American Economic Journal: Macroeconomics 2021, 13(2): 254–291
https://doi.org/10.1257/mac.20180116

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

By Joshua Abel and Andreas Fuster*

We use quasi-random access to the Home Affordable Refinance Program (HARP) to identify the causal effect of refinancing into a lower-rate mortgage on borrower balance sheet outcomes. Refinancing substantially reduces borrower default rates on mortgages and other debt. Refinancing also causes borrowers to expand their use of debt instruments, such as auto loans, home equity lines, and other consumer debts that are proxies for spending. Borrowers that appear more constrained ex ante grow these debts more strongly after refinancing but also pay down credit card balances by more. These borrowers also have lower take-up of the refinancing opportunity. (*JEL* G51, G21, E52)

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 lower-rate mortgages. 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.

*Abel: Analysis Group, 111 Huntington Avenue, Boston, MA 02199 (email: joshuadabel@alumni.harvard.edu); Fuster: Swiss National Bank, Börsenstrasse 15, 8022 Zurich, Switzerland (email: andreas.fuster@gmail.com). Giorgio Primiceri was coeditor for this article. We thank Valentin Bolotnyy, John Campbell, Gabriel Chodorow-Reich, Michael Fratantoni, Edward Glaeser, John Mondragon, Joseph Tracy, Paul Willen, three anonymous referees, and seminar audiences at Harvard University, FRB Chicago, the Federal Reserve Board, George Washington University, JPMorgan Chase Institute, ASSA 2018 Meeting (Philadelphia), EEA (Cologne), SSES (St. Gallen), and VfS Committee on Monetary Theory and Policy (Zurich) for helpful comments. Abel was a graduate fellow at the Federal Reserve Bank of Boston, and Fuster was an employee at the Federal Reserve Bank of New York while most of this paper was written. The views expressed in this paper are solely those of the authors and not necessarily those of the Federal Reserve Banks of Boston and New York, the Federal Reserve System, or the Swiss National Bank.

†Go to https://doi.org/10.1257/mac.20180116 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

[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).

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 2016. 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 percent, 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. 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 be better at budgeting, making a default less likely. Another element of selection comes from the fact that refinancing requires a borrower to fulfill 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 refinancing activity surges later in the sample period, suggesting that the two groups would not otherwise have evolved differently.

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 percent 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 25 percent of the savings resulting from the decreased mortgage payments. This net effect combines larger increases (new debts) corresponding to about 55 percent of mortgage payment savings and decreases (pay-downs of existing debt) of about 30 percent 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 30 percent.

Our analysis puts a particular focus on heterogeneity in the effects of refinancing across different subgroups of the sample. To do so, we rely on summary indicators of an individual's financial health and liquidity, specifically 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 the strong reductions in mortgage default are almost entirely concentrated in borrowers who appear to be struggling: those with low credit scores, high utilization rates of existing revolving credit lines, and high CLTVs. These borrowers reduce their mortgage default rates around 50 percent. For borrowers in stronger financial positions, mortgage default is not a relevant margin: they default at low rates and experience small treatment effects. Similarly, reductions in non-mortgage default are concentrated in the constrained borrowers.

There is also substantial heterogeneity across borrower types in how refinancing affects the accumulation and pay-downs of different kinds of debt. We find, for instance, that low credit score borrowers more strongly increase their borrowing through auto loans, HELOCs, and retail consumer debt, which points toward a stronger spending impact for these borrowers. At the same time, they more strongly pay down credit card debt than borrowers with higher credit scores. Similarly, borrowers with high credit utilization prior to refinancing increase their auto debt by more, but they also use the newly available cash flow to aggressively pay off credit card debt. We find little heterogeneity in debt responses across borrowers depending on their CLTV. Overall, our results highlight that conclusions on heterogeneous responses to refinancing can vary depending on what debt types are studied, and that it is thus important to study multiple dimensions to get a full picture.

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. Within the sample of HARP-eligible borrowers, take-up is highest in areas with higher incomes, while there is no strong relationship with either local education levels or mortgage market concentration. Furthermore, individual indicators of good financial health—high credit score, high levels of untapped credit, low CLTV—predict a higher likelihood of take-up. Note, however, that this does not imply that these borrowers also increased their refinancing the most *as a result of HARP*. In fact, HARP eligibility had a stronger effect on the refinancing propensity of more financially constrained borrowers (especially those with high CLTV).

Most closely related to our work are papers by Karamon, McManus, and Zhu (2016) and Ehrlich and Perry (2015). Karamon, McManus, and Zhu (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 they do not study other outcomes.[3]

Agarwal et al. (2017b) also study HARP, comparing GSE-securitized (and therefore HARP-eligible) loans to privately securitized/nonagency (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

[2] In contrast to Karamon, McManus, and Zhu (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, McManus, and Zhu (2016) but focus primarily on the "intensive margin," comparing HARP refinances with payment reductions of different sizes. 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.

complementary, and it is 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 (2019); Beraja et al. (2019); Berger et al. (2018); Di Maggio, Kermani, and Palmer (2016); Eichenbaum, Rebelo, and Wong (2018); Greenwald (2018); or Wong (2019).[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, Moll, and Violante (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, Rosen, and Yao 2015; Andersen et al. 2019; Bond et al. 2017; Campbell 2006; Johnson, Meier, and Toubia 2018; or Keys, Pope, and Pope 2016). Finally, our results also inform the recent literature on mortgage design. Campbell (2013), Eberly and Krishnamurthy (2014), and Guren, Krishnamurthy, and McQuade (2018) all make the point that mortgages that automatically lower payments in downturns (such as ARMs, assuming 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, such as HARP, that make payment reductions easier to achieve for constrained households in downturns.

I. 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

[5] A number of papers study the effects of equity withdrawal or “cash-out” refinancing, including Hurst and Stafford (2004); Bhutta and Keys (2016); and Chen, Michaux, and Roussanov (2020). The HARP refinances we study involve at most very limited cash-out.

[6] Remy, Lucas, and Moore (2011) point out that, of course, this is not free—it is a transfer from mortgage investors to these households. However, since a nontrivial 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, Giglio, and Pathak 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 Agarwal et al. (2017a), Ganong and Noel (2018), and Scharlemann and Shore (2016) for studies of the effects of this program.

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 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 percent 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 percent—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 percent. In September 2009, this was raised to 125 percent, and in June 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 III 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 take-up was initially weak, so it was significantly reformed in mid-2012—the so-called "HARP 2.0." In addition to eliminating the LTV cap, HARP 2.0

[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 percent LTV at the time of the refinance. Refinancing with an LTV above 80 percent 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. (2019) 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.

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 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 were refinanced through HARP before its conclusion at the end of 2018, 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.

II. Data

Our analysis relies on "Credit Risk Insights Servicing McDash" (CRISM 2019), 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.

A. *CRISM*

CRISM covers about 60 percent 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.

[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).

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 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 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 (2019). For 83 percent of borrowers we have zip code-level HPIs, while for the rest we use either county-, MSA-, or state-level HPIs. Selection into our sample is 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.

B. *Sample Selection Criteria*

HARP's design—along with our empirical strategy—dictates three main sample selection criteria.[17] We select borrowers who: (i) have a fixed-rate mortgage[18] guaranteed by a GSE (Fannie Mae or Freddie Mac); (ii) originated that mortgage between January and June 2009; and (iii) were current on payments and had an LTV of at least 80 percent on that mortgage in March 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.[19]

[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.

[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 six 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 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] Borrowers with ARMs could refinance through HARP, but only about 1 percent of originations over January–June 2009 were ARMs. We drop them in order to focus exclusively on FRMs, though results are essentially unchanged if ARMs are included (as in an earlier draft).

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

Second, our sample only includes borrowers who originated a mortgage between January and June 2009. Our key instrument (discussed in detail in Section III) 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 III.

Third, we only include borrowers who were current on their payments and whose estimated updated LTV was above 80 percent in March 2010. As previously discussed, HARP was targeted at borrowers with LTVs above 80 percent. 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 2010—to measure LTV and decide whether to include the borrower in the sample. Note that March 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 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 rerun 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.

C. *Summary Statistics*

Table 1 provides summary statistics and compares our sample to three different populations: all GSE loans with high LTVs in March 2010; all loans originated in our window, January–June 2009; and all loans outstanding. The summary statistics are as of March 2010, the observation month when inclusion in our sample is determined, and as before, we restrict ourselves to borrowers who were current on their payments then.[20]

Relative to the high-LTV GSE 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 one year, loan balances tend to be

[20] Online Appendix Table A-1 presents the geographic breakdown of our sample, showing that no single state accounts for more than 7.2 percent of observations and that home price growth during the period of 2009–2010 varied substantially across the most common states.

Table 1—Moments of the Distributions of Key Observables for Borrowers with FRMs Whose Payments were Current as of March 2010

	Panel A. Sample				*Panel B. High-LTV GSE loans*			
	Mean	Median	10th pctl.	90th pctl.	Mean	Median	10th pctl.	90th pctl.
CLTV	87.9%	85.3%	80.8%	96.3%	115.6%	97.5%	82.6%	137.5%
Orig. LTV	84.1%	80.0%	77.9%	93.5%	86.5%	80.0%	69.9%	95.0%
Interest rate	4.95%	4.88%	4.50%	5.38%	6.03%	6.00%	5.00%	6.88%
Mortgage principal	$225k	$212k	$113k	$360k	$184k	$166k	$80k	$319k
Credit score	762	775	698	809	717	741	596	803
Credit utilization	24.6%	12.8%	0.1%	71.7%	40.1%	32.0%	0.6%	93.6%
30-year? (share)	92.5%				93.4%			
Has junior lien? (share)	9.4%				26.7%			
Purchase mortgage? (share)	27.6%				47.2%			
Ever default? (share)	3.6%				11.6%			
FHA/VA? (share)	-				-			
Observations	219,407				5.83 million			
	Panel C. All January 2009–June 2009 origs.				*Panel D. All outstanding loans*			
	Mean	Median	10th pctl.	90th pctl.	Mean	Median	10th pctl.	90th pctl.
CLTV	75.4%	77.3%	44.5%	100.2%	87.6%	75.9%	29.4%	112.7%
Orig. LTV	73.5%	76.4%	44.0%	98.2%	82.2%	75.8%	37.5%	97.5%
Interest rate	4.98%	4.88%	4.50%	5.50%	5.96%	5.88%	4.88%	7.13%
Mortgage principal	$213k	$188k	$91k	$380k	$152k	$124k	$42k	$297k
Credit score	744	768	650	810	722	750	602	807
Credit utilization	29.3%	15.4%	0.3%	84.7%	36.8%	25.6%	0.4%	92.0%
30-year? (share)	85.3%				73.2%			
Has junior lien (share)	13.5%				28.5%			
Purchase mortgage (share)	26.4%				37.2%			
Ever default (share)	5.7%				8.9%			
FHA/VA? (share)	30.1%				15.3%			
Observations	2.98 million				29.61 million			

Notes: Summary statistics as of March 2010. Panel A uses our sample: borrowers with GSE-guaranteed FRMs with LTVs above 80 percent, originated in the first half of 2009. Panel B uses a 5 percent sample of GSE FRMs with LTVs above 80 percent in March 2010. Panel C uses a 5 percent sample of all borrowers originating FRMs in the first half of 2009 and outstanding in March 2010. Panel D uses a 5 percent sample of all FRMs outstanding in March 2010. Credit utilization is measured as the sum of HELOC and bank card balances divided by the sum of HELOC and bank card limits. "Ever default" is an indicator for whether a loan ever entered 90+ days delinquency over its lifetime (after March 2010). Observations for panels B–D result from multiplying by 20 the number of loans used to generate the statistics.

Source: Data is from CRISM.

larger. Similarly, because they 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 likely due to the tightening in credit standards following the onset of the financial crisis.

Comparing our sample with panel C (all originations over January–June 2009), we mechanically have higher LTVs, as high LTV is one of our selection criteria. On the other characteristics our sample once again appears to feature relatively strong borrowers (i.e., higher credit scores, lower overall default rate), though differences are much smaller than with panel B. In this case, the differences are largely due to the fact that a significant share of mortgages originated in this window (30 percent) are FHA/VA loans, which tend to be riskier but are excluded from our sample due to our empirical design. Comparing to all outstanding loans (panel D), the conclusions are similar to the ones for panel B, except that panel D is not restricted to high-LTV borrowers and thus shows an average LTV closer to the one in our sample.

Based on panel B, it is likely that results from our sample understate the impact of a 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.[21]

D. *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. (2019), 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 four 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 percent of originations in McDash) but the mortgage is in the same zip code as the terminated loan.[22]

[21] Online Appendix Table A-2 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 score are much smaller than in Table 1.

[22] For about 25 percent 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 six months, as CRISM tracks borrowers for six 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 six months following the refinance, and their behavior is extremely similar, allaying concerns that the uncensored refinancers that we track are not representative.

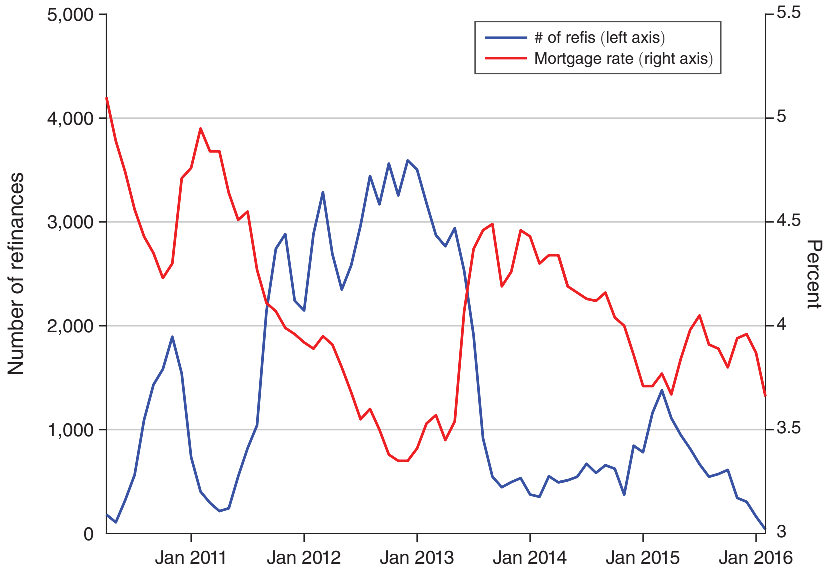


Figure 1. Refinancing Activity and Mortgage Rates

Notes: This figure plots the number of refinances by month in CRISM sample against the average interest rate on a new 30-year fixed-rate mortgage, as reported in the Freddie Mac (2020) Primary Mortgage Market Survey.

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 219,407 borrowers in our sample, 97,670 ($\approx$45%) 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 percent, on average. The remaining panels of the figure show an event study of changes in debt balances around the time of a non-cashout refinance.[23] This provides initial evidence that borrowers respond to

[23] Specifically, we focus on borrowers whose first mortgage balances do not change by more than $5,000 from month -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 for which balances do not change by more than $5,000 but is not predictive of refinances with larger balance shifts.

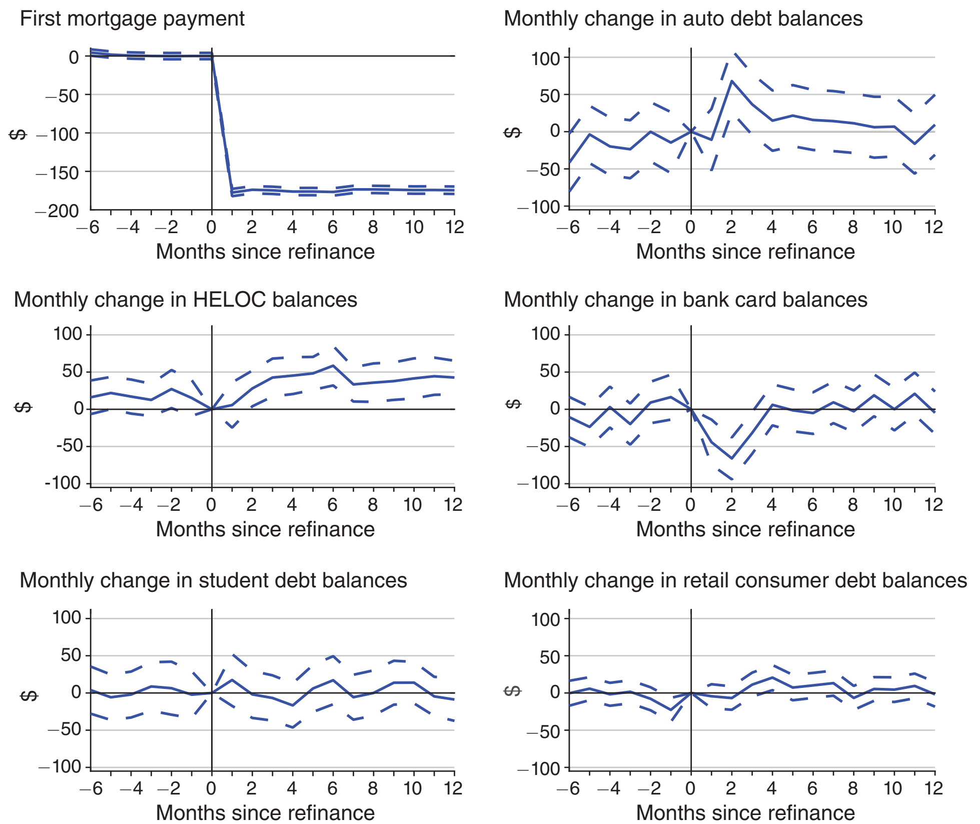


Figure 2. Event Study of Balance Sheet around the Time of a Refinance

Notes: Month of refinance corresponds to 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 code-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 percent confidence intervals of the estimates, based on standard errors that are clustered at the borrower level.

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 initially 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.

III. 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 then discuss how different sets of controls affect the strength of the identifying assumptions and conclude by presenting our workhorse econometric model.

A. *Was There Nonrandom Sorting around the Cutoff Date?*

HARP was unveiled in March 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 those 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

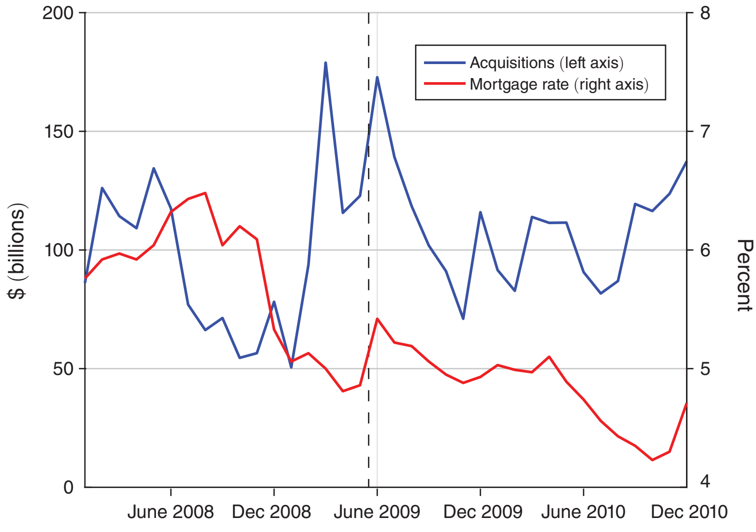


Figure 3. GSE Acquisitions and Mortgage Rates around the Eligibility Cutoff Date

Sources: Monthly GSE acquisitions data are from Fannie Mae (2019) and Freddie Mac (2019). Interest rate on new conforming 30-year fixed-rate mortgages is from Freddie Mac (2020) Primary Mortgage Market Survey.

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, McManus, and Zhu's (2016) 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 online Appendix A 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 we find no systematic evidence for this.

B. *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 six months of originations comprises 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, credit 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 the differences

[24] Online Appendix Table A-3 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 one 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.

Table 2—Comparison of HARP-Eligible versus HARP-Ineligible Borrowers in Our Sample

	Mean: Eligible ($N = 115{,}130$)	Mean: Ineligible ($N = 104{,}277$)	Ineligible − Eligible	SE on Diff.
CLTV (%)	87.3	88.5	1.2	0.1
Orig. CLTV (%)	86.6	87.4	0.8	0.1
Credit score	763.6	759.6	−4.0	0.2
Interest rate (%)	4.98	4.93	−0.05	0.00
Credit utilization (%)	24.5	24.7	0.2	0.1
First mortgage balance (1,000$)	225.9	225.1	−0.8	0.4
All other debt balances (1,000$)	24.3	25.0	0.8	0.1
Purchase mortgage (%)	25.8	29.6	3.8	0.2
Refinanced (%)	53.17	34.96	−18.21	0.21
Defaulted (%)	3.53	3.64	0.11	0.08
Servicing transferred from McDash (%)	6.15	8.22	2.07	0.11
Refinanced out of McDash (%)	11.37	11.82	0.45	0.14
Non-refi prepay (%)	21.10	19.27	−1.83	0.17
Active (%)	57.85	57.05	−0.79	0.21

Notes: The 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 2010. The 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.

are small.[25] The similarity of the two groups on observable characteristics is reassuring that they are similar on unobservable characteristics as well, lending validity to the eligibility instrument.[26]

C. *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 pretreatment.[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

[25] A notable difference is that the eligible subsample includes fewer purchase loans. This reflects the refinancing wave that was strongest in early 2009, after the Fed's initial asset purchase announcement, and also seasonality in home purchase activity. We control for all of these observable characteristics in the analysis below, but these controls are not very important, since the two groups are so similar. The one exception is that controlling for the 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. Online Appendix Figure A-1 shows that this is not an issue, as only about 1 percent of borrowers left the sample within 14 months of originating their mortgage, and this attrition was quite balanced across the eligible and ineligible groups.

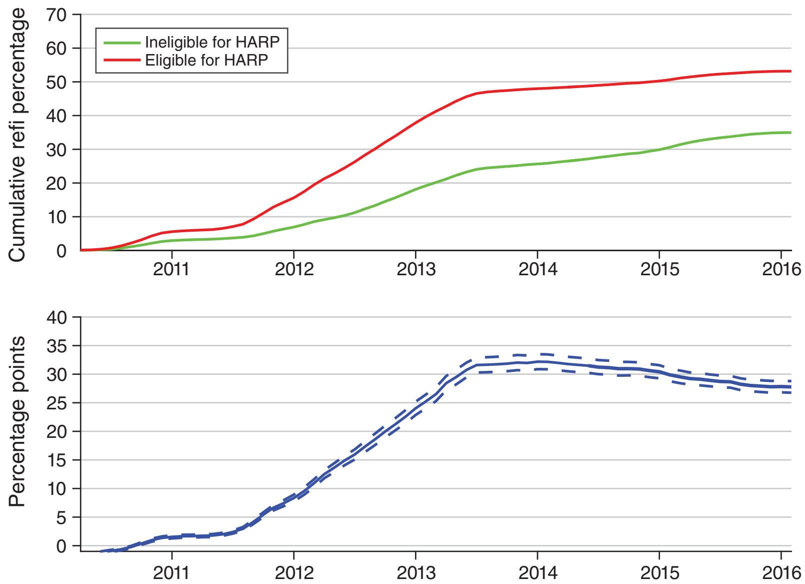


Figure 4. Effect of HARP Eligibility (Based on Cutoff Date) on Refinancing

Notes: The top panel shows the raw cumulative share refinanced, split by HARP eligibility. For the bottom panel, we estimate $E\left[Refied_{it} \middle| Eligible_i, X_{it}\right] = \sum_{2010:3}^{2016:2} \gamma_\tau \left(Eligible_i \times I_{\{\tau=t\}}\right) + X_{it}\beta$, 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 code FEs, guarantee lag FEs, and ten equally sized bins for each of: CLTV (lagged three months and at origination), credit score (lagged three months and at origination), credit utilization (lagged three months and at origination), initial mortgage interest rate, remaining principal balance, and initial debt balances. We also include indicators for mortgage "purpose" (e.g., purchase, cash-out refi, etc.). Dashed lines represent the 95 percent 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 percent of the sample exits before February 2016, as shown in Table 2).

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:

$$(1) \qquad E\left[Refied_{it} \middle| t, Eligible_i, X_{it}\right] = \sum_{\tau=2010:3}^{2016:2} \gamma_\tau \left(Eligible_i \times I_{\{t=\tau\}}\right) + X_{it}\theta,$$

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 code fixed effects (FEs). The bottom panel of Figure 4 plots $\{\gamma_t\}_{t=2010:3}^{2016:2}$, 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 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:

$$E\left[Y_{it} | t, Eligible_i, X_{it}\right] = \sum_{\tau=2009:10}^{2016:2} \delta_\tau \left(Eligible_i \times I_{\{t=\tau\}}\right) + X_{it}\lambda, \tag{2}$$

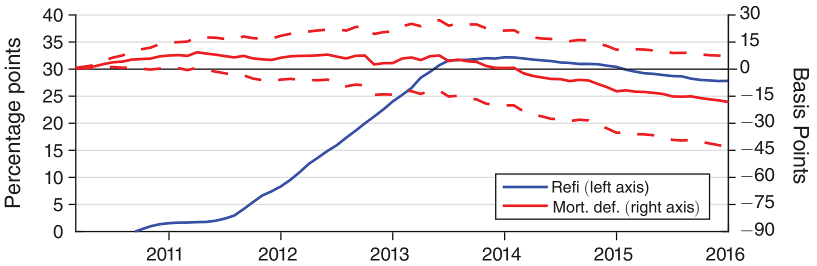
where Y_{it} is the outcome of interest—either a default indicator or a first difference of debt.[28] Note that the sample begins in October 2009 rather than March 2010. For the first stage, there is mechanically no effect prior to March 2010 because inclusion in the sample required the original loan 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, $\left\{\sum_{\tau=2009:10}^{t} \delta_\tau\right\}_{t=2009:10}^{2016:2}$, 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.

D. *Identification Assumptions*

The previous subsections have argued that HARP-eligible borrowers are comparable to those that are ineligible, but in a regression framework, they must be

[28] We trim observations from a regression if the balance in that debt category in that month or the one prior is greater than the ninety-ninth percentile of nonzero balances in the sample. We have also winsorized based on this criteria, and we have trimmed and winsorized extreme *changes* as well, and results are similar.



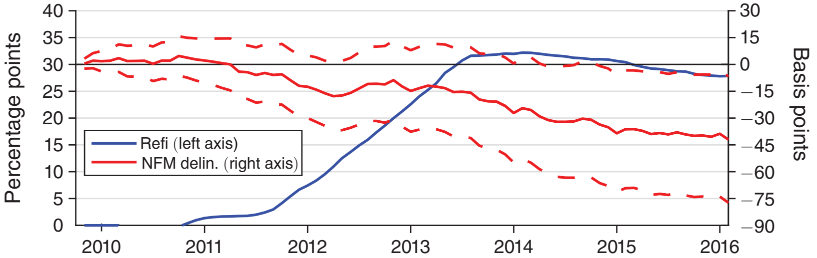


FIGURE 5. DYNAMIC REDUCED-FORM EFFECT OF HARP ELIGIBILITY ON MORTGAGE DEFAULT AND NON-FIRST-MORTGAGE SERIOUS DELINQUENCY

Notes: The top panel shows cumulative effect of HARP eligibility (based on cutoff date) on mortgage default; bottom panel shows the effect on non-first-mortgage serious delinquency. We estimate $E[Default_{it}|Eligible_i, X_{it}] = \sum_{t_1}^{2016:2} \delta_\tau (Eligible_i \times I_{\{\tau=t\}}) + X_{it}\lambda$, and we report $\{\sum_{\tau=t_1}^{t} (\delta_\tau)\}$. X_{it} contains quarter-by-zip code FEs, guarantee lag FEs, and ten equally sized bins for each of: CLTV (lagged three months and at origination), credit score (lagged three months and at origination), credit utilization (lagged three months and at origination), and initial debt balances. We also include indicators for mortgage "purpose" (e.g., purchase, cash-out refi, etc.). The top panel also controls for mortgage interest rate and mortgage loan balance. $t_1 = 2010{:}3$ for mortgage default (top panel) and $t_1 = 2009{:}10$ 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 blue for comparison. Dashed lines represent the 95 percent confidence intervals of the estimates, based on standard errors that are clustered at the county level.

comparable only conditional on the other covariates included. Here we outline different sets of covariates and discuss the different 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 three months and at origination), credit score (lagged three months and at origination), credit utilization (lagged three months and at origination), the loan purpose of the first mortgage (purchase, cash-out refinance, etc.), remaining principal balance on the first mortgage, initial interest rate on the mortgage, and initial balances of the other debt categories. This means the identifying assumption for these regressions, which will be denoted by a (1) in most of the tables in the results sections, follows.

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.*

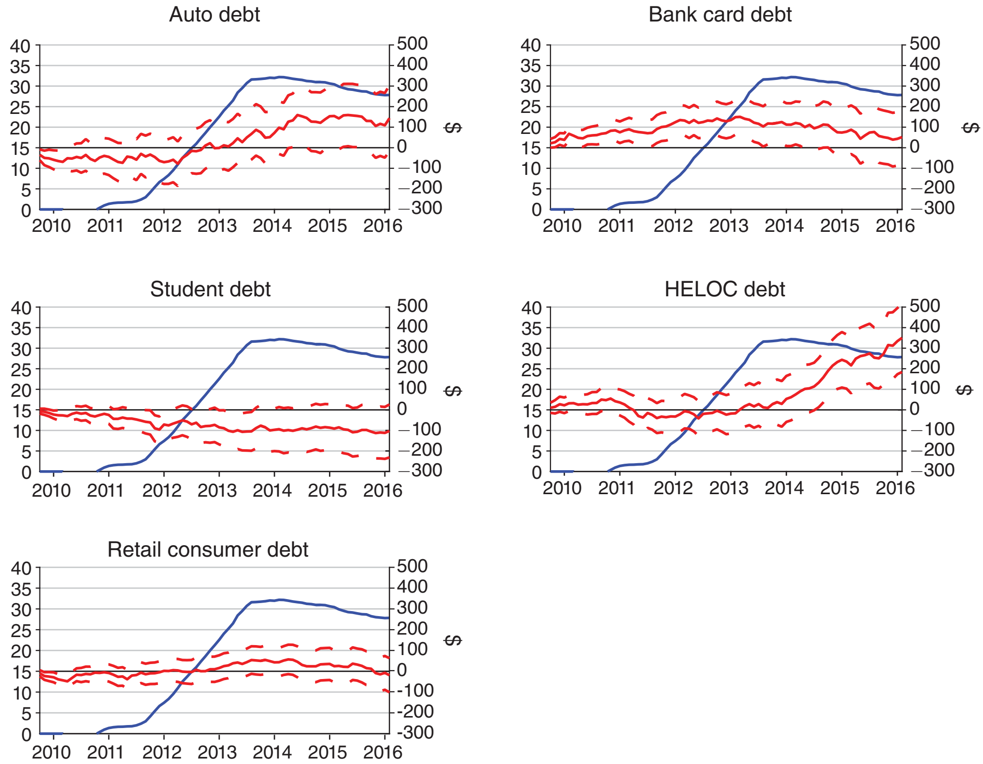


FIGURE 6. DYNAMIC REDUCED-FORM EFFECT OF HARP ELIGIBILITY ON ACCUMULATION OF DIFFERENT TYPES OF DEBTS

Notes: Cumulative effects of HARP eligibility (based on cutoff date) are shown. We estimate $E\left[Y_{it} \middle| Eligible_i, X_{it}\right] = \sum_{2009:10}^{2016:2} \delta_\tau \left(Eligible_i \times I_{\{\tau=t\}}\right) + X_{it}\beta$, and we report $\left\{\sum_{\tau=2009:10}^{t} \left(\delta_\tau\right)\right\}_{t=2009:10}^{2016:2}$. X_{it} contains quarter-by-zip code FEs, guarantee lag FEs, and ten equally sized bins for each of: CLTV (lagged three months and at origination), credit score (lagged three months and at origination), credit utilization (lagged three months and at origination), initial mortgage interest rate, remaining principal balance, and initial debt balances. We also include indicators for mortgage "purpose" (e.g., purchase, cash-out refi, etc.). Dashed lines represent the 95 percent confidence intervals of the estimates, based on standard errors that are clustered at the county level. The cumulative effect on refinancing (left y-axis) is shown in blue for comparison.

We have provided evidence in support of that assumption throughout Section III; 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.[29] In this specification, which will be denoted by

[29] 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.

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 a 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, two 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. However, our discussion 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.[30]

E. *Econometric Model*

To quantify the relationship between refinancing and an outcome of interest, Y, we estimate

$$E\left[Y_{it}|\widetilde{Refied}_{it}, X_{it}\right] = \beta\widetilde{Refied}_{it} + X_{it}\kappa, \tag{3}$$

where 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). We compare OLS estimates to those from 2SLS (IV). In the OLS specifications, $\widetilde{Refied}_{it}$ indicates whether borrower i has refinanced in or before month 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:

$$E\left[Refied_{it}|t, Eligible_i, X_{it}\right] = \sum_{\tau=q_1}^{2016:\mathrm{I}} \gamma_\tau\left(Eligible_i \times I_{t=\tau}\right) + X_{it}\theta. \tag{4}$$

In the 2SLS results, $\widetilde{Refied}_{it}$ 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

[30] In our sample, only approximately 25 percent 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.

Figure 4, this basically means we are identified off of post-2011 differences between the groups.

When we study default outcomes, Y_{it} corresponds to a binary indicator for default (and equation (3) thus is a linear probability model). We follow the literature in defining a mortgage default to have occurred when a borrower reaches 90+ days delinquent, and we say a borrower has reached serious delinquency on non-mortgage debt if she has three consecutive months with delinquent balances on debts other than the first mortgage. For those outcomes, we censor borrowers after the hazard occurs for the first time or when they exit the sample due to a nonrefinance prepayment or servicing transfer. When we study debt accumulation, Y_{it} is the first difference of the debt balance in question, and we censor after a borrower exits due to a nonrefinance prepayment or servicing transfer. All regressions cover October 2009 through February 2016, with the exception of the mortgage default regressions; inclusion in our sample requires that a borrower not default prior to March 2010, so we begin the estimation sample for that outcome in that month, rather than October 2009. Thus, q_1 in equation (4) is 2009:IV for all regressions other than for mortgage default, for which $q_1 = 2010$:I.

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

(i) quarter fixed effects (FEs), zip code FEs, observables;[31]

(ii) + zip code × quarter FEs;

(iii) + guarantee lag FEs;

(iv) + cohort FEs.

Specifications (i) and (iv) correspond to Identifying Assumptions 1 and 2, respectively. We show specifications (ii) and (iii) 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.

We also study heterogeneity in treatment effects. To do so, we create an indicator for an observation being above the median C, where C is either credit score, credit utilization, or CLTV, and then estimate coefficients on $Refied_{it}$ and $Refied_{it} \times HighC_{it}$. We instrument for these two variables with $Eligible_i$, $Eligible_i \times HighC_{it}$, and the joint interaction of these variables with a full set of quarter dummies.

A few notes are in order regarding this 2SLS model. First, as usual in IV settings, β in equation (3) corresponds to a local average treatment effect (LATE) that is specific to people who are "compliers" in the sense defined by Angrist, Imbens, and Rubin (1996)—borrowers who would refinance if eligible for HARP and not

[31] We partition observables into ten equally sized bins and allow for an unconstrained step function of those bins. We have also run specifications allowing for those step functions to vary arbitrarily across different years. In those specifications, the "average" treatment effect regressions are extremely similar to what we present in this paper, but the standard errors on the heterogeneity regressions increase on the order of 100 percent, and coefficients can be quite sensitive to functional forms of the controls (e.g., number of bins, which coefficients vary over time, etc.). For this reason, we present results with coefficients that do not vary over time.

otherwise. This is a beneficial feature of our approach, as the response of the group of borrowers who are sensitive to the option of refinancing is an important one for policymakers to consider, arguably more so than the average response over the entire population.

Second, one may be concerned that while HARP eligibility may be as good as randomly assigned, the timing of a refinance is still endogenous. Importantly, however, this is not a threat to identification: random assignment of eligibility still means that differences between the two groups can be causally attributed to refinancing through HARP. If certain types of borrowers refinance more quickly, this could mean that the LATE being recovered is weighted more toward that group's specific LATE, as they spend more time in treatment. Alternatively, some borrowers refinance at more opportune times than others and obtain larger rate reductions, and may therefore respond more strongly. However, this is not problematic from the perspective of identification, since the implicit control group for the treated borrowers are ones that would have done the same (i.e., refinanced at the same time), *had they been eligible for HARP*. We also show below that the average payment reductions from refinancing do not vary much across borrowers with high or low credit score or credit utilization (both of which plausibly correlate with borrower sophistication), so that this particular form of heterogeneity may not be too pronounced in our application.

Third, note that equation (3) 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 online Appendix B we explore more flexible specifications that allow the effects to vary with time since the refinance, but those specifications are largely underpowered, so we do not feature them in our main results.

IV. Effects of Refinancing on Mortgage Default

This section studies the impact of refinancing on mortgage default. We find a strong reduction in the mortgage default hazard, in excess of the implied sensitivity of default to payment size found in previous studies of ARMs. We also document that effects are much larger for borrowers that are more financially constrained.

Average Effects.—Panel A of Table 3 shows the results for different sets of controls 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 larger point estimate.[32] The estimated reduction from the other specifications is centered around 2.5 bp, or roughly 40 percent of the sample average.[33] The OLS results are substantially larger: the reductions range from 4.1–4.3 bp, or about 70 percent. This

[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 6.6 bp. However, we prefer a weighted average of monthly default rates, where the weight of month τ is equal to the number of borrowers with $Refied_{i\tau} = 1$. This weighted measure

Table 3—Regression Estimates of the Treatment Effect of Refinancing on Monthly Probabilities of Mortgage Default and Non-mortgage Serious Delinquency

	Panel A. Mortgage default				Panel B. Non-mtg. serious delinq.			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
IV								
Basis points	−2.68	−2.46	−2.17	−3.94	−3.48	−3.15	−3.33	−3.35
(SE)	(0.77)	(0.77)	(0.77)	(1.03)	(1.28)	(1.33)	(1.32)	(3.50)
OLS								
Basis points	−4.30	−4.11	−4.10	−4.27	−2.43	−2.32	−2.31	−2.26
(SE)	(0.25)	(0.27)	(0.26)	(0.29)	(0.27)	(0.27)	(0.27)	(0.27)
Quarter FEs	✓	✓	✓	✓	✓	✓	✓	✓
Zip-code FEs	✓	✓	✓	✓	✓	✓	✓	✓
Observables	✓	✓	✓	✓	✓	✓	✓	✓
Q-by-zip FEs		✓	✓	✓		✓	✓	✓
Guar. lag FEs			✓	✓			✓	✓
Cohort FEs				✓				✓
N (mill.)	11.5	11.5	11.5	11.5	13.9	13.9	13.9	13.9

Notes: 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 three 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 5.9 bp, while it is 11.7 bp 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 ten equally-sized bines for each of: CLTV (lagged three months and at origination), credit score (lagged three months and at origination), credit utilization (lagged three months and at origination), initial mortgage rate, initial debt balances, and remaining principal balance. We also include indicators for mortgage "purpose" (e.g., purchase, cash-out refi, etc.). Standard errors (in parentheses) are clustered at the county level.

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.

These findings are in line with recent studies that show large effects of refinancing FRMs on mortgage default. Karamon, McManus, and Zhu (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 percent. 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 percent. 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 percent payment reduction (the average in our sample) is about 15 percent, much smaller than the 40 percent we find.[34]

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 5.9 bp.

[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 percent payment reduction decreasing default rates by about

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) estimate a LATE, as explained above. 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.

Heterogeneity.—Panel A of Table 4 shows that the strength of the estimated effect varies substantially with key borrower characteristics. High-credit score 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 credit score borrowers, on the other hand, saw a reduction in their monthly default hazard of 5.8 bp (their average default rate is 10.3 bp), and high credit utilization borrowers reduced their default rates by 4.5 bp (compared to an average of 8.8 bp). These groups, then, experienced a reduction in their default hazard of over 50 percent. 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.[35] 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.

One possible driver of heterogeneity in treatment effects could be differential treatment intensity. However, online Appendix Table A-4 shows that the average payment reduction was essentially uniform across the groups, in the 11–12 percent range, so any heterogeneity comes from differential responses rather than different treatments. Furthermore, it is interesting to ask whether the differential treatment effects we identify here also result in differential effects of program eligibility, which combines the estimated LATE with the "first stage" of refinancing take-up. In online Appendix C, we provide evidence on this question (for all outcomes) based on a

25 percent. This is larger than the effects found by Tracy and Wright (2016) and Fuster and Willen (2017) but still substantially smaller than the effect we find.

[35] We find very similar results if we make the high-CLTV determination at 100 percent rather than the sample median.

Table 4—Regression Estimates of Heterogeneous LATEs of Refinancing on Mortgage Default and Non-mortgage Serious Delinquency

Characteristic:	Credit score			Credit utilization			CLTV		
	Low (L)	High (H)	H − L	Low	High	H − L	Low	High	H − L
Panel A. Mortgage default									
Basis points	−5.76	0.94	6.69	0.26	−4.49	−4.75	−1.33	−4.44	−3.10
(SE)	(1.08)	(0.67)	(0.84)	(0.75)	(0.99)	(0.87)	(0.84)	(1.04)	(1.15)
Avg. default rt. (bp)	10.3	1.8		2.4	8.8		4.1	9.1	
Panel B. Non-mortgage serious delinquency									
Basis points	−8.56	0.46	9.02	−0.45	−6.20	−5.75	−2.67	−5.02	−2.35
(SE)	(1.91)	(1.04)	(1.35)	(1.14)	(1.69)	(1.19)	(1.51)	(1.45)	(1.63)
Avg. default rt. (bp)	24.0	1.7		3.7	19.2		10.1	15.0	

Notes: 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 three 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" CLTV (or credit score or credit utilization) if her score in month $t - 3$ 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 CLTV) 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. Additional controls include ten equally-sized bins for each of: CLTV (lagged three months and at origination), credit score (lagged three months and at origination), credit utilization (lagged three months and at origination), initial mortgage rate, remaining principal balance, and initial debt balances. We also include indicators for mortgage "purpose" (e.g., purchase, cash-out refi, etc.). Average default rates (bottom row) are monthly and weighted by take-up. Standard errors (in parentheses) are clustered at the county level.

simpler "triple differences" strategy.[36] We find that, indeed, financially constrained borrowers did have their incidence of mortgage default lowered more by HARP eligibility than did more secure borrowers. We also return to the question of how the differential treatment effects relate to refinance take-up below in Section VI.

Taken together, our 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) 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.

V. Effects of Refinancing on Non-Mortgage Debt

In this section, we turn to consequences of refinancing a first mortgage for other items on households' balance sheets. The main focus is on changes in

[36] As discussed in online Appendix C, the triple differences approach allows for more straightforward interpretations of this question than the workhorse econometric framework we use in the main text, although at the cost of lower statistical power.

Table 5—Summary Statistics of Non-mortgage Debt Balances and Monthly Changes

Levels:	*Balance*			$Bal \mid Bal > 0$	
	Mean	SD	% pos.	Mean	SD
Auto debt	8,835	12,305	51	17,426	12,204
HELOC	2,272	11,999	7	32,557	32,819
Bank card debt	6,078	8,592	85	7,114	8,890
Student debt	4,384	15,300	17	26,394	28,780
Retail cons. debt	2,091	6,208	43	4,898	8,748

Changes:	Δ*Balance*				$\Delta Bal \mid \Delta Bal > 0$		$\Delta Bal \mid \Delta Bal < 0$	
	Mean	SD	% pos.	% neg.	Mean	SD	Mean	SD
Auto debt	17	3,142	3	46	11,989	12,050	−693	1,860
HELOC	9	2,321	1	4	4,617	13,306	−1,283	7,503
Bank card debt	23	2,246	43	42	1,189	1,915	−1,165	2,332
Student debt	−5	1,898	2	12	3,269	9,194	−549	3,814
Retail cons. debt	2	1,405	15	30	808	2,906	−384	1,415

Note: We trim observations from a calculation if their balance in that debt category is greater than the ninety-ninth percentile of nonzero balances in the sample.

debt balances, though we also study effects on delinquencies. As discussed in Section IIIE, our main dependent variable of interest is the monthly change in debt balances. In addition to studying effects on the net change in balances, we also decompose the effect into positive and negative changes, by rerunning the regressions with negative (positive) values censored to 0. 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. We show that different borrower characteristics are strongly predictive of which behavior prevails.

Table 5 shows summary statistics for debt levels and changes by debt category. About half of the observations have positive auto debt, which is typically being amortized, though in about 3 percent 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, the most common type of non-mortgage debt in 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. 2019; Di Maggio, Kermani, and Palmer 2016; Di Maggio et al. 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

Table 6—Regression Estimates of the LATE of Refinancing on Monthly Debt Accumulation

	Net change				Positive	Negative
Auto debt	23	20	20	20	44	−25
(SE)	(8)	(8)	(8)	(13)	(10)	(8)
HELOC	34	36	35	38	34	1
	(7)	(7)	(7)	(10)	(9)	(8)
Bank card debt	−3	−5	−6	−20	23	−30
	(5)	(5)	(5)	(7)	(10)	(10)
Student debt	0	−2	−3	9	−11	8
	(5)	(5)	(5)	(7)	(6)	(5)
Retail consumer debt	−1	−1	0	0	10	−10
	(4)	(4)	(4)	(6)	(5)	(4)
Initial bal.	✓	✓	✓	✓	✓	✓
Quarter FEs	✓	✓	✓	✓	✓	✓
Zip-code FEs	✓	✓	✓	✓	✓	✓
Observables	✓	✓	✓	✓	✓	✓
Q-by-zip FEs		✓	✓	✓	✓	✓
Guar. lag FEs			✓	✓	✓	✓
Cohort FEs				✓		
Observations (mill.)	14.5	14.5	14.5	14.5	14.5	14.5

Notes: 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 ten equally sized bins for each of: CLTV (lagged three months and at origination), credit score (lagged three months and at origination), credit utilization (lagged three months and at origination), initial mortgage rate, remaining principal balance, and initial debt balances. We also include indicators for mortgage "purpose" (e.g., purchase, cash-out refi, etc.). Standard errors (in parentheses) are clustered at the county level.

auto debt accumulation of around \$20/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 (\$44), and decreases as they pay down their debt (−\$25). 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 percent of the magnitude of the payment reduction, very much in line with the $\$44/\$173 = 25\%$ increase that we find. Di Maggio et al. (2017) look at ARMs and find the increase in auto debt is closer to 10 percent 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. (2017) sample resulted in much larger payment reductions (of about \$940/month, or 50 percent).

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 somewhat larger than auto debt—about \$35/month—and similarly stable. Splitting this net effect into positive and negative components, it appears that HELOC draws (the positive change) are smaller than for auto debt, but 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 \$23 of bank card debt and \$10 of retail consumer debt per month, while focusing on negative changes show these balances being paid off by \$30 and \$10 per month. The estimates for student debt show less debt accumulation and slower paydowns, but the evidence is weak and marginally significant.[37]

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 55 percent of the monthly savings on mortgage payments (\$100/\$173), more than twice the response one would see by looking at just autos. The reductions in debt were equal to about 30 percent of the mortgage payment savings (\$56/\$173), leading to a net increase in debt accumulation of about 25 percent 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 (2018). 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 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 specification described in Section IIIE to simultaneously instrument for $Refied_{it}$ and $Refied_{it} \times HighCreditScore_{it}$ (or $Refied_{it} \times HighCreditUtilization_{it}$, or $Refied_{it} \times HighCLTV_{it}$).[38]

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

[37] As opposed to delinquency outcomes, where it is intuitive that the OLS magnitude should be biased upward relative to the true causal effect, for these debt accumulation outcomes it is not obvious which way the bias would go. Online Appendix Table A-5 shows the OLS results. The IV effects are larger in magnitude for the auto and HELOC measures, smaller for bank card debt, and similar for the student and retail consumer debt categories.

[38] As for mortgage default, online Appendix C again presents reduced-form evidence on heterogeneity in the effects of HARP eligibility on debt changes across categories.

Table 7—Regression Estimates of Heterogeneous LATEs of Refinancing on Monthly Debt Accumulation

Characteristic:	Credit score			Credit utilization			CLTV		
	Low (L)	High (H)	H − L	Low	High	H − L	Low	High	H − L
Auto debt	32	9	−22	10	27	17	28	10	−18
(SE)	(10)	(8)	(10)	(8)	(9)	(9)	(10)	(9)	(11)
HELOC	55	17	−37	61	16	−45	42	39	−3
	(9)	(7)	(9)	(7)	(8)	(8)	(9)	(8)	(11)
Bank card debt	−17	3	19	18	−30	−48	−4	−12	−8
	(7)	(5)	(7)	(5)	(7)	(7)	(7)	(6)	(8)
Student debt	−8	2	11	−4	−2	2	−6	−1	5
	(6)	(5)	(5)	(5)	(6)	(4)	(6)	(5)	(6)
Retail cons. debt	15	−15	−30	−2	0	2	−4	4	8
	(5)	(4)	(4)	(4)	(5)	(4)	(5)	(4)	(6)

Notes: 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 $t - 3$ 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 CLTV) 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. Additional controls include ten equally-sized bins for each of: CLTV (lagged three months and at origination), credit score (lagged three months and at origination), credit utilization (lagged three months and at origination), initial mortgage rate, remaining principal balance, and initial debt balances. We also include indicators for mortgage "purpose" (e.g., purchase, cash-out refi, etc.). Standard errors (in parentheses) are clustered at the county level.

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 much less so 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 paydown of that debt in the last column of Table 6 is driven by high credit score borrowers. On the other hand, bank card debt gets paid down by low credit score borrowers only.

Looking at the split by credit utilization, the auto debt response is much stronger for the constrained group. But low-utilization borrowers are more prone to add to their HELOC and bank card debts, a sensible result given that borrowers with high credit utilization have little remaining debt available on those instruments, by definition. Interestingly, those with high credit utilization use the increased cash flow to pay down their bank card debt. This, too, 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 of debt accumulation to refinances. 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. 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 these differences 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 of refinancing on the likelihood of serious delinquency on a debt other than the first mortgage. The takeup-weighted monthly hazard rate in the sample is 11.7 bp for this outcome, twice that of mortgage default.

Panel B of Table 3 shows that refinancing reduced the monthly serious delinquency hazard by about 3.3 bp, or 30 percent. Panel B of Table 4 again looks for heterogeneity along a handful of observable dimensions. As with mortgage default, 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 8.6 bp (compared to a base of 24.0 bp), while those with high credit utilization lowered their default rate by 6.2 bp (compared to an average default rate of 19.2 bp). 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.

These findings imply 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.

VI. Take-up

The preceding analysis finds that refinancing 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, Pope, and Pope 2016; Andersen et al. 2019; Agarwal, Rosen, and Yao 2015).

In this section, we ask which borrower characteristics predict take-up. In addition to the financial variables available in our microdata (CLTV, credit score, credit utilization), we will include some local (zip code- 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 2012) data to measure local mortgage market concentration, following Scharfstein and Sunderam (2016). We also merge in zipcode-level US Census Bureau (2019) 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 Cox proportional hazard model in the monthly data, restricting ourselves to the HARP-eligible population, beginning in March 2010. We treat a borrower as censored after she refinances for the first time (our outcome of interest), and we also censor observations in which the borrower would have been ineligible for HARP due to past delinquencies—in particular if she has any delinquencies in the previous six months or more than one in the previous 12 months. About 53 percent of borrowers refinance in this sample, which includes around 115 thousand borrowers and 4.5 million borrower-month observations.

Results are shown in Table 8. We report the estimates in terms of hazard ratios, so all coefficients should be compared to 1.00. For instance, a coefficient of 1.90 on the indicator for credit score above 800 implies that a borrower with a credit score above 800 is 90 percent more likely to refinance than one with a credit score below 675, the omitted category. (A coefficient of 0.65 conversely would imply that a refinance was 35 percent less likely than in the omitted category.) All columns include flexible controls for loan balance and contract rate, as well as a full set of indicators for loan purpose and quarter FEs.[39] Column 6 is the most restrictive specification as it includes state-by-quarter FEs. Results are quite consistent across all columns.

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 top quartile of credit utilization is only about 7 percent less likely to refinance than the bottom quartile. The credit score and CLTV results are quite a bit stronger. Borrowers with credit scores above 750 are almost twice as likely to refinance as those below 675, while those with CLTVs above 95 percent are over 30 percent less likely to refinance than those with CLTVs below 85 percent. The credit score result may be picking up a lack of financial sophistication on

[39] As expected, borrowers with higher balances and higher interest rates are more likely to refinance; for the sake of space, we suppress these results.

Table 8—Cox Proportional Hazard Estimation of Refinancing Take-up among HARP-Eligible Borrowers

	(1)	(2)	(3)	(4)	(5)	(6)
Credit score (≤675 omitted)						
676–725	1.61	1.61	1.61	1.61	1.61	1.60
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
726–775	1.83	1.83	1.82	1.82	1.82	1.79
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
776–800	1.98	1.98	1.97	1.98	1.98	1.92
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
>800	1.90	1.90	1.90	1.90	1.90	1.83
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Credit utilization (first quartile omitted)						
Second quartile	1.05	1.05	1.05	1.05	1.05	1.05
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Third quartile	1.00	1.00	1.00	1.00	1.00	1.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Fourth quartile	0.93	0.93	0.93	0.93	0.93	0.93
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Cred. lim. = 0	0.85	0.85	0.86	0.86	0.86	0.86
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
CLTV (≤85% omitted)						
85%–90%	0.86	0.86	0.86	0.86	0.86	0.85
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
90%–95%	0.76	0.76	0.76	0.76	0.76	0.75
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
>95%	0.68	0.68	0.69	0.69	0.69	0.67
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
Borrower age (middle tercile omitted)						
Bottom tercile	0.99	0.99	0.99	0.99	0.99	0.97
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Top tercile	0.98	0.98	0.99	0.99	0.99	0.99
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Junior Lien	1.01	1.01	1.01	1.01	1.01	0.99
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Mkt. share of top 4 banks (middle tercile omitted)						
Bottom tercile		0.98			0.97	0.99
		(0.03)			(0.03)	(0.02)
Top tercile		0.99			1.01	1.02
		(0.03)			(0.03)	(0.02)
% ≥ Bachelor's deg. in zip code (middle tercile omitted)						
Bottom tercile			0.95		0.99	0.99
			(0.01)		(0.02)	(0.01)
Top tercile			1.02		1.00	0.99
			(0.02)		(0.02)	(0.02)
Median income in zip code (middle tercile omitted)						
Bottom tercile				0.93	0.93	0.94
				(0.01)	(0.01)	(0.01)
Top tercile				1.05	1.05	1.08
				(0.02)	(0.02)	(0.01)
Quarter FEs	✓	✓	✓	✓	✓	✓
State-by-quarter FEs						✓
Balance, interest rate, loan purp. controls	✓	✓	✓	✓	✓	✓

Notes: The average monthly refinancing rate in the sample was 132 bp. Borrowers are censored after refinancing or defaulting, and we also exclude a borrower in the 6 months following a delinquency, when they are ineligible for HARP (or if they have more than one delinquency in the previous 12 months). Credit score, credit utilization, CLTV, and the indicator for having a junior lien are lagged three months. We do not include zip code FEs for computational tractability, but in a linear model analogous to column 1, we find that the inclusion of zip code FEs has negligible effects on the coefficients. Balance and interest rate are controlled for with step functions estimated on ten equally sized bins. Estimates are reported in terms of the hazard ratio. Standard errors (in parentheses) are clustered at the county level. Observations = 4,526,312.

the part of the borrower, 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 I).[40]

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. In online Appendix C, we study a subtly different question: what borrower characteristics are associated with greater responsiveness to HARP access? In that analysis, we find that in fact struggling borrowers, particularly as measured by CLTV, are more likely to be induced to refinance by HARP eligibility. Simply put, while high-CLTV borrowers are less likely than low-CLTV borrowers to refinance when eligible for HARP (as in Table 8), that difference widens substantially in the ineligible group.[41] This emphasizes why programs like HARP can be so important: in their absence, the benefits of declining interest rates flow largely to a subset of borrowers that is in a stronger financial position and less responsive to the payment reduction, while policies that specifically target struggling borrowers can help them take advantage of favorable mortgage terms as well (even if they still remain less inclined to do so than better-off borrowers).

We find only weak evidence of a hump-shaped pattern in age, as documented in Agarwal et al. (2009). Specifically, the middle tercile in our data includes ages 37–48, and we find that borrowers in this age range are slightly more likely than older or younger borrowers to refinance, though this is at best marginally significant. We also find little effect from having a junior lien associated with the first mortgage (controlling for CLTV and other characteristics), and we find no evidence that concentrated markets (measured by the share of mortgages in 2012 made by a county's four largest lenders) are associated with less refinancing activity, as in Scharfstein and Sunderam (2016).

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. We also find evidence that borrowers in zip codes with more educated populations are more likely to refinance, though this association seems to be driven by the income result above, as it disappears when income is also included in the regression.[42]

[40] Relatedly, online Appendix Table A-6 shows a version of this analysis allowing for coefficients on these individual financial characteristics to vary between the early (until mid-2012) part of the sample and the later part. We find that for credit score and, especially, credit utilization, these coefficients do not change substantially, but the negative impact of CLTV on refinancing propensity is substantially stronger in the early period, before HARP 2.0.

[41] This is because low-CLTV borrowers who are HARP-ineligible can often still refinance through other channels, while there are essentially no alternatives for the ineligible high-CLTV borrowers.

[42] Keys, Pope, and Pope (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.

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.[43] As discussed, though, HARP itself did disproportionately increase refinancing probabilities of struggling borrowers, as intended by the policy.

VII. 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 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 other debt instruments and different proxies for distress. The leading example is bank card debt, which borrowers with low credit scores and high credit utilization pay down using their improved cash flow. The cash flow, then, does not merely boost spending but is also used to repair borrower 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. From the perspective of monetary policy, 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, Krishnamurthy, and McQuade 2018) that mortgages that have payment reductions triggered automatically in downturns (such as ARMs) can be a useful stabilizing tool for the macroeconomy.

[43] 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. (2019) 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.

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.

REFERENCES

Abel, Joshua, and Andreas Fuster. 2021. "Replication data for: How Do Mortgage Refinances Affect Debt, Default, and Spending? Evidence from HARP." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E116461V1.

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, Tomasz Piskorski, and Amit Seru. 2017a. "Policy Intervention in Debt Renegotiation: Evidence from the Home Affordable Modification Program." *Journal of Political Economy* 125 (3): 654–712.

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

Agarwal, Sumit, John C. Driscoll, Xavier Gabaix, and David Laibson. 2009. "The Age of Reason: Financial Decisions over the Life Cycle and Implications for Regulation." *Brookings Paper on Economic Activity* 39 (2): 51–101.

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel. 2018. "Do Banks Pass through Credit Expansions to Consumers Who Want to Borrow?" *Quarterly Journal of Economics* 133 (1): 129–90.

Agarwal, Sumit, Richard J. Rosen, and Vincent Yao. 2015. "Why Do Borrowers Make Mortgage Refinancing Mistakes?" *Management Science* 62 (12): 3494–3509.

Amromin, Gene, and Caitlin Kearns. 2014. "Access to Refinancing and Mortgage Interest Rates: HARPing on the Importance of Competition." FRB Chicago Working Paper 2014-25.

Andersen, Steffen, John Y. Campbell, Kasper Meiser Nielsen, and Tarun Ramadorai. 2019. "Sources of Inaction in Household Finance: Evidence from the Danish Mortgage Market." NBER Working Paper 21386.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444–55.

Auclert, Adrien. 2019. "Monetary Policy and the Redistribution Channel." *American Economic Review* 109 (6): 2333–67.

Beraja, Martin, Andreas Fuster, Erik Hurst, and Joseph Vavra. 2019. "Regional Heterogeneity and the Refinancing Channel of Monetary Policy." *Quarterly Journal of Economics* 134 (1): 109–83.

Berger, David W., Konstantin Milbradt, Fabrice Tourre, and Joseph Vavra. 2018. "Mortgage Prepayment and Path-Dependent Effects of Monetary Policy." NBER Working Paper 25157.

Bhutta, Neil, and Benjamin J. Keys. 2016. "Interest Rates and Equity Extraction during the Housing Boom." *American Economic Review* 106 (7): 1742–74.

Bond, Philip, Ronel Elul, Sharon Garyn-Tal, and David K. Musto. 2017. "Does Junior Inherit? Refinancing and the Blocking Power of Second Mortgages." *Review of Financial Studies* 30 (1): 211–44.

Campbell, John Y. 2006. "Household Finance." *Journal of Finance* 61 (4): 1553–1604.

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

Campbell, John Y., and João F. Cocco. 2015. "A Model of Mortgage Default." *Journal of Finance* 70 (4): 1495–1554.

Campbell, John Y., Stefano Giglio, and Parag Pathak. 2011. "Forced Sales and House Prices." *American Economic Review* 101 (5): 2108–31.

Chen, Hui, Michael Michaux, and Nikolai Roussanov. 2020. "Houses as ATMs? Mortgage Refinancing and Macroeconomic Uncertainty." *Journal of Finance* 75 (1): 323–75.

CoreLogic. 2019. "Home Price Indexes (HPI)." Federal Reserve System RADAR Interface (accessed August 2019).

CRISM. 2019. "Credit Risk Insights Servicing McDash." Federal Reserve System RADAR Interface (accessed August 2019).

Deaton, Angus. 1991. "Saving and Liquidity Constraints." *Econometrica* 59 (5): 1221–48.

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

Di Maggio, Marco, Amir Kermani, and Christopher Palmer. 2016. "How Quantitative Easing Works: Evidence on the Refinancing Channel." NBER Working Paper 22638.

Dudley, William C. 2012. "Housing and the Economic Recovery." Speech, New Jersey Bankers Association Economic Forum, Iselin, New Jersey, January 6, 2012.

Eberly, Janice, and Arvind Krishnamurthy. 2014. "Efficient Credit Policies in a Housing Debt Crisis." *Brookings Papers on Economic Activity* 54 (2): 73–136.

Ehrlich, Gabriel, and Jeffrey Perry. 2015. "Do Large-Scale Refinancing Programs Reduce Mortgage Defaults? Evidence from a Regression Discontinuity Design." US Congressional Budget Office (CBO) Working Paper 2015-06.

Eichenbaum, Martin, Sergio Rebelo, and Arlene Wong. 2018. "State Dependent Effects of Monetary Policy: The Refinancing Channel." NBER Working Paper 25152.

Elul, Ronel, Nicholas S. Souleles, Souphala Chomsisengphet, Dennis Glennon, and Robert Hunt. 2010. "What 'Triggers' Mortgage Default?" *American Economic Review* 100 (2): 490–94.

Fannie Mae. 2019. "Monthly Summary." Fannie Mae. http://fanniemae.com/portal/about-fm/investor-relations/monthly-summary.html (accessed August 2019).

Federal Housing Finance Agency Office of Inspector General (FHFAOIG). 2013. *Home Affordable Refinance Program: A Mid-Program Assessment—Evaluation Report EVL-2013-006*. Washington, DC: FHFAOIG.

Frame, W. Scott, Andreas Fuster, Joseph Tracy, and James Vickery. 2015. "The Rescue of Fannie Mae and Freddie Mac." *Journal of Economic Perspectives* 29 (2): 25–52.

Freddie Mac. 2019. "Monthly Volume Summaries." Freddie Mac. https://web.archive.org/web/20110302140033/http://www.freddiemac.com/investors/volsum/ (accessed August 2019).

Freddie Mac. 2020. "30-Year Fixed Rate Mortgage Average in the United States (MORTGAGE30US)." FRED Economics Data, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/MORTGAGE30US (accessed February 2020).

Fuster, Andreas, and Paul S. Willen. 2010. "\$1.25 Trillion Is Still Real Money: Some Facts about the Effects of the Federal Reserve's Mortgage Market Investments." Federal Reserve Bank of Boston Public Policy Discussion Paper 10-4.

Fuster, Andreas, and Paul S. Willen. 2017. "Payment Size, Negative Equity, and Mortgage Default." *American Economic Journal: Economic Policy* 9 (4): 167–91.

Ganong, Peter, and Pascal Noel. 2018. "Liquidity vs. Wealth in Household Debt Obligations: Evidence from Housing Policy in the Great Recession." University of Chicago, Becker Friedman Institute Working Paper 2018-58.

Greenspan, Alan. 2004. "Testimony Before the Committee on Financial Services, U.S. House of Representatives." February 11, 2004. https://www.federalreserve.gov/boarddocs/hh/2004/february/testimony.htm.

Greenwald, Daniel L. 2018. "The Mortgage Credit Channel of Macroeconomic Transmission." http://www.dlgreenwald.com/uploads/4/5/2/8/45280895/mortgage_credit_channel_feb_2018.pdf.

Guren, Adam M., Arvind Krishnamurthy, and Timothy J. McQuade. 2018. "Mortgage Design in an Equilibrium Model of the Housing Market." NBER Working Paper 24446.

HMDA. 2012. "Home Mortgage Disclosure Act Records for the Year 2012." HMDA. https://www.consumerfinance.gov/data-research/hmda/ historic-data/ (accessed February 2020).

Hubbard, R. Glenn, and Christopher J. Mayer. 2009. "The Mortgage Market Meltdown and House Prices." *B.E. Journal of Economic Analysis and Policy* 9 (3): Article 8.

Hurst, Erik, and Frank Stafford. 2004. "Home Is Where the Equity Is: Mortgage Refinancing and Household Consumption." *Journal of Money, Credit and Banking* 36 (6): 985–1014.

Jappelli, Tullio, and Luigi Pistaferri. 2014. "Fiscal Policy and MPC Heterogeneity." *American Economic Journal: Macroeconomics* 6 (4): 107–36.

Johnson, Eric J., Stephan Meier, and Olivier Toubia. 2018. "What's the Catch? Suspicion of Bank Motives and Sluggish Refinancing." *Review of Financial Studies* 32 (2): 467–95.

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

Karamon, Kadiri, Douglas McManus, and Jun Zhu. 2016. "Refinance and Mortgage Default: A Regression Discontinuity Analysis of HARP's Impact on Default Rates." *Journal of Real Estate Finance and Economics* 55: 457–75.

Kau, James B., Donald C. Keenan, Walter J. Muller, III, and James F. Epperson. 1992. "A Generalized Valuation Model for Fixed-Rate Residential Mortgages." *Journal of Money, Credit and Banking* 24 (3): 279–99.

Keys, Benjamin J., Devin G. Pope, and Jaren C. Pope. 2016. "Failure to Refinance." *Journal of Financial Economics* 122 (3): 482–99.

Remy, Mitchell, Deborah Lucas, and Damien Moore. 2011. "An Evaluation of Large-Scale Mortgage Refinancing Programs." US Congressional Budget Office (CBO) Working Paper 2011-4.

Scharfstein, David, and Adi Sunderam. 2016. "Market Power in Mortgage Lending and the Transmission of Monetary Policy." www.hbs.edu/faculty/Pages/download.aspx?name=Market%20Power%20in%20Mortgage%20Lending%20and%20the%20Transmission%20of%20Monetary%20Policy.pdf.

Scharlemann, Therese C., and Stephen H. Shore. 2016. "The Effect of Negative Equity on Mortgage Default: Evidence from HAMP's Principal Reduction Alternative." *Review of Financial Studies* 29 (10): 2850–83.

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

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

US Census Bureau. 2019. "American Community Survey." American FactFinder. https://factfinder.census.gov/faces/nav/jsf/pages/download_center.xhtml (accessed August 2019).

Wong, Arlene. 2019. "Refinancing and the Transmission of Monetary Policy to Consumption." https://static1.squarespace.com/static/576576adbe659449f97e0d35/t/5ce1f9f23fdc1e00014f3d22/1558313459256/Paper_AERrevision_0519.pdf.

Zeldes, Stephen P. 1989. "Consumption and Liquidity Constraints: An Empirical Investigation." *Journal of Political Economy* 97 (2): 305–46.

Zhu, Jun, Jared Janowiak, Lu Ji, Kadiri Karamon, and Douglas McManus. 2015. "The Effect of Mortgage Payment Reduction on Default: Evidence from the Home Affordable Refinance Program." *Real Estate Economics* 43 (4): 1035–54.

This article has been cited by:

1. David Berger, Konstantin Milbradt, Fabrice Tourre, Joseph Vavra. 2021. Mortgage Prepayment and Path-Dependent Effects of Monetary Policy. *American Economic Review* **111**:9, 2829-2878. [Abstract] [View PDF article] [PDF with links]