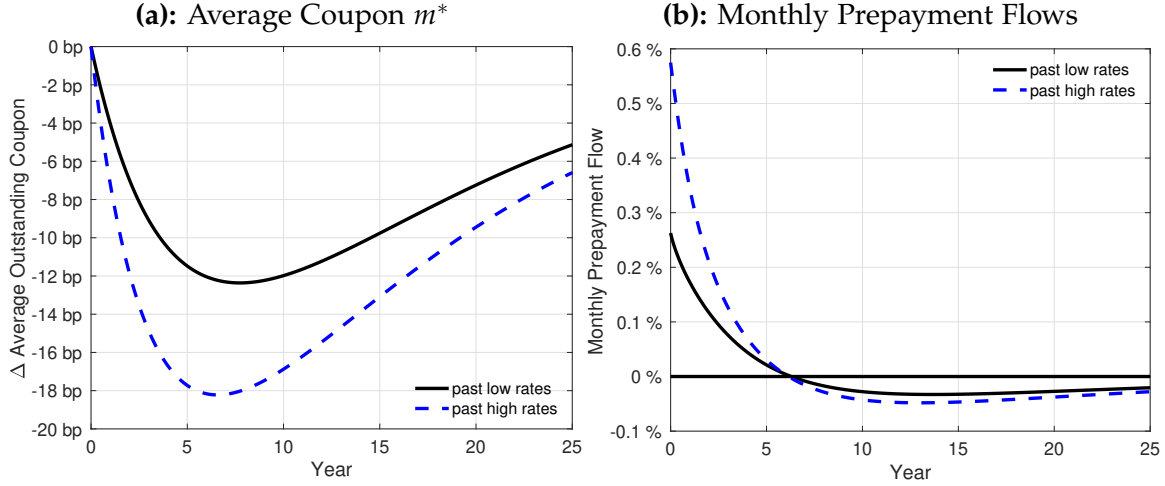
"ammunition" after raising interest rates and the speed with which it uses it up after cutting interest rates. Appendix Figure A-11, shows how IRFs evolve over time after the regime shift.

However, visually unpacking the dynamic evolution of a sequence of functions is hard. Thus, in order to more cleanly illustrate the dynamics we wish to emphasize, we compute for each IRF its cumulative discounted value $CIRF_t = \int_t^\infty e^{-\delta(s-t)} IRF_{m^*}(s)ds$, which we view as a convenient statistic summarizing the potency of monetary policy at each date t. Figure 15 then plots $CIRF_t$ in the regime-shift economies relative to $CIRF_t$ in the baseline economy, for various years t after the regime shift.

Figure 15: Regime Shift: Cumulative Discounted Average Coupon IRF decline in r to 100bps shock
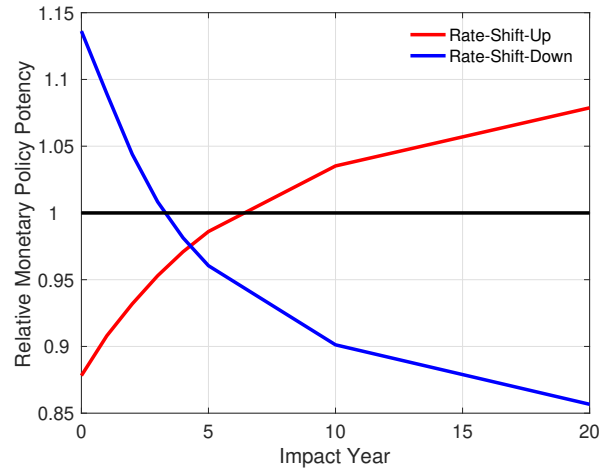
We see that in the "Rate-Shift-Down" economy, monetary policy at first is more potent than in the baseline, as the decline in rates initially boosts $frac > 0$. However, the potency of monetary policy then declines as households refinance and $frac > 0$ falls. The "Rate-Shift-Up" economy exhibits opposite patterns. However, the economies exhibit asymmetric dynamics: the potency of monetary policy falls more quickly in the "Rate-Shift-Down" economy than it rises in the "Rate-Shift-Up" economy, so that relative monetary policy potency crosses 1 (its baseline effectiveness) almost twice as fast. In this sense, monetary policy uses up ammunition when cutting rates more rapidly than it reloads ammunition when

raising rates.[56] Indeed, when rates fall, households actively refinance and the coupon distribution converges more rapidly to the new long run average, whereas rate increases affect the coupon distribution only slowly through exogenous moves, which force prepayment.[57]

8 Path-Dependent Effects on Consumption

The previous section demonstrates the path-dependent monetary transmission to mortgage prepayment and average coupons. However, we ultimately care about whether this matters for monetary transmission to aggregate spending. In this section, we embed these effects into the consumption block of our model to show that they are indeed important. We begin by showing that refinancing in our model has a large effect on individual household spending, closely in line with our empirical event study in Figure 6. Specifically, we again expose our model to actual interest rates from 1992-2017, identify those households that refinance and compute their consumption before and after refinancing. Figure 16, shows that in the model event study, upon refinancing households' average consumption jumps from $54.4k per year to $56.1 per year, a 3.1% jump in the consumption rate. The size and dynamics of these responses are very similar to the data, after accounting for the fact that the event study in the data measures investment and not service flows. Converting the empirical estimates to implied auto service flows yields an increase of 2.61% in the year after refinancing and 2.15% two years after refinancing. This model event study shows that refinancing indeed leads to increased spending for those households who refinance.

Figure 16: Event study: consumption response to refinancing in benchmark model
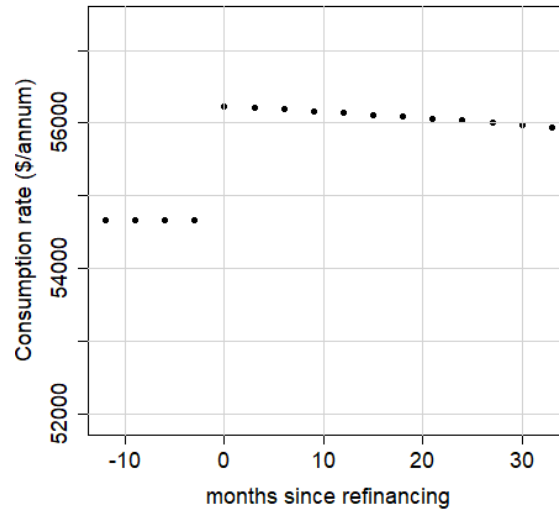
We next turn to aggregate spending responses to shocks to r. These responses cannot be characterized analytically but will clearly depend on household refinancing patterns, equilibrium relationships between r and m, and on more standard substitution, income and wealth effects of changing rates. Figure 17 shows the aggregate consumption IRF to the 100bps decline in r and the "max" rate decline in the two scenarios studied in Section 7.3: (a) the "baseline" economy, and (b) the "secular decline" economy.

---

[56]To be clear, this exercise focuses on stimulus from the first 100bps cuts. When considering "max" cuts, there is always more "policy space" in the high rate than low rate regime, but maximum stimulus power increases very slowly after raising rates.

[57]Importantly, results modeling inflows of new mortgages from young households do not change this conclusion.

31

This figure shows that our conclusions for mortgage market transmission are echoed in consumption: consumption responds much more to rate cuts in the "secular decline" economy than in the "baseline" economy. For example, on impact the aggregate spending semi-elasticity is 67bps in the baseline economy vs. 96bps in the secular decline economy (i.e. a 43% increase over the baseline).

Figure 17: IRF of Consumption – Baseline vs. Secular Decline

(a): 100bps shock   **(b):** max shock
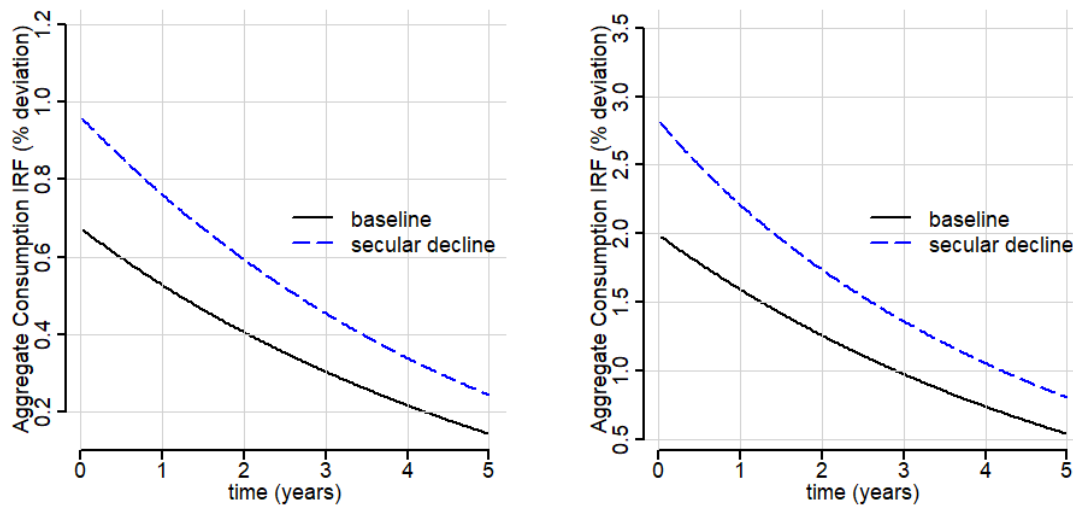
Figure 18, similarly shows the effects of past easing and tightening explored in Section 7.4 but now for consumption. Our conclusions again mirror those in mortgage space: after a tightening cycle with past high rates, monetary stimulus is more powerful than after an easing cycle with past low rates.

Figure 18: IRF of Consumption – Past High Rates vs. Past Low Rates

(a): 100bps shock   **(b):** max shock
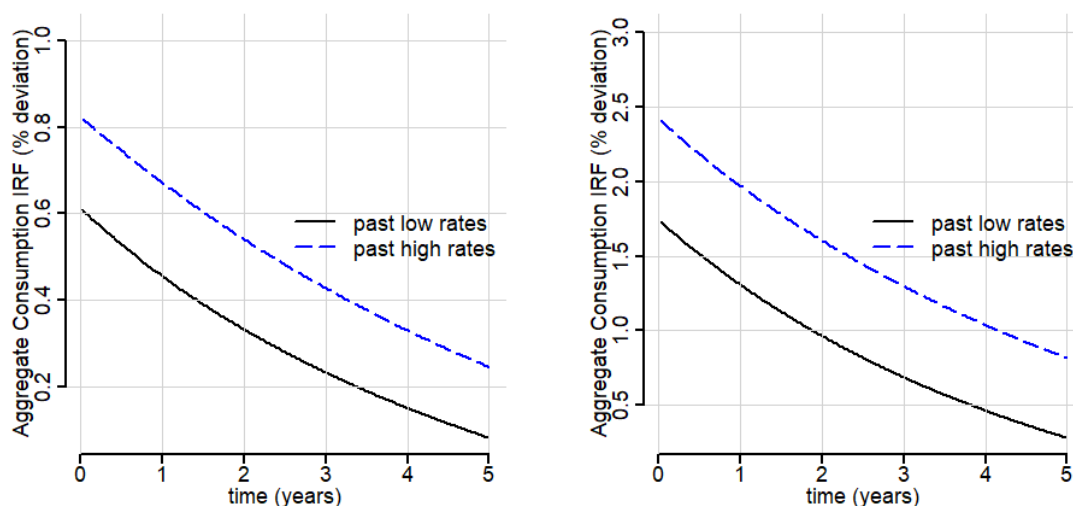
These results show numerically that our model generates large and path-dependent aggregate spend-

ing responses to rate cuts, at frequencies relevant for Fed stimulus. Path-dependent spending effects mimic path-dependent refinancing effects, and we already showed in Section 7.1 that these effects can be explained through $frac > 0$. This observation, together with the large spending responses to individual refinancing shown in Figure 16 shows that the path-dependent spending responses to monetary policy shown in this section reflect the path-dependent responses of mortgage refinancing to monetary policy. As a final piece of evidence to better understand the path-dependent transmission of interest rates to consumption, in Appendix A.3.8 we explore a simple complete markets, partial equilibrium version of our baseline model in which these effects can be characterized analytically.

In this deterministic, partial equilibrium, complete markets environment, we assume household preferences, mortgages and refinancing frictions are identical to those in our benchmark model. The household is endowed with a constant income $\bar{y}$ per unit of time, time-zero savings w, a mortgage with constant balance F, and r_t and m_t are perfectly predictable and converge asymptotically to their long-run average.[58] Steady-state consumption in this model is equal to permanent income: $c_0 = \bar{y} + \delta(w - F)$. We then analyze a small mean-reverting shock to the nominal rate in the neighborhood of the steady state, which also impacts mortgage market interest rates:

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

While we endogeneized the relationship between short rates and mortgage rates in our benchmark model, in this simplified environment we assume that the pass-through π between short rates and mortgage rates is fixed exogenously. Our analytical characterization of the impulse response of expected prepayment rates and expected coupon rates to this short rate shock (see Appendix A.3.8) shows that the average coupon response is amplified by (i) the size of the initial rate shock, (ii) the pass-through π, (iii) the persistence of the shock, and (iv) the attention rate χ_c. More importantly, we show that the semi-elasticity of consumption to a time-zero short rate shock is:

$$\left. \frac{\partial \ln c}{\partial r} \right|_{t=0} = \frac{-1}{\eta_r + \delta} \left[\frac{1}{\gamma} - \frac{\delta w}{c_0} + \mathbb{I}_{\{\epsilon_0 < 0\}} \left(\frac{\eta_r + \delta}{\eta_r + \delta + \chi_c} \right) \frac{\chi_c \pi F}{c_0} \right] \tag{6}$$

The first two terms in brackets in (6) capture standard income, substitution and wealth effects, while the third term captures the role of prepayable mortgage debt. Absent mortgage debt ($F = 0$), our result is identical to the result obtained in Kaplan, Moll and Violante (2018): the direct, partial equilibrium, effect on consumption of a small interest rate shock is higher if (a) the rate of time preference is small, (b) the persistence of the rate shock is high, and (c) the inter-temporal elasticity of substitution $1/\gamma$ is high. The consumption response is slightly muted (for reasonable asset-to-income levels) by the presence of initial savings w. The strength of the prepayment channel is increasing with (a) pass-through π, (b) the attention rate χ_c, (c) the persistence of the monetary shock, and (d) the mortgage balance F.

Our formula allows us to quantify the strength of different transmission channels in this simplified environment. Substituting the numerical parameters of our benchmark model, together with that model's average pass-through of 0.57 and $w = \$125,000$ (to match the same c_0 as our benchmark model) delivers a semi-elasticity of -1.38. The first two-terms, which capture standard transmission effects, generate a semi-elasticity of -0.66. This implies that prepayment accounts for 52% of total monetary

---

[58]To ensure finite, non-zero steady-state consumption, long-run short and mortgage rates must equal the discount rate, δ.

transmission in this model. Thus mortgage refinancing can play a large role in monetary transmission. Of course, this steady-state calculation assumes that the rate shock pushes all households from zero to positive gaps. In practice, the share of households with positive gaps after a rate cut will depend on the previous history of rates. This path-dependence is the central insight of our paper.

9 Robustness

9.1 Cash-out Refinancing

While our benchmark model is focused on the effects of payment reductions through rate refinancing, we analyze here an extension to explore whether cash-out refinancing alters our conclusions about monetary policy path-dependence. At "Calvo" attention times, households have the ability to refinance their mortgage at the then-market interest rate and to extract home equity, subject to (a) a debt-to-income limit of 43%, and (b) an LTV limit of 80%. We assume that households only extract home-equity if their interest rate gap is positive, and when they do so, we assume that they borrow up to the maximum allowed amount. The assumption that households only extract home equity when their rate gap is positive is broadly supported by the data (see Section 4.2). Our assumption that households extract the maximum possible amount of equity when refinancing is meant to be conservative: if path-dependence holds both in an environment with no cash-out refinancing and in an environment with maximum possible home equity extraction, then it will likely hold in intermediate environments where households choose mortgage balances endogenously.

To generate a non-trivial stationary cross-sectional distribution of mortgage debt, we assume mortgages amortize at a constant rate $\alpha = 3\%$ (as in Gorea and Midrigan (2017)), to match a mortgage duration (excluding prepayments) of approximately 30 years, the prevalent mortgage contract in the US. This calibration choice leads to an ergodic average mortgage balance of around $154,000, consistent with the average mortgage debt in our sample and with our benchmark model.

We then recompute IRFs to a 100bps rate decline under different rate histories. Appendix Figure A-12 shows that cashout refinancing strengthens monetary transmission and amplifies our earlier conclusions. Interestingly, in the model with cashout refinancing, IRFs exhibit a mild hump shape. Cashout refinancing amplifies the consumption response to rate declines, and refinancing occurs more gradually than standard income, substitution and wealth responses to short rate changes, which occur immediately on impact.

9.2 Refinancing and the Life-cycle

Our baseline model assumes infinitely lived agents and abstracts from the life-cycle. Life-cycle effects are important for explaining certain housing patterns, and Wong (2019) shows that these effects matter for the average transmission of monetary policy through refinancing. It is less obvious that these channels should matter for the path-dependence and time-varying effects of monetary policy that we emphasize in this paper. To explore this, we analyze an extension of our benchmark model that includes life-cycle elements. Households transition stochastically between "young", "middle-aged" and "old" at Poisson arrival times. We calibrate labor income, retirement income and model amortizing mortgages with balances matched to data counterparts in the 2001 SCF. Households also have a standard bequest motive

$W^{1-\gamma}/(1-\gamma)$. This model thus matches broad life-cycle debt, income and consumption patterns.[59]. To preserve stationarity, whenever an old household dies, a young household is borne. Appendix Figure A-13 shows that including life-cycle elements indeed leaves our conclusions about path-dependent effects of monetary policy unchanged.

9.3 Hybrid Refinancing Frictions

Figure A-14 shows that repeating our exercise using the "hybrid" model with both Calvo and menu cost frictions, set to match the empirical prepayment hazard yields similar conclusions.[60]

9.4 Alternative Monetary Policy Persistence

While we estimated our interest rate process (3) using MLE, monetary policy induced rate changes might have different persistence from the general time-series of interest rates we observe in the data. To ensure our results do not depend materially on this persistence parameter, we re-solve our benchmark model with both higher and lower persistence of the short-rate process. Appendix Figure A-15 shows results. As implied by Equation 6, average consumption responses increase with the persistence of interest rates: a shock with greater persistence is a "bigger shock" in cumulative terms and mechanically leads to bigger responses. However, changing the persistence of our rate process leaves our conclusions about path-dependence unchanged.

9.5 The Role of Foreigners

We calibrate our baseline model so that on average, when interest rates decline, mortgage payments fall by the same amount as capital income, implying that capital income declines are borne on average by domestic savers. While this holds on average, the financial intermediary still faces profit shocks under different rate paths. In our baseline model, we assume that the intermediary is owned by foreigners who absorb these shocks. However, the quantitative magnitude of this assumption is modest. To show this, we assume instead that financial intermediaries are owned by wealthy, risk-neutral domestic agents who smooth consumption out of permanent income, and then recalculate how aggregate spending responds to rate cuts in the secular decline vs. the ergodic economy. In this model with domestic owners, average household spending rises by \$146 more in the secular decline economy than in the ergodic economy on impact after a 100bp rate cut. This difference is \$171 in our baseline model with foreign owned intermediaries. Thus, we preserve 85% of our benchmark model's path-dependence if we assume fully domestic-owned rather than fully foreign-owned intermediaries.

10   Conclusion

The Fed kept interest rates low for a long period of time after the Great Recession to stimulate the economy. Rates have now risen but are still low by historical standards, leaving relatively little room

---

[59]We define young households in the data as 25-40, middle as 40-62 and old as 62-75. We match average labor income for the young and middle aged households to SCF data and assume income risk as in Floden and Linde (2001). The old have constant retirement income matched to SCF data. We also match mortgage balances by age group, so young households (resp. middle-age and old) have 1.11 (resp. 1.00 and 0.65) times the household average mortgage balance.

[60]For numerical feasibility, we continue to use the mortgage pricing function $m(r)$ from the Calvo model.

for future cuts if they are needed to stimulate the economy. In this paper, we argue that looking solely at the level of *current* rates provides an incomplete picture of Fed "policy space", because the presence of significant US household debt in fixed-rate pre-payable mortgage contracts leads to path-dependent consequences of monetary policy. In fact, we argue that the path of interest rates in recent years leaves the Fed with even less room for stimulus than suggested by current low rates.

We argue that monetary policy "reloads" stimulus power very slowly after raising rates. Furthermore, the long secular decline in mortgage rates is unlikely to continue forever, and monetary policy potency will be weaker in a stable or increasing rate environment. Finally, the extended period with zero rates allowed households to lock in low mortgage rates, which will dampen future Fed stimulus power. We highlight these observations in our modeling exercise, but it is important to note that our modeling framework can provide transparent policy guidance more generally: $frac > 0$ is a straightforward statistic to measure the Fed's power to stimulate mortgage markets at a moment in time. Moreover, our micro-data consistent model allows us to easily calculate how $frac > 0$ and thus future stimulus power evolves in response to policy actions today.

Our insight that rate histories matter for current monetary policy transmission through the mortgage market is not necessarily specific to the US. While mortgage contracts with fixed rates and no prepayment penalty are uncommon outside of the US and Denmark, similar path-dependence forces can arise due to the timing of home purchases in countries with limited refinancing and can induce echo effects several years in the future in countries where mortgage contracts must be refinanced at fixed intervals.[61]

More broadly, our point that the Fed faces an intertemporal trade-off between current and future policy effectiveness extends beyond the mortgage context which is the focus of our paper. Similar forces likely exist for any rate-sensitive decision with irreversibility, such as auto and home purchases and firm investment decisions (McKay and Wieland, 2019). Lowering rates can stimulate the economy by encouraging adjustment today but at the cost of lowering future effectiveness because there are fewer agents left to adjust in the future. In our paper, we focus on a context in which this broad mechanism is most easily observed and disciplined.

Finally, we provide a positive analysis of monetary policy effectiveness at a point in time, and we show how current actions affect future effectiveness. We do not deliver a normative analysis of optimal policy. For example, our results show that it will take a long time for monetary policy to recover ammunition after raising rates. This means that leaving rates low for a long time may have negative consequences for future stimulus ability, but this does not on its own imply that the Fed should raise rates earlier in order to regain policy space. We characterize these intertemporal effects in this paper, but we think exploring implications of prepayable mortgage debt for optimal policy is an interesting area for future research.

References

Abel, Joshua, and Andreas Fuster. 2018. "How Do Mortgage Refinances Affect Debt, Default, and Spending? Evidence from HARP."

---

[61]For example, the Bank of Canada Financial System Review in November 2017 expressed concern that 47% of all Canadian mortgages would refinance in 2018 since many households took out new mortgages when rates were low in 2013.

Agarwal, Sumit, Gene Amromin, Souphala Chomsisengphet, Tim Landvoigt, Tomasz Piskorski, Amit Seru, and Vincent Yao. 2017. "Mortgage Refinancing, Consumer Spending, and Competition: Evidence from the Home Affordable Refinancing Program." NBER Working Paper 21512.

Aiyagari, S Rao. 1994. "Uninsured idiosyncratic risk and aggregate saving." *The Quarterly Journal of Economics*, 109(3): 659–684.

Andersen, Steffen, John Y. Campbell, Kasper Nielsen, and T. Ramadorai. 2019. "Sources of Inaction in Household Finance: Evidence from the Danish Mortgage Market." *NBER Working Paper No. 21386*.

Auclert, Adrien. 2019. "Monetary policy and the redistribution channel." *AER*, 109(6).

Auerbach, Alan, and Yuriy Gorodnichenko. 2012. "Fiscal Multipliers in Recession and Expansion." In *Fiscal Policy after the Financial Crisis*. University of Chicago Press.

Beraja, Martin, Andreas Fuster, Erik Hurst, and Joseph Vavra. 2019. "Regional Heterogeneity and the Refinancing Channel of Monetary Policy." *Quarterly Journal of Economics*.

Berger, David, Veronica Guerrieri, Guido Lorenzoni, and Joseph Vavra. 2018. "House Prices and Consumer Spending." *Review of Economic Studies*, 85(3): 1502–1542.

Bhutta, Neil, and Benjamin J. Keys. 2016. "Interest Rates and Equity Extraction During the Housing Boom." *American Economic Review*, 106: 1742–74.

Caballero, Ricardo, and Eduardo Engel. 2007. "Price stickiness in Ss models: New interpretations of old results." *Journal of Monetary Economics*, 54.

Chen, Hui, Michael Michaux, and Nikolai Roussanov. 2019. "Houses as ATMs: Mortgage Refinancing and Macroeconomic Uncertainty." *Journal of Finance*, 75: 542–588.

Coibion, Olivier. 2012. "Are the Effects of Monetary Policy Shocks Big or Small?" *American Economic Journal: Macroeconomics*, 4: 1–32.

Cox, John C, Jonathan E Ingersoll Jr, and Stephen A Ross. 2005. "A theory of the term structure of interest rates." In *Theory of Valuation*. 129–164. World Scientific.

Crawley, Edmund, and Andreas Kuchler. 2018. "Consumption Heterogeneity: Micro Drivers and Macro Implications." *Danish National Bank Working Paper 129*.

DeFusco, Anthony, and John Mondragon. 2020. "No Job, No Money, No Refi: Frictions to Refinancing in a Recession." mimeo.

Deng, Yongheng, John M. Quigley, and Robert Order. 2000. "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options." *Econometrica*, 68(2): 275–307.

Di Maggio, Marco, Amir Kermani, Benjamin J. Keys, Tomasz Piskorski, Rodney Ramcharan, Amit Seru, and Vincent Yao. 2017. "Interest Rate Pass-Through: Mortgage Rates, Household Consumption, and Voluntary Deleveraging." *American Economic Review*, 107(11): 3550–88.

Doepke, Matthias, and Martin Schneider. 2006. "Inflation and the Redistribution of Nominal Wealth." *Journal of Political Economy*, 114(6): 1069–1097.

Doepke, Matthias, Martin Schneider, and Veronika Selezneva. 2015. "Distributional Effects of Monetary Policy."

Eichenbaum, Martin, Sergio Rebelo, and Arlene Wong. 2019. "State Dependency and the Efficacy of Monetary Policy: The Refinancing Channel." NBER Working Paper 25152.

Floden, Martin, and Jesper Linde. 2001. "Idiosyncratic Risk in the United States and Sweden: Is There a Role for Government Insurance?" *Review of Economic Dynamics*, 4(2): 406–437.

Gertler, Mark, and Peter Karadi. 2015. "Monetary Policy Surprises, Credit Costs, and Economic Activity." *American Economic Journal: Macroeconomics*, 7(1): 44–76.

Gilchrist, Simon, David Lopez-Salido, and Egon Zakrajsek. 2015. "Monetary Policy and Real Borrowing Costs at the Zero Lower Bound." *American Economic Journal: Macroeconomics*, 7(1): 77–109.

Gorea, Denis, and Virgiliu Midrigan. 2017. "Liquidity constraints in the US housing market." NBER.

Green, Jerry, and John B Shoven. 1986. "The Effects of Interest Rates on Mortgage Prepayments." *Journal of Money, Credit and Banking*, 18(1): 41–59.

Greenwald, Daniel. 2018. "The Mortgage Credit Channel of Macroeconomic Transmission." MIT Sloan Working Paper.

Guren, Adam M, Alisdair McKay, Emi Nakamura, and Jon Steinsson. 2018. "Housing Wealth Effects: The Long View." National Bureau of Economic Research Working Paper 24729.

Justiniano, Alejandro, Giorgio E Primiceri, and Andrea Tambalotti. 2017. "The Mortgage Rate Conundrum." National Bureau of Economic Research Working Paper 23784.

Kaplan, Greg, Benjamin Moll, and Giovanni L Violante. 2018. "Monetary policy according to HANK." *American Economic Review*, 108(3): 697–743.

Keys, Benjamin J., Devin G. Pope, and Jaren C. Pope. 2016. "Failure to Refinance." *Journal of Financial Economics*, 122: 482–499.

Krivenko, Pavel. 2018. "Unemployment and the US housing market during the Great Recession." *Mimeo*.

Landvoight, Tim, Monika Piazzesi, and Martin. Schneider. 2015. "The Housing Market(s) of San Diego." *American Economic Review*, 105: 1371–1407.

McKay, Alisdair, and Johannes F Wieland. 2019. "Lumpy Durable Consumption Demand and the Limited Ammunition of Monetary Policy." NBER Working Paper 26175.

Mian, Atif, Kamelesh Rao, and Amir Sufi. 2013. "Household Balance Sheets, Consumption, and the Economic Slump." *Quarterly Journal of Economics*, 1687–1726.

Nakamura, Emi, and Jon Steinsson. 2018. "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect." *The Quarterly Journal of Economics*, 133(3): 1283–1330.

Ramey, Valerie. 2016. "Macroeconomic Shocks and Their Propagation." *Handbook of Macroeconomics*.

Romer, Christiana, and David. Romer. 2004. "A new measure of monetary shocks: Derivation and implications." *American Economic Review*, 94: 1055–1084.

Scharfstein, David, and Adi Sunderam. 2016. "Market Power in Mortgage Lending and the Transmission of Monetary Policy."

Schwartz, Eduardo S., and Walter N. Torous. 1989. "Prepayment and the Valuation of Mortgage-Backed Securities." *The Journal of Finance*, 44(2): 375–392.

Vavra, Joseph. 2014. "Inflation Dynamics and Time-Varying Volatility: New Evidence and an Ss Interpretation." *Quarterly Journal of Economics*, 129(1).

Wong, Arlene. 2019. "Refinancing and the Transmission of Monetary Policy to Consumption." Princeton University Working Paper.

11 Tables

Table 1: Predictions of Alternative Summaries for Rate Incentives

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
frac > 0	2.051***						
	(0.386)						
frac > 50bp		1.989***					
		(0.389)					
frac > 100bp			2.249***				
			(0.450)				
Mean Gap				0.759***			
				(0.132)			
frac > $250					1.926***		
					(0.384)		
frac > $500						1.844***	
						(0.399)	
frac > $1000							1.662***
							(0.437)
Constant	0.0416	0.365**	0.673***	0.993***	0.304*	0.516***	0.941***
	(0.196)	(0.145)	(0.107)	(0.0558)	(0.156)	(0.131)	(0.0913)
Adj. R^2	0.282	0.277	0.246	0.263	0.274	0.247	0.142
N	304	304	304	304	304	304	304
Date Range	92-17m4	92-17m4	92-17m4	92-17m4	92-17m4	92-17m4	92-17m4

Newey-West standard errors in parantheses. *=10%, **=5%, ***=1% significance. frac > 50bp is the fraction of loans with gaps greater than 50 basis points and frac > 100bp is the fraction with gaps greater than 100 basis points. Mean gap is the average gap in a month. frac > $250 ($500, $1000) is the fraction of loans with annual savings greater than $250 ($500, $1000), which we compute by multiplying the current gap times the outstanding balance. 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.

Table 2: Relationship Between Rate Incentives and Prepayment

	(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	2.05***	2.82***		2.06***	2.70***	2.07***	2.16***	4.82***	2.30***
	(0.39)	(0.50)		(0.36)	(0.87)	(0.39)	(0.33)	(0.86)	(0.42)
LTV		-6.49***	-2.35				-2.36*	-12.3***	-0.89
		(1.66)	(1.49)				(1.25)	(2.53)	(0.83)
Constant	0.042	3.58***	2.88***	-0.045	-0.026	-0.40	1.43*	5.86***	-0.0034
	(0.20)	(0.80)	(1.00)	(0.19)	(0.40)	(0.30)	(0.85)	(1.07)	(0.50)
Adj. R^2	0.28	0.47	0.027	0.56	0.30	0.61	0.57	0.67	0.62
N	304	304	304	108	120	76	108	120	76

Newey-West standard errors in parantheses. *=10%, **=5%, ***=1% significance. 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.

Table 3: Auto Sales Growth Responses to Refinancing and Mortgage Rate Changes

	(1) OLS	(2) OLS	(3) IV	(4) IV	(5) OLS	(6) OLS	(7) IV	(8) IV
Δ freq	0.0190***	0.0299***	0.0809***	0.135***				
	(0.00725)	(0.00942)	(0.0277)	(0.0397)				
freq					-0.00130	0.0329***	0.00365	-0.124
					(0.00339)	(0.0111)	(0.0146)	(0.0968)
freq $\times\Delta$ FRM					-0.0343*	-0.0446*	-0.203***	-0.386**
					(0.0191)	(0.0229)	(0.0661)	(0.179)
Quarter $\times$ MSA FE	No	Yes	No	Yes	No	Yes	No	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	85328	84585	85328	84585	85328	84585	85328	84585
Date Range	98-17m4	98-17m4	98-17m4	98-17m4	98-17m4	98-17m4	98-17m4	98-17m4

Standard errors two-way clustered by MSA and month. *=10%, **=5%, ***=1% significance. Prepayment is measured using loan level data from McDash Performance data. Δ freq is the change in prepayment and Δ FRM is the change in the 30-year fixed rate mortgage between month $t - 1$ and t. To account for lags between origination and spending, the outcome in all regressions is auto sales growth from R.L. Polk measured between t and $t + 1$. IV specifications instrument for freq and Δ freq using frac > 0 and Δ frac >0.

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. 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 prepaying. 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 prepaying loan to a newly originating loan so that we can measure exactly why each individual loan is prepaying. 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
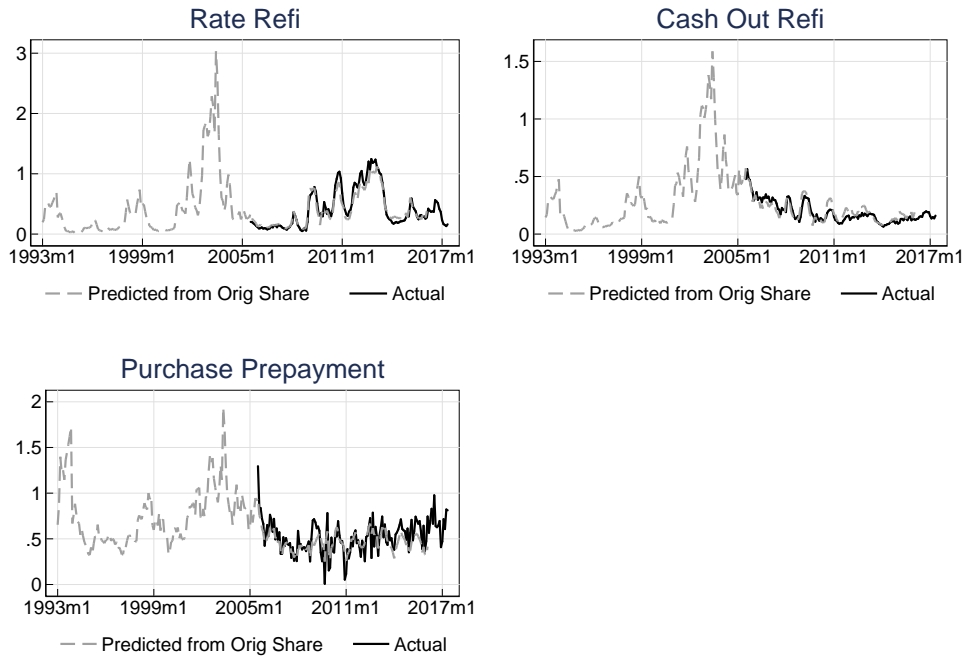
Overall, the McDash Performance data set contains information on approximately 180 million loans. After 2005, the McDash Performance data set covers roughly 50% of total U.S. mortgage debt as measured by the Federal Reserve. Prior to 2005, coverage is somewhat lower, ranging from around 10% market coverage 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.[62] 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.[63]

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 next. Nevertheless, we have repeated our analysis of prepayment using series derived from CoreLogic

---

[62]Note 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.

[63]Note 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.

LLMA Performance data and arrive at similar conclusions.

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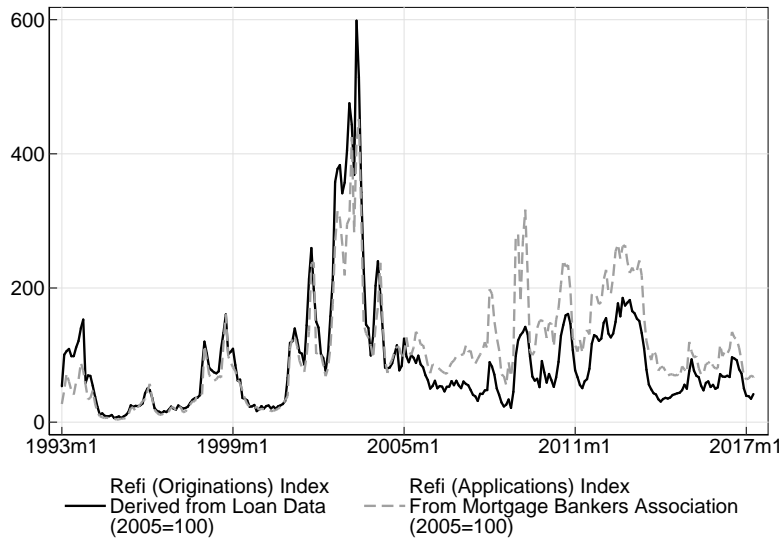

Figure shows an index of refinancing computed using McDash Performance data and CoreLogic Origination Purpose data compared to the Mortgage Banker's Association Refinancing Application Index. Note that the loan-level index measures originations while the MBA index measures applications. Indices are normalized 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 Mc-Dash 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 individuals[64] 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.[65] These sample restrictions leave

---

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

[65]Equifax 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 be 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.[66] 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)[67], 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

---

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

[67]To 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.[68] 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.

---

[68]We 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.[69] 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.

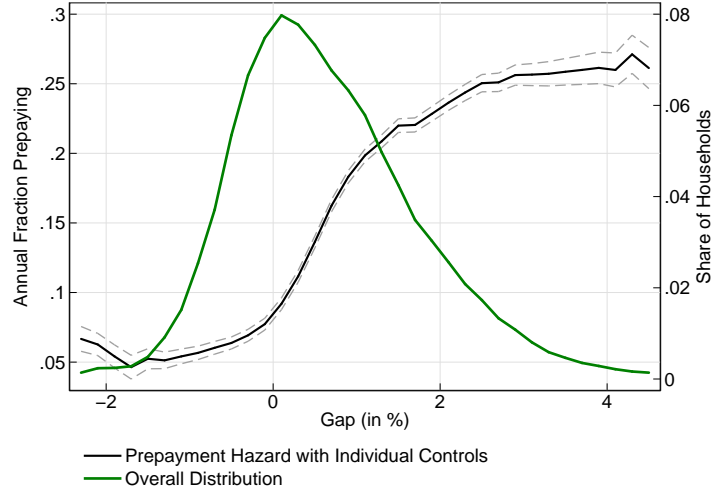**Figure A-3: Robustness of Prepayment Hazard to Annual Instead of Monthly Frequency**

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.

---

[69]We 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
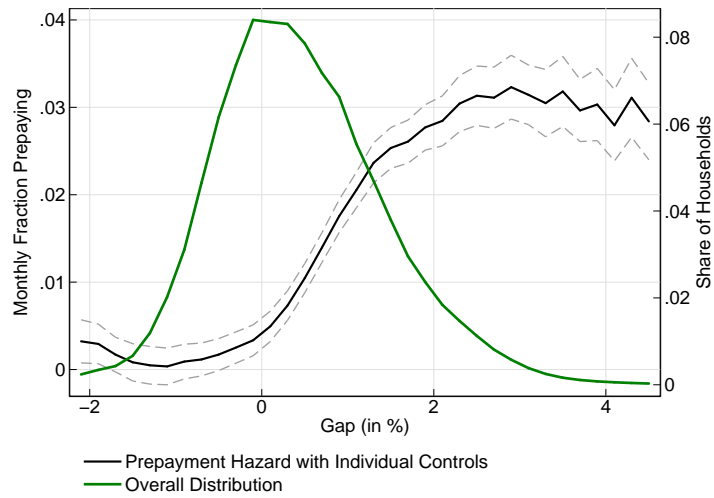

Figure shows point estimates and 95% confidence intervals for coefficients on the 20bp bin dummies in regression 1. We exclude loan-month observations with balances less than $100,000. Standard errors are two-way clustered by household and month using data from 2005m6-2017m4.

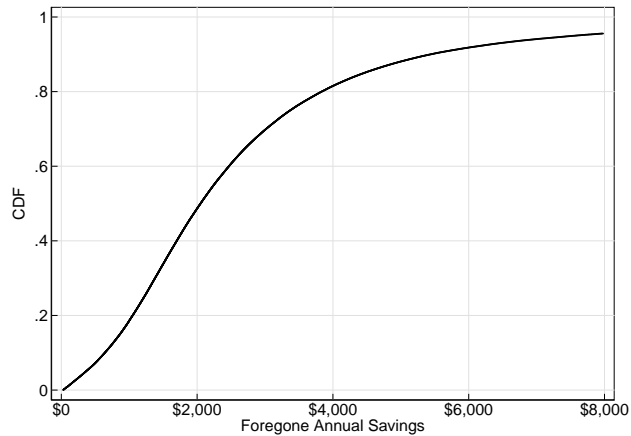**Figure A-5: Annual Payment Savings Lost By Not Refinancing**

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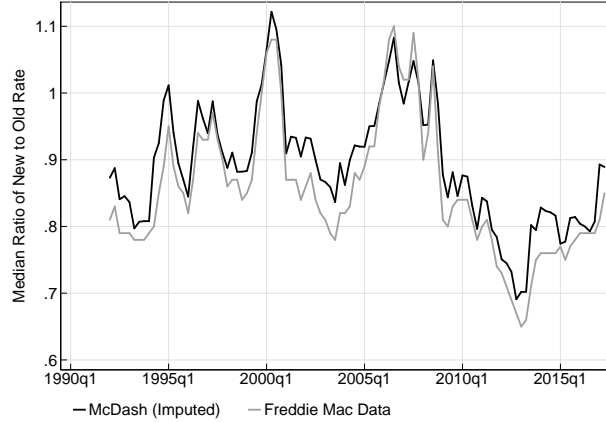
For each month, we estimate the ratio of new rates relative to outstanding old rates for refinancing loans in our data. We then calculate the median of this ratio for each month and compare this to published data for the same object from Freddie Mac.

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
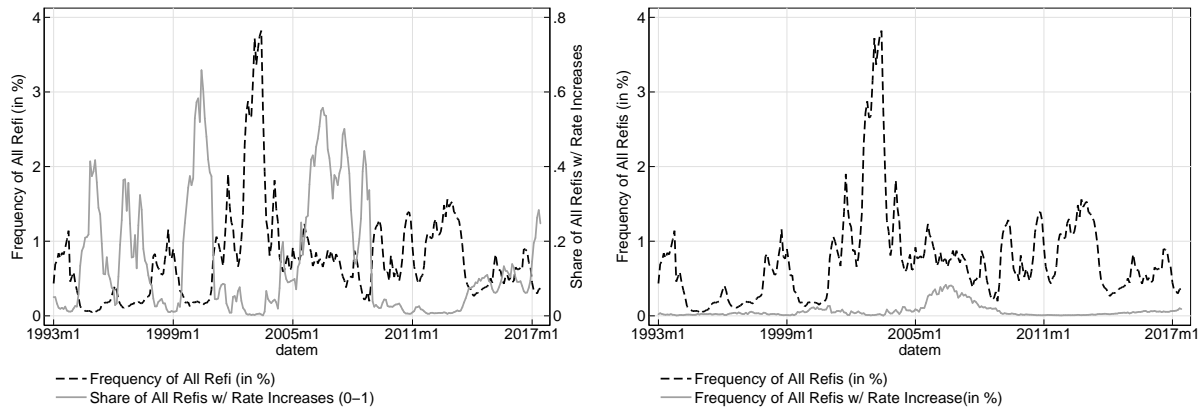
In particular, for each each month and for all prepaid GSE loans in the McDash data, we compute the estimated gap between the outstanding interest rate and the estimated new rate for the loan. We then apply the refinancing hazard computed in CRISM data after 2005 to allocate all of these prepayments 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.
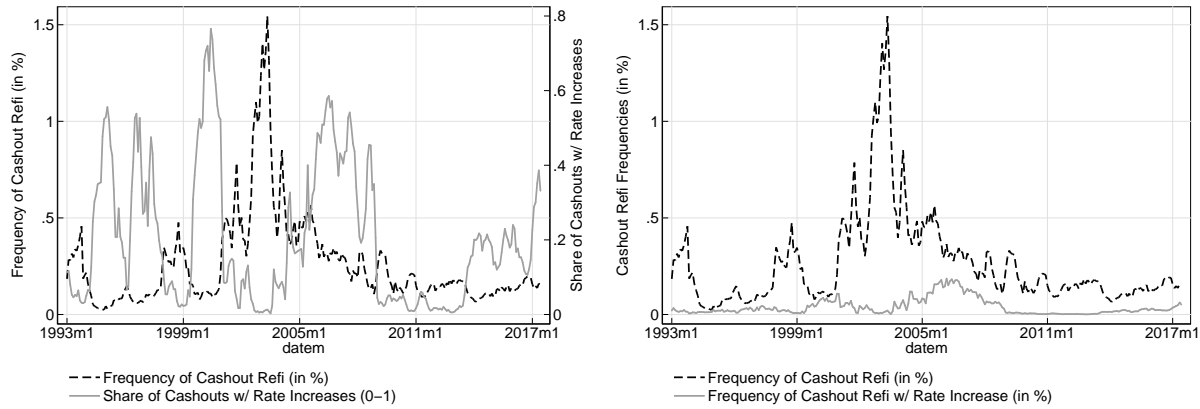
Furthermore, Figure A-7 shows that changes in refinancing frequency are key to reconciling this time-series pattern with our cross-sectional fact that refinancing into higher rates is very unusual. Panel (a)

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

shows that the times when most refis result in rate increases are precisely the times when the frequency 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, Michaux and Roussanov (2019).

Figure A-9: Distribution of Gaps at Two Dates
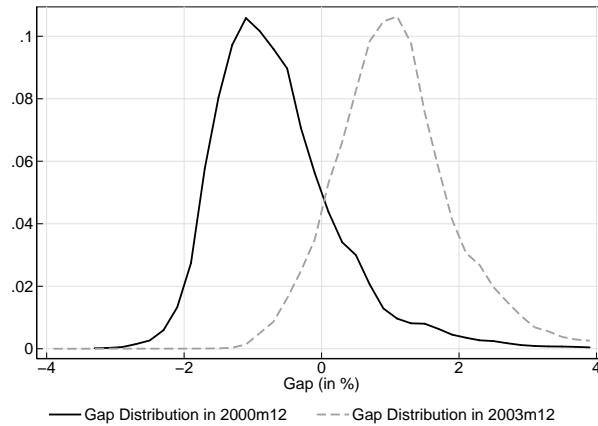
Figure A-9 shows the distribution of interest rate gaps at two points in time. The solid black line shows the distribution of gaps in December of 2000 while the dotted gray line shows the distribution of gaps in December of 2003. The key takeaway is that there is significant variation in these distributions 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.

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

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***	2.82***	2.82***	2.82***	2.83***	2.48***	2.53***	2.14***
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.36)	(0.38)	(0.39)
LTV	-6.49***	-6.49***	-6.49***	-6.49***	-6.64***	-4.96***	-4.96***	0.53
	(1.66)	(1.66)	(1.66)	(1.66)	(1.70)	(1.15)	(1.26)	(5.30)
Constant	3.58***	3.58***	3.58***	3.58***	3.64***	2.79***	2.80***	-0.97
	(0.80)	(0.80)	(0.80)	(0.80)	(0.81)	(0.57)	(0.61)	(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.

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, Primiceri and Tambalotti (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.[70] 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 (Agarwal 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 × 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 $frac > 0$.

---

[70]We 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-2: Effects of Rate Gaps on Prepayment Propensities by MSA

	(1) Tot Prepay	(2) Rate-Refi	(3) Cashout	(4) Purchase
frac > 0	3.36***	2.06***	0.63***	0.65***
	(0.36)	(0.25)	(0.13)	(0.25)
Quarter X MSA FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Adj. R^2	0.98	0.93	0.94	0.79
N	87219	87219	87219	87219
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***	3.01***	3.34***
	(0.83)	(0.72)	(0.74)
LTV		-7.83***	-6.32***
		(1.57)	(1.33)
Constant	-0.073	4.43***	3.50***
	(0.55)	(0.73)	(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

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.[71]

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.[72] $frac > 0$ 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.

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***		1.465***	0.296**		0.599***	0.400***		0.670***
	(0.187)		(0.255)	(0.128)		(0.118)	(0.132)		(0.124)
LTV		0.343	-1.787**		-1.786***	-2.657***		-1.385***	-2.358***
		(0.820)	(0.844)		(0.303)	(0.399)		(0.442)	(0.401)
Constant	-0.411***	0.212	0.566	0.0670	1.375***	1.520***	0.380***	1.507***	1.669***
	(0.0948)	(0.533)	(0.413)	(0.0672)	(0.206)	(0.189)	(0.0690)	(0.281)	(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.

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 constraint 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.

---

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

[72]After 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 Table **??**.

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

	(1)	(2)	(3)	(4)
frac > 0	-0.0486***		-0.0470***	-0.0531***
	(0.00579)		(0.00531)	(0.00654)
Δ FRM		0.0188***	-0.0286**	0.0369
		(0.00516)	(0.0124)	(0.0356)
Δ FRM $\times$ (frac > 0)			0.0542**	0.0679***
			(0.0231)	(0.0244)
Δ FRM $\times$ mean LTV				-0.118*
				(0.0673)
mean LTV				0.0521**
				(0.0217)
Constant	0.0180***	-0.0142***	0.0175***	-0.0109
	(0.00284)	(0.00178)	(0.00273)	(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. Δ FRM 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, $frac > 0$ and LTV is measured as of month t, Δm^* is measured between month t and month $t+1$ and Δ FRM 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 $frac > 0$ 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 $frac > 0$ and Δ FRM: interest rate pass-through into average coupons is much stronger when $frac > 0$ is large. As we discuss below, this increase in rate pass-through with $frac > 0$ 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 $frac > 0$ and Δ FRM 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***	2.90***		2.48***	2.86***	2.51***	2.55***	5.06***	2.53***
	(0.37)	(0.50)		(0.39)	(0.88)	(0.39)	(0.37)	(0.92)	(0.43)
LTV		-5.85***	-1.09				-1.58	-12.7***	-0.064
		(1.83)	(1.62)				(1.70)	(2.89)	(0.90)
Constant	0.074	3.31***	2.16**	-0.086	0.024	-0.65**	0.90	6.18***	-0.62
	(0.16)	(0.93)	(1.07)	(0.19)	(0.35)	(0.29)	(1.14)	(1.28)	(0.52)
Adj. R^2	0.33	0.45	0.0022	0.61	0.33	0.68	0.62	0.67	0.67
N	303	303	303	108	120	75	108	120	75

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***	2.63***		1.77***	2.69***	1.40***	1.85***	4.55***	1.98***
	(0.43)	(0.45)		(0.33)	(0.92)	(0.39)	(0.30)	(0.82)	(0.38)
LTV		-6.65***	-4.24***				-2.23**	-11.3***	-2.15***
		(1.30)	(1.36)				(1.08)	(1.91)	(0.71)
Constant	-0.12	3.60***	4.01***	-0.024	-0.29	-0.039	1.38*	4.95***	0.88*
	(0.26)	(0.62)	(0.91)	(0.19)	(0.56)	(0.30)	(0.76)	(0.74)	(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***	3.22***		2.41***	3.18***	2.69***	2.50***	5.19***	2.64***
	(0.40)	(0.54)		(0.39)	(0.90)	(0.38)	(0.37)	(0.95)	(0.42)
LTV		-6.98***	-1.88				-2.23	-12.6***	0.25
		(1.95)	(1.76)				(1.51)	(2.93)	(1.06)
Constant	0.092	3.96***	2.82**	-0.037	0.084	-0.61**	1.37	6.31***	-0.73
	(0.19)	(0.97)	(1.17)	(0.20)	(0.37)	(0.27)	(1.02)	(1.29)	(0.59)
Adj. R^2	0.33	0.48	0.011	0.58	0.36	0.71	0.59	0.66	0.71
N	303	303	303	108	120	75	108	120	75

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 to our main regression results when we only include loans with a fixed interest rate. 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 ARM share, while time-varying, is typically low in our sample reflecting the fact that the 30-yr fixed rate mortgage is the dominant mortgage contract in the U.S. over this sample period and also because even when focusing on ARMs, many are still in the initial fixed rate period when gaps and refi incentives can emerge.

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***	2.79***		2.01***	2.40***	1.90***	2.09***	5.09***	2.24***
	(0.40)	(0.50)		(0.30)	(0.91)	(0.39)	(0.27)	(0.88)	(0.40)
LTV		-6.89***	-3.04**				-2.32**	-13.6***	-1.24
		(1.58)	(1.38)				(1.08)	(2.39)	(0.77)
Constant	0.22	3.97***	3.34***	0.069	0.27	-0.20	1.53**	6.60***	0.35
	(0.20)	(0.76)	(0.93)	(0.16)	(0.41)	(0.27)	(0.75)	(0.97)	(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 $P(r, r^*)$:

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

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

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

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:

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

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 $m^* \neq m(r)$):

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

$\mathcal{L}_y^*$ (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 $m(r)$:

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

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:

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

In the above, $C\left(r_t, W_t, m_t^*, Y_t\right)$ 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 artifical 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 $\vec{g}_t$ be 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:

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

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:

$$\frac{\partial}{\partial r_t} \left(\frac{d\bar{m}^*}{dt} \right) = m'(r_t) [\nu + \chi [1 - F_t(m(r_t))]] - \chi \left(\frac{\partial}{\partial r_t} \mathbb{E}_i[m_{it}^*|m_{it}^* > m(r_t)] \right) [1 - F(m(r_t), t)]$$
$$+ \chi (\mathbb{E}_i[m_{it}^*|m_{it}^* > m(r_t)] + m(r_t)) f_t(m(r_t)) m'(r_t)$$

Note that

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

so that

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

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

Plugging in, we are left with

$$\frac{\partial}{\partial r_t}\left(\frac{d\bar{m}^*}{dt}\right) = m'\left(r_t\right)\left[\nu + \chi\left[1 - F\left(m\left(r_t\right), t\right)\right]\right] \tag{A5}$$

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

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

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:

$$V(w, \epsilon; \bar{y}) := \max_c \int_0^{+\infty} e^{-\delta t} \frac{c_t^{1-\gamma}}{1 - \gamma} dt$$

$$\text{s.t. } dw_t = \left[(\delta + \epsilon_t)w_t + \bar{y} - c_t\right] dt$$

$$d\epsilon_t = -\eta_r \epsilon_t dt$$

$$(w_0, \epsilon_0) = (w, \epsilon)$$

21

The value function V satisfies the following HJB equation

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

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

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

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

$$c_0^{-\gamma} = V_0'$$
$$-\gamma \frac{c_1}{c_0} = \frac{V_1'}{V_0'}$$

The terms of order zero in our asymptotic expansion must satisfy

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

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

$$V_1(w;\bar{y}) = \frac{w V_0'(w;\bar{y})}{\eta_r + \delta} = \frac{w (\delta w + \bar{y})^{-\gamma}}{\eta_r + \delta}$$
$$\frac{c_1(w;\bar{y})}{c_0(w;\bar{y})} = \frac{-1}{\eta_r + \delta} \left[\frac{1}{\gamma} - \frac{\delta w}{c_0(w;\bar{y})} \right]$$

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, Moll and Violante (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

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

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

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

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 Additional Figures and Tables

Figure A-10 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 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-11 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-11 (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

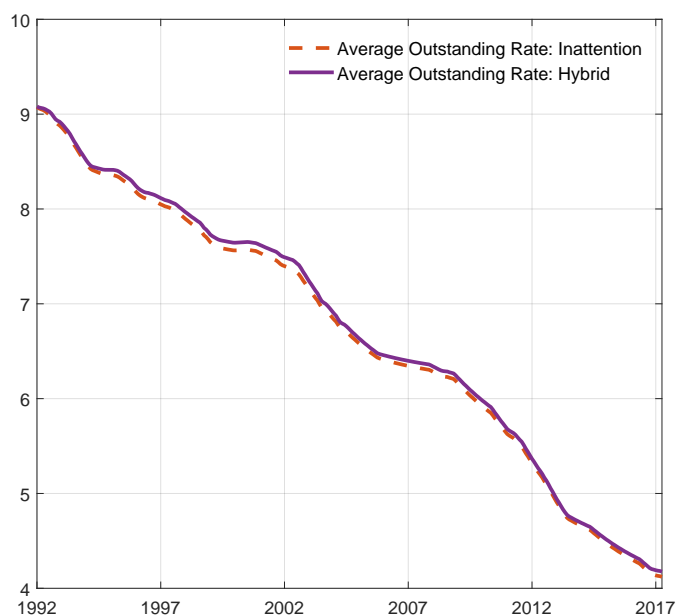**Figure A-10:** Hybrid Menu Cost+Inattention Model vs. Inattention Model: Time-Series Fit
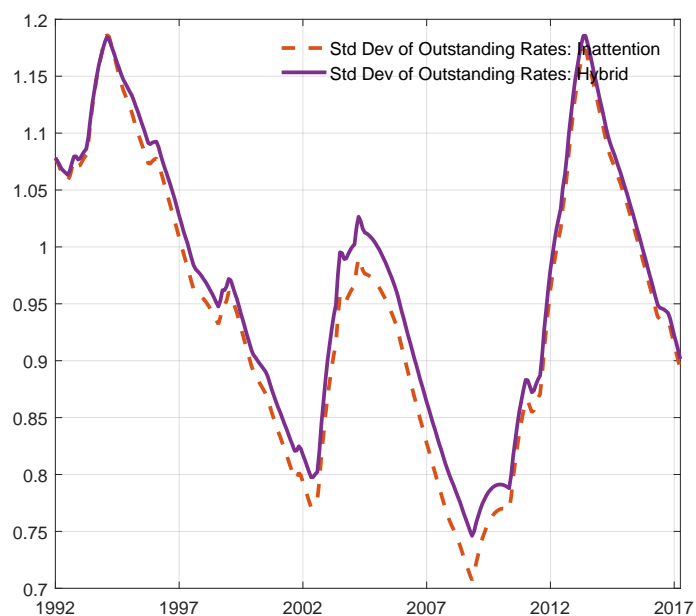
(a): Distribution of Gaps
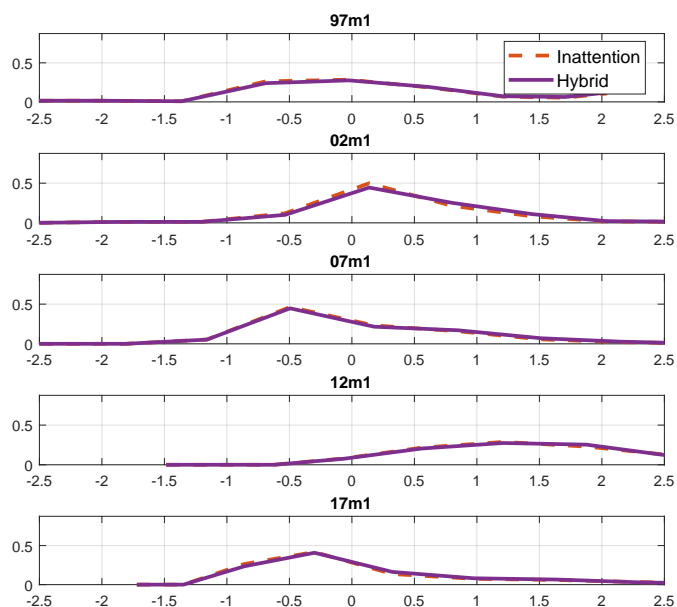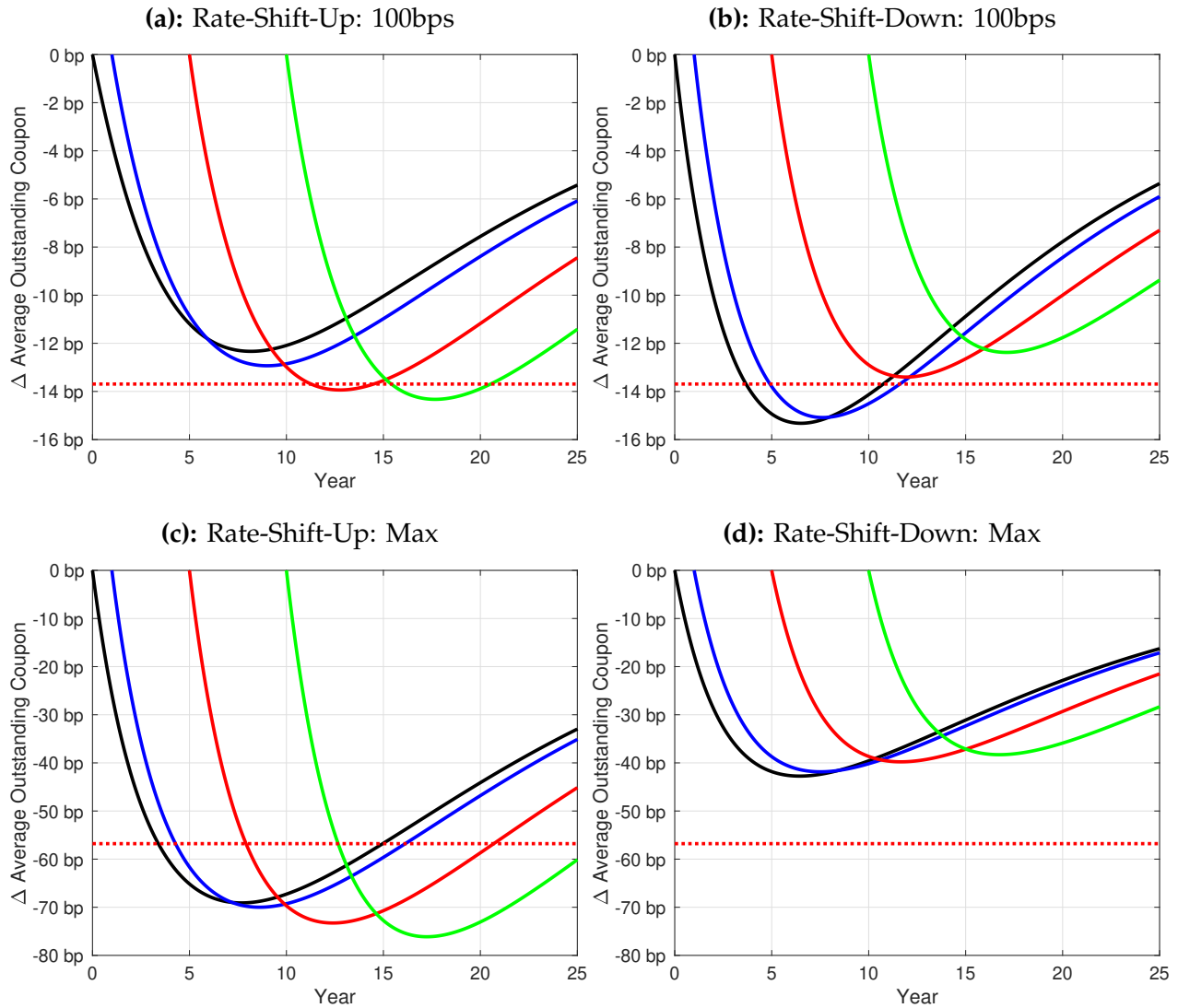
(b): Frequency
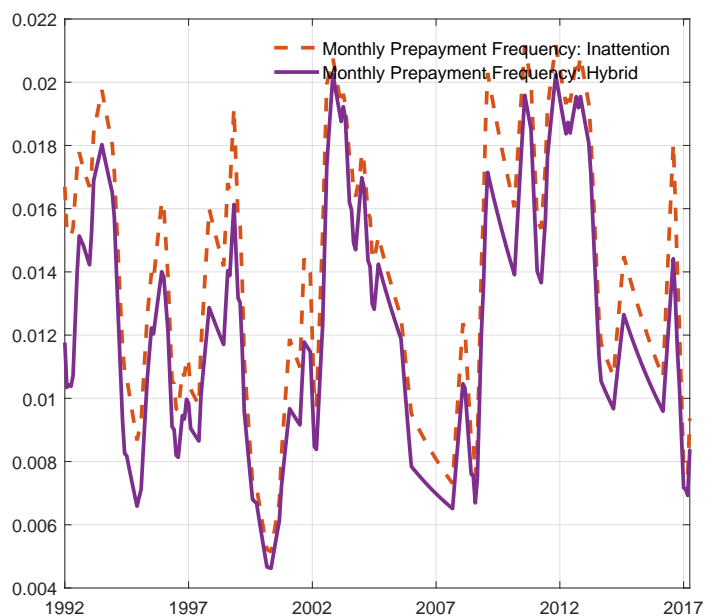
(c): Average Rates

(d): Std Dev of Rates

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.

Figure A-11: Regime Shift: Average Coupon m^* to 100bps & Max decline in r

(a): Rate-Shift-Up: 100bps

(b): Rate-Shift-Down: 100bps

(c): Rate-Shift-Up: Max

(d): Rate-Shift-Down: Max

There are a number of takeaways from Figure A-11. 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-12: Robustness of Consumption IRFs to Including Cashout Refi
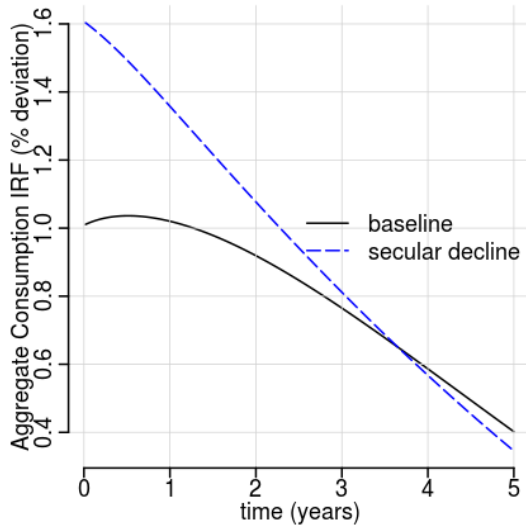
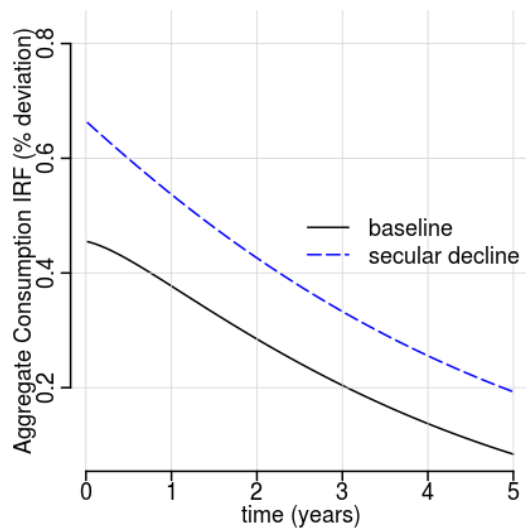**(a):** baseline vs. secular decline   **(b):** past high vs. past low rates
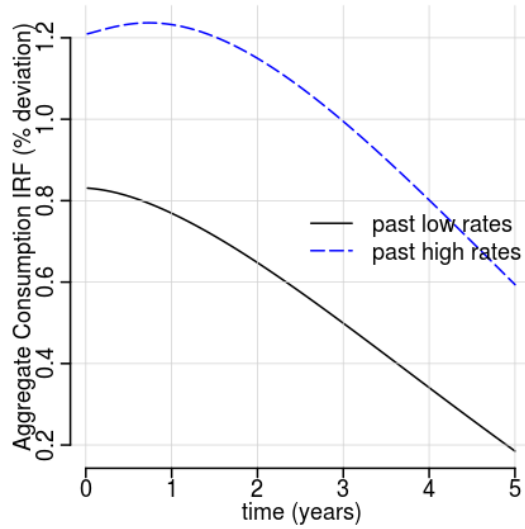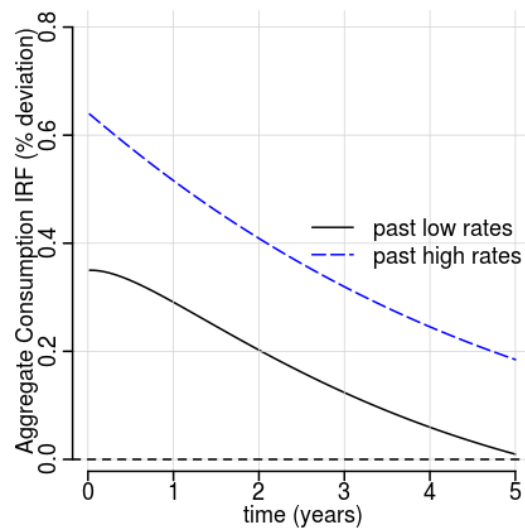

Figure A-13: Robustness of Consumption IRFs to Life-Cycle

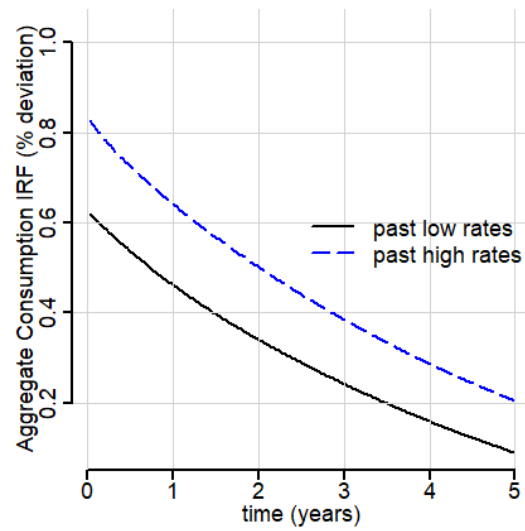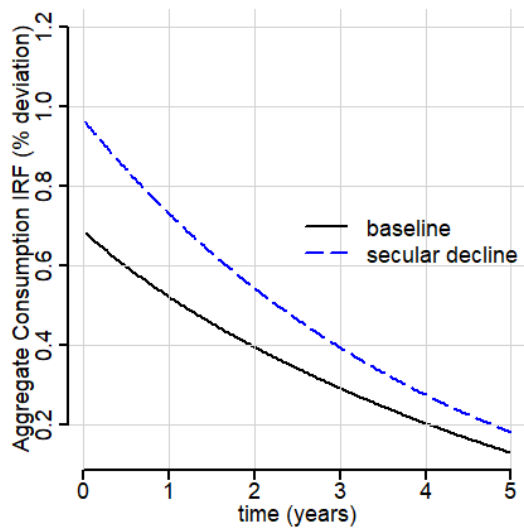**(a):** baseline vs. secular decline   **(b):** past high vs. past low rates

Figure A-14: Robustness of Consumption IRFs to Hybrid Frictions
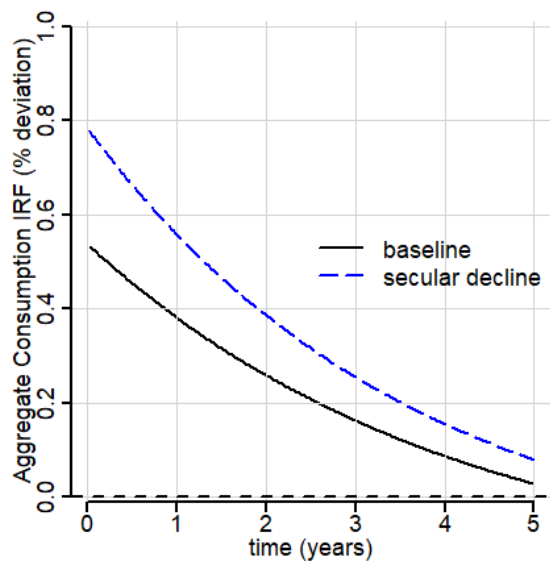
(a): baseline vs. secular decline   **(b):** past high vs. past low rates
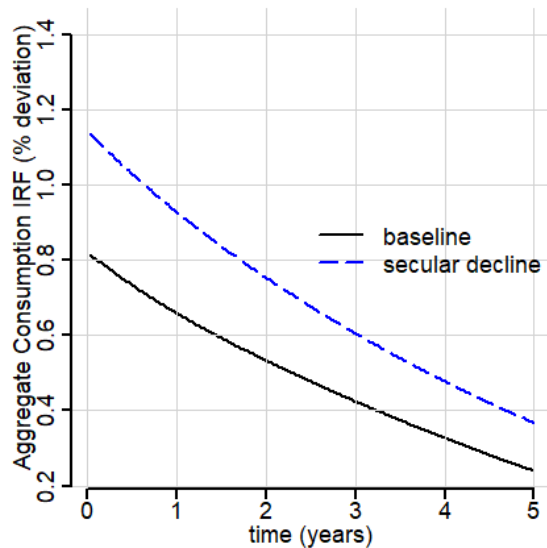
Figure A-15: Robustness of Consumption IRFs to Interest Rate Persistence

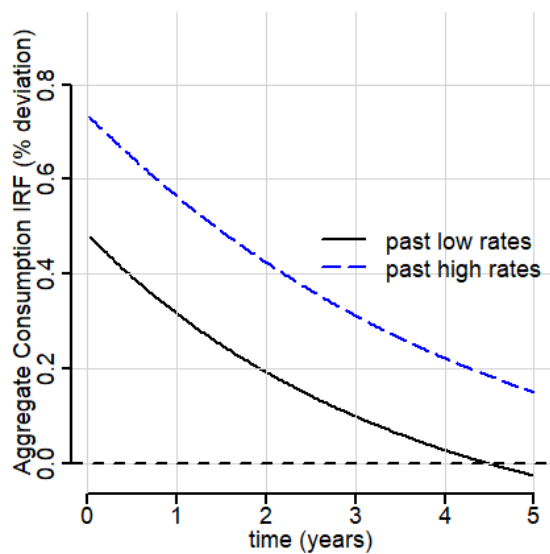**(a):** Half-life $0.66 \times h$ **(b):** Half-life $1.33 \times h$

(c): Half-life $0.66 \times h$ **(d):** Half-life $1.33 \times h$