MEYER
FELLOWSHIP PAPER

How Do Mortgage Refinances Affect Debt, Default, and Spending? Evidence from HARP

MAY 2019 | JOSHUA ABEL AND ANDREAS FUSTER

JCHS
JOINT CENTER FOR
HOUSING STUDIES
OF HARVARD UNIVERSITY

How Do Mortgage Refinances Affect Debt, Default, and Spending? Evidence from HARP

Joshua Abel Andreas Fuster

May 3, 2019

Abstract

We use quasi-random access to the Home Affordable Refinance Program (HARP) to identify the causal effect of refinancing a mortgage on borrower balance sheet outcomes. We find that on average, refinancing into a lower-rate mortgage reduced borrowers' default rates on mortgages and non-mortgage debts by about 40% and 25%, respectively. Refinancing also caused borrowers to expand their use of debt instruments, such as auto loans, home equity lines of credit (HELOCs), and other consumer debts that are proxies for spending. All told, refinancing led to a net increase in debt equal to about 20% of the savings on mortgage payments. This number combines increases (new debts) of about 60% of the mortgage savings and decreases (pay-downs) of about 40% of those savings. Borrowers with low credit scores or low levels of unused revolving credit grow their auto and HELOC debt more strongly after a refinance but also reduce their bank card balances by more. Finally, we show that take-up of the refinancing opportunity was strongest among borrowers that were in a relatively better financial position to begin with.

1 Introduction

This paper seeks to refine our understanding of how refinancing a mortgage affects household outcomes. This issue has attracted particular attention in recent years, as US monetary policy in the wake of the Great Recession worked to an important extent through large-scale purchases of mortgage-backed securities, with the goal of lowering mortgage rates. This in turn was supposed to stimulate the housing market and enable households to refinance into a lower-rate mortgage. The resulting reduction in debt service costs should both reduce default risk and increase consumers' ability to spend.[1] Reflecting this policy importance, there has also been a recent surge in academic interest in the "refinancing channel" of monetary policy. However, there is still not much clean evidence on the *causal* effects of refinancing on borrower outcomes, nor on the heterogeneity of these effects across different borrower types.

We use quasi-random access to a refinancing opportunity during the recovery from the Great Recession to study how refinancing a mortgage affects households' financial decisions and outcomes. Specifically, we exploit the fact that the Home Affordable Refinance Program (HARP), which was introduced in early 2009 to enable borrowers to refinance even if they had little equity in their home, was only available to borrowers whose loans had been securitized before a certain cutoff date. We focus on borrowers who originated loans in a six-month window near that cutoff date and show that those that are eligible are subsequently much more likely to refinance over the period from 2010 to 2015. Based on this source of variation, we first confirm some findings in the previous literature: refinancing (which lowers the monthly payment by about $175, or 11%, on average) substantially reduces mortgage default and spurs borrowers to take on auto debt, a proxy for buying a car. We then show that the effects—increasing balances and decreasing defaults—extend to other debt instruments, such as home equity lines of credit and retail consumer debt. These effects tend to be strongest among borrowers who appear to be more financially constrained. And while we find that on average, refinancing causes households to take on new debts, we also show that for some groups and some debt categories, the improved cash flow is instead used to pay down debts.

Much of the existing evidence on the effects of changes in required debt payments comes from "resets" of adjustable-rate mortgages (or ARMs). In this market segment, one can compare borrowers who originated their loans at the same time, but with different initial fixed-rate periods, so that the payment resets occur at different times. Based on such a design, Fuster and Willen (2017) and Tracy and Wright (2016) find that payment reductions lead to relatively large reductions in default probabilities. Di Maggio et al. (2017) extend the studied outcomes beyond mortgage default and show that ARM payment reductions also increase new auto debt originations and are furthermore used by some borrowers to accelerate the amortization of their mortgages.

Refinances of fixed-rate mortgages (or FRMs) often result in similar payment reductions as the ARM resets studied in these papers, and it is therefore plausible that they would have similar effects.

[1] See e.g. Hubbard and Mayer (2009), Dudley (2012), or Stiglitz and Zandi (2012). Policymakers, of course, considered refinancing an important driver of consumer spending long before that (e.g. Greenspan, 2004).

However, there are some reasons why the effects might differ. First, ARM borrowers, who constitute a relatively small part of the US mortgage market, could be different from FRM borrowers along observable or unobservable dimensions. Second, ARM downward resets are (potentially) temporary, while FRM refinances result in permanent payment reductions, which could lead their effects to be larger. Third, and more subtly, the selection of borrowers who benefit from the payment reduction is different: among ARM borrowers with the same loan characteristics, all will benefit without requiring an active decision (after loan origination). In contrast, for FRMs, refinancing is an active choice, and those borrowers that refinance may react differently to the resulting payment reduction than the average borrower would.

Establishing the causal effect of a refinancing on borrower outcomes is complicated precisely because of the selection element due to the active decision. For example, a more financially sophisticated household may be more likely to refinance after a drop in mortgage rates and also better at budgeting, making a default less likely. Another element of selection comes from the fact that a refinancing requires that the borrower fulfills underwriting criteria such as sufficiently high income—a borrower who just lost their job may be unable to refinance but likely to default on their loan. As a consequence, to cleanly establish the causal effect of refinancing, one needs exogenous variation in the probability that two otherwise similar households will refinance.

Design details of HARP provide such variation. The program, which was only accessible to borrowers with mortgages guaranteed by Fannie Mae or Freddie Mac (the government-sponsored enterprises, or GSEs), was further restricted to borrowers whose mortgage the GSE had purchased before June 1, 2009. We will argue that this cutoff date quasi-randomly caused a group of borrowers to be eligible and another to be ineligible. Since the program was announced in March 2009, a couple of months before the cutoff date, a potential worry is that borrowers or servicers acted to affect the probability of later eligibility. We examine this possibility in a variety of ways, but find little evidence that suggests that this threatens the validity of our empirical strategy. In particular, our outcomes of interest only start differing once HARP refinance activity surges later in the sample period, suggesting that the two groups would not otherwise have evolved differentially.

Our dataset, which combines mortgage servicing records (from McDash) with consumer credit records (from Equifax), allows us to track the monthly evolution of balances and delinquencies across various debt categories. Focusing first on average effects across borrowers, we find that a refinance is followed by a roughly 40% reduction in the likelihood of mortgage default. Related to the discussion above, this effect is quite a bit larger than what existing ARM studies have found for comparable payment reductions, and suggests that relying on results from ARM studies may lead one to *underestimate* the default-reducing effects of FRM refinances.

In addition to reducing mortgage defaults, refinancing increases the monthly accumulation of non-mortgage debt by about 20% of the savings resulting from the decreased mortgage payments. This net effect combines larger increases (new debts) corresponding to about 60% of mortgage payment savings and decreases (pay-downs of existing debt) of about 40% of payment savings.

Debt increases are most pronounced for auto debt and HELOCs; pay-downs are concentrated in credit cards. Refinancing furthermore reduces the likelihood of becoming seriously delinquent on non-mortgage debts by around 25%.

Our data also provides us with useful summary indicators of an individual's financial health and liquidity, such as their updated credit score (FICO® Score) and their utilization rate of revolving credit. We use these indicators, along with our estimate of a borrower's combined loan-to-value ratio (CLTV), to study how the effects of a refinancing vary across different types of borrowers. We find that borrowers that appear less constrained prior to the refinance increase their auto debt by less, suggesting a smaller response in durable consumption for those borrowers. However, less constrained borrowers modestly increase their credit card balances, in stark contrast to more constrained borrowers who tend to pay those down with the newly available cash flow.

We supplement our causal analysis by looking at what observable characteristics predict take-up of a refinancing opportunity, as even among our HARP-eligible group, half the sample does not take advantage of the historically low interest rates that prevailed during the sample period. We find that individual indicators of good financial health—high credit score, high levels of un-tapped credit, low CLTV—predict a higher likelihood of take-up. Furthermore, take-up is highest in areas with higher incomes, while there is no strong relationship with either local education levels or mortgage market concentration.

Most closely related to our work are papers by Karamon et al. (2016) and Ehrlich and Perry (2015). Karamon et al. (2016) exploit the same HARP cutoff as we do, using Freddie Mac internal data, to study the effect of a HARP refinance on mortgage default.[2] Ehrlich and Perry (2015) similarly rely on a date-based eligibility cutoff embedded in a streamlined refinance program of the Federal Housing Administration. Both papers find effects on mortgage defaults that are similarly large as ours, but do not study other outcomes.[3]

Agarwal et al. (2017b) also study HARP, comparing GSE-securitized (and therefore HARP-eligible) loans to privately-securitized/non-agency (ineligible) loans. They show that (over 2009 to 2013) eligible borrowers had a substantially higher refinancing probability. They then also study the effects of refinances on individuals' auto debt accumulation (though not on the other debt outcomes we look at) and find positive effects.[4] We instead focus on eligibility variation *within* GSE-securitized mortgages only, which should maximize comparability of the eligible and ineligible groups. Given the different identification strategies we view our papers as complementary, and it is

[2]In contrast to Karamon et al. (2016), we retain non-HARP-eligible borrowers who refinanced outside the program in our sample. Since we use market-wide data, we also allow for "cross-GSE" refinances (e.g. from Freddie Mac to Fannie Mae), or cases where the new loan remains in the lender's portfolio.

[3]Another related study is by Zhu et al. (2015), who use Freddie Mac mortgages like Karamon et al. (2016) but focus primarily on the "intensive margin," comparing HARP refinances with payment reductions of different size. They find that larger payment reductions result in lower default probabilities.

[4]Agarwal et al. (2017b) furthermore show that ZIP codes with more eligible borrowers see higher car sales, credit card spending, and house price growth, and lower foreclosures; this is consistent with HARP refinances having local aggregate effects. In addition, they show that lenders were able to exploit their market power when originating HARP loans; see also Amromin and Kearns (2014) for related evidence.

reassuring that different sources of variation lead to similar conclusions. Furthermore, we advance the literature by exploring a richer set of outcomes as well as additional dimensions of heterogeneity.

Aside from these papers, our work contributes to a rapidly growing literature studying the refinancing channel (or more broadly, the redistribution channel) of monetary policy, such as Auclert (2017), Beraja et al. (2017), Di Maggio et al. (2016), Greenwald (2017), or Wong (2016).[5] Our results on heterogeneous effects on different types of borrowers relate to the broader literature that emphasizes the importance of such heterogeneity for monetary and fiscal policy, including for instance Agarwal et al. (2018), Jappelli and Pistaferri (2014), or Kaplan et al. (2018).

Our take-up analysis ties us to the household finance literature that has sought to understand why many households fail to refinance despite what appear to be clear benefits from doing so (e.g. Agarwal et al. 2015, Andersen et al. 2017, Bond et al. 2017, Campbell 2006, Johnson et al. 2015, or Keys et al. 2016). Finally, our results also inform the recent literature on mortgage design. Campbell (2013), Eberly and Krishnamurthy (2014), and Guren et al. (2017) all make the point that mortgages that automatically lower payments in downturns (such as ARMs, assuming nominal interest rates fall) could offer large benefits by freeing up cash flow for constrained households and therefore spurring consumer spending in periods of inadequate demand.[6] Our results directly speak to these arguments, as we show that the households whose spending appears most responsive to a payment reduction are also least likely to pursue one. This negative relationship between the propensity to refinance and the responsiveness to doing so strengthens the case for policies that make payment reductions easier to achieve in downturns.

2 The Home Affordable Refinance Program (HARP)

In the years following the peak of the US housing boom in 2006, home prices and interest rates fell dramatically. As a result, millions of homeowners had a strong financial incentive to refinance their fixed-rate mortgages into these low interest rates but were unable to do so, as the fall in home prices had erased much or all of their home equity—the collateral for a new loan. In response, HARP was announced by the Department of the Treasury on March 4, 2009, as part of its "Making Home Affordable" (MHA) initiative.[7] The purpose of HARP was to allow "homeowners who have a solid payment history on an existing mortgage owned by Fannie Mae or Freddie Mac [...] to refinance their loan to take advantage of today's lower mortgage rates or to refinance an

[5] A number of papers study the effects of equity withdrawal or "cash-out" refinancing, including Hurst and Stafford (2004), Chen et al. (2013), and Bhutta and Keys (2016). The HARP refinances we study involve at most very limited cash-out.

[6] Remy et al. (2011) point out that, of course, this is not free—it is a transfer from mortgage investors to these households. However, since a non-trivial share of investors are outside of the US economy or may otherwise not fully adjust their spending, this still provides aggregate stimulus, although the total magnitude is challenging to assess. The reduction in defaults due to payment reductions likely also has substantial positive aggregate effects, in part because negative externalities from foreclosures (e.g. Campbell et al., 2011) are avoided.

[7] The second main component of MHA was the Home Affordable Modification Program (HAMP), which was targeted primarily at borrowers already in delinquency or at immediate risk of becoming delinquent. See Ganong and Noel (2017), Agarwal et al. (2017a) and Scharlemann and Shore (2016) for studies of the effects of this program.

adjustable-rate mortgage into a more stable mortgage, such as a 30-year fixed rate loan" even if "these borrowers would [normally] be unable to refinance because their homes have lost value, pushing their current loan-to-value ratios above 80%."[8] The additional stated goals of the program were to reduce the government's exposure to mortgage credit risk and to stabilize housing markets.

HARP allowed borrowers with LTVs above 80% to refinance, but the program was restricted to mortgages that had been guaranteed either by Fannie Mae or Freddie Mac (the "government-sponsored enterprises," or GSEs), which by this time were under the conservatorship of the Federal Housing Finance Administration (FHFA). Prior to HARP, the GSEs did not purchase mortgages with low borrower equity—loans with LTVs greater than 80%—unless the borrower had purchased private mortgage insurance (PMI) to limit the GSEs' credit loss in case of a borrower default.[9]

HARP imposed a handful of additional eligibility criteria. Since the program was targeted at "responsible homeowners," borrowers had to be current on their payments, with no late payments in the prior six months and no more than one in the previous year. There was also initially an LTV cap. For the first few months of HARP's existence, it was restricted to borrowers with LTVs below 105%. In September of 2009, this was raised to 125%, and in June of 2012, the cap was lifted entirely. An additional restriction was that a borrower was only able to use the program once.

We give special attention to one final eligibility criterion: only loans guaranteed by a GSE before June 1, 2009, were eligible for HARP.[10] No official justification was given for this date—or the existence of a cutoff at all—though an unofficial (rumored) rationale seems to be that households entering the housing market after that time should be well aware of the risks associated with the market and presumably are therefore less worthy of government assistance.[11] Whatever the reason, the practical consequence of the requirement was to limit the program's pool of eligible borrowers substantially. This requirement is of particular importance in this paper: we will argue in Section 4 that it can be used as an instrument for refinancing, as it somewhat arbitrarily allowed some homeowners to refinance while restricting other similar homeowners from doing so.

HARP takeup was initially weak, so it was significantly reformed in 2012—the so-called "HARP 2.0." In addition to eliminating the LTV cap, HARP 2.0 facilitated the use of the program for borrowers with existing PMI. There were also substantial reductions in "loan-level pricing adjustments," fees charged by the GSEs when acquiring the mortgage. The need for a manual appraisal

[8] See the program announcement at https://www.treasury.gov/press-center/press-releases/Pages/200934145912322.aspx. More information about HARP is available at https://www.fhfa.gov/PolicyProgramsResearch/Programs/Pages/Home-Affordable-Refinance-Program.aspx.

[9] Of course a borrower with sufficient liquidity could always pay down the loan balance to reduce it to an 80% LTV at the time of the refinance. Refinancing with an LTV above 80% at the additional cost of obtaining PMI had historically been possible, but over this period, PMI supply was severely restricted due to insurers' financial distress (e.g. http://www.nytimes.com/2009/03/01/realestate/01Mort.html). Beraja et al. (2017) show that during 2009, prior to HARP, locations where borrowers had higher LTVs saw substantially less refinancing.

[10] In late 2013, this rule was changed so that the June 1, 2009, cutoff date applied to the date of *origination* rather than the date of *guarantee*. However, this was after the large bulk of HARP activity had already occurred, so we will treat the cutoff as applying to the date of guarantee throughout the paper.

[11] See http://mortgageporter.com/2012/03/why-is-june-1-2009-the-cut-off-date-for-home-affordable-refis.html.

was largely eliminated, and required documentation for things like borrower income was also relaxed. As a result, HARP borrowers had access to a streamlined refinancing opportunity after 2012 that was easier and, in many cases, had fewer upfront costs than the standard process for borrowers with higher equity.[12] Finally, in January 2013, FHFA relaxed the representations and warranties ("reps and warrants") rules lenders commit to when selling loans to the GSEs.[13] This likely increased lender HARP participation and enhanced competition (Agarwal et al., 2017b).

Around 3.5 million mortgages have been refinanced through HARP since early 2009, with most of the refinances occurring in 2012 and 2013, owing to a combination of HARP 2.0's more relaxed rules and the concurrent plunge in interest rates.

3 Data

Our analysis relies on "Credit Risk Insights Servicing McDash" (CRISM), a dataset that merges Equifax's credit bureau data on consumer debt liabilities with mortgage servicing data from McDash (owned and licensed by Black Knight). This section proceeds in four steps: first, we describe CRISM's features and why it is well-suited to our study; second, we discuss how HARP's eligibility criteria guide our sample selection; third, we provide summary statistics for our sample and compare it to the mortgage population at large; and finally, we describe refinancing activity in our sample and look at suggestive evidence of the effect of a refinance by performing an event study.

3.1 CRISM

CRISM covers about 60% of the US mortgage market during our sample period, providing McDash's mortgage data merged with Equifax's credit bureau data at a monthly frequency. CRISM is well-suited to studying refinances—and HARP in particular. Mortgage servicer data alone typically does not include unique borrower identifiers, making it impossible to track borrowers through a refinance, as one loan terminates and a new one originates. The credit bureau data, however, does include an identifier, allowing us to link loans through a refinance. But the credit bureau data does not report whether and when a loan has been guaranteed by a GSE. This information is essential for our identification strategy and is contained in the McDash data. The complementary attributes of the two datasets make CRISM uniquely suited for this study.

Both parts of CRISM are useful for tracking outcomes of interest, as well. For mortgage default, we will use McDash's reporting of delinquency status.[14] Equifax allows us to track balances on a wide range of other debt instruments. Specifically, we look at auto loans,[15] bank cards, student

[12] A caveat, studied in depth by Bond et al. (2017), is that in many cases, HARP borrowers with junior liens had to obtain a resubordination agreement from the lender of the junior lien. This could be hard if the lenders were difficult to contact or used their ability to hold up the process in an attempt to extract surplus.

[13] For further details, see Federal Housing Finance Agency Office of Inspector General (2013).

[14] We follow much of the literature (e.g. Tracy and Wright, 2016) in using 90+ days delinquency according to the Mortgage Bankers Association (MBA) measure as our flag for mortgage default.

[15] Equifax actually has separate categories for auto loans from banks and auto loans from auto finance companies, but we simply add them together for the entirety of the analysis.

loans, home equity lines of credit (HELOCs), and finally a set of smaller categories that we will refer to as "retail consumer debt."[16] Equifax, in addition to reporting overall balances for each category of debt, reports the amount of debt in each category on which borrowers are current on their payments. We use this to back out a measure of delinquency on non-first-mortgage debts. The Equifax data also contains a borrower's updated credit credit score in each month.

To measure borrowers' updated LTVs, we use their remaining principal balance in the numerator, while for the denominator (home value) we follow standard practice and assume that the value of the property (whose appraisal we observe at the time of loan origination) evolves according to a local home price index (HPI) from CoreLogic. For 83% of borrowers we have ZIP-level HPIs, while for the rest we use either county-, MSA-, or state-level HPIs. Our sample selection will be based on first-lien LTVs only, since HARP eligibility is based on those, but in other parts of the analysis we will use combined LTV (CLTV) ratios, where junior liens are added to the numerator.

3.2 Sample Selection Criteria

HARP's design—along with our empirical strategy–dictates three main sample selection criteria.[17] We select borrowers who: a) have a mortgage guaranteed by a GSE (Fannie Mae or Freddie Mac); b) originated that mortgage between January and June of 2009; c) were current on payments and had an LTV of at least 80% on that mortgage in March of 2010. We now explain these criteria.

First, we only study borrowers who have mortgages guaranteed by a GSE because only these were eligible for HARP. This ensured that the federal government, which had taken the GSEs into conservatorship, was not being exposed to additional credit risk by HARP. As non-GSE mortgages were categorically ineligible, we exclude them.[18]

Second, our sample only includes borrowers who originated a mortgage between January and June of 2009. Our key instrument (discussed in detail in Section 4) is based on the eligibility criterion that the mortgage must have been guaranteed before June 1, 2009. Since, as we show later, GSEs typically guarantee mortgages within a few months of origination, this window gives us a sample that is fairly balanced on eligibility. In addition, the window is wide enough to generate a large sample while being narrow enough to be fairly homogeneous, as we will argue in Section 4.

Third, we only include borrowers who are current on their payments and whose estimated updated LTV was above 80% in March 2010. As previously discussed, HARP was targeted at

[16] This combines three separate Equifax categories: retail debt, consumer finance debt, and "other" debt.

[17] In addition to the three main criteria discussed here, we drop borrowers whose credit files indicate that they have multiple first mortgages for more than 6 months or who have multiple active McDash first mortgages in a single month. We do this in order to be confident that the second mortgage balances are collateralized by the property we observe and not a second home. We also drop loans if we cannot determine whether they were guaranteed before the HARP eligibility cutoff. This occurs if a loan is listed as having originated before June of 2009 but does not appear in CRISM until after and is listed as being guaranteed by a GSE in its initial observation. In that case, we cannot know whether the guarantee occurred before or after June 1, 2009.

[18] As of end-2008, about 43% of outstanding mortgages were guaranteed by the GSEs; among mortgages originated in 2009, the share securitized through the GSEs was higher, at 63%, since the private securitization market had disappeared (source: Frame et al., 2015).

borrowers with LTVs above 80%. For our purposes, there is some ambiguity about how to impose that requirement, because LTV is a dynamic variable. As a result, a mortgage with a high LTV in one month could have a much lower LTV months later, depending on how home prices evolve. For simplicity, we choose a point in time—March of 2010—to measure LTV and decide whether to include the borrower in the sample. Note that March of 2010 is not a relevant date for the program; it is a date that we as researchers use to create our sample. We choose this date because it allows the mortgages to age a year on average (as they originated between January and June of 2009), while at the same time being before the drop in long-term rates that accelerated in the summer of 2010 as a consequence of the European sovereign debt crisis and financial markets' anticipation of the Federal Reserve's QE2 action. For robustness, we have re-run the analysis choosing a different point in time, March 2011, and while this does change the composition of borrowers in the sample, the results are extremely similar to what we report below using the March 2010 LTV. Finally, we drop borrowers who are not current as of March 2010 because HARP was restricted to borrowers with no missed payments in the previous 6 months and no more than 1 missed payment over the past 12 months. While a borrower who was delinquent in March 2010 could cure and use HARP later, we exclude them in order to focus on the group most likely to be able to use the program.

3.3 Summary Statistics

Table 1 provides summary statistics for our sample and compares it to the population of all borrowers in CRISM that were guaranteed by a GSE and had an LTV above 80%, as well as to all GSE borrowers in CRISM. The summary stats are for March of 2010, the observation month when inclusion in our sample is determined.[19]

Relative to the full high-LTV population (panel B), our sample (panel A) has a handful of differences, most of which can be traced back almost mechanically to the different origination windows (early 2009 versus all). As our sample originated in early 2009 and had therefore aged for only about one year, loan balances tend to be larger. Similarly, because they all originated after the first steep decline in mortgage rates in late 2008, they have lower interest rates as well. And while our sample has high LTVs by construction, they are not nearly as bad at that point as the broader high-LTV population, as our sample did not experience the (large) portion of the price decline that occurred before 2009. Credit utilization, defined as the sum of HELOC and bank card balances divided by the sum of HELOC and bank card limits, is also lower in the sample, perhaps because these borrowers, if they did acquire HELOCs, had simply not had as much time to utilize them. The sample also has higher credit scores, which is less mechanical and likely due to the tightening in credit standards following the onset of the financial crisis. The comparison with panel C, which shows all GSE borrowers regardless of LTV, is similar, though this group has LTVs that are a bit *better* (lower) than in our sample, and the difference in average credit scores is not as large.

Note that, based on Table 1, it is likely that results from our sample understate the impact of a

[19] As before, we restrict ourselves to borrowers who are current on their payments.

typical HARP refinance. Because borrowers in our sample already have low contract interest rates, the payment relief they receive from refinancing is relatively small. Furthermore, we will show later that the effects of refinancing tend to be larger for borrowers with low credit scores and high utilization rates of their revolving credit lines. As discussed above, our sample tends to have high credit scores and low credit utilization rates compared to the population. Succinctly, it is likely that our population received a relatively small treatment from refinancing and was also somewhat less sensitive to the treatment than the overall HARP population. While this may limit the validity of our sample for a full HARP program evaluation, such an evaluation is not our goal—rather, we are interested in the effects of a "typical" refinance. In that regard, while still not fully representative, our sample is closer to the broader population.[20]

3.4 Refinancing Activity

We now describe refinancing in our sample over the observation period, which runs from 2009 through February 2016. Following the procedure of Beraja et al. (2017), we consider a refinance to have occurred if a McDash mortgage terminates in a "voluntary payoff" and a new loan appears for the same borrower within 4 months, so long as either (i) the listed "purpose" of the new loan is a refinance, or (ii) the purpose of the new loan is not known (which is the case for about 25% of originations in McDash) but the mortgage is in the same ZIP code as the terminated loan.[21]

In the early years of the program (i.e. through the middle of 2011), refinancing activity in our sample was relatively weak, as Figure 1 shows. This is unsurprising, since interest rates had not yet fallen much since 2009, meaning most borrowers in our sample had no incentive to refinance. There was a brief spike in 2010 when interest rates temporarily fell by about 50 basis points (bp), but participation did not pick up in earnest until at least a year later, when interest rates fell in a more sustained manner, due to macroeconomic developments and monetary policy interventions. Refinancing activity declined sharply in the middle of 2013 when interest rates rose rapidly following the so-called "taper tantrum."

Of the 220,160 borrowers in our sample, 97,928 ($\approx$ 44%) completed a refinance. The first panel of Figure 2 shows that these borrowers were able to cut their mortgage payments by $173, or about 11%, on average. The remaining panel of the figure show an event study of changes in debt balances around the time of a non-cashout refinance.[22] This provides initial evidence that borrowers respond

[20] Appendix Table A-1 compares borrowers who refinance in our sample to all borrowers who refinance over the same time period. The gaps in interest rate, CLTV, and especially credit are much smaller than in Table 1.

[21] For about 25% of refinances identified, the new loan does not appear in McDash—it only appears in the Equifax data. We do count these as refinances so long as the ZIP code does not change, but we are limited in our ability to track borrower outcomes following the refinance. In particular, we are unable to track whether they are making their mortgage payments, since the McDash data is missing, and we are only able to track their balances on other debts (auto loans, etc) for 6 months, as CRISM tracks borrowers for 6 months after they exit the McDash sample. Our analysis is conducted at the borrower-month level, so we simply treat these refinancers who exit the sample as censored. However, we are able to compare the debt balances of these censored refinancers to those of the uncensored refinancers (whose new loans do appear in McDash) for the 6 months following the refinance, and their behavior is extremely similar, allaying concerns that the uncensored refinancers that we track are not representative.

[22] Specifically, we focus on borrowers whose first mortgage balances do not change by more than $5,000 from month

to the positive cash flow shock associated with the refinance. Both auto and HELOC debt show large increases in the months following the refinance. The initial response appears somewhat larger for auto debt, while the effect is more sustained for HELOCs. The event study also suggests that borrowers initally pay down some of their bank card debt, while there is no response on average for student debt or retail consumer debt.

While the event study is suggestive of borrowers responding to the refinances, it is difficult to interpret causally. As the timing of both the refinance and, say, an auto purchase are flexible, the relationship in the event study could be driven by borrowers timing their refinance to slightly precede their planned purchase of car. This is just one possibility that would generate the pattern in Figure 2 without there being a true effect of the refinance. The remainder of this paper works to identify more credibly causal estimates.[23]

4 Empirical Strategy

Mortgages guaranteed by a GSE before June 2009 were eligible to be refinanced through HARP, while those guaranteed afterward were not. We now argue that eligibility based on this criterion is a valid instrument for identifying the impact of refinancing.

It is important to have an instrument to answer this question because refinancing is endogenous, as it is an active choice being made by a household. Consider the case of mortgage default. If a household is underwater on its mortgage and assigns a high likelihood to moving soon (perhaps due to labor market outcomes or a change in family structure), it is unlikely to refinance—as the benefits of doing so will be short-lived—and likely to default. This will generate a negative correlation between refinancing and mortgage default that is not causal: those who refinance were less likely to default, regardless of any treatment effect of lower mortgage payments. By restricting ourselves to only the variation in refinancing activity that is predicted by the cutoff date, we are able to identify the causal impact of a refinance, so long as we are confident that borrowers guaranteed before the cutoff date are not systematically different from those guaranteed after (conditional on observables), other than in their eligibility to use HARP.

We will now show evidence in favor of this identifying assumption. This proceeds in three steps. First, we argue that there is no strategic sorting of GSE guarantees around the cutoff date. Second, we show that, at the beginning of the sample period, the ineligible and eligible groups are balanced on key observables. Third, we show that the two groups display little difference in key balance sheet outcomes before late 2011, when refinancing activity surges, suggesting the groups would have continued on parallel trends in the absence of the eligibility requirement. We conclude

-1 to month 1. By dropping apparent cash-out and "cash-in" refinances, this allows us to focus on the group most likely to be using HARP. In the IV regressions in subsequent sections, we do not drop any refinances (as we do here) because our instrument, HARP eligibility, drives variation only in HARP refinances. Notably, HARP eligibility is highly predictive of refinances whose balances do not change by more than $5,000 but is not predictive of refinances with larger balance shifts.

[23] Furthermore, an event study is not capable of identifying an effect on mortgage default, an important outcome, since a borrower cannot have previously defaulted on a mortgage she is refinancing.

Section 4 by discussing different specifications and how they affect the strength of the identifying assumptions.

4.1 Was There Non-Random Sorting Around the Cutoff Date?

HARP was unveiled in March of 2009, a few months before the eligibility cutoff date. This allows the possibility that strategic behavior, either by borrowers, lenders, or the servicers of the loans, could cause the eligible and ineligible groups to differ in ways that may be difficult to observe. For instance, if some servicers wanted to ensure that their loans were eligible to be refinanced and these same servicers have, say, unobservably higher-quality borrowers, then the eligibility instrument will not be valid, as the eligible group will be of higher unobserved quality. Any difference in outcomes between the two groups could then be attributed to that difference, rather than HARP eligibility.

One consequence of this kind of behavior could be a spike in GSE guarantees just before the cutoff date, as the strategic actors hurry their loans through the process. Figure 3 shows that this did not occur. There is a large increase in GSE acquisitions in the six months or so before the cutoff date, but this is almost certainly a result of the large increase in refinancing activity caused by the decline in interest rates toward the end of 2008. There are two months with particularly large spikes in guarantee activity, but these are March and June of 2009—the former being far in advance of the cutoff date (and essentially concurrent with the announcement of the program) and the latter being too late to maintain eligibility.[24] Furthermore, while our data does not contain information on the specific *day* that a loan was guaranteed, Karamon et al. (2016)'s Freddie Mac data does, and they show that guarantee volume is smooth through the cutoff date, as are the observables of the borrowers. Finally, in Appendix A.1 we further test whether securitization speeds for loans in our sample vary by LTV (since HARP eligibility was differentially valuable for different LTV ranges) but find no systematic evidence for this.

4.2 Balance on Observable Characteristics

The previous subsection argued that the eligible and ineligible groups do not differ due to strategic sorting, but of course it is possible that they differ for other reasons. However, our window of 6 months of originations comprise a fairly homogeneous group of borrowers, meaning that eligibility for HARP is plausibly the only systematic difference between the eligible and ineligible groups. Table 2 shows that the two groups are very similar on key observable characteristics: CLTV, FICO® Score, interest rate, credit utilization, and debt balances. While the groups can be distinguished in a statistical sense for most of the variables, the economic magnitudes of all the differences are small.[25] The similarity of the two groups on observable characteristics is reassuring

[24] Appendix Table A-2 supplements this with the CRISM micro-data, where we show the breakdown of guarantee month for each origination cohort. This confirms that it typically takes about 1 month for a loan to be guaranteed and that guarantee volume is essentially driven by the previous month's origination volume, with no evidence that loans were rushed to the GSEs before the cutoff date.

[25] We control for all of these observable characteristics in the analysis below, but as this analysis suggests, these controls are not very important, since the two groups are so similar. The one exception is that controlling for the

that they are similar on unobservable characteristics as well, lending validity to the eligibility instrument.[26]

4.3 Pre-trends

The top panel of Figure 4 shows how cumulative refinancing propensities evolve for the two groups over our sample period: the gap between the two groups starts widening in late 2011. We now show that the timing of the emergence of differences in important outcomes between the two groups follows a similar pattern. Specifically, we find little difference between the eligible and ineligible groups when we examine how balance sheet variables of interest evolved prior to late 2011—a period of time we can think of as pre-treatment.[27] These parallel trends prior to the wave of refinancing lend credibility to the assumption that, in the absence of this refinancing option, they would have continued along similar paths, allowing us to attribute the differences in the post-2011 period to a causal effect of refinancing. This gives us further confidence that, in addition to being similar along observable dimensions as shown in the previous subsection, the groups are similar along unobservable dimensions as well.

To evaluate pre-trends, we must first establish when the "pre-period" is. This setting differs from the textbook example of program evaluation because participants are not treated at a fixed point in time—rather, each eligible individual could choose to refinance at any time. We let the data tell us when the "treatment period" begins. Specifically, we look to see when the eligibility instrument starts to become predictive of refinancing. Then, we can check pre-trends by looking at other outcomes prior to that time.

To this end, we consider a "dynamic first stage" regression. We estimate the following equation:

$$\mathbb{E}[\text{Refied}_{it}|t, \text{Eligible}_i, X_{it}] = \sum_{\tau=201004}^{201602} \gamma_\tau(\text{Eligible}_i \times I_{\{t=\tau\}}) + X_{it}\theta, \tag{1}$$

where Refied_{it} indicates that borrower i refinanced in some month $\tau \leq t$ and X_{it} has the observables discussed in the following subsection, including quarter-by-ZIP fixed effects (FEs). The bottom panel of Figure 4 plots $\{\gamma_t\}_{t=201004}^{201602}$, the dynamic first stage effect of eligibility on the likelihood of having refinanced. As discussed earlier, there was a small flurry of refinancing activity in late 2010, and this is reflected in a small first stage effect early in the sample period. However, it is not until the more sustained drop in interest rates beginning in late 2011 that refinancing picked up substantially, and this is exactly when the HARP eligibility instrument begins to predict strong

interest rate does substantially strengthen the results for mortgage default.

[26] Table 2 also shows that while attrition from the sample is sizable, it is balanced across the two groups. Our analysis will be done at the monthly level, so we will simply censor borrowers as they leave the sample.

[27] A related concern is that there was differential attrition for the two groups between the time of origination and March 2010. Figure A-2 shows that this is not an issue, as only about 1% of borrowers left the sample within 14 months of originating their mortgage (we choose 14 because that is the number of months between the first originations—January 2009—and the March 2010 sample selection date). Furthermore, this attrition was quite balanced across the eligible and ineligible groups.

differences between the two groups. By the time interest rates rise again in mid-2013 and refinancing dries up, the eligible group remaining in the sample is about 30 percentage points (pp) more likely to have refinanced, a difference that persists nearly undiminished through the rest of the sample period.

We now show that eligibility was not predictive of different balance sheet outcomes prior to this surge in HARP activity. We first estimate a dynamic reduced form regression, similar to the dynamic first stage:

$$\mathbb{E}[Y_{it}|t, \text{Eligible}_i, X_{it}] = \sum_{\tau=200910}^{201602} \delta_\tau (\text{Eligible}_i \times I_{\{t=\tau\}}) + X_{it}\lambda, \tag{2}$$

where Y_{it} is the outcome of interest—either a default indicator or a first difference of debt.[28] Note that the sample begins in October of 2009 rather than April of 2010. For the first stage, there is mechanically no effect prior to March of 2010 because inclusion in the sample required the original loan to be active as of that date, so it cannot have been refinanced. However, for these other outcomes (with the lone exception of mortgage default), we are able to look back further to assess the pre-trends. Also note that, because the outcomes are noisier than the first stage, we will report the cumulative effect, $\left\{ \sum_{\tau=200910}^{t} \delta_\tau \right\}_{t=200910}^{201602}$, as this smooths out some of the noise.

Figure 5 shows results for default on the first mortgage (top panel) and serious delinquency on other debts (bottom panel). The first thing to notice is that HARP-eligible borrowers default less on both types of debt. In subsequent sections, we will formalize that observation into a quantitative treatment effect of refinancing. For this section, the key observation is that, while somewhat noisy, the effects occur in the later part of the sample, after there has been refinancing activity in the sample (the dynamic first stage is plotted, too, for reference): before 2011, the two groups appear identical, and then as more of the eligible borrowers refinance, a gap emerges. Figure 6 shows the same analysis for debt balances, where we look at auto debt, bank card debt, student debt, HELOC debt, and retail consumer debt. Of the five categories, only auto and HELOC debt show strong differences between the eligible and ineligible groups, and both of those emerge sharply late in the sample. All of this reassures us that the eligible and ineligible groups are similar except for their ability to refinance their mortgage, so our HARP eligibility instrument allows us to identify a truly causal effect of refinancing.

4.4 Identification Assumptions and Regression Specifications

The previous subsections have argued that HARP-eligible borrowers are comparable to those that are ineligible, but in a regression framework, they must be comparable only conditional on the other covariates included. Here we outline different sets of covariates and discuss the different

[28] We trim observations from a regression if the balance in that debt category is greater than the 99^{th} percentile of non-zero balances in the sample. We have also winsorized based on this criteria, and we have trimmed and winsorized extreme *changes* as well, with all methods of dealing with outliers delivering very similar results.

identifying variation they depend on.

Our baseline specification will control for the following observable characteristics, in addition to a full set of indicators for ZIP code and time (quarter): CLTV (lagged 3 months and as of March 2010), credit score (lagged 3 months and as of March 2010), credit utilization (lagged 3 months and as of March 2010), remaining principal balance on the first mortgage, initial interest rate on the mortgage, and initial balances of the other debt categories.[29] This means the identifying assumption for these regressions, which will be denoted by a (1) in most of the tables in the results section, is:

Identifying Assumption 1 *Conditional on ZIP code, quarter, and observable covariates, the only systematic difference between borrowers guaranteed before and after June 1, 2009, is eligibility for HARP.*

We have provided evidence in support of that assumption throughout Section 4; in fact, the evidence in Table 2 suggests that identification may be achieved even without conditioning on observables.

However, an even more restrictive source of variation can be exploited. Because the cutoff date is based on the month of guarantee and not month of origination, we can include controls for origination cohort and still have variation in eligibility. That specification allows for the eligible and ineligible groups to systematically differ, just not within cohorts. In fact, we can do even better by also including a full set of indicators for guarantee lag.[30] In this specification, which will be denoted by a (4) in most of the tables in the results section, the identifying assumption is even weaker:

Identifying Assumption 2 *Conditional on observables, all differences between eligible and ineligible borrowers within the cohort—other than access to HARP—are driven by time-invariant differences between borrowers with different guarantee lags.*

To get an intuitive sense of how this has relaxed our assumption, consider an example. Suppose that borrowers whose guarantee lag is greater than, say, 2 months have unobservable qualities that make them more likely to default, and this is true for all cohorts. Then the ineligible group will be unobservably more likely to default, even within a cohort. However, we are still able to identify our causal effect because those within-cohort differences are driven by effects associated with different guarantee lags, which we remove with the guarantee lag indicators because they are consistent across cohorts.

Our results are robust to this weaker assumption. In our discussion, we will put less emphasis on the results corresponding to this assumption because those specifications tend to have large standard errors, since there is less variation to be exploited.[31]

[29] Allowing the effects of these covariates to change over time does not alter our qualitative findings.

[30] Note that the cohort indicators and guarantee lag indicators are separate. We cannot use cohort-by-guarantee lag controls, since within those bins, there is no variation in eligibility.

[31] In our sample, only approximately 25% of the variation in eligibility is within-cohort, so the standard errors for regressions corresponding to Identifying Assumption 2 are substantially larger than for the other specifications, as seen below.

5 Effects of Refinancing on Mortgage Default

This section studies the impact of refinancing on mortgage default. In line with other studies that have used a similar approach, we find a strong reduction in the mortgage default hazard, in excess of the implied sensitivity of default to payment size found in most previous studies of ARMs.

5.1 Econometric Model

To assess the relationship between refinancing and mortgage default, we estimate a linear probability model. Following the literature, we define default to be the event of going 90+ days delinquent on the mortgage, and we censor borrowers once they default for the first time or leave the dataset via a non-refinance prepayment or servicing transfer. We then estimate

$$\mathbb{E}[\text{Default}_{it}|\widetilde{\text{Refied}}_{it}, X_{it}] = \beta\widetilde{\text{Refied}}_{it} + X_{it}\kappa, \tag{3}$$

where Default_{it} indicates whether borrower i defaulted in period t, X_{it} includes fixed effects for quarter and ZIP code, flexible controls for the observables described above, and other covariates that depend on the specification (see below). In this section we compare OLS estimates to those from 2SLS (IV). In the OLS specifications, $\widetilde{\text{Refied}}_{it}$ indicates whether borrower i has refinanced in or before period t. For the IV specifications, we run a first-stage regression that captures the differential refinancing activity between the eligible and ineligible groups and how that evolved over time, as we saw earlier in Figure 4:

$$\mathbb{E}[\text{Refied}_{it}|t, \text{Eligible}_i, X_{it}] = \sum_{\tau=2010Q2}^{2016Q1} \gamma_\tau(\text{Eligible}_i \times I_{t\in\tau}) + X_{it}\theta. \tag{4}$$

In the 2SLS results, $\widetilde{\text{Refied}}_{it}$ is the fitted value from regression (4). Intuitively, we will find an IV effect if the eligible and ineligible groups behave differently in the same months when the (cumulative) refinance shares differ the most. As we saw in Figure 4, this basically means we are identified off of post-2011 differences between the groups.

Note that this specification imposes the assumption that the treatment effect of refinancing does not depend on how long ago the refinance occurred. As refinances occurred about 3.5 years before the end of our sample on average, we can think of the estimated effects as the average over 3.5 years. In Appendix A.2, we explore more flexible specifications that allow the effects to vary with time since refinance, but those specifications are largely under-powered so we do not feature them in our main results.

We will run four specifications with an expanding set of controls in X_{it}:

1. quarter fixed effects (FEs), ZIP FEs, observables;
2. + ZIP $\times$ quarter FEs;

3. + guarantee lag FEs;

4. + cohort FEs.

Specifications 1 and 4 correspond to Identifying Assumptions 1 and 2, respectively. We show specifications 2 and 3 to give more of a sense of the sensitivity of the results to the conditioning variables. Note that standard errors are clustered at the county level throughout.

5.2 Results

The effects on mortgage default are very large. Panel A of Table 3 shows the results for the different Xs across OLS and IV. The IV results are quite stable across specifications, though—as we will see throughout the results—inclusion of cohort effects increases the standard errors substantially and, in this case, leads to an appreciably higher point estimate.[32] The estimated reduction of the monthly default probability ranges from 2.1bp to 3.2bp, or roughly 40% of the sample average.[33] The OLS results are substantially larger: the reductions range from 4.6-4.8bp, or about 80%. This suggests that borrowers who refinance are unobservably less likely to default, even in the absence of refinancing, highlighting the need for an instrument to study this question.

As mentioned above, these results—as well as those to follow—are robust to using March 2011 as the sample selection date, as opposed to March 2010. Furthermore, while our sample includes both ARMs and FRMs and loans of different maturities, the results are essentially unchanged if we restrict to 30-year FRMs.

These findings are in line with recent studies that show large effects of refinancing FRMs on mortgage default. Karamon et al. (2016) use a regression discontinuity approach to study Freddie Mac loans guaranteed around the cutoff date and estimate that a HARP refinance reduces the default likelihood by about 50%. Ehrlich and Perry (2015) use a similar approach to study FHA loans going through a streamlined refinance program and find a reduction in the default probability of 40%. Our approach yields very similar results. Interestingly, these effects are larger than what one would predict based on results from papers studying ARMs. Fuster and Willen (2017) use ARM borrowers who mostly receive large payment reductions, while Tracy and Wright (2016) use a set of ARMs with smaller resets. Using the estimates of either study, the implied reduction in default rates caused by a 11% payment reduction (the average in our sample) is about 15%, much smaller than the 40% we find.[34]

[32] Given the strong evidence of a first-stage effect of HARP eligibility on refinancing shown in Figure 4, it is not surprising that first-stage F-statistics (not reported) are well above any level that would lead one to worry about a weak instruments problem.

[33] The simple default rate in our sample is 6.6bp. However, we prefer a weighted average of monthly default rates, where the weight of month τ is equal to the number of borrowers with $\text{Refied}_{i\tau} = 1$. This weighted measure is more reflective of the true default hazard that borrowers faced in the "treatment period," which as discussed above is not a binary partition of time but a more gradual transition. This "takeup-weighted monthly default rate" is 5.9bp.

[34] Di Maggio et al. (2017) also briefly study the effects of ARM resets on defaults in their sample, though since this is not the main focus of their paper, they do not provide much detail for their analysis of that outcome. The effect size they report would correspond to an 11% payment reduction decreasing default rates by about 25%. This is larger than the effects found by Tracy and Wright (2016) and Fuster and Willen (2017) but still substantially smaller

This speaks to the potential pitfalls of applying results from studies of ARMs, a relatively small segment of the US mortgage market, to the dominant FRM segment. One explanation for the discrepancy is simply that ARM borrowers differ from FRM borrowers, and so their sensitivities to payment reductions may differ as well. A second possibility is that a payment reduction achieved by refinancing is permanent, whereas a reset on an ARM could be reversed if interest rates rise in the future, dampening the effect. Yet another potential explanation is that on average FRM and ARM borrowers are similar, but the causal effects identified in these studies are for different subsets of these groups. Specifically, as ARM borrowers receive their payment changes passively—when interest rates fall, the reduction goes to all ARM borrowers—these studies identify the population's average treatment effect (ATE). In contrast, the FRM studies (this one included) identify a local average treatment effect (LATE) that is specific to people who are "compliers" in the sense defined by Angrist et al. (1996)—borrowers who would refinance if eligible for HARP and not otherwise. As such, this average is taken not over the whole population but over this more select group that is willing to actively pursue a refinance. It makes sense that this group would be more responsive, as they are evidently quite engaged with the details of their mortgages. This is also the group whose behavior is most relevant for understanding the refinancing channel.

Panel A of Table 4 shows that the strength of the estimated effect varies substantially depending on a handful of salient borrower characteristics. We create an indicator for an observation being above or below the median, say, credit score and then estimate coefficients on Refied_{it} and $\text{Refied}_{it} \times \text{HighCreditScore}_{it}$. We instrument for these two variables with Eligible_i, $\text{Eligible}_i \times \text{HighCreditScore}_{it}$, and the joint interaction of these variables with a full set of quarter dummies. Note that the average payment reduction was essentially uniform across the groups, in the 10-12% range, so any heterogeneity comes from differential responses rather than different treatments. The results show that high-CreditScore borrowers and low-credit utilization borrowers had no discernible response to their refinance in terms of default behavior, in part because they default at very low rates anyway. Low-CreditScore borrowers, on the other hand, saw a reduction in their monthly default hazard of 4.8bp (their average default rate is 10.5bp), and high-credit utilization borrowers reduced their default rates by 4.5bp (compared to an average of 8.3bp). These groups, then, experienced a reduction in their default hazard of around 50%. Borrowers with high CLTVs also experienced a larger effect than those with lower values, though the difference is less stark than with the other two variables, perhaps because the sample is, by construction, made up of borrowers with high LTVs, so the differences are more muted. All told then, we find evidence not only of a strong treatment effect but one that is considerably stronger for borrowers with characteristics that indicate financial distress and higher risk of default.

These results suggest that liquidity constraints are a key driver of mortgage default, in line with existing evidence (e.g. Elul et al., 2010) and as featured in state-of-the-art models (e.g. Campbell and Cocco, 2015). In contrast, frictionless models of mortgage default (e.g. Kau et al., 1992)

than the effect we find.

would not predict heterogeneity in default elasticity to payment reductions, since all that matters to borrowers in such models is the present cost of the mortgage discounted at the risk-free rate. However, in our data essentially only ex-ante more constrained borrowers default, and only those borrowers' default propensity is reduced by the refinance into a lower-rate mortgage.

6 Effects of Refinancing on Non-Mortgage Debt

In this section, we look at the consequences of refinancing a first mortgage for other items on households' balance sheets. The main focus is the change in debt balances, though we also study effects on delinquencies. The econometric methods employed here parallel those used to study mortgage default in Section 5 quite closely. One key difference is that, as opposed to the mortgage default regressions where we censored following the borrower's first instance of mortgage default, in this section we do not censor at all when looking at debt balances, and we censor after the first instance of serious delinquency on non-mortgage debt when studying that outcome. Also note that, when looking at the first difference of debt balances, our focus is on the net change in balances, but we will decompose the effect into positive and negative changes. We do this by re-running the regressions with negative (positive) values censored to 0 to evaluate the impact on positive (negative) changes. This is important because increased cash flow could cause borrowers to take on more debt, as their ability to service that debt has increased, but it could also be used to pay down outstanding debts.

Table 5 shows summary statistics for debt levels and accumulation by debt category. About half of the observations have positive auto debt, which is typically being amortized, though in about 3% of months, borrowers take out large auto loans, so that on net the average borrower adds $17 of auto debt per month. HELOCs, student loans and retail consumer debt follow the same basic pattern—relatively rare and large positive accumulation followed by a longer, slower process of pay-downs—though they are less prevalent and smaller than auto loans. Bank cards, which are held by 85% of the sample, show a somewhat different pattern, as increases and decreases are essentially equally likely, with the increases being somewhat larger, leading to an average increase of $23/month.

Average Effects. The top row of Table 6 contains our main results on auto debt accumulation. This is a useful place to begin because other studies (e.g. Agarwal et al., 2017b; Beraja et al., 2017; Di Maggio et al., 2016, 2017) have looked at this outcome and argued that reducing mortgage payments leads to increased auto debt, as households take out auto loans to buy new cars. We, too, find evidence of a significant causal effect of a mortgage refinance on a consumer's auto debt. Refinancing leads to a net increase in auto debt accumulation of around $25/month. This estimate is stable across specifications. As the final two columns show, this net change combines two competing effects: increases to debt as households finance car purchases ($43), and decreases as they pay down their debt (-$19). The former, which proxies for auto spending, can be compared to other studies.

Agarwal et al. (2017b) do event studies of HARP refinances and find that the increase in auto debt is 20-22% of the magnitude of the payment reduction, very much in line with the $43/173 = 24\%$ increase that we find. Di Maggio et al. (2017) look at ARMs and find the increase in auto debt is closer to 10% of the payment reduction. This suggests that, as with mortgage default, the elasticity of the response may be larger for FRMs than ARMs. The studies, though, are difficult to compare directly, since the ARM resets in the Di Maggio et al. sample resulted in much larger payment reductions (of about $940/month, or 50%).

Looking beyond auto debt, Table 6 shows that HELOC debt is the one other instrument that shows a net effect on balances. The magnitude of the effect is similar to that of auto debt—$20-24/month—and similarly stable. Splitting this net effect into positive and negative components, it appears that HELOC draws (the positive change) are a bit less responsive than auto debt, and less is done to pay the debt down as well. These results suggest that the spending response induced by refinancing is larger and more broad-based than previously shown, as we are the first to look at debt balances beyond auto debt.

Bank card debt and retail consumer debt show little evidence of a net effect, but this is the result of combining increases in the balance with decreases. Specifically, when looking only at positive changes, refinancing caused borrowers to accumulate an additional $26 of bank card debt and $13 of retail consumer debt per month, while focusing on negative changes show these balances being paid off by $31 and $11 per month. Student debt shows no meaningful response, either in a statistical sense or in terms of magnitude.[35]

Looking across all these debt categories, we see that while auto debt, the focus of related previous literature, is an important component of households' responses to refinancing, it is only one of the active margins. In particular, the positive accumulation of debt is about 60% of the monthly savings on mortgage payments ($106/$173), more than twice the response one would see by looking at just autos. The reductions in debt were equal to about 40% of the mortgage payment savings ($64/$173), leading to a net increase in debt accumulation of about 20% of the savings caused by refinancing, on average.

These strong results stand in interesting contrast to the effects from reducing principal balances for highly-leveraged borrowers, studied in Ganong and Noel (2017). Using quasi-experimental variation generated by HAMP, they show that writing down the debt of underwater borrowers to a lower—but still underwater—level has no effect on consumer spending, because this intervention does not provide borrowers with collateral that can be borrowed against. In contrast, we find that lowering mortgage payments, which increases cash on hand, can be quite effective in spurring consumer spending. These results support the argument of Eberly and Krishnamurthy (2014) that payment relief is likely a more efficient intervention for policymakers in a downturn than principal

[35] As opposed to delinquency outcomes, where it is intuitive that the OLS magnitude should be biased upwards relative to the true causal effect, for these debt accumulation outcomes it is not obvious which way the bias would go. Table A-4 shows the OLS results. The IV effects are stronger for the auto and HELOC measures, smaller for bank card debt, and very comparable for the student and retail consumer debt categories.

reduction.

Heterogeneity. The previous discussion emphasized that, in some of the debt categories, refinancing caused sizable increases of debt balances for some borrowers but decreases for others. We now look at whether certain attributes predict whether a borrower will react strongly, either in the positive or negative direction. We use the same specification as described toward the end of Section 5 to simultaneously instrument for Refied_{it} and $\text{Refied}_{it} \times \text{HighCreditScore}_{it}$ (or $\text{Refied}_{it} \times \text{HighCreditUtilization}_{it}$, or $\text{Refied}_{it} \times \text{HighCLTV}_{it}$).

Table 7 shows the results for heterogeneity along these dimensions. Consistent with Di Maggio et al. (2017) and Agarwal et al. (2017b), we find that a low credit score is predictive of strong positive debt growth. The large average net effects on auto and HELOC debt appear to be driven by low-credit score borrowers and hardly at all by those with high credit scores. In addition, the positive accumulation of retail consumer debt discussed above seems to be driven by low-credit score households, whereas the de-accumulation of that debt in the last column of Table 6 is driven by high-credit score borrowers.

The split by credit utilization is a bit less dramatic, as both high- and low-utilization borrowers respond along the auto and HELOC margins. Interestingly, there is a big split along this dimension when looking at bank card debt. Those with low credit utilization respond to their refinance by accumulating more debt on their bank card, while those with high credit utilization use the increased cash flow to pay down their bank card debt. This is sensible given that bank card debt is not collateralized and comes with high interest rates.

Of the three characteristics, CLTV does the least to predict responsiveness to refinances across the debt categories. The difference between high- and low-CLTV borrowers is not statistically significant for any category, and it appears that both groups acted broadly like the average refinancer whose results are shown in Table 6. This could be because our identifying variation is all coming from borrowers with high CLTVs (as these effects are attributable to HARP eligibility). As a result, even if there are important differences between how truly low- and high-CLTV borrowers respond to a refinance, we may be unable to detect thes differences because there is little CLTV variation in our sample.

These heterogeneity results are consistent with standard models of consumption in the presence of liquidity constraints (e.g. Zeldes, 1989; Deaton, 1991). If a borrower would like to consume some of her future income in the present but lacks the access to credit required to facilitate this, she will be inclined to immediately spend a large portion of any windfall, as present marginal utility is quite high. Our results show that, indeed, borrowers with low credit scores and high utilization of their current lines of credit—people who likely face relatively tight liquidity constraints—respond most strongly to the positive cash flow shock generated by refinancing their mortgage.

In sum, we have shown that the effect of a refinance on consumer debt balances is more nuanced than previously demonstrated. As in previous studies, we find a strong increase in auto debt. We have shown that, in addition, increased HELOC debt is an important margin. Both of these effects are especially evident for borrowers with low credit scores, and this group also draws on retail consumer debt after refinancing. Borrowers with sufficient unused credit on their bank cards use them to take on debt, as well, while those closer to their limits use the increased cash flow to pay down these balances.

Delinquency. Finally, we look at effects on the likelihood of serious delinquency on a debt other than the first mortgage. We define a serious delinquency to have occurred if a borrower has had 3 consecutive months with positive delinquent non-mortgage balances. We censor borrowers after their first serious delinquency or when they exit McDash due to a non-refinance prepayment or servicing transfer. The takeup-weighted monthly hazard rate in the sample is 11.7bp for this outcome, twice that of mortgage default.

Panel B of Table 3 shows the treatment effect on serious delinquency on non-mortgage debt. Though this is one outcome that is quite sensitive to the inclusion of cohort effects, the weight of the evidence suggests that refinancing reduced this hazard by about 3bp, or 25%. As with mortgage default, we find that the OLS estimate is larger, further evidence that borrowers who refinance are be unobservably less likely to default.

Panel B of Table 4 again looks for heterogeneity along a handful of observable dimensions. As with most of the other outcomes, we find that borrowers with low credit scores and borrowers with high credit utilization are responsible for most of the overall treatment effect. Those with low credit scores reduced their non-mortgage delinquency rate by 6.1bp (compared to a base of 24.4bp), while those with high credit utilization lowered their default rate by 5.5bp (compared to an average default rate of 20.6bp). Again, we find some evidence that high-CLTV borrowers were more responsive than their low-CLTV counterparts as well, but the difference is less clear. While the absolute effect is stronger for the low-credit score and high-credit utilization borrowers, as with mortgage default, for this outcome the proportional effect does not seem much different for these borrowers than for the overall sample—a reduction of about 25%.

This provides evidence that refinancing not only helps households and the owners of their mortgage credit risk, but the households' other creditors, too—particularly for households with worse observable financial characteristics. This positive externality further illustrates the importance of looking at the entirety of households' balance sheets when assessing the effects of mortgage refinancing.

7 Take-up

The preceding analysis finds that refinances had beneficial effects on borrower finances, reducing the probability of default on mortgages and other debt, and allowing for additional spending. Nevertheless, half of HARP-eligible borrowers in the sample did not refinance. This is in line with other studies finding that many borrowers fail to refinance when it appears to be in their financial interest to do so (e.g., Campbell, 2006; Keys et al., 2016; Andersen et al., 2017; Agarwal et al., 2015). In this section, we ask which borrower characteristics predict take-up. In addition to the financial variables available in the CRISM microdata (CLTV, credit score, etc.), we will include some local (ZIP- and county-level) characteristics as well. These were irrelevant for the previous analysis because we always included ZIP-code FEs to get the cleanest possible variation. However, in this section, we are less interested in producing convincing causal identification and more interested in simply describing who refinances. These additional variables come from a few sources. We use Home Mortgage Disclosure Act (HMDA) data to measure local mortgage market concentration, following Scharfstein and Sunderam (2016). We also merge in ZIP code-level Census data on income and education. One final addition to the set of variables is borrower age. This is available in CRISM, and we use it to see if we find life-cycle variation in the use of the refinancing opportunity, in the spirit of Agarwal et al. (2009).

We run a linear probability model to correlate refinancing activity with the micro variables in CRISM and the local aggregate variables from the Census and HMDA. For this section only, we do not use the panel feature of CRISM—instead, each borrower is a single observation. The outcome variable is an indicator for whether the borrower refinances at any point in the sample. All time-varying covariates are measured in March of 2010. Because HARP eligibility is such a strong predictor of refinancing behavior in our sample, we keep only HARP-eligible borrowers, reducing our sample size from 220 thousand to 106 thousand. About 53% of this sample refinanced.

Results are shown in Table 8. Column (1) includes ZIP-code FEs while column (2) does not. Comparing the two demonstrates that the estimated effects of individual characteristics are barely altered. Subsequent columns include various local characteristics, as discussed below.

Looking at individual-level financial variables first, we see that refinancing was more likely among borrowers with high credit scores, low utilization of revolving credit, and low CLTVs, all indicators of financial health. The effect of credit utilization is intuitive, since highly liquidity-constrained borrowers should have high effective discount rates and so be less inclined to undertake a refinance, which has up-front costs and delayed benefits. Having said that, the effect is not very strong—the difference in the refinancing probabilities for very high and very low values of credit utilization is only a couple of percentage points. The credit score and CLTV results are quite a bit stronger, the former showing a +15pp effect of being in the top of the credit score distribution, while there is a -9pp effect from having a very high CLTV. The credit score result may be picking up a lack of financial sophistication or a wariness by lenders of working with bad credit risks (either because of high servicing costs or because they feared future liability due to reps and warrants). The same reason could explain the CLTV result. Alternatively, deeply underwater borrowers may have believed they were ineligible for a refinance, despite HARP. This is all the more plausible given that when HARP was first rolled out, borrowers with very high LTVs were not eligible (see Section 2). Broadly speaking, these results suggest that the borrowers who we found earlier to respond most strongly to a refinance across a broad set of measures were the least likely to actually undertake one.

We also find a hump-shaped pattern in age, broadly consistent with Agarwal et al. (2009)'s findings across household financial behaviors. Specifically, the middle tercile in our data includes ages 37-48, and we find that borrowers in this age range are significantly more likely than older or younger borrowers to refinance, though the effect size is modest. Having a junior lien associated with the first mortgage seems to have no effect on the likelihood of refinancing (controlling for CLTV and other characteristics); we do, however, find that ARM borrowers were much less likely to use HARP.

Looking at local mortgage market concentration (measured by the share of mortgages in 2010 made by a county's four largest lenders), we find only weak evidence consistent with Scharfstein and Sunderam (2016) that concentrated markets are associated with less refinancing activity—being in the top concentration tercile appears to be correlated with a smaller likelihood of a refinance, but the effect is not statistically significant. Turning finally to the variables from the Census, we find a monotonic, positive, and significant relationship between ZIP code median income and the likelihood of refinancing. This result could reflect the direct effect of having more resources (similar to the credit utilization result above), or could be driven by easier access to finance for borrowers in higher-income areas. Surprisingly, we find that borrowers in ZIP codes with highly-educated populations are less likely to refinance, although differences are small and not very significant.[36]

Even without a causal interpretation, this analysis gives a better sense of who refinances their mortgages when eligible for a program that makes refinancing easily accessible. In particular, the evidence suggests that consumers in stronger financial positions are quite a bit more likely to refinance, and these are exactly the ones whom we find the weakest responses for in the causal estimation. This finding is reminiscent of other settings where it is difficult for public policies to reach the recipients that might derive particularly large benefits from them.[37]

8 Conclusion

Using a cutoff date in the eligibility criteria for HARP, we find that lowering monthly mortgage payments by refinancing decreases the likelihood of default—on mortgages as well as other debts—substantially, particularly for borrowers with low credit scores and high credit utilization rates. We also find strong effects on the accumulation of other debts, with some borrowers using the payment

[36] The unconditional relationship (unreported) goes in the expected direction. Keys et al. (2016) study local covariates of the share of borrowers who fail to refinance despite it being in their interest to do so (not restricting to HARP-eligible borrowers). Similar to us, they find little association with mortgage market concentration or education levels. Unlike us, they do not find much of a relation with incomes.

[37] In the context of monetary policy operating through household credit markets, Agarwal et al. (2018) find that as banks' cost of funds decreases, they increase credit card limits the most for borrowers whose spending responds least strongly. Fuster and Willen (2010) and Beraja et al. (2017) show that the Federal Reserve's first round of large-scale asset purchases triggered increased refinancing activity, but mostly by high-credit score borrowers and in locations that were already doing relatively better.

relief to take on new debts and others paying down their existing debts. In particular, borrowers with low credit scores add to their auto, HELOC, and retail consumer debts substantially, while borrowers with high credit scores react less strongly. Borrowers with high credit utilization rates originate auto debt more than do their low utilization counterparts, but they are quite a bit less aggressive in their use of HELOCs and bank cards.

These results broaden our understanding of how cash flow shocks affect borrower behavior. As discussed, the related literature on mortgage payment changes has primarily focused on auto debt accumulation (e.g. Di Maggio et al., 2017; Agarwal et al., 2017b) and finds that borrowers with low credit scores, a proxy for financial constraints, are more responsive. While we find this also, we show that the pattern is different for different debt instruments and different proxies for distress. The leading example is bank card debt, which borrowers with high credit utilization actually pay down using their improved cash flow. The cash flow, then, does not merely boost spending but is also used to repair borrowers' balance sheets in sensible ways.

We further provide evidence that those tags for strong responses to a refinance—low credit score, high credit utilization—also predict a relatively *low* probability of refinancing. One potential explanation for this is that these households are not prepared to pay the immediate costs of the refinancing process due to their current distress, while another is that these borrowers are simply less financially sophisticated, which is the underlying cause of both the financial distress and the lower likelihood of refinancing. Either way, from the perspective of a policymaker, this relationship weakens the refinancing channel, as lower mortgage rates are less likely to translate into payment reductions for exactly the borrowers who are most responsive to them. This strengthens the point made by the mortgage design literature (e.g. Eberly and Krishnamurthy, 2014; Guren et al., 2017) that mortgages that have payment reductions triggered automatically in downturns (such as ARMs) can be a useful stabilizing tool for the macroeconomy.

Finally, from the point of view of evaluating HARP, our evidence implies that the program had large beneficial effects on the households that were able to benefit from it. At the same time, as we discussed, the initial use of the program prior to HARP 2.0 was rather limited, due to various frictions. Had the program reached more borrowers in 2009-2010, after the initial drop in mortgage rates due in part to the Federal Reserve's asset purchases, this could have supported the housing market and aggregate demand during the worst phase of the recession. Thus, it appears important to learn from this experience and ensure that refinancing frictions are minimized especially in times when the effects we document are most valuable.

Bibliography

Agarwal, S., G. Amromin, I. Ben-David, S. Chomsisengphet, T. Piskorski, and A. Seru (2017a): "Policy Intervention in Debt Renegotiation: Evidence from the Home Affordable Modification Program," *Journal of Political Economy*, 125, 654–712.

Agarwal, S., G. Amromin, S. Chomsisengphet, T. Landvoigt, T. Piskorski, A. Seru, and V. Yao

——— (2017b): "Mortgage Refinancing, Consumer Spending, and Competition: Evidence From the Home Affordable Refinance Program," Working Paper 21512, NBER.

Agarwal, S., S. Chomsisengphet, N. Mahoney, and J. Stroebel (2018): "Do Banks Pass Through Credit Expansions to Consumers Who Want to Borrow?" *Quarterly Journal of Economics*, 133, 129–190.

Agarwal, S., J. Driscoll, X. Gabaix, and D. Laibson (2009): "The Age of Reason: Financial Decisions over the Life Cycle and Implications for Regulation," *Brookings Paper on Economic Activity*, 2, 51–101.

Agarwal, S., R. J. Rosen, and V. Yao (2015): "Why Do Borrowers Make Mortgage Refinancing Mistakes?" *Management Science*, 62, 3494–3509.

Amromin, G. and C. Kearns (2014): "Access to Refinancing and Mortgage Interest Rates: HARPing on the Importance of Competition," Working Paper Series WP-2014-25, FRB Chicago.

Andersen, S., J. Campbell, K. Nielsen, and T. Ramadorai (2017): "Inattention and Inertia in Household Finance: Evidence from the Danish Mortgage Market," Working Paper 21386, NBER.

Angrist, J., G. Imbens, and D. Rubin (1996): "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 434, 444–55.

Auclert, A. (2017): "Monetary Policy and the Redistribution Channel," Working Paper 23451, NBER.

Beraja, M., A. Fuster, E. Hurst, and J. Vavra (2017): "Regional Heterogeneity and Monetary Policy," Staff Report No. 731, Federal Reserve Bank of New York.

Bhutta, N. and B. J. Keys (2016): "Interest Rates and Equity Extraction during the Housing Boom," *American Economic Review*, 106, 1742–1774.

Bond, P., R. Elul, S. Garyn-Tal, and D. K. Musto (2017): "Does Junior Inherit? Refinancing and the Blocking Power of Second Mortgages," *Review of Financial Studies*, 30, 211–244.

Campbell, J. and J. Cocco (2015): "A Model of Mortgage Default," *Journal of Finance*, 70, 1495–1554.

Campbell, J., S. Giglio, and P. Pathak (2011): "Forced Sales and Home Prices," *American Economic Review*, 5, 2108–31.

Campbell, J. Y. (2006): "Household Finance," *Journal of Finance*, 61, 1553 – 1604.

——— (2013): "Mortgage Market Design," *Review of Finance*, 17, 1 – 33.

Chen, H., M. Michaux, and N. Roussanov (2013): "Houses as ATMs? Mortgage Refinancing and Macroeconomic Uncertainty," Working Paper 19421, NBER.

Deaton, A. (1991): "Saving and Liquidity Constraints," *Econometrica*, 59, 1221–1248.

Di Maggio, M., A. Kermani, B. J. Keys, T. Piskorski, R. Ramcharan, A. Seru, and V. Yao (2017): "Interest Rate Pass-Through: Mortgage Rates, Household Consumption, and Voluntary Deleveraging," *American Economic Review*, 107, 3550–3588.

Di Maggio, M., A. Kermani, and C. Palmer (2016): "How Quantitative Easing Works: Evidence on the Refinancing Channel," Working Paper 22638, NBER.

Dudley, W. C. (2012): "Housing and the Economic Recovery," Remarks at the New Jersey Bankers Association Economic Forum, January 6, 2012.

Eberly, J. and A. Krishnamurthy (2014): "Efficient Credit Policies in a Housing Debt Crisis," *Brookings Papers on Economic Activity,*, Fall 2014.

Ehrlich, G. and J. Perry (2015): "Do Large-Scale Refinancing Programs Reduce Mortgage Defaults? Evidence From a Regression Discontinuity Design," Working Paper Series 2015-06, Congressional Budget Office.

Elul, R., N. S. Souleles, S. Chomsisengphet, D. Glennon, and R. Hunt (2010): "What 'Triggers' Mortgage Default?" *American Economic Review*, 100, 490–494.

Federal Housing Finance Agency Office of Inspector General (2013): "Home Affordable Refinance Program: A Mid-Program Assessment," .

Frame, W. S., A. Fuster, J. Tracy, and J. Vickery (2015): "The Rescue of Fannie Mae and Freddie Mac," *Journal of Economic Perspectives*, 29, 25–52.

Fuster, A. and P. Willen (2017): "Payment Size, Negative Equity, and Mortgage Default," *American Economic Journal: Economic Policy*, 9, 167–191.

Fuster, A. and P. S. Willen (2010): "$1.25 Trillion is Still Real Money: Some Facts About the Effects of the Federal Reserve's Mortgage Market Investments," Public Policy Discussion Paper 10-4, Federal Reserve Bank of Boston.

Ganong, P. and P. Noel (2017): "The Effect of Debt on Default and Consumption: Evidence from Housing Policy in the Great Recession," Working Paper, Harvard University.

Greenspan, A. (2004): "Testimony Before the Committee on Financial Services, U.S. House of Representatives," February 11, 2004.

Greenwald, D. L. (2017): "The Mortgage Credit Channel of Macroeconomic Transmission," Working Paper, MIT Sloan.

Guren, A., A. Krishnamurthy, and T. McQuade (2017): "Mortgage Design in an Equilibrium Model of the Housing Market," Working Paper.

Hubbard, R. G. and C. J. Mayer (2009): "The Mortgage Market Meltdown and House Prices," *The B.E. Journal of Economic Analysis & Policy*, 9, Article 8.

Hurst, E. and F. Stafford (2004): "Home Is Where the Equity Is: Mortgage Refinancing and Household Consumption," *Journal of Money, Credit, and Banking*, 36, 985–1014.

Jappelli, T. and L. Pistaferri (2014): "Fiscal Policy and MPC Heterogeneity," *American Economic Journal: Macroeconomics*, 6, 107–136.

Johnson, E., S. Meier, and O. Toubia (2015): "Leaving Money on the Kitchen Table: Exploring Sluggish Mortgage Refinancing Using Administrative Data, Surveys, and Field Experiments," Working Paper, Columbia University.

Kaplan, G., B. Moll, and G. L. Violante (2018): "Monetary Policy According to HANK," *American Economic Review*, 108, 697–743.

Karamon, K., D. McManus, and J. Zhu (2016): "Refinance and Mortgage Default: A Regression Discontinuity Analysis," *The Journal of Real Estate Finance and Ecnonomics*, 1–19.

Kau, J. B., D. C. Keenan, W. J. Muller, and J. F. Epperson (1992): "A Generalized Valuation Model for Fixed-Rate Residential Mortgages," *Journal of Money, Credit and Banking*, 24, 279–299.

Keys, B., D. Pope, and J. Pope (2016): "Failure to Refinance," *Journal of Financial Economics*, 122, 482 – 499.

Remy, M., D. Lucas, and D. Moore (2011): "An Evaluation of Large-Scale Mortgage Refinancing Programs," Working Paper 2011-4, Congressional Budget Office.

Scharfstein, D. and A. Sunderam (2016): "Market Power in Mortgage Lending and the Transmission of Monetary Policy," Working Paper, Harvard University.

Scharlemann, T. C. and S. H. Shore (2016): "The Effect of Negative Equity on Mortgage Default: Evidence From HAMP's Principal Reduction Alternative," *Review of Financial Studies*, 29, 2850–2883.

Stiglitz, J. E. and M. Zandi (2012): "The One Housing Solution Left: Mass Mortgage Refinancing," *New York Times*, August 13, A17.

Tracy, J. and J. Wright (2016): "Payment Changes and Default Risk: The Impact of Refinancing on Expected Credit Losses," *Journal of Urban Economics*, 93, 60–70.

Wong, A. (2016): "Population Aging and the Transmission of Monetary Policy to Consumption," Working Paper, Northwestern University.

Zeldes, S. (1989): "Consumption and Liquidity Constraints: An Empirical Investigation," *Journal of Political Economy*, 97, 305–346.

Zhu, J., J. Janowiak, L. Ji, K. Karamon, and D. McManus (2015): "The Effect of Mortgage Payment Reduction on Default: Evidence from the Home Affordable Refinance Program," *Real Estate Economics*, 43, 1035–1054.

	A. Sample (N=220k)				B. High-LTV, GSE (N=6.5 mill)				C. GSE (N=21.3 mill)			
(As of March 2010)	Mean	Median	10^{th} Pctl	90^{th} Pctl	Mean	Median	10^{th} Pctl	90^{th} Pctl	Mean	Median	10^{th} Pctl	90^{th} Pctl
CLTV	87.9%	85.3%	80.8%	96.3%	111.1%	98.5%	82.7%	140.8%	74.5%	72.3%	29.9%	113.4%
Interest Rate	4.95%	4.88%	4.50%	5.38%	5.97%	6.00%	4.88%	6.88%	5.73%	5.75%	4.75%	6.75%
Mortgage Principal	$226k	$212k	$114k	$360k	$189k	$171k	$83k	$326k	$155k	$131k	$47k	$298k
FICO® Score	762	775	698	809	717	741	595	802	736	766	632	809
Credit Utilization	24.6%	12.8%	0.1%	71.7%	40.7%	33.1%	0.7%	94.0%	33.4%	20.4%	0.4%	88.6%
ARM? (share)	0.4%				10.4%				6.5%			
30-year? (share)	93.1%				94.0%				72.7%			
Junior lien? (share)	10.4%				28.1%				30.6%			

Table 1: Moments of the distributions of key observables in the CRISM dataset for borrowers with mortgages whose payments were current as of March of 2010. Panel A looks at our sample: borrowers with GSE-guaranteed mortgages with LTVs above 80%, originated in the first half of 2009. Panel B drops the origination cohort filter, leaving all high-LTV GSE borrowers. Panel C drops the LTV requirement, so all GSE borrowers are included. Credit utilization is measured as the sum of HELOC and bank card balances divided by the sum of HELOC and bank card limits.

	Mean: Eligible (N = 115,298)	Mean: Ineligible (N = 104,862)	Diff
CLTV	87.3%	88.5%	$-1.2\%^{***}$
FICO® Score	763.6	759.6	4.0^{***}
Interest rate	4.98%	4.92%	$0.05\%^{***}$
Credit Utilization	24.6%	24.7%	-0.1%
First Mortgage Balance	$225.3k	$225.9k	$-\$0.6k$
All Other Debt Balances	$24.3k	$25.1k	$-\$0.8k^{***}$
Refinanced	53.16%	34.93%	$18.23\%^{***}$
Defaulted	3.53%	3.65%	-0.12%
Servicing transferred from McDash	6.20%	8.23%	$-2.04\%^{***}$
Refinanced out of McDash	11.36%	11.80%	$-0.44\%^{***}$
Non-refi prepay	21.09%	19.33%	$1.75\%^{***}$
Active	57.82%	56.98%	$0.84\%^{***}$

Table 2: Top half of the table checks for balance in our sample between borrowers with mortgages purchased by a GSE before the cutoff date (eligible) and mortgages purchased after (ineligible). Variables in the top half are measured in March of 2010. Bottom half of the table shows the fraction that ever refinanced in the sample and the termination status of the mortgages as of February 2016. $*$: significant at 10% level; $**$: significant at 5% level; $***$: significant at 1% level.

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	A. Mortgage Default				**B. Non-Mtg Serious Delinquency**			
	IV				IV			
Basis Points	-2.41	-2.21	-2.09	-3.21	-2.81	-2.74	-3.03	-1.17
Std. Err.	$(0.74)^{***}$	$(0.74)^{***}$	$(0.74)^{***}$	$(1.02)^{***}$	$(1.18)^{**}$	$(1.24)^{**}$	$(1.24)^{**}$	(1.71)
	OLS				OLS			
Basis Points	-4.78	-4.58	-4.59	-4.77	-3.42	-3.33	-3.35	-3.30
Std. Err.	$(0.26)^{***}$	$(0.27)^{***}$	$(0.27)^{***}$	$(0.29)^{***}$	$(0.26)^{***}$	$(0.27)^{***}$	$(0.27)^{***}$	$(0.27)^{***}$
Quarter FEs	✓	✓	✓	✓	✓	✓	✓	✓
ZIP-code FEs	✓	✓	✓	✓	✓	✓	✓	✓
Observables	✓	✓	✓	✓	✓	✓	✓	✓
Q-by-ZIP FEs		✓	✓	✓		✓	✓	✓
Guar. Lag FEs			✓	✓			✓	✓
Cohort FEs				✓				✓
N (mill.)	11.6	11.6	11.6	11.6	11.8	11.8	11.8	11.8

Table 3: Regression estimates of the treatment effect of refinancing on monthly probabilities of mortgage default and non-mortgage serious delinquency. For borrower i in month t, the refinancing indicator is turned on if she has completed a refinance in some month $\tau \leq t$. The mortgage default indicator is turned on if she is at least 90 days delinquent on her mortgage in month t, and the non-mortgage serious delinquency indicator is turned on if she has had 3 consecutive months with delinquent balances on non-first-mortgage debt. Borrowers are censored after their first month in default/serious delinquency. The takeup-weighted monthly mortgage default rate is 5.9bp, while it is 11.7bp for non-mortgage serious delinquency. IV estimates result from instrumenting for the refinance indicator with HARP eligibility, interacted with a full set of quarter indicators. "Observables" include CLTV (lagged 3 months and as of March 2010), credit score (lagged 3 months and as of March 2010), credit utilization (lagged 3 months and as of March 2010), initial mortgage rate, initial debt balances, and remaining principal balance. Standard errors (in parentheses) are clustered at the county level. $*$: significant at 10% level; $**$: significant at 5% level; $***$: significant at 1% level.

Borrower Characteristic:	FICO® Score			Credit Utilization			CLTV		
	Low	High	Diff	Low	High	Diff	Low	High	Diff
				A. Mortgage Default					
Basis Points	-4.68	0.02	-4.71	-0.06	-4.50	4.34	-1.49	-4.05	2.57
(Std. Err.)	(1.07)***	(0.64)	(0.86)***	(0.73)	(1.00)***	(0.91)***	(0.82)*	(0.99)***	(1.10)**
Average Default Rate (bp)	10.5	1.8		3.6	8.3		4.1	9.1	
				B. Non-Mortgage Serious Delinquency					
Basis Points	-6.39	-0.97	-5.42	-0.76	-5.83	5.07	-2.57	-4.28	1.71
(Std. Err.)	(1.80)***	(0.99)	(1.33)***	(1.08)	(1.63)***	(1.20)***	(1.46)*	(1.36)***	(1.56)
Average Default Rate (bp)	24.4	1.7		4.1	20.6		10.1	15.1	

Table 4: Regression estimates of heterogeneous LATEs of refinancing on mortgage default and non-mortgage serious delinquency. For borrower i in month t, the refinancing indicator is turned on if she has completed a refinance in some month $\tau \leq t$. The mortgage default indicator is turned on if she is at least 90 days delinquent on her mortgage in month t, and the non-mortgage serious delinquency indicator is turned on if she has had 3 consecutive months with delinquent balances on non-first-mortgage debt. Borrowers are censored after their first month in default/serious delinquency. The borrower is considered to have a "high" credit score (or CLTV or credit utilization) if her score in month t is larger than the median value in the entire sample. IV estimates result from instrumenting for the refinance indicator and its interaction with the borrower characteristic of interest (e.g. High Credit Score) with HARP eligibility, interacted with the variable of interest, further interacted with a full set of quarter indicators. The set of controls corresponds to column (3) in the previous table: Quarter FEs, ZIP-code FEs, observables, and guarantee lag FEs. "Observables" include CLTV (lagged 3 months and as of March 2010), credit score (lagged 3 months and as of March 2010), credit utilization (lagged 3 months and as of March 2010), initial mortgage rate, remaining principal balance, and initial debt balances. Default rates (bottom row) are monthly and weighted by take-up. Standard errors (in parentheses) are clustered at the county level. $*$: significant at 10% level; $**$: significant at 5% level; $***$: significant at 1% level.

	$Balance$			$Bal \| Bal > 0$		$\Delta Balance$				$\Delta Bal \| \Delta Bal > 0$		$\Delta Bal \| \Delta Bal < 0$	
	Mean	SD	% Pos	Mean	SD	Mean	SD	% Pos	% Neg	Mean	SD	Mean	SD
Auto	8,832	12,304	51	17,426	12,205	17	3,142	3	51	11,990	12,054	-693	1,861
HELOC	2,276	12,016	7	32,588	32,855	9	2,327	1	4	4,624	13,322	-1,287	7,521
Bank Card	6,077	8,592	85	7,112	8,890	23	2,246	43	42	1,189	1,915	-1,165	2,333
Student	4,836	15,307	17	26,408	28,796	-5	1,899	2	12	3,271	9,202	-550	3,818
Retail Cons. Dt.	2,088	6,205	43	4,895	8,746	2	1,405	15	30	808	2,905	-384	1,416

Table 5: Summary statistics of debt balances in CRISM sample. We trim observations from a calculation if their balance in that debt category is greater than the 99^{th} percentile of non-zero balances in the sample.

	Net Change				Positive	Negative
Auto	27	24	24	21	43	-19
(Std. Err.)	(8)***	(8)***	(8)***	(12)*	(7)***	(7)***
HELOC	23	24	23	20	33	-10
	(6)***	(6)***	(6)***	(9)**	(8)***	(6)
Bank Card	-3	-4	-5	-18	26	-31
	(5)	(5)	(5)	(7)**	(9)***	(9)***
Student	1	-1	-2	8	-9	7
	(5)	(5)	(5)	(6)	(6)	(5)
Retail Consumer Debt	1	1	2	4	13	-11
	(3)	(3)	(3)	(5)	(4)***	(4)***
Initial Bal.	✓	✓	✓	✓	✓	✓
Quarter FEs	✓	✓	✓	✓	✓	✓
ZIP-code FEs	✓	✓	✓	✓	✓	✓
Observables	✓	✓	✓	✓	✓	✓
Q-by-ZIP FEs		✓	✓	✓	✓	✓
Guar. Lag FEs			✓	✓	✓	✓
Cohort FEs				✓		
N (mill.)	13.6	13.6	13.6	13.6	13.6	13.6

Table 6: Regression estimates of the LATE of refinancing on monthly debt accumulation. For borrower i in month t, the refinancing indicator is turned on if she has completed a refinance in some month $\tau \leq t$. Outcomes are the first difference in debt balances. "Net Change" is the simple difference, while the column labeled "Positive" censors negative changes to 0, and the column labeled "Negative" censors positive changes to 0. IV estimates result from instrumenting for the refinance indicator with HARP eligibility, interacted with a full set of quarter indicators. "Observables" include CLTV (lagged 3 months and as of March 2010), credit score (lagged 3 months and as of March 2010), credit utilization (lagged 3 months and as of March 2010), initial mortgage rate, remaining principal balance, and initial debt balances. Standard errors (in parentheses) are clustered at the county level. $*$: significant at 10% level; $**$: significant at 5% level; $***$: significant at 1% level.

Characteristic:	FICO® Score			Cred. Util.			CLTV		
	Low	High	Diff	Low	High	Diff	Low	High	Diff
Auto	35	14	21	17	29	-13	30	17	12
(Std. Err.)	(10)***	(8)*	(10)**	(9)*	(9)***	(9)	(10)***	(9)**	(11)
HELOC	42	7	35	32	18	13	32	21	12
	(8)***	(6)	(8)***	(6)***	(8)**	(8)*	(8)***	(6)***	(10)
Bank Card	-15	4	-18	14	-26	40	-3	-2	-2
	(7)**	(5)	(7)***	(5)***	(7)***	(7)***	(6)	(6)	(7)
Student	-4	0	-5	-5	0	-5	-3	-2	0
	(6)	(5)	(5)	(5)	(6)	(5)	(6)	(5)	(6)
Retail Consumer Debt	16	-12	28	-1	3	-4	-2	6	-7
	(5)***	(4)***	(5)***	(4)	(5)	(4)	(5)	(4)	(5)

Table 7: Regression estimates of heterogeneous LATEs of refinancing on monthly debt accumulation. For borrower i in month t, the refinancing indicator is turned on if she has completed a refinance in some month $\tau \leq t$. The outcome variable is the first difference in the debt balance. The borrower is considered to have a "high" credit score (or CLTV or credit utilization) if her score in month t is larger than the median value in the entire sample. IV estimates result from instrumenting for the refinance indicator and its interaction with the variable of interest (e.g. High Credit Score) with HARP eligibility, interacted with the variable of interest, further interacted with a full set of quarter indicators. The set of controls corresponds to column (3) in the previous tables: Quarter FEs, ZIP-code FEs, observables, and guarantee lag FEs. "Observables" include CLTV (lagged 3 months and as of March 2010), credit score (lagged 3 months and as of March 2010), credit utilization (lagged 3 months and as of March 2010), initial mortgage rate, remaining principal balance, and initial debt balances. Standard errors (in parentheses) are clustered at the county level. $*$: significant at 10% level; $**$: significant at 5% level; $***$: significant at 1% level.

Dependent variable: Refinance at some point over March 2010 - February 2016 (0/100)

FICO® Score ($\leq$ 675 omitted)						
676-725	8.3	8.8	8.8	8.8	8.8	8.8
	(0.9)***	(0.8)***	(0.8)***	(0.8)***	(0.8)***	(0.8)***
726-775	12.0	12.4	12.4	12.4	12.4	12.4
	(0.8)***	(0.7)***	(0.7)***	(0.7)***	(0.7)***	(0.7)***
776-800	14.1	15.0	15.0	15.0	14.9	15.0
	(0.8)***	(0.7)***	(0.7)***	(0.7)***	(0.7)***	(0.7)***
$>$ 800	13.8	14.8	14.8	14.8	14.8	14.8
	(0.9)***	(0.8)***	(0.8)***	(0.8)***	(0.8)***	(0.8)***
Credit Utilization (1^{st} quartile omitted)						
2^{nd} quartile	0.9	0.9	0.9	0.8	0.8	0.8
	(0.5)**	(0.4)**	(0.4)**	(0.4)**	(0.4)*	(0.4)*
3^{rd} quartile	0.5	0.5	0.5	0.5	0.4	0.4
	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)
4^{th} quartile	-1.7	-1.9	-1.9	-1.9	-1.9	-1.9
	(0.6)***	(0.5)***	(0.5)***	(0.5)***	(0.5)***	(0.5)***
Cred. Lim. = 0	-2.2	-2.7	-2.7	-2.7	-2.6	-2.6
	(0.8)***	(0.7)***	(0.7)***	(0.7)***	(0.7)***	(0.7)***
CLTV ($\leq$ 85% omitted)						
85%-90%	-2.6	-2.8	-2.8	-2.8	-2.7	-2.7
	(0.4)***	(0.4)***	(0.4)***	(0.4)***	(0.4)***	(0.4)***
90%-95%	-4.8	-4.5	-4.6	-4.5	-4.4	-4.5
	(0.5)***	(0.5)***	(0.5)***	(0.5)***	(0.5)***	(0.5)***
$>$ 95%	-8.6	-9.2	-9.2	-9.2	-9.1	-9.1
	(0.7)***	(0.6)***	(0.6)***	(0.6)***	(0.6)***	(0.6)***
Borrower age (middle tercile omitted)						
Bottom tercile	-1.2	-1.3	-1.3	-1.3	-1.3	-1.3
	(0.4)***	(0.4)***	(0.4)***	(0.4)***	(0.4)***	(0.4)***
Top tercile	-1.4	-2.4	-2.4	-2.4	-2.3	-2.3
	(0.4)***	(0.4)***	(0.4)***	(0.4)***	(0.4)***	(0.4)***
Junior Lien	-1.4	-0.0	-0.0	-0.0	-0.1	-0.1
	(0.7)*	(0.7)	(0.7)	(0.7)	(0.7)	(0.7)
ARM	-5.2	-6.7	-6.8	-6.8	-6.8	-6.9
	(4.6)	(4.0)*	(4.0)*	(4.0)*	(4.1)*	(4.1)*
Mkt Share of Top 4 Banks (middle tercile omitted)						
Bottom tercile			-0.5			-0.6
			(0.9)			(0.8)
Top tercile			-0.8			-1.2
			(0.8)			(0.8)
% $\geq$ Bachelor's Deg. in ZIP (middle tercile omitted)						
Bottom tercile				-0.8		0.1
				(0.5)		(0.5)
Top tercile				-0.3		-1.0
				(0.4)		(0.6)*
Median Income in ZIP (middle tercile omitted)						
Bottom tercile					-1.6	-2.0
					(0.5)***	(0.5)***
Top tercile					0.9	1.5
					(0.6)	(0.6)**
N (thousands)	106	106	106	106	106	106
ZIP fixed effects	Y	N	N	N	N	N

Table 8: Borrower-level, linear probability model regressions of refinancing. The outcome variable is an indicator for whether the borrower ever refinanced. Effects are expressed in percentage points. All time-varying covariates are measured in March 2010. For context, 53.4% of the sample refinanced. Flexible controls for contract rate and mortgage balance are included in the regression but excluded from the table for space concerns. As expected, they show that borrowers with higher interest rates and mortgage balances are more likely to refinance. Standard errors (in parentheses) are clustered at the county level. $*$: significant at 10% level; $**$: significant at 5% level; $***$: significant at 1% level.

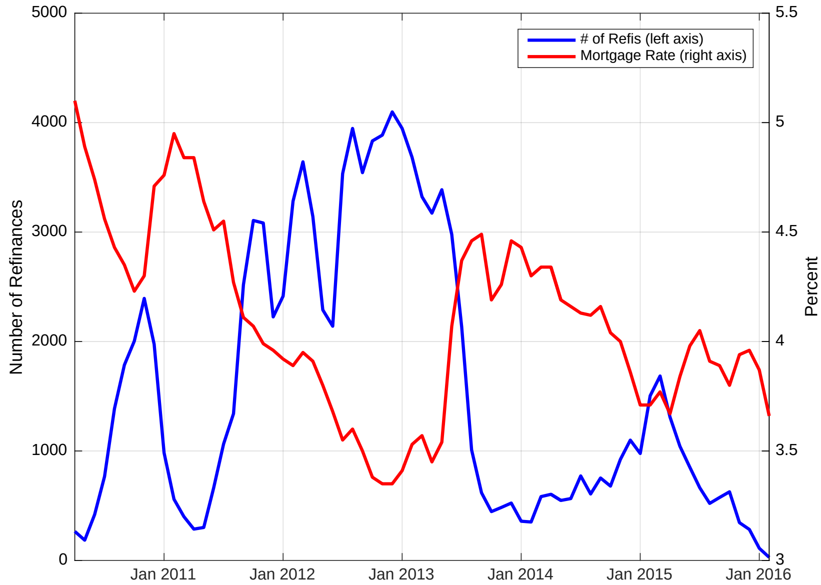

Figure 1: Refinancing activity by month in CRISM sample. We plot this series against the average interest rate on a new 30-year fixed-rate mortgage, as reported in the Freddie Mac Primary Mortgage Market Survey.

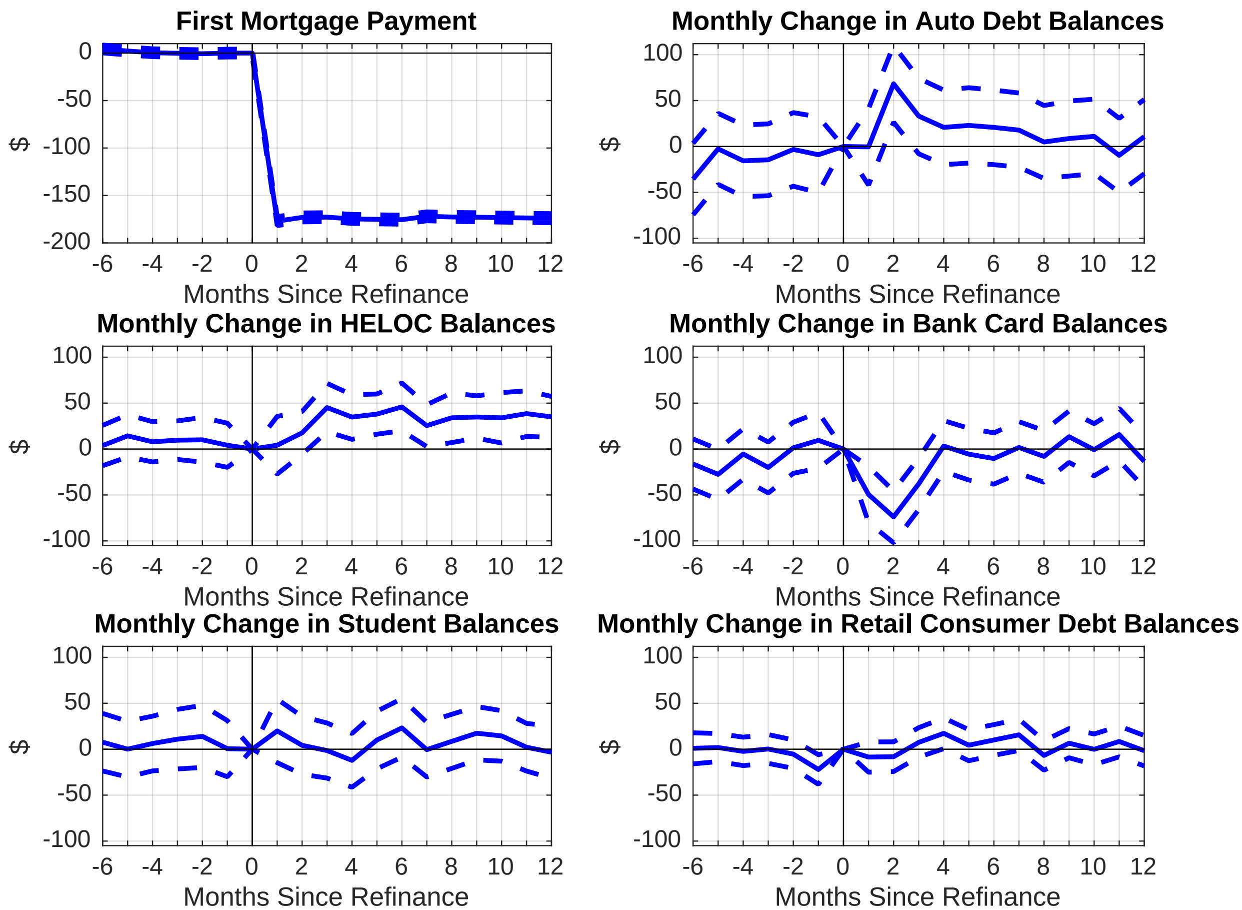

Figure 2: Event study of balance sheet around the time of a refinance, which we define as month 0. We regress the first difference of each debt balance (except in the first panel, which looks at the level of the first mortgage payment) on a full set of indicators for event time (omitting month 0) and plot the coefficients associated with those months. The regressions shown include no controls, though they are hardly changed by adding in borrower FEs, or ZIP-by-month FEs. This is a balanced panel of borrowers who are observed 6 months prior to and 12 months after the refinance. We further restrict to borrowers whose first mortgage balances did not change by more than $5,000 from month -1 to month 1. Dashed lines represent the 95% confidence intervals of the estimates, based on standard errors that are clustered at the borrower level.

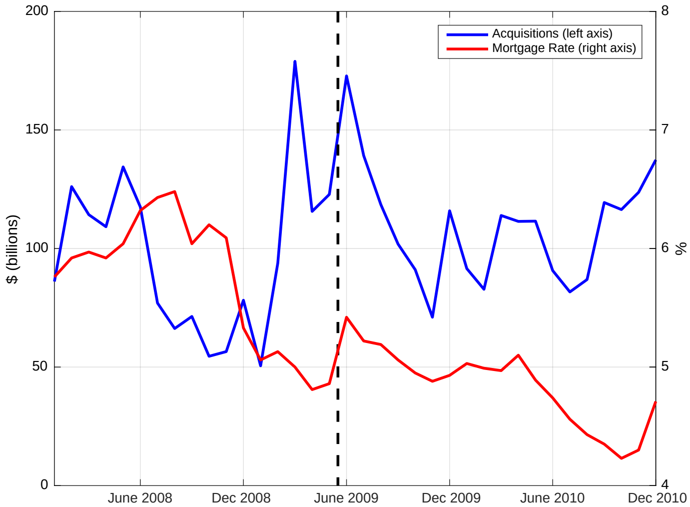

Figure 3: GSE acquisitions by month around the eligibility cutoff date, plotted against the interest rate on new conforming 30-year fixed-rate mortgages. The acquisitions data can be found at `http://www.freddiemac.com/investors/financials/monthly-volume-summaries.html` and `http://fanniemae.com/portal/about-fm/investor-relations/monthly-summary.html`. Mortgage rate is from Freddie Mac Primary Mortgage Market Survey.

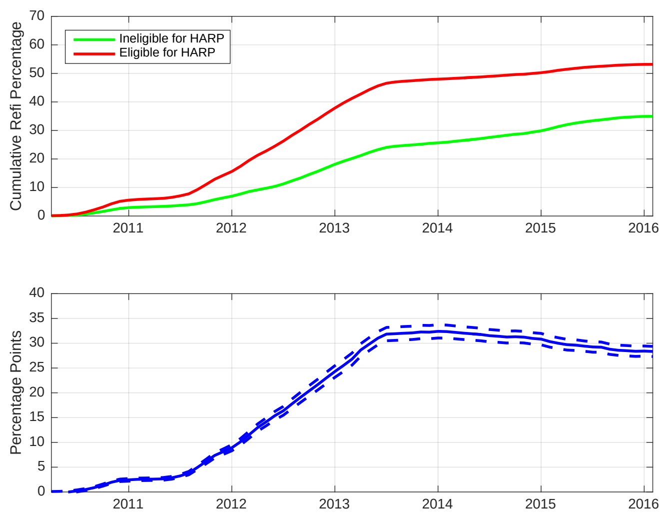

Figure 4: Effect of HARP eligibility (based on cutoff date) on refinancing. The top panel shows the raw cumulative share refinanced, split by HARP eligibility. For the bottom panel, we estimate $E[Refied_{it}|Eligible_i, X_{it}] = \sum_{201004}^{201602} \gamma_\tau (Eligible_i * I_{\{\tau=t\}}) + X_{it}\beta$, where $Refied_{it}$ is an indicator for whether borrower i refinanced at some $\tau \leq t$. We report $\{\gamma_t\}$. X_{it} contains quarter-by-ZIP FEs, guarantee lag FEs, and flexible controls for CLTV (lagged 3 months and as of March 2010), credit score (lagged 3 months and as of March 2010), credit utilization (lagged 3 months and as of March 2010), initial mortgage interest rate, remaining principal balance, and initial debt balances. Dashed lines represent the 95% confidence intervals of the estimates, based on standard errors that are clustered at the county level. Note that the bottom panel not only conditions on a set of X variables but also conditions on not having exited the sample before month t (approximately 40% of the sample exits before February 2016, as shown in Table 2).

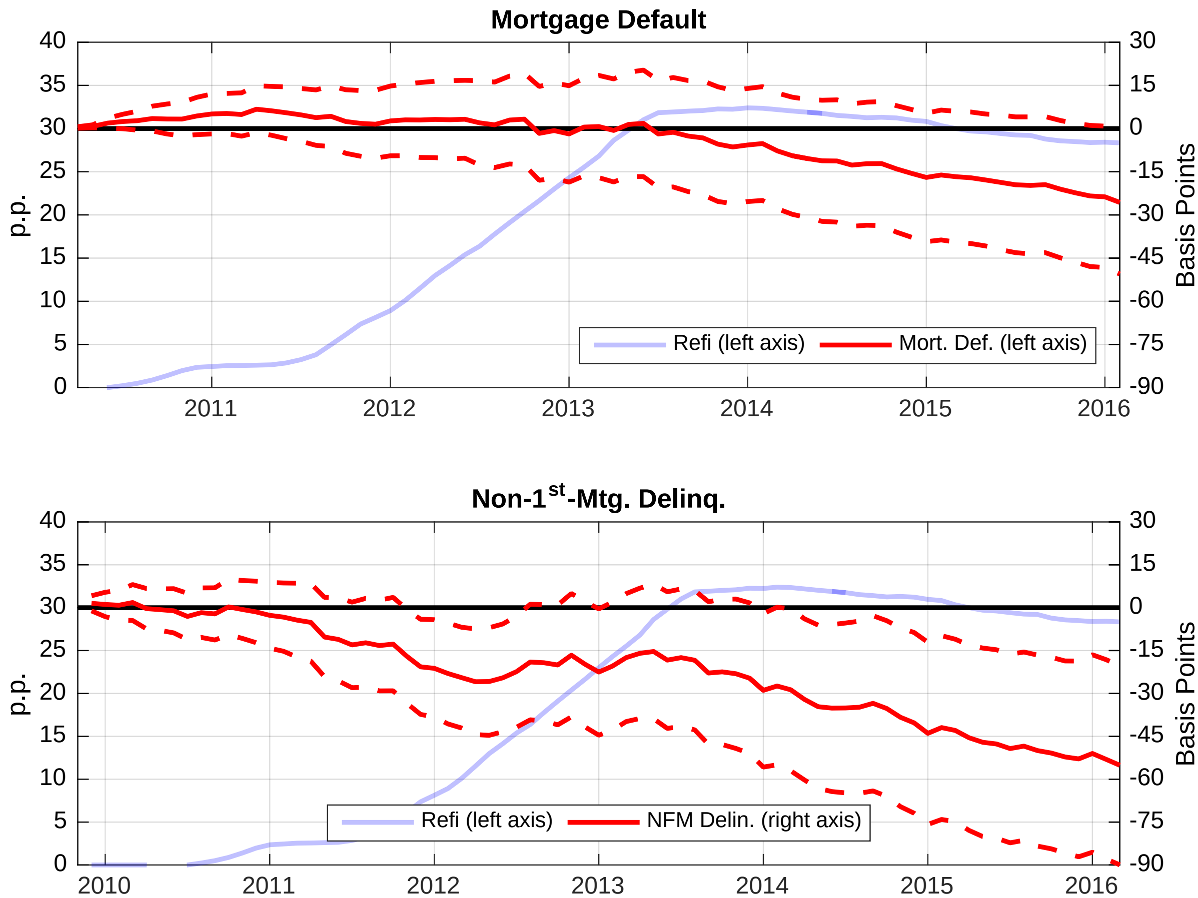

Figure 5: Dynamic reduced form effect of HARP eligibility (based on cutoff date) on mortgage default (top panel) and non-first-mortgage serious delinquency (bottom panel). The cumulative effect is shown. In particular, we estimate $E[Default_{it}|Eligible_i, X_{it}] = \sum_{t_1}^{201602} \delta_\tau(Eligible_i * I_{\{\tau=t\}}) + X_{it}\lambda$, and we report $\left\{\sum_{\tau=t_1}^{t}(\delta_\tau)\right\}$. X_{it} contains quarter-by-ZIP FEs, guarantee lag FEs, and flexible controls for CLTV (lagged 3 months and as of March 2010), credit score (lagged 3 months and as of March 2010), credit utilization (lagged 3 months and as of March 2010), and initial debt balances. The top panel also controls for mortgage interest rate and mortgage loan balance. $t_1 = 201004$ for mortgage default (top panel) and $t_1 = 200910$ for non-first-mortgage delinquency (bottom panel). Because inclusion in the sample requires the mortgage to be current on payments as of March 2010, we do not include prior months in the top panel. The cumulative effect on refinancing (dynamic first stage) is shown in the background for comparison. Dashed lines represent the 95% confidence intervals of the estimates, based on standard errors that are clustered at the county level.

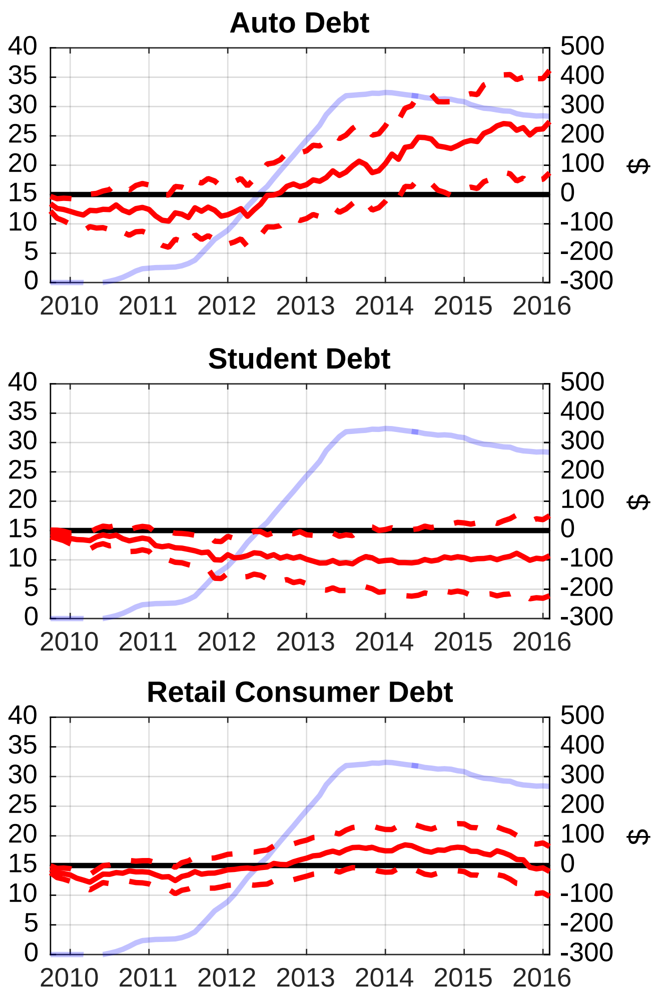

Figure 6: Dynamic reduced form effect of HARP eligibility (based on cutoff date) on accumulation of different types of debts. The cumulative effect is shown. In particular, we estimate $E[Y_{it}|Eligible_i, X_{it}] = \sum_{200910}^{201602} \delta_\tau (Eligible_i * I_{\{\tau=t\}}) + X_{it}\beta$, and we report $\left\{ \sum_{\tau=200910}^{t} (\delta_\tau) \right\}_{t=200910}^{t=201602}$. X_{it} contains quarter-by-ZIP FEs, guarantee lag FEs, and flexible controls for CLTV (lagged 3 months and as of March 2010), credit score (lagged 3 months and as of March 2010), credit utilization (lagged 3 months and as of March 2010), initial mortgage interest rate, remaining principal balance, and initial debt balances. Dashed lines represent the 95% confidence intervals of the estimates, based on standard errors that are clustered at the county level. The cumulative effect on refinancing is shown in the background for comparison.

Appendix – for Online Publication

A.1 Additional Test for Strategic Behavior by Lenders/Servicers

In Section 4.1 of the main text, we argued that there was little evidence for lenders or servicers trying to alter securitization speeds in response to the HARP announcement in March 2009. In this section, we present an additional test, based on the idea that the strategic incentives to alter securitization speeds varied depending on LTV: the HARP option is relatively more valuable when a borrower has a high LTV as opposed to a low one, as HARP is superfluous to low-LTV borrowers. This implies that even if there is no evidence of an overall shift of guarantees from after the cutoff date to before, strategic behavior may be masked if, say, servicers ensure that their high-LTV borrowers are guaranteed on time and delay the process for low-LTV borrowers. As a result, we would see borrowers in this window being guaranteed relatively quickly if their origination LTV is high.[38]

To investigate this, we look at all originations in the January-June 2009 window that are eventually guaranteed by a GSE. Table A-3 shows that there is essentially no difference in the guarantee lag (months between origination and GSE guarantee) depending on LTV bins. While some bins have statistically significant effects, there is no discernible pattern and the magnitudes are minuscule. Additionally, the low values for R^2 tell us that LTV (as well as the other variables) has essentially no role in determining guarantee lag. Similarly, if we look at the binary eligibility indicator itself as the outcome rather than the guarantee lag, there is no evidence of strategic manipulation. Column (4) shows significant results (with high-LTV loans being less likely to be eligible), but this is driven purely by the fact that high-LTV loans tended to be originated a bit later and so were less likely to be eligible. Columns (5) and (6) show that when origination cohort is controlled for, there is no meaningful difference in eligibility across the LTV distribution. The R^2 row suggests that only the origination cohort is important for determining eligibility, not LTV.

Figure A-1 provides further evidence that there was no manipulation by showing that not only do the different LTV bins have the same guarantee lag on average (as shown in Table A-3), but in fact the entire distribution is essentially identical across the different bins.

A.2 Estimating Dynamic Effects of Refinancing

In Section 5.1, we discussed how our main empirical specification imposes the restriction that the treatment effect of refinancing does not depend on how long ago the refinance occurred. In this section, we discuss how one can relax that restriction.

Recall that the treatment variable, Refied_{it}, is an indicator for whether borrower i refinanced in some month $\tau \leq t$. We instrument for this using HARP eligibility interacted with time (quarter) to

[38] One could also imagine the incentive going in the other direction, if servicers want to minimize the likelihood that their borrowers refinance (since that may lead to a loss of servicing fees). In this case as well, however, we would expect guarantee speeds to vary systematically with LTV.

estimate the causal effect. Intuitively, we find a treatment effect if the eligible and ineligible groups differ substantially in some outcome (e.g. first difference of auto debt) in the months when the eligible group is far more likely to have refinanced (basically, post-2011). However, in principle, the empirical framework allows for estimation of a more flexible response, where the treatment effect can differ depending on how long ago the borrower refinanced. This is done by simply including multiple treatment variables of the form:

$$\text{Refied}_{it}^{j_1,j_2} = \begin{cases} 1 & \text{if refinanced between } j_1 \text{ and } j_2 \text{ months ago} \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

For example, we can define two such variables, $\text{Refied}_{it}^{0,12}$ and $\text{Refied}_{it}^{13,\infty}$. By jointly estimating these, we allow the treatment effect in the first year to differ from its longer-term impact. The same set of instruments is used to estimate the causal effects: HARP eligibility interacted with time. Intuitively, since the refinancing boom occurred in 2012, the estimator will assign a strong effect to $\text{Refied}_{it}^{0,12}$ if the outcome variable is starkly different between the eligible and ineligible groups in 2012/2013, and it will assign a strong effect to $\text{Refied}_{it}^{13,\infty}$ if the difference between the groups is strong later in the sample period.

Table A-5 shows the results of estimating this more flexible specification for our main outcomes. The first takeaway is simply that the standard errors on the difference between the coefficients for the two periods (≤ 12 months, > 12 months) are very large. Intuitively, we are able to identify a statistically significant effect in the main results of the paper because we are using all of the post-refinancing-boom data (2012-2016) to identify a single coefficient—the one on Refied_{it}. The more flexible specification forces each coefficient to be estimated based on only some of the post-2011 variation, and so it is under-powered. This is why we do not focus on these results in the main text, and refrain from attempting to allow for even more flexible models.

Nonetheless, the results in Table A-5 are suggestive of some interesting effects. In particular, while the differences are not statistically significant, it seems that the reduction in defaults (both for mortgages and other debts) was not strong in the first year after a borrower refinanced but quite a bit stronger thereafter. This may be because, as discussed at length in the text, refinancing borrowers tended to be in relatively good financial health and so were not at heavy risk of default at the time of the refinance. However, the resulting payment reduction may have helped them withstand negative shocks that subsequently befell them.

Furthermore, it seems as though borrowers responded to their improved cash flow by immediately drawing on their retail consumer debt instruments, and this may be the case for auto loans as well, though the difference is not statistically significant. HELOC debt seems to show the opposite pattern, with the strong response being delayed by at least a year.

	A. Sample				B. Population			
	Mean	Median	10^{th} Pctl	90^{th} Pctl	Mean	Median	10^{th} Pctl	90^{th} Pctl
CLTV	81.6%	81.9%	67.4%	94.5%	71.2%	71.8%	36.1%	102.9%
Interest Rate	4.96%	4.88%	4.63%	5.38%	5.40%	5.38%	4.25%	6.5%
Mortgage Principal	$231k	$220k	$119k	$357k	$200k	$164k	$67k	$367k
FICO® Score	768	780	708	808	751	769	672	806
Credit Utilization	23.7%	12.5%	0.4%	68.0%	31.5%	19.2%	0.7%	83.7%
ARM? (share)	1.3%				6.0%			
30-year? (share)	91.5%				78.6%			
Junior lien? (share)	6.7%				24.1%			

Table A-1: Moments of the distributions of key observables in the CRISM dataset for borrowers who refinanced between March 2010 and February 2016. Panel A looks at our sample, while Panel B looks at a 1% random sample of all borrowers in CRISM. Statistics are calculated as of 1 month prior to the refinance.

	GSE Purchase Month								
Origination Mo.	Jan 09	Feb 09	Mar 09	Apr 09	May 09	Jun 09	Jul 09	Aug 09 +	
Jan 09	867	16,135	15,194	531	176	86	30	1,391	34,410
	(3%)	(47%)	(44%)	(2%)	(1%)	(0%)	(0%)	(4%)	
Feb 09	0	1,053	27,693	3,383	678	244	52	1,775	34,878
	(0%)	(3%)	(79%)	(10%)	(2%)	(1%)	(0%)	(5%)	
Mar 09	0	0	6,159	15,441	7,913	1,733	138	1,244	32,628
	(0%)	(0%)	(19%)	(47%)	(24%)	(5%)	(0%)	(4%)	
Apr 09	0	0	0	1,401	17,619	11,614	729	1,389	32,752
	(0%)	(0%)	(0%)	(4%)	(54%)	(35%)	(2%)	(4%)	
May 09	0	0	0	0	1,055	23,440	2,090	1,059	27,644
	(0%)	(0%)	(0%)	(0%)	(4%)	(85%)	(8%)	(4%)	
Jun 09	0	0	0	0	0	7,543	30,722	19,583	57,848
	(0%)	(0%)	(0%)	(0%)	(0%)	(13%)	(53%)	(34%)	
	867	17,188	49,046	20,756	27,441	44,660	33,761	26,441	220,160

Table A-2: Month of GSE purchase, by cohort, in the CRISM sample. Parentheses show percentage of origination cohort c purchased in month t. The vertical line between May and June indicates the eligibility cutoff.

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.:	Guarantee Lag (months)			Eligible for HARP (p.p.)		
Init. LTV (< 60% omitted)						
60-65%	-0.014	-0.021	-0.022	1.708	0.334	0.163
	(0.007)*	(0.007)***	(0.007)***	(0.178)***	(0.091)***	(0.099)*
65-70%	-0.006	-0.014	-0.022	1.344	0.105	-0.011
	(0.008)	(0.008)*	(0.008)***	(0.188)***	(0.100)	(0.110)
70-75%	0.033	0.021	0.011	1.589	-0.152	-0.259
	(0.008)***	(0.008)**	(0.008)	(0.203)***	(0.096)	(0.107)**
75-80%	0.046	0.036	0.022	-0.128	-0.659	-0.761
	(0.009)***	(0.009)***	(0.009)**	(0.214)	(0.093)***	(0.091)***
80-85%	-0.005	-0.019	-0.037	1.575	0.977	0.937
	(0.013)	(0.013)	(0.012)***	(0.246)***	(0.130)***	(0.136)***
85-90%	0.029	0.020	-0.015	-4.666	-0.875	-0.832
	(0.012)**	(0.012)*	(0.012)	(0.434)***	(0.147)***	(0.143)***
$\geq$ 90%	-0.006	-0.014	-0.038	-6.137	0.465	0.351
	(0.020)	(0.020)	(0.017)**	(0.807)***	(0.257)*	(0.218)
N (mill.)	1.5	1.5	1.5	1.5	1.5	1.5
R^2	0.000	0.008	0.046	0.001	0.678	0.679
Cohort FEs		✓	✓		✓	✓
Other controls			✓			✓

Table A-3: Regressions to test whether HARP's eligibility requirement induced servicers to guarantee high-LTV mortgages more quickly. Columns (1)-(3) use guarantee lag in months as the left-hand side variable (its sample average is 1.65 months). Columns (4)-(6) use HARP eligibility—an indicator for whether the guarantee occurred before June 2009—as the left-hand side variable (which equals 1 for 56.24% of the sample). The sample includes all GSE loans that were originated between January and June 2009. "Other controls" include flexible controls for credit score, credit utilization, interest rate and mortgage balance, as well as ZIP-code FEs. Standard errors (in parentheses) are clustered at the county level. $*$: significant at 10% level; $**$: significant at 5% level; $***$: significant at 1% level.

	Net Change				Positive	Negative
Auto	10	11	11	11	33	-21
(Std. Err.)	(2)***	(2)***	(2)***	(2)***	(2)***	(2)***
HELOC	8	6	6	5	18	-12
	(2)***	(2)***	(2)***	(2)***	(2)***	(2)***
Bank Card	-17	-16	-16	-17	37	-53
	(1)***	(1)***	(1)***	(1)***	(2)***	(2)***
Student	-1	-1	-1	-1	-2	1
	(1)	(1)	(1)	(1)	(1)	(1)
Retail Consumer Debt	2	3	3	3	14	-12
	(1)***	(1)***	(1)***	(1)***	(1)***	(1)***
Initial Bal.	✓	✓	✓	✓	✓	✓
Quarter FEs	✓	✓	✓	✓	✓	✓
ZIP-code FEs	✓	✓	✓	✓	✓	✓
Observables	✓	✓	✓	✓	✓	✓
Q-by-ZIP FEs		✓	✓	✓	✓	✓
Guar. Lag FEs			✓	✓	✓	✓
Cohort FEs				✓		
N (mill.)	13.6	13.6	13.6	13.6	13.6	13.6

Table A-4: Regression estimates of the relationship between refinancing and monthly debt accumulation by OLS. For borrower i in month t, the refinancing indicator is turned on if she has completed a refinance in some month $\tau \leq t$. Outcomes are the first difference in debt balances. "Net Change" is the simple difference, while the column labeled "Positive" censors negative changes to 0, and the column labeled "Negative" censors positive changes to 0. "Observables" include CLTV (lagged 3 months and as of March 2010), credit score (lagged 3 months and as of March 2010), credit utilization (lagged 3 months and as of March 2010), initial mortgage rate, remaining principal balance, and initial debt balances. Standard errors (in parentheses) are clustered at the county level. $*$: significant at 10% level; $**$: significant at 5% level; $***$: significant at 1% level.

Time Since Refi:	≤ 12 months	> 12 months	Diff
	A. Debt Default		
Mortgage Default	0.11	-2.57	2.69
	(2.77)	(0.79)***	(3.02)
Non-Mtg Default	-1.06	-3.48	2.42
	(3.47)	(1.31)***	(3.73)
	B. Debt Accumulation		
Auto	47	18	29
	(23)**	(10)*	(28)
HELOC	-7	31	-38
	(17)	(8)***	(22)*
Bank Card	7	-8	15
	(13)	(6)	(15)
Student	-17	2	-19
	(13)	(5)	(15)
Retail Consumer Debt	30	-6	36
	(10)***	(5)	(13)***

Table A-5: Regression estimates of the LATE of refinancing on monthly debt accumulation and the likelihood of default, allowing the effects to differ between the first 12 months since refinancing and after. Outcomes in Panel A are default indicators (effect reported in bp), and outcomes in Panel B are the first difference in debt balances (effect reported in \$). IV estimates result from instrumenting for the refinance indicators (one indicating whether borrower i refinanced in the past 12 months, the other indicating whether borrower i refinanced more than 12 months ago) with HARP eligibility interacted with a full set of quarter indicators. All regressions also control for CLTV (lagged 3 months and as of March 2010), credit score (lagged 3 months and as of March 2010), credit utilization (lagged 3 months and as of March 2010), initial mortgage rate, remaining principal balance, and initial debt balances. Standard errors (in parentheses) are clustered at the county level. $*$: significant at 10% level; $**$: significant at 5% level; $***$: significant at 1% level.

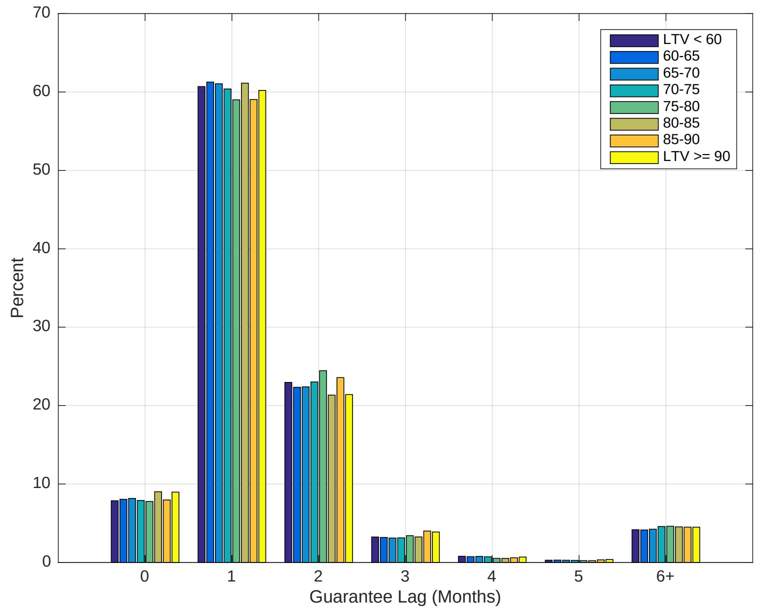

Figure A-1: Histogram of guarantee lag in CRISM. Guarantee lag is defined as the number of months between the mortgage being originated and being purchased by a GSE. We take the set of loans originated between January and June of 2009 and split them between the bins shown.

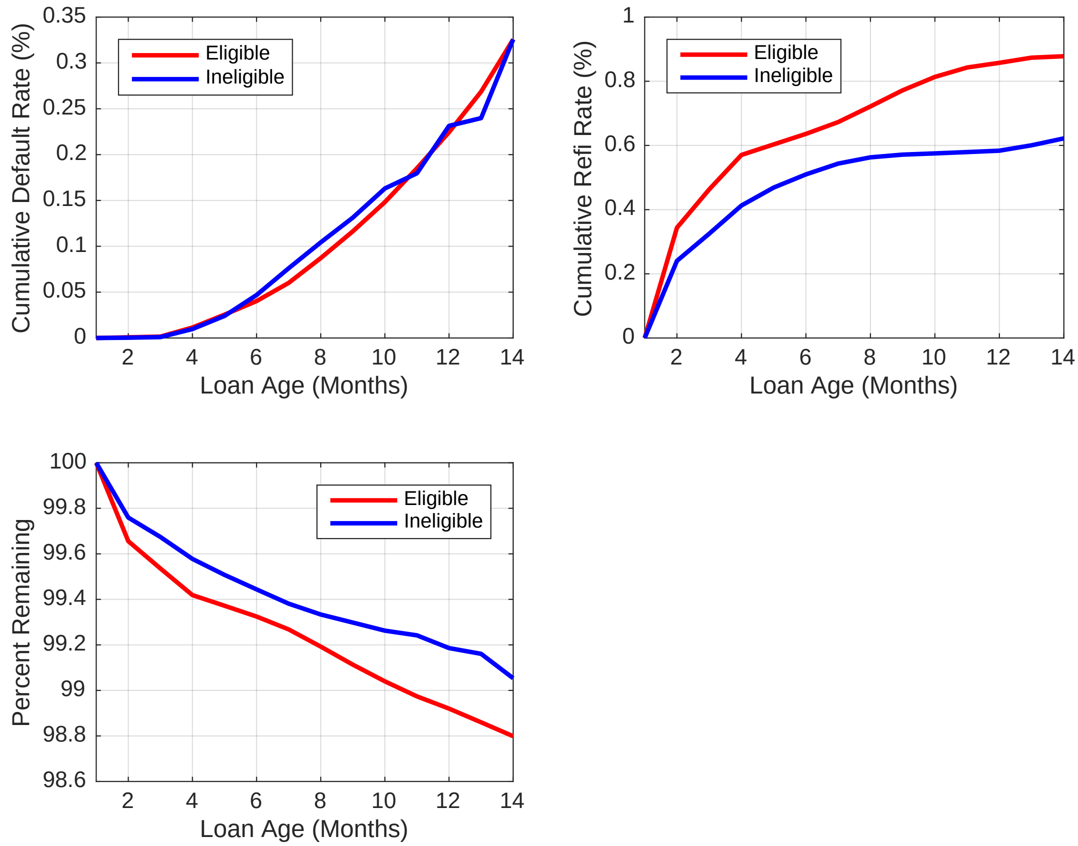

Figure A-2: Cumulative refinance and default rates for GSE loans originated between January and June 2009 with initial LTVs above 70%. This is to check whether attrition from the data prior to sample selection (which occurs in March 2010) could affect the results. Less than 1% of the loans exited from the sample in intervening months, so any induced selection is likely to be minimal. Furthermore, HARP eligibility was not very predictive of this form of attrition—default rates across the eligible and ineligible groups were nearly identical. While the eligible group was more likely to refinance in this period (and thus drop from the sample), this is entirely due to time effects (a regression that controls for month effects finds no significant predictive value of HARP eligibility on the probability of refinancing in this period). Thus, our decision to wait ≈ 1 year before selecting our sample seems unlikely to have caused any meaningful selection problems due to attrition.