Mortgage Prepayment and Path-dependent Effects of Monetary Policy

Berger, David W.; Milbradt, Konstantin; Tourre, Fabrice; Vavra, Joseph

Link to publication in CBS Research Portal

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

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

March 2021

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-dependence of prepayment rates to rate gaps (cf. Schwartz and Torous (1989)) but also the fact that most households nevertheless do not refinance even with strong rate incentives (cf. Keys et al. (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 like 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. Our model addresses three limitations of our empirical analysis for understanding aggregate monetary policy transmission through mortgage prepayment. First, we have rich data on how prepayment responds to interest rates, but our data on how borrower spending then responds to these changes in mortgage payments is comparatively limited. Second, even if high quality data linking mortgages and spending existed, this would only measure the response of borrowers to refinancing and would miss related spending by lenders.[8] Third, our empirical evidence does not allow us to perform policy counterfactuals or study how monetary policy potency would vary under alternative policy actions.

To address these issues, we embed a model of fixed rate mortgage prepayment into an otherwise standard incomplete markets consumption-savings model with labor income risk. We explore the spending implications of mortgage prepayment in this intentionally standard quantitative consumption-savings model since we have limited micro data on consumption outcomes with which to discipline more complicated modeling environments. The novelty of our analysis stems from the way in which we integrate mortgage prepayment: we assume that households face interest rate fluctuations and finance home purchases with fixed rate mortgage debt that can be refinanced subject to some frictions, which we discipline using the micro evidence from the first half of the paper. We then additionally include 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

---

[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.

time lower returns for lenders.

The goal of this model is thus two-fold. First, we want to show that through the lens of a workhorse quantitative consumption model the micro-level prepayment patterns we document indeed matter for aggregate monetary transmission, even after accounting for equilibrium effects of borrower refinancing on lender income. Second, we want to provide guidance about the efficacy of the prepayment channel at a moment in time and about the future evolution of monetary potency given current policy choices. Importantly, the goal of the model is not to quantify all channels of monetary policy or even all effects working through housing and mortgage markets. In order to provide as much intuition about the role that this prepayment channel plays in the transmission of monetary policy, we purposely keep several elements simple: all households have identical constant mortgage balances and we abstract from cashout refinancing, lifecycle effects, and house price dynamics. This simplifies equilibrium calculations dramatically and allows us to isolate and characterize the independent influence of refinancing frictions, which we can directly discipline using our micro evidence. However, we relax many of these assumptions in various robustness exercises. Unsurprisingly, exact quantitative predictions change somewhat as we alter various model elements, but our key qualitative insights about the strong path-dependent effects of the prepayment channel of monetary policy remain unchanged.

Refinancing frictions are central to all of our model conclusions. We allow for different refinancing frictions emphasized in the extant 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 a crucial advantage over this hybrid model (and other menu-cost based models of prepayment) when exploring *aggregate* spending implications of monetary policy. In the pure Calvo model, prepayment decisions are orthogonal to consumption-savings decisions, which means that mortgage prices can be computed without introducing the endogenous joint distribution of households as a state-variable. [9] In practice, this means that it is feasible to solve for equilibrium interactions between borrowers and

---

[9]This separation between liquidity and rate motives for refinancing might seem like a disadvantage, but it is useful to 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 et al. (2019)).

lenders in the Calvo model but not in the hybrid or other menu cost models. This in turn potentially matters for assessing the aggregate implications of monetary policy, since rate changes redistribute resources between borrowers and lenders. Without an equilibrium link between the mortgage rates paid by borrowers and the capital income received by lenders, models may not capture the full effect of rate changes on aggregate spending. Given these dramatic computational advantages of the Calvo model over the hybrid model despite similar observable implications, we focus primarily on the Calvo 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 in calculating equilibrium, our simple setup also 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 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]

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 et al. (2019). They make similar arguments

---

[10]Cf. Green and Shoven (1986), Schwartz and Torous (1989), Deng et al. (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.

[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 et al. (2015)

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 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 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]

---

[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.

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 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:

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

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]

---

[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.

[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.

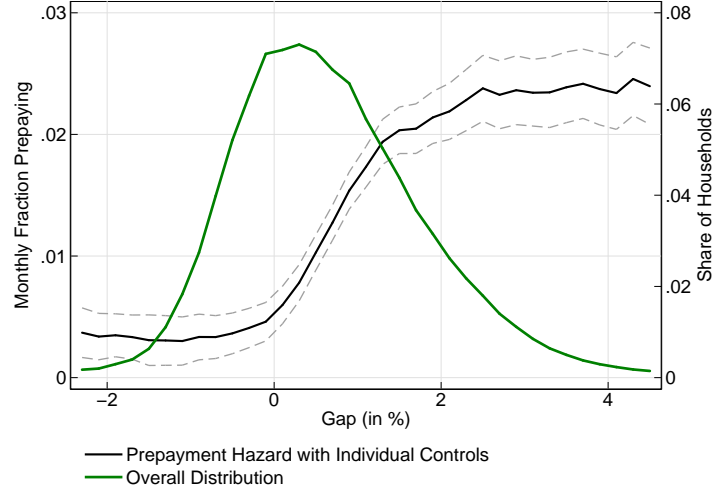**Figure 1: Prepayment Hazard with Individual Controls**

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 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.[24] 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 et al. (2016)), we argue that these patterns are more plausibly explained by inattention to mortgage rates and suboptimal refinancing decisions.

---

[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 et al. (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.

[24]The stable higher hazard for gaps above 150 basis points is quantitatively relevant for a large share of loans: more than a quarter of balance-weighted loan-months have gaps greater than 150 basis points and half of all loans have a maximum gap above 150 basis point.

4.1.1 Lender-based Refinancing Constraints?

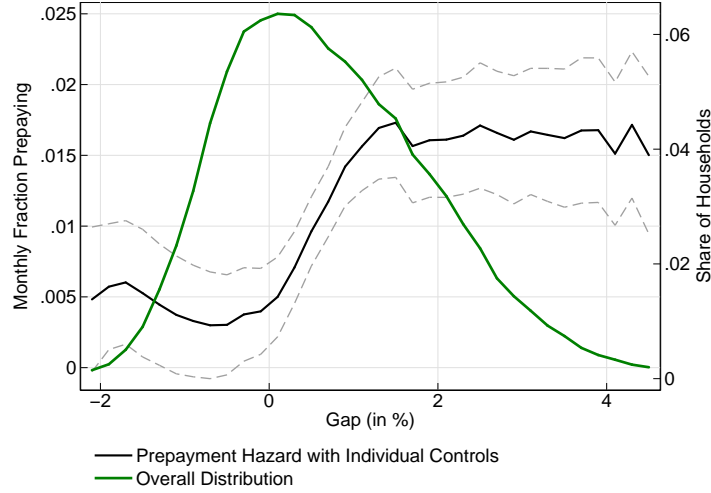**Figure 2: Prepayment Hazard with Individual Controls: Unconstrained Households**

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 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.[25] 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.

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.

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 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

---

[25]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.

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%).[26] 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.[27] 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 as a non-parametric function of a loan's current LTV ratio rather than interest rate gaps.[28] 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.

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

---

[26]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.

[27]See for example Chen et al. (2019) Figure 1.

[28]We also control non-parametrically for rate gaps, but the hazard as a function of LTV is insensitive to these controls.
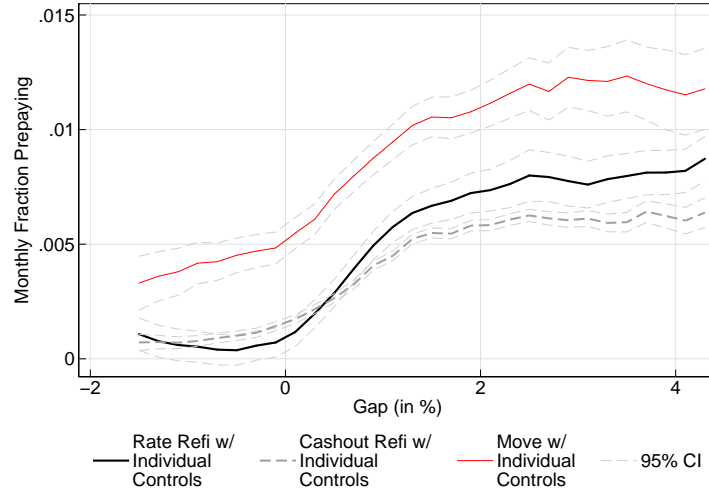
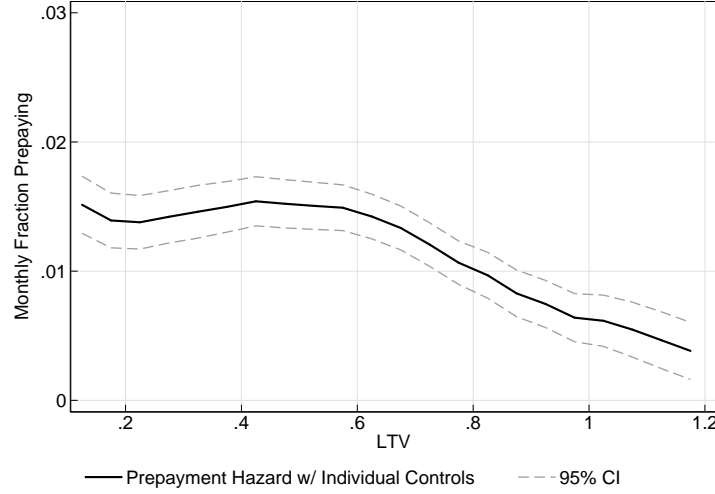## Figure 4: Prepayment Hazard by Leverage with Individual Controls

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.

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.

Figure 5: Prepayment vs. Fraction with Positive Rate Gap Time-Series
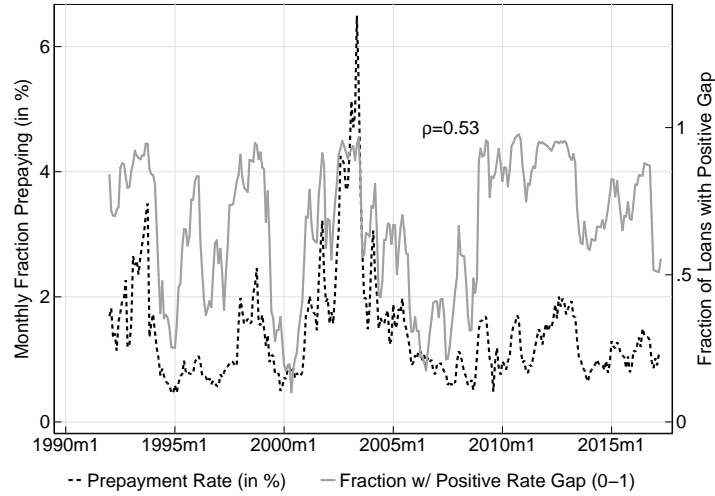
Figure shows the fraction of loans in McDash Performance data with positive rate gaps in each month as well as the fraction of loans prepaying in each month. The time sample is 1992m1-2017m4.

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 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

13

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.[29] 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

---

[29]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.

that our results are very similar when we control for one hundred calendar-quarter fixed effects and so identify off only aggregate time-series variation within-quarter. 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 do not change at high frequencies. In Appendix Table A-2 we additionally exploit cross-sectional variation by showing similar relationships hold at the MSA level, even after including both calendar-month and MSA × calendar-quarter fixed effects. MSA fixed effects absorb differences like age and education which vary across space, and calendar-month fixed effects absorb high frequency aggregate time-series effects.

However, even with these detailed controls, these relationships may still not reflect a causal relationship 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, 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 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. On average this sample covers just over 80% of all loans and is fairly stable over time except for a moderate decline during the housing boom when ARMs became more popular.[30] However, 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

---

[30]This number for 30-year fixed rate loans understates the broader importance of fixed rates, since most ARMs in the US have some fixed period of 3-10 years.

year.[31] 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.[32] 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 $300 ($78) billion.[33] 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 $2,320 ($603). 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, 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.[34]

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).[35] 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.[36] For example,

---

[31]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.

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

[33]This discounts payments over 30 years at a 15% rate to account for a 12% annual prepayment rate and risk free rate of 3%.

[34]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.

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

[36]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).

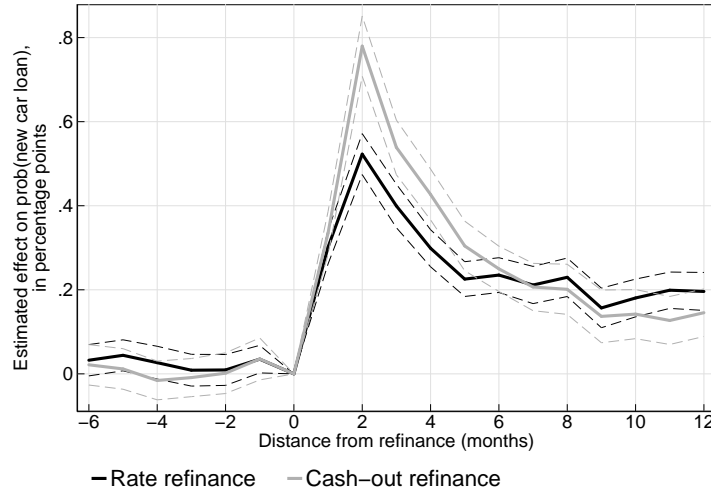**Figure 6: Individual Level response of Car Purchase to Refinancing**

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.

while time fixed effects absorb aggregate shocks (like interest rate movements) which simultaneously affect demand for car buying and refinancing, and household fixed effects absorb permanent differences across households, these specifications cannot control for idiosyncratic shocks which might lead to non-causal correlations. 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: Heterogeneity by Annual Interest Rate Savings

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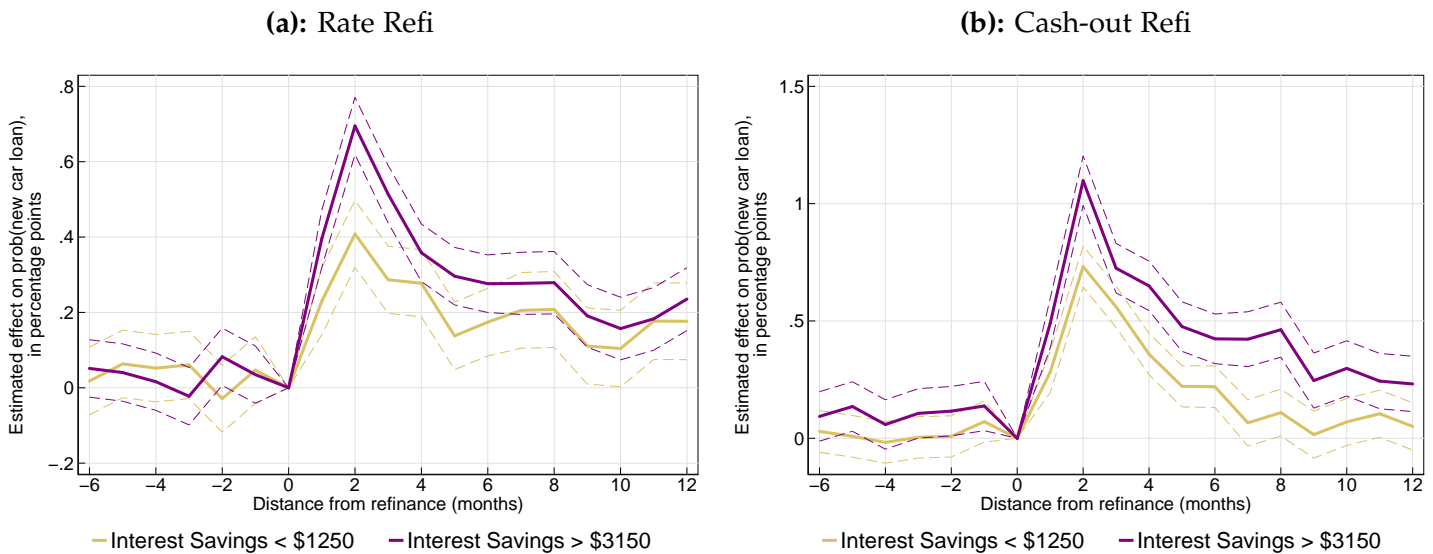
This figure repeats the analysis in Figure 6 but splitting also by the highest and lowest quartile of interest rate savings.

Figure 7 explores the relationship between how much households save on interest payments when

17

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.[37] 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.[38] Panel (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.[39]

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.[40] 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, 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

---

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

[38]This means results are not driven by the obvious potential confounding factor in months when mortgage and auto interest rates are low, mortgage savings are high and it is also more attractive to buy a car.

[39]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).

[40]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%.

18

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.[41] 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 2007 by Wieland and Yang (2020), so our sample period is 1969q1-2007q4. 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 $h = 0, ..., H$:

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

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.[42] The IRFs are the sequences $\{\beta_h^1, \beta_h^2\}_{h=0}^H$, representing output and consumption responses at future horizons to shocks today when $frac > 0$ is above or below the median, respectively.

Figure 8 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.[43] The red-dashed line

---

[41]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.

[42]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%.

[43]This exercise is closest to our model, which features shocks to the short-rate not to the mortgage rate, however Appendix Figure A-10 shows that we obtain similar albeit noisier results if we instead run a regression on mortgage rates instrumented using monetary policy shocks.

Figure 8: Response to Identified Monetary Policy Shocks

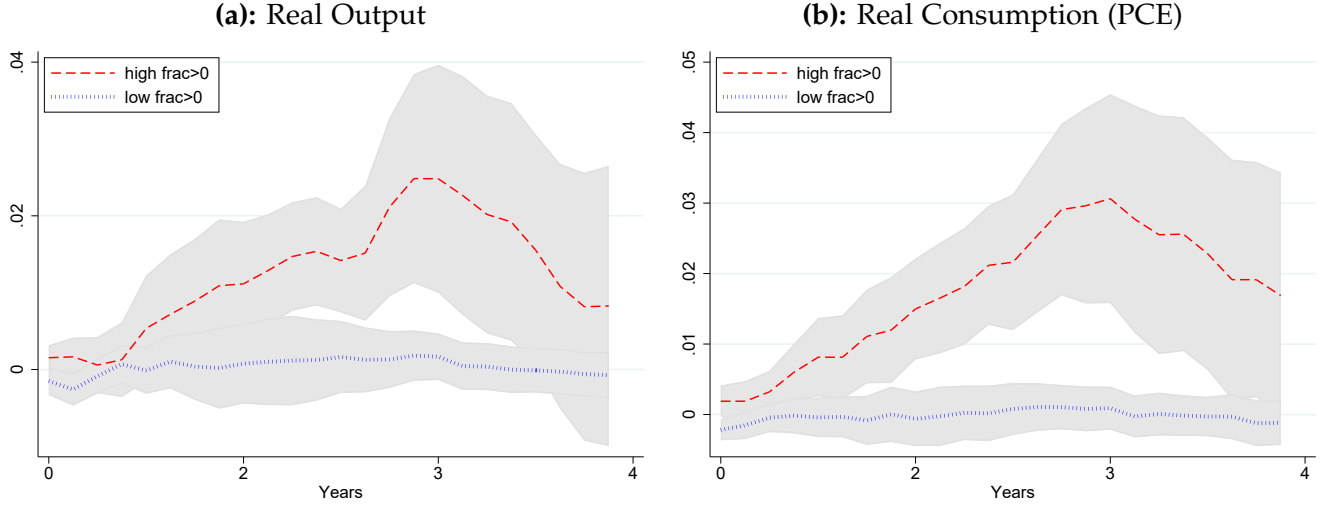**(a):** Real Output   **(b):** Real Consumption (PCE)

Figure shows the response of log real output and non-durable consumption to the identified monetary shocks from Romer and Romer (2004), extended through 2007 by Wieland and Yang (2020). Gray areas 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.

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.

Importantly, this variation in responses to monetary policy shocks with $frac > 0$ is distinct from state-dependence in the response to monetary policy which has previously been documented in the literature. For example, Tenreyro and Thwaites (2016) provide evidence that monetary policy is less effective during recessions. We confirm that their result also holds in our data. Furthermore, there is a small negative correlation between $frac > 0$ and their boom indicator, so the larger response to $frac > 0$ that we find is not driven by larger responses during booms. More generally, there is lots of movement in $frac > 0$ at higher frequencies than business cycles, which helps for separate identification. Finally, we also note that we emphasize state-dependent responses to monetary policy shocks, as distinct from asymmetric responses to positive and negative monetary policy shocks. As we discuss in Section 7.1, our model generates spending responses which depend crucially on $frac > 0$, but it implies nearly symmetric responses to empirically realistic interest rate increases and decreases.

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 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 et al. (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 prepayable mortgage with balance F and coupon rate m_t^*.[44] In our benchmark model we assume that all households

---

[44]We assume the house has constant price and so abstract from interest rate effects on consumption arising through house price effects documented in Mian et al. (2013), Berger et al. (2018) and Guren et al. (2018). Modeling these effects would likely

have identical, constant mortgage balances. This is clearly an important simplification. However, in Section 9.1, we introduce cash-out refinancing and heterogeneous mortgage balances. In the model with cashout, the average effects of mortgage prepayment are somewhat more front-loaded than in the model with only rate refinancing. However, cashout has little effect on the overall quantitative importance of *path-dependence* of monetary policy at business cycle frequencies, which is the main focus of our paper.

We focus exclusively on rate refinancing in our benchmark model for several reasons. 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.[45]

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.[46]

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:

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

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

---

amplify our conclusions: interest rate histories leading to greater house price growth, prepayment activity and cash-out should lead to greater resulting consumption.

[45]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.

[46]Landvoight et al. (2015) provides empirical evidence for these exogenous moving shocks. Krivenko (2018) documents that they are important for matching the dynamics of house prices.

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:

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

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):[47]

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

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

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

with Y_t following (3), r_t following (2), and $m_t = m(r_t, S_t)$ 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:

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

At any "fixed-cost" attention time τ_f, it is optimal for the household to refinance whenever the value of

---

[47]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.

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.[48] 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.[49]

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 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.[50]

---

[48] 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.

[49] 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.

[50] These conclusions are robust to different fixed costs as well as permanent heterogeneity in fixed costs. Of course, a random

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
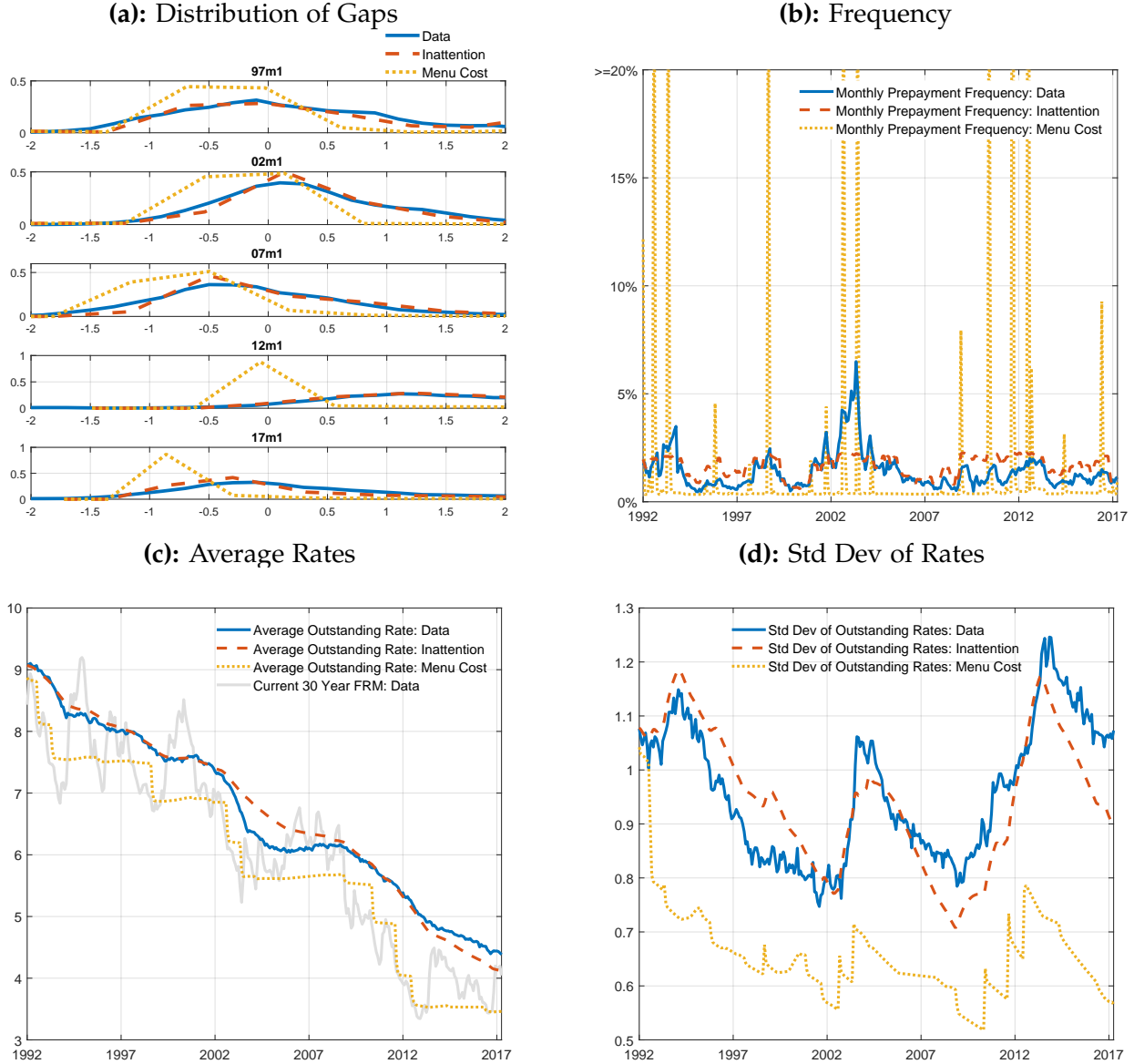
This poor time-series fit arises because the menu cost model generates a prepayment hazard at odds with the micro data, as shown in Panel (a) of Figure 10. The menu cost model implies a hazard which is too low for moderate positive gaps but which then becomes too high for large positive gaps. This occurs 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, 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

_____

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
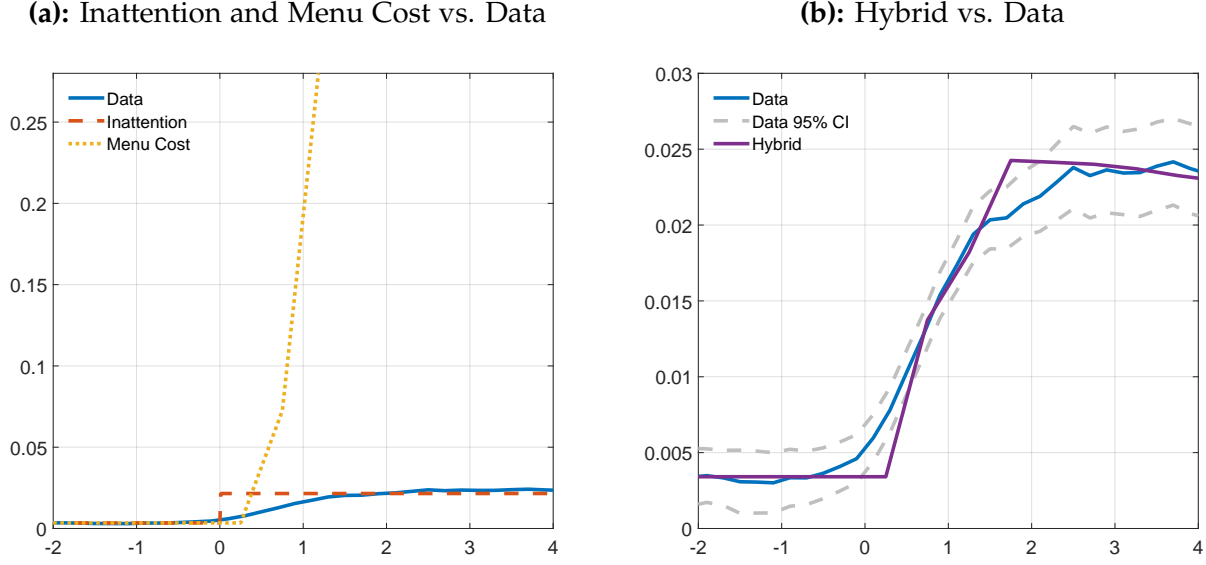
jointly pick χ_c, χ_f and κ to match the entire empirical hazard. Panel (b) of Figure 10 shows that a model with both frictions indeed better fits the prepayment hazard.[51] However, Appendix Figure A-11 shows that this hybrid model generates time-series implications almost identical to the pure Calvo 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.[52]

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 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

---

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

[52]That is, we solve the household model with general frictions but equilibrium mortgage pricing from the pure Calvo model since calculating mortgage pricing in the case with general frictions is infeasible.

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.[53] One can interpret this calibration strategy as targeting a closed economy "on average" but with stochastic foreign capital flows. Since δ does not enter the mortgage pricing equation, this in essence just picks δ to target a value of financial wealth which ensures this balancing. This requires $\delta = 14\%$ per annum, and generates an ergodic average liquid savings $\mathbb{E}[W_t] = \$36,000$.

We calibrate the speed of mean reversion and the volatility parameter of our rate process with MLE using daily data for 3-months treasury yields from 1982 to 2018, and obtain $\eta_r = 13\%$ (corresponding to a half life of 5.3 years) and $\sigma_r = 6\%$ (corresponding to an ergodic standard deviation of short rates of 2.2%). Given (η_r, σ_r), we then set the ergodic mean of the process to $\bar{r} = 3.5\%$ so that the corresponding initial model implied mortgage rate at the mean is equal to its empirical counterpart in 2017 when we start our counterfactual experiments, i.e., $m(\bar{r}) = m_{2017m04}$.

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.[54] Our calibration of the short rate process thus means that we match the broad behavior of short rates and resulting pass-through to mortgage rates over our sample period. We explore the robustness of our results to alternative interest rate calibrations with lower persistence in Section 9.

While we already showed that our model is a good fit to refinancing patterns, it is important to show that our full model also has reasonable implications for spending, since this is one of our key model outcomes. To provide some additional confidence in our model results, we thus show that our calibrated model fits a number of untargeted consumption and savings patterns in the data. Figure A-12 Panel (a) shows that the model does a good job of reproducing the distribution of consumption and of liquid assets from the Consumer Expenditure Survey and the Survey of Consumer Finance. In particular, the cross-sectional distribution of consumption in the model lines up extremely closely with the data. The model also replicates the fact that assets are substantially more concentrated than consumption, although they are somewhat less concentrated than the data. This is not particularly surprising, since it is well known that consumption-savings models struggle to match the very top of the wealth distribution without adding elements like entrepreneurship.

More importantly, the model does a good job of matching the conditional response of consumption to income shocks. The model implies a mean MPC to transitory income shocks of 0.27. Kaplan and Violante (2014) find an MPC to tax rebates of 0.22-0.24 using the methodology of Johnson et al. (2006) but trimming outliers to account for noise in the CEX. Misra and Surico (2014) estimate an MPC of 0.24.

---

[53]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.

[54]Gilchrist et al. (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.

Lewis et al. (2020) estimate an entire distribution of MPCs, and Figure A-12 Panel (b) shows that our model does a good job of matching this data. For example, they estimate that the 90th percentile of the MPC distribution is 56% larger than the mean MPC, which is identical to what our model delivers. We also find that households in the lower 50 percentile of the liquid savings distribution in our model have MPCs roughly twice that of households in the upper 50 percentile of the liquid savings distribution. This heterogeneity is directly in line with recent estimates in Crawley and Kuchler (2018).

Although panel data linking consumption to refinancing is very limited, the fact that our model does a good job of matching the consumption responses to other disposable income changes gives us confidence in its predictions along dimensions which are less easily measured. In addition, although clearly not a perfect test, we show below that our model implies spending responses to refinancing events which closely mirror the patterns we found for auto spending in Section 5. Finally, our model also replicates the empirical finding in Wong (2019) that homeowners who refinanced their mortgages at some point in the last year have consumption that is more responsive to lagged interest rate shocks than those who did not adjust their mortgages over the last year.[55]

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, $frac > 0$.

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

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

$[1 - F_t(m_t)]$ 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 $t + dt$.

---

[55]Specifically, we find an annual semi-elasticity of consumption to r of -1.47% for households who refinance at some point in the year after the shock versus a semi-elasticity of -0.54% for those who do not. These numbers are indistinguishable from the estimates in Wong (2019) at standard significance values.

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:

$$\lim_{t \to 0} \frac{dIRF_{m^*}^{\varepsilon}(t)}{dt} \approx \varepsilon \frac{\partial}{\partial r_0}\left(\frac{d\bar{m}_0^*}{dt}\right) = \varepsilon \cdot m'(r_0)(\nu + \chi_c[1 - F_0(m_0)]), \tag{6}$$

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. This formula tells us the effect of changes in the market mortgage rate on average coupons will depend on ($frac > 0$) but not on any additional characteristic of the distribution of mortgage gaps. This is a surprising result for a model with a state-dependent adjustment hazard and it arises precisely because the only state-dependence occurs at a gap of zero, meaning that the extensive margin of refinancing plays no role in the response to small interest rate shocks. This in turn implies that our model exhibits state-dependent responses, which depend on $frac > 0$, but it does not exhibit asymmetric responses to positive and negative interest rate changes.[56] That is, the sign of the IRF in Equation 6 depends on whether interest rates increase or decrease, but the magnitude of the response does not.

As a simple numerical example using this formula, 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.[57] 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,

---

[56]This symmetry in Equation 6 only holds exactly in response to infinitesimal shocks, but responses are nearly symmetric in our full model with 100 basis point shocks.

[57]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
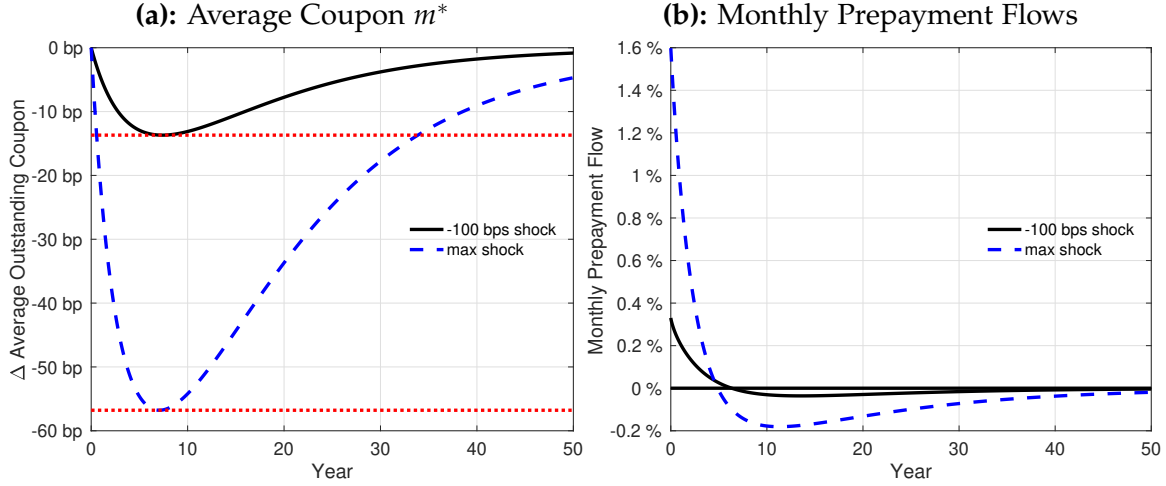
with the 30-year mortgage rate ultimately falling from roughly 9% to 4%. To explore the effects of this trend, we follow the procedure of Section 6.5.1 by initializing our model with the 1992 empirical coupon distribution and exposing it to the observed sequence of interest rates from 1992 to 2017 to 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.[58] 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.[59]

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

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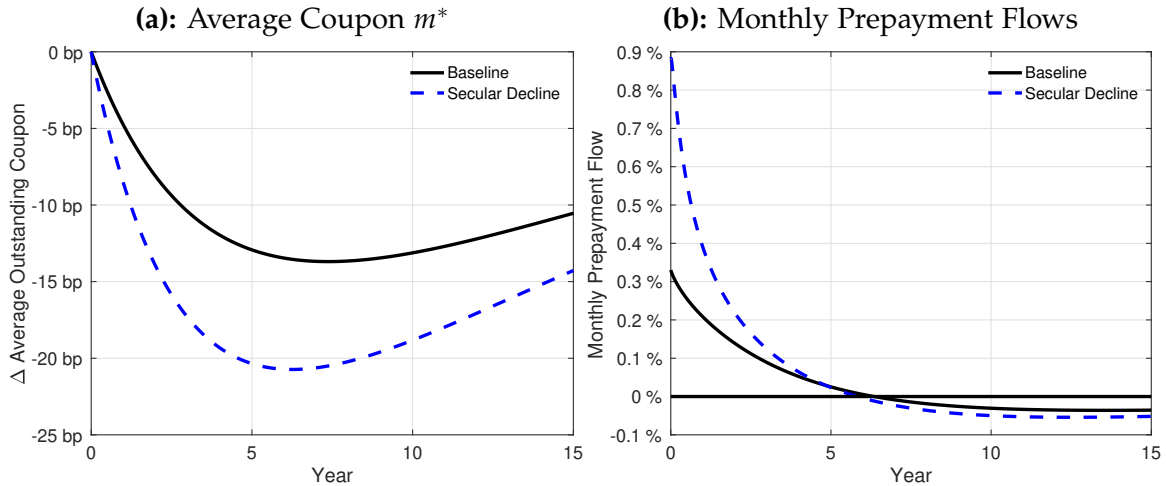
This is because the economy with a history of declining rates starts with larger $frac > 0$ today, as

---

[58]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.

[59]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
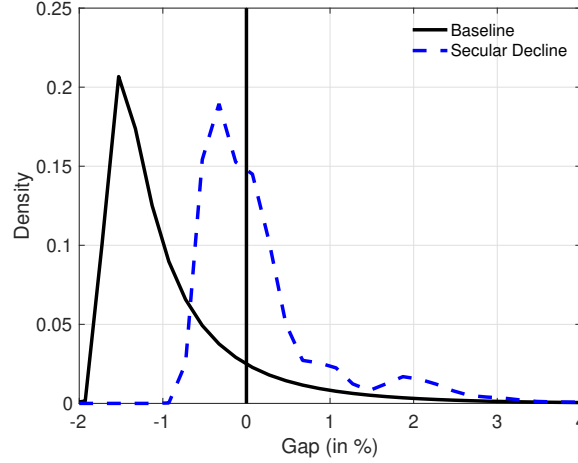
7.4 Effects of Past Easing and Tightening Cycles

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.[60] 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 $\{r_t\}_{t < 0}$. Furthermore, we note that these exercises focused on situations in which $\{r_t\}_{t < 0}$ differed for an extended period of time. In this section, we illustrate the dynamic effects of actions today on future monetary policy potency in order to show that these effects also matter for interest rate movements at all frequencies. To do this, we compare two economies which are initially identical but that differ in the future path of $\{r_t\}_{t \geq 0}$, and we explore how the potency of monetary policy varies across time as the path of interest rates evolves.

---

[60]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
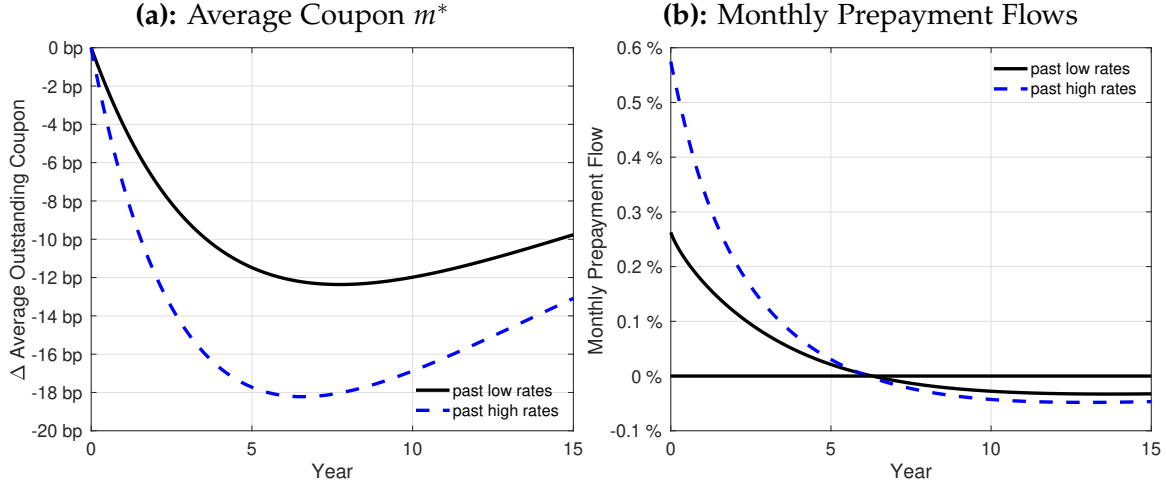
In particular, we consider two economies that are initially in the "baseline" environment described in Section 7.2, with identical cross-sectional distributions of outstanding mortgage rates. In one of them (the "Rate-Shift-Up" economy), the ergodic mean short rate permanently increases by 100bps at time zero, while in the other (the "Rate-Shift-Down" economy), it permanently decreases by 100bps.[61] One can interpret these regime shifts as permanent changes in the stance of monetary policy (which becomes either more hawkish or dovish).[62] We then compute the effects of rate cuts occurring at different points in time $t > 0$ *after* the initial regime shift.

While the time zero shift in interest rate regimes is permanent, assessing the dependence of the IRFs on the timing of a future rate cut provides information on how quickly the potency of monetary policy is altered after a change in its stance. For example, this exercise allows us to assess the speed with which the Fed regains "ammunition" after switching to a hawkish stance with higher rates and the speed with which it uses it up after switching to a dovish stance with lower rates. Appendix Figure A-14 shows IRFs to rate cuts which occur at various points $t > 0$ after the shift in stance.

However, visually unpacking the dynamic evolution of a sequence of functions is difficult. Thus, in order to more cleanly illustrate the dynamics we wish to emphasize, we compute for each shock impact time t the cumulative discounted value of the outstanding coupon IRF, i.e., $CIRF_t \equiv \int_t^\infty e^{-\delta(s-t)} IRF_{m^*}(s)ds$. We view this 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.

Specifically, this figure directly illustrates how the potency of a 100 basis point rate cut varies with the time $t > 0$ since the change in monetary policy stance. Immediately after the start of the dovish regime in the "Rate-Shift-Down" economy the potency of monetary policy rises relative to the economy with no regime shift ($CIRF_{0,shift-down}$ is 13.6% bigger than $CIRF_{0,no-shift}$).[63] However, the later the timing $t > 0$ of the rate cut after the switch to a dovish stance, the more monetary policy potency is diminished. This is

---

[61] To remove mechanical drift in the rate process, we respectively increase and decrease r_0 to its new future ergodic mean.

[62] We discuss also some implications for temporary shifts in monetary policy stance below.

[63] 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.

Figure 15: Regime Shift: Cumulative Discounted Average Coupon IRF decline in r to 100bps shock
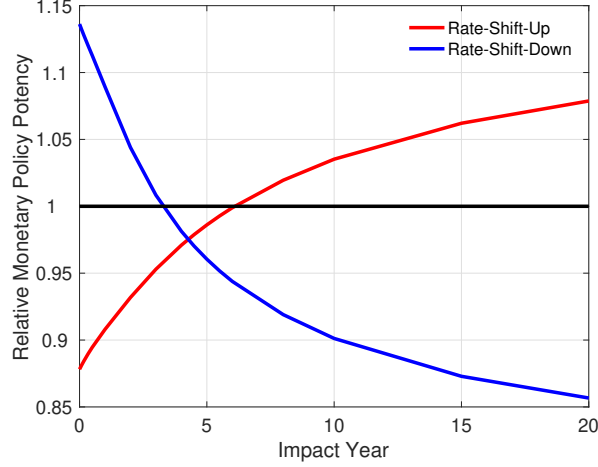
because in the dovish regime households actively refinance into lower rates and $frac > 0$ falls over time. If the rate cut occurs 3 years or later after the initial regime shift, monetary policy is less potent than under the baseline scenario ($CIRF_{t,shift-down}$ crosses 1 at approximately $t = 3$). Conversely, a 100 basis point rate cut which occurs immediately after the switch to a more hawish regime in the "Rate-Shift-Up" economy has a more limited effect than in the economy with no regime change ($CIRF_{0,shift-up}$ is 12.2% less than $CIRF_{0,no-shift}$). Potency then slowly increases with $t > 0$, eventually surpassing the no-shift baseline scenario at horizons beyond 6 years.

Importantly, there is an asymmetry in the dynamics of monetary policy potency after an initial shift towards either a hawkish or a dovish stance: monetary policy uses up ammunition more quickly after transitioning to a dovish stance than it regains it after transitioning to a hawkish stance. As just noted, the initial effects of lower rates are reversed almost twice as rapidly as the initial effects of higher rates (i.e. the relative monetary policy potency crosses 1, its baseline effectiveness, almost twice as fast in the "Rate-Shift-Down" than in the "Rate-Shift-Up" scenario). In this sense, monetary policy uses up ammunition when cutting rates more rapidly than it reloads ammunition when raising rates. 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.

We can also see these dynamics show up even at higher frequencies by comparing the slope of $CIRF_t$ at different points in the two economies. The absolute value of the slope of $CIRF_t$ in the first year after a rate increase is 57% larger than in the first year after a rate decrease. This type of asymmetry leads to important path-dependence at business cycle frequencies. For example, cutting interest rates for five years and then returning them to normal leads to a decline in $CIRF_0$ after normalization of 13% while raising rates for five years and then returning them to normal leads to an increase in $CIRF_0$ of only 5%.

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 $55.1k per year to $57.0 per year, a 3.4% 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.[64] 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
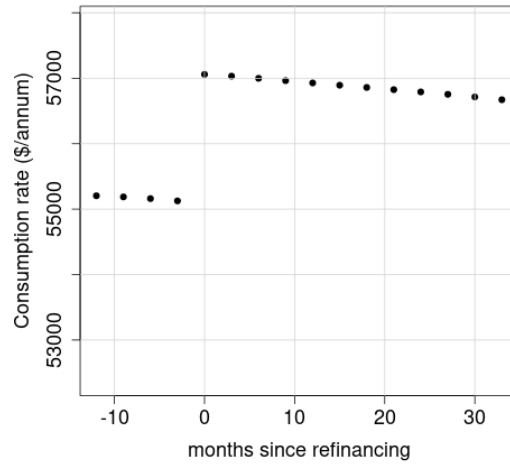
We next turn to aggregate spending responses to shocks to r. These responses cannot be characterized analytically but will clearly depend on household refinancing patterns, equilibrium relationships between r and m, and on more standard substitution, income and wealth effects of changing rates. Figure 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. 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 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.

These results show numerically that our model generates aggregate spending responses to rate cuts at frequencies relevant for Fed stimulus but that these effects vary importantly with the past path of

---

[64]Specifically, we assume that auto service flows are proportional to the size of the auto stock A whose evolution is given by $A' = (1 - \delta_a)A + \pi_t P$, where δ_a is the depreciation rate of autos, π_t is the probability of purchasing and P is a constant new car purchase price. We assume the auto stock is initially in steady-state $A = P$ with $\delta_a = \overline{\pi}_t$ and then compute the evolution of A after an increase in purchase probability π_t as in the empirical event study.

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

(a): 100bps shock   **(b):** max shock
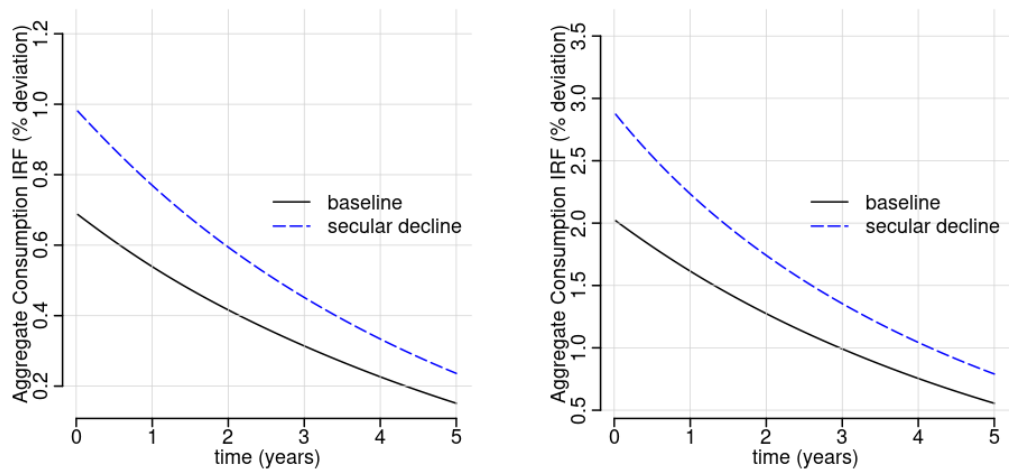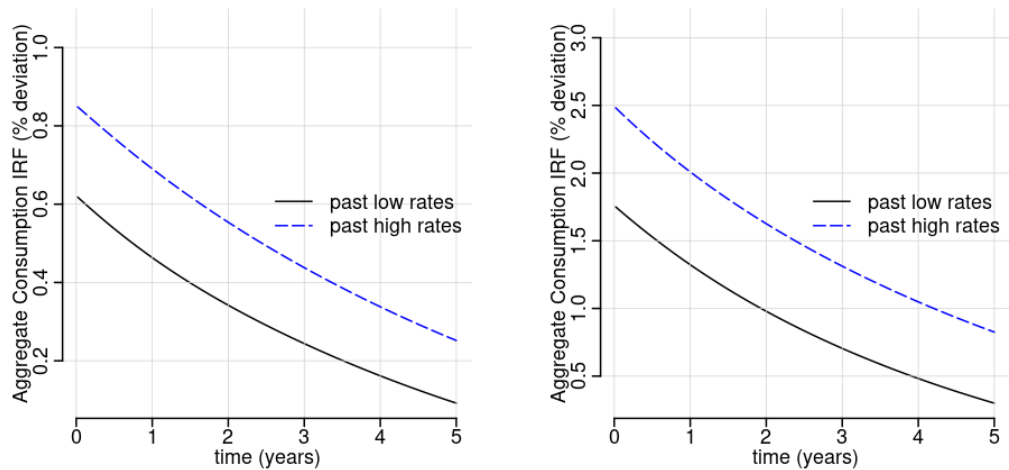
Figure 18: IRF of Consumption – Past High Rates vs. Past Low Rates

(a): 100bps shock   **(b):** max shock

interest rates.[65] 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.[66] 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:

$$\left. \frac{\partial \ln c}{\partial r} \right|_{t=0} = \frac{-1}{\eta_r + \delta} \left[\frac{1}{\gamma} - \frac{\delta w}{c_0} + \mathbb{I}_{\{\epsilon_0 < 0\}} \left(\frac{\eta_r + \delta}{\eta_r + \delta + \chi_c} \right) \frac{\chi_c \pi F}{c_0} \right] \tag{7}$$

The first two terms in brackets in (7) 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 et al. (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 $\pi = 0.57$ and $w = \$125,000$ (to match the same c_0 as our benchmark

---

[65] We note here that the past path in these exercises differs in a very persistent way but that this is not essential for these conclusions: any past path which leads to a materially different distribution of rate gaps will lead to different effects of current rate cuts.

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

model) delivers a semi-elasticity of -1.44. The negative value implies that aggregate consumption jumps upwards in response to a downward short rate shock. This arises in part because as is typical in most quantitative macro models, the substitution effect dominates the income effect.

The first two-terms in (7), which capture standard transmission effects, generate a semi-elasticity of -0.66. This implies that prepayment accounts for 54% of total monetary transmission in this model. Thus mortgage refinancing plays a large role in monetary transmission in this complete markets setup. Our baseline incomplete markets model features additional cross-sectional heterogeneity in MPCs which further affects monetary transmission.[67] In particular, this MPC heterogeneity means that increased spending by net borrowers after refinancing exceeds spending declines by net savers.

Of course, these complete markets calculations assume 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 use a common assumption in the literature and assume mortgages amortize at a constant rate α which we calibrate to 2.25% per annum to match the average amortization rate in our CRISM data. This calibration choice leads to average mortgage balance of around $156,000 under the interest rate series observed in the data, consistent with 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-15 shows that introducing cashout refinancing does not change our basic conclusions about how the past path of interest rates matters for responses to monetary policy. Although responses to rate cuts in the model with cashout are somewhat less persistent, the quantitative magnitude of path-dependence is similar at business cycle frequencies relevant for monetary policy.[68]

---

[67]We find substantial MPC heterogeneity in response to transitory changes in disposable income but also in response to changes in disposable income with longer half-lives of 7.3 years, similar to the typical realized duration of mortgages.

[68]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. At the same time, effects have less persistence at very long horizons

For example, in the cashout model, the cumulative discounted IRF of consumption over the first year is 55% larger in the secular trend economy than it is in the ergodic economy, while it is 44% larger in the baseline Calvo model. Over the first 2 years, path-dependence is nearly identical with a cumulative IRF that is 42% larger in the secular economy than in the ergodic economy for the cashout model and 45% larger for the baseline model. Over the first three years, the ratio is 45% for the baseline model vs. 33% for the cashout model.[69] At longer horizons, path-dependence in the baseline model more substantially exceeds that in the cashout model since the baseline model exhibits more persistence. However, nominal demand effects at five to ten year horizons are essentially irrelevant for analyzing the power of monetary policy to stimulate the economy during recessions, and they are unlikely to have real consequences anyway since nominal frictions matter little at those horizons.

Overall these results suggest that cashout refinancing is not of key importance for our argument that the refinancing of fixed rate mortgages leads to path-dependent consequences of monetary policy. However, it is important to note that even in this cashout model, refinancing decisions continue to be determined by Calvo information frictions interacting with rate incentives. In future work it would be useful to explore how models with more complicated refinancing frictions like in the "hybrid" model with both Calvo and menu cost frictions would interact with cashout refinancing.

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 which we describe in more detail in Appendix A.4. 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. To preserve stationarity, whenever an old household dies, a young household is borne. Appendix Figure A-16 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-17 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.[70]

since most households have refinanced and cashed out after 3-4 years.

[69]Our paper focuses primarily on path-dependence, but it is worth noting that average responses in the ergodic distribution are moderately amplified in the cashout model since this model introduces an additional force for monetary transmission. However even this difference is not huge: the cumulative IRF in the cashout model over the first year is 86 basis points while it is 57 basis points in the baseline model.

[70]As noted in Section 6.5.1, the Calvo model allows us to solve for mortgage pricing, which is infeasible with general frictions. Thus, we continue to use the mortgage pricing function $m(r)$ from the Calvo model in this exercise.

9.4 Alternative Monetary Policy Persistence

While we estimated our interest rate process (3) using MLE to fit the general time-series of interest rates we observe in the data, monetary policy induced rate changes might have a different persistence from that of more general rate movements. Furthermore, we might be interested in interest rate movements which occur just at business cycle frequencies rather than also at even longer horizons even though effects at both horizons affect refinancing incentives. To ensure our results do not depend materially on the persistence parameter, we re-solve our benchmark model with both higher and lower persistence of the short-rate process. Appendix Figure A-18 shows results. As implied by Equation 7, 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 remained low through the start of the pandemic, leaving little room for substantial rate cuts. 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 left the Fed with even less room for stimulus than suggested by the low level of 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.

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.

While we think these conclusions are quite robust, it is useful to highlight some limitations of our analysis and ongoing areas for future research. First, household data containing linked information on spending, interest rates and refinancing is extremely limited. Collecting better data on this front would substantially inform further modeling exercises. In addition, to obtain a model which can be solved in equilibrium and characterized in a transparent way, we focus on a relatively simple environment. We explore robustness to introducing various more complicated features like lifecycle effects, more complicated refinancing frictions and cashout refinancing, but we do not combine all of these elements together in a single model. While we argue that a simple menu cost model is at odds with prepayment data, it is possible that some more sophisticated menu cost model which combines these features might fit the data without having to rely on household inattention.

Second, our insight that rate histories matter for current monetary policy transmission through the mortgage market focuses on the US, but related forces could also matter in countries with different mortgage market institutions. Mortgage contracts with fixed rates and no prepayment penalty are uncommon outside of the US and Denmark, but 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.[71]

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."

Agarwal, Sumit, Gene Amromin, Souphala Chomsisengphet, Tim Landvoigt, Tomasz Piskorski, Amit

---

[71]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.

Seru, and Vincent Yao. 2017. "Mortgage Refinancing, Consumer Spending, and Competition: Evidence from the Home Affordable Refinancing Program." Working Paper 21512, NBER.

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.

Berger, David, Konstantin Milbradt, Fabrice Tourre, and Joseph Vavra. 2021. "Data and code for: Mortgage Prepayment and Path-Dependent Effects of Monetary Policy."Technical report, American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor], `http://doi.org/10.3886/E134161V1`.

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.

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. `10.1257/aer.20141313`.

Doepke, Matthias, and Martin Schneider. 2006. "Inflation and the Redistribution of Nominal Wealth." *Journal of Political Economy* 114 (6): 1069–1097, `https://EconPapers.repec.org/RePEc:ucp:jpolec:v:114:y:2006:i:6:p: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."Technical report, 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.

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, https://EconPapers.repec.org/RePEc:mcb:jmoncb:v:18:y:1986:i:1:p:41-59.

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

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

Johnson, David S., Jonathan A. Parker, and Nicholas S. Souleles. 2006. "Household Expenditure and the Income Tax Rebates of 2001." *American Economic Review* 96 (5): 1589–610.

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

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

Kaplan, Greg, and Giovanni L. Violante. 2014. "A Model of the Consumption Response to Fiscal Stimulus Payments." *Econometrica* 82 (4): 1199–1239. 10.3982/ECTA10528.

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.

Lewis, Daniel J, Davide Melcangi, and Laura Pilossoph. 2020. "Latent heterogeneity in the marginal propensity to consume." *FRB of New York Staff Report* (902): .

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

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

Misra, Kanishka, and Paolo Surico. 2014. "Consumption, Income Changes, and Heterogeneity: Evidence from Two Fiscal Stimulus Programs." *American Economic Journal: Macroeconomics* 6 (4): 84–106. 10.1257/mac.6.4.84.

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."Technical report.

Schwartz, Eduardo S., and Walter N. Torous. 1989. "Prepayment and the Valuation of Mortgage-Backed Securities." *The Journal of Finance* 44 (2): 375–392. 10.1111/j.1540-6261.1989.tb05062.x.

Tenreyro, Silvana, and Gregory Thwaites. 2016. "Pushing on a String: US Monetary Policy Is Less Powerful in Recessions." *American Economic Journal: Macroeconomics* 8 (4): 43–74. 10.1257/mac.20150016.

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

Wieland, Johannes, and Mu-jeung Yang. 2020. "Financial Dampening." *Journal of Money, Credit and Banking* 52 (1): 79–113. 10.1111/jmcb.12681.

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

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.