Inequality During the COVID-19 Pandemic: The Case of Savings from Mortgage Refinancing

Sumit Agarwal[*], Souphala Chomsisengphet[†], Hua Kiefer[‡], Leonard C. Kiefer[§], and Paolina C. Medina[¶]

January 24, 2021

Abstract

We study the distribution of savings from mortgage refinancing across income groups during the COVID-19 pandemic. Between February and June 2020, the difference in savings from refinancing between high- and low-income borrowers was 10 times higher than before. This was the result of two factors: individuals in the top quintile of the income distribution increased their refinancing activity more than comparable borrowers in the bottom quintile and, conditional on refinancing, they also captured the largest improvements in interest rates. Exploiting county-by-month variation in COVID-19 case rates we tie these results to the pandemic and explain up to 74% of increases in refinancing inequality through local economic conditions. Using data on refinancing applications, we find suggestive evidence that approval rates were affected by capacity constraints. However capacity constraints do not explain refinancing inequality as they did not lead lenders to sort approvals on income. Instead, our results are better explained by the under-representation of low-income borrowers in the pool of applications. We estimate a difference of $5 billion in savings from refinancing between the top quintile of the income distribution and the rest of the market. This has implications for the transmission of monetary policy and for the evolution of wealth inequality.

[*]National University of Singapore, ushakri@yahoo.com
[†]Office of the Comptroller of the Currency, souphala.chomsisengphet@occ.treas.gov
[‡]Federal Deposit Insurance Corporation, hkiefer@fdic.gov
[§]Freddie Mac, leonard_kiefer@freddiemac.com
[¶]Texas A&M University, pmedina@tamu.edu
We thank seminar participants at AEI Housing Collaborative, OCC, FHFA Fall Econ Summit, Freddie Mac Economic Seminar and Texas A&M for valuable comments. The views expressed in this paper do not necessarily reflect the views of Freddie Mac or its management, the Federal Deposit Insurance Corporation, the Office of the Comptroller of the Currency, the U.S. Department of the Treasury, or any federal agency and do not establish supervisory policy, requirements, or expectations.

1 Introduction

Mortgage refinancing is one of the channels through which expansionary monetary policy affects individual consumption (Di Maggio et al., 2017; Boyce et al., 2012; Agarwal et al., Forthcoming; Berger et al., 2018, 2019; Eichenbaum, Rebelo and Wong, 2018; Agarwal et al., 2017). However, the magnitude of the consumption response depends on the characteristics of those who take advantage of refinancing opportunities (Wong, 2019), including their marginal propensity to consume. In this paper, we document that during the COVID-19 pandemic, savings from refinancing were concentrated in the top segments of the income distribution, and this concentration was higher than in previous periods of large interest rate reductions. This is an important result to evaluate the effectiveness of monetary policy during the pandemic, as high-income individuals have lower marginal propensities to consume (Baker et al., 2020; Karger and Rajan, 2020; Di Maggio, Kermani and Palmer, 2020).

We study refinancing decisions of individuals for whom interest rates were sufficiently low to justify refinancing efforts. Our analysis uses a rich dataset of refinanced mortgages originally funded by Freddie Mac and matched to new refinancing loans also funded by Freddie Mac. For these loans, we observe the contract terms of both the old and new loan, as well as detailed origination records for both mortgages, including borrowers' income and purpose of the loan. In contrast to previous work using prevailing market average interest rates from the Primary Mortgage Market Survey (PMMS) to approximate savings from refinancing, we are able to observe the interest rates that specific borrowers receive when they refinance their mortgages. We document that savings from refinancing are in the order of \$8,800, conditional on refinancing. This value is smaller than the \$11,700 that result from the standard PMMS-based calculations. Furthermore, these savings are highly concentrated in the top segments of the income distribution. While savings from refinancing naturally vary across the income distribution (as they depend on unpaid balances, and interest rate differentials), we find that before 2020, most of these savings were explained by off-the-shelf control variables (FICO, unpaid balance, original interest rate, LTV, loan age). However in 2020, the same analysis reveals that differences in savings from refinancing between the top and bottom quintiles of the income distribution increased 10 times.

The increased inequality in the distribution of savings from refinancing is the result of two factors: individuals in the top quintile of the income distribution increased their refinancing activity more than their counterparties in the bottom quintile, and conditional on refinancing, they also captured the largest improvements in interest rates. Before 2020, individuals in the top and bottom quintile of the income distribution had basically the same probability of refinancing estimated at 1.14% after controlling for observable characteristics. During 2020, the bottom quintile of the income distribution increased its refinancing activity by 1.25 percentage points (pp), whereas the top quintile of the income distribution increased its refinancing activity by more than 8 pp. In addition, higher-income individuals also captured the largest improvements in interest rates. Before 2020, individuals in the bottom quintile of the income distribution received a 1.66 pp reduction in interest rates, conditional on refinancing. This reduction reached 1.83 pp in 2020 (i.e. a 0.17 pp improvement). In contrast, individuals in the top quintile of the income distribution who refinanced their mortgages received reductions of only 1.49 pp before the pandemic, but reached an average 1.87 bps reduction in 2020 (i.e., a 0.38 pp improvement).

Overall, we estimate a $5 billion gap in savings from refinancing between the top quintile of the income distribution and the rest of the market. If individuals in lower segments of the income distribution were receiving the same savings from refinancing as individuals in the top quintile of the income distribution, they would capture an additional $5 billion in refinance savings.

We complement the analysis with loan servicing data from McDash Analytics covering about 60% of the US mortgage market. In this dataset, we observe when a mortgage in the portfolio is prepaid. Using prepayments as a proxy for refinancing activity we confirm that refinancing activity, was concentrated in the top sections of the income distribution and that the differences in refinancing activity across the income distribution were significantly sharper than in previous periods of large interest rate reduction.

We then test the link between increases in refinancing inequality and the COVID-19 pandemic using county-by-month variation in COVID-19 case rates (cases per 100,000 people). We provide within zip-code estimates showing that the refinancing income gap increases 707% as we move from the bottom to the top quintile of the distribution of case rates. To explore the channels through which the pandemic affects refinancing inequality, we use a subset of our data for which we have information on local economic conditions shown to be affected by the pandemic and that could affect refinancing inequality. Specifically, we estimate the effect of increases in time spent at home (using GPS data from Google, at the county-by-month level), increases in unemployment (using initial unemployment insurance claim rates, measured at the county-by-month level), and increases in the the fraction of mortgages on forbearance (using state-by-month data from TransUnion monthly reports).

When analyzing the effect of each of these variables separately, we find a positive correlation between refinancing inequality and time spent at home, as well as between refinancing inequality and forbearance rates. We do not find any relation between unemployment insurance claims and refinancing inequality, which could be explained by the high replacement rate of unemployment benefits to original wages observed during the pandemic (Ganong, Noel and Vavra, 2020). However, we recognize that these three variables are highly correlated, are not measured at the same geographic level and their coarseness introduce measurement error on inferences at the individual level. We thus interpret the effect of these variables jointly measuring the impact of local economic conditions on refinancing inequality. We find that time spent at home, unemployment insurance claims, and forbearance rates can jointly explain between 39% and 74% of the impact of COVID-19 case rates on refinancing inequality, i.e., controlling for these three variables and their interactions with income reduces the magnitude of the coefficients of COVID-19 case rates on refinancing inequality by up to 74%.

We investigate if the increases in refinancing inequality are driven by changes in behavior of lenders or borrowers using a proprietary data set with information on refinance applications submitted to Freddie Mac's underwriting system, regardless of whether they were ultimately approved by lenders or funded by Freddie Mac. We find that lenders who experienced the largest growth rates in applications, show a decrease in the number of applications that were eventually funded by Freddie Mac, which we interpret as suggesting that capacity constraints became binding for some lenders during the period. However, this effect is present across the income distribution and therefore, it cannot explain the increases in refinancing inequality we document. Similarly, when we focus on applications that resulted in loans ultimately funded by Freddie Mac and, we find that applications of low-income borrowers are processed in about the same time (or slightly faster) as applications of high-income borrowers.

While we do not find evidence that lenders prioritized the processing of applications of high income borrowers, we find that high-income borrowers are overrepresented in the pool of applications. More than 17.5% of refinance applications come from borrowers in the top decile of the income distribution of portfolio mortgages for which the option to refinance becomes in-the-money during the observation period (as defined by Agarwal, Driscoll and Laibson (2013)). This suggests that lower-income borrowers are less likely to apply than their high-income counterparts.

Our paper contributes to a large literature studying mortgage refinancing and its consequences for the economy. First, in terms of the transmission of monetary policy, previous work has documented the importance of the mortgage refinancing channel (Campbell, 2006; Scharfstein and Sunderam, 2016; Bhutta and Keys, 2016; Koijen, Van Hemert and Van Nieuwerburgh, 2009; Chen, Michaux and Roussanov, 2020; Di Maggio et al., 2017; Di Maggio, Kermani and Palmer, 2020; Eichenbaum, Rebelo and Wong, 2018), highlighting the importance of the distribution of savings across areas with different local economic conditions or across borrowers with different characteristics (Beraja et al., 2017; Wong, 2019; Laibson, Maxted and Moll, 2020). To our knowledge, our paper is the first to provide an estimate of the variation in actual savings from refinancing (with mortgage specific interest rates) across the income distribution, which arguably captures variation in marginal propensities to consume (D'Amico, Kurakula and Lee, 2020; Boyce et al., 2012; Agarwal et al., Forthcoming, 2017; Berger et al., 2018, 2019; Di Maggio et al., 2017; Karger and Rajan, 2020). [1]

In doing so, we also contribute to the literature studying the distributional consequences of monetary policy. Previous work has focused on the wealth effect of inflation through changes in the value of nominal assets (Doepke and Schneider, 2006) or the medium to long term impacts of monetary policy on labor markets Dynarski et al. (1997), Sterk and Tenreyro (2013). In contrast, we study a direct immediate impact of interest rate reductions on household wealth: savings from refinancing, which turn out to be larger for individuals with higher income. In this context, the closest work to ours is Beraja et al. (2017) who finds that monetary policy amplifies local regional disparities, through its effect on mortgage refinancing and spending.

Our results build on the work of Agarwal, Driscoll and Laibson (2013); Keys, Pope and Pope (2016); Bennett, Peach and Peristiani (2000); Johnson, Meier and Toubia (2015); Agarwal, Ben-David and Yao (2017); Agarwal, Rosen and Yao (2016); Andersen et al. (2020); DeFusco and Mondragon (2020), who explain low refinancing activity, even when facing sufficiently low interest rates, as a result of limited financial literacy, the presence of behavioral biases, strict documentation requirements, or other frictions in the mortgage market. We expand on this work by focusing on differences in refinancing activity across the income distribution, both in the extensive and intensive margins, with special focus on the pandemic period. Our

[1] To our knowledge, the only other paper using matched refinancing transactions with information on old and new interest rates for every transaction is Berger et al. (2019). Their focus is different from ours, as they study path dependence effects of monetary policy.

results are consistent with (Nothaft and Chang, 2005; Willen and Zhang, 2020; Goodstein, 2014) who find that propensities to refinance vary with income and race.

Finally, our paper also contributes to a fast-growing literature studying the economic impact of the COVID-19 pandemic. Recent work documents strong decreases in consumption (Baker et al., 2020; Coibion, Gorodnichenko and Weber, 2020; Chen, Qian and Wen, 2020; Cox et al., 2020; Dunn, Hood and Driessen, 2020) and disruptions to labor markets (Kurmann, Lale and Ta, 2020; Bartik et al., 2020; Mathy, 2020; Ganong, Noel and Vavra, 2020). In both cases, the impact has disproportionally affected individuals in the lowest segment of the income distribution (Chetty et al., 2020; Molly and Martha, 2020; Mongey, Pilossoph and Weinberg, 2020; Adams-Prassl et al., 2020) thus increasing income inequality. We argue that unequal access to refinancing opportunities is an important channel through which inequality increased during the pandemic.

2 Data Description

Our analysis is based on several sources of data. First, we use a unique administrative loan-level dataset for conventional single-family loans funded by Freddie Mac. This dataset includes all outstanding single-family 30-year fixed-rate mortgages that were funded by Freddie Mac and were active during the period of analysis. We followed those loans through time and observed whether or not the loan was prepaid during the wave. In addition, for a subset of loans that were prepaid, we matched a new loan also funded by Freddie Mac that was originated at the same property address within a 45-day window of the closure of the prepaid loan. For those matched transactions, we collected loan-level attributes of the newly originated loan at the same address. In cases where the loan was refinanced, we observed the new loan product and loan attributes, including the new interest rate. We also identified cases where the prepayment was not for a refinance, but rather a home purchase.

The second dataset consists of loan-level information provided by residential mortgage servicers and collected by Black Knight. This data provides extensive information on loan, property, and borrower characteristics at the time of origination as well as dynamically updated loan information subsequent to origination. Following the literature, we restrict our sample to owner-occupied, single-family, first-lien loans. We focus on 30-year fixed-rate conventional mortgages (i.e., FHA, VA, other government-insured loans are excluded). Exotic loans (e.g., loans with balloon payment, negative amortization, or prepayment penalty) are excluded from our sample. Loans that are in foreclosure, bankruptcy and REO status or are less than 2 months old are also removed from our sample. For each loan in the portfolio, we observe if the loan was prepaid or not. One drawback of this data is that we do not know the reason of prepayment, i.e. sold property, or refinance. We use prepayments as a proxy for refinancing, and use the terms prepayment and refinancing interchangeably. Section 7 discusses in detail the potential bias introduced by the use of this proxy, and provides a series of robustness tests with complementary data-sets to show that our results are indeed driven by refinancing activity and not by other prepayments.

We use these two data sets to show that the gap in refinancing activity and in savings from refinancing between high and low-income borrowers is significantly higher in 2020, compared to other periods of large reductions in interest rates. To do so, we focus the analysis on five periods characterized by interest rate drops: 2015, 2016, 2017, 2019 and 2020. Figure 1 shows the evolution of 30-year fixed mortgage rates between 2014 and 2020 (top panel). The highlighted periods correspond to 5-month windows with the largest drops in interest rates. During the period of October 2014 through February 2015, there was a 0.64 percentage point drop in interest rates between the high and low points of the period. Between May and September of 2016 and 2017 there was a 0.25 and 0.32 percentage point decrease, respectively. The period of May 2019 through September 2019 experienced a drop of 0.71 percentage points. Finally, between February 2020 and June 2020 there was a decrease of 0.77 percentage points. The bottom panel of Figure 1 shows refinancing activity. We can see that most of the periods with the largest drops in interest rates were also characterized by the largest spikes in refinancing activity.

In each of these periods, we identify mortgages for which the refinancing option was not in-the-money before the window of observation and becomes in-the-money during the window of observation. To do so, we use the model of optimal refinancing proposed by Agarwal, Driscoll and Laibson (2013), which provides a closed-form solution for the problem of optimal refinancing. Under some assumptions, this model identifies a threshold for which it is optimal to trade-in an old in-the-money refinancing option, for a new out-of-the-money refinancing option that is acquired, taking into account closing costs, mortgage size, taxes, the standard deviation of the mortgage interest rate, and a calibrated Poisson rate of exogenous repayment capturing the combined effects of moving events, principal repayment and inflation-driven depreciation of the mortgage obligation. For calibrated choices of these parameters, the optimal refinancing differentials range typically from 100 to 200 basis points.

Table 1 describes the set of newly-in-the-money mortgages before and after the pandemic. During the most recent wave of low interest rates, 13.77% of active mortgages became in-the money. This number is almost twice as large as the fraction of active mortgages that became in-the money in previous waves of low interest rates, and is reflective of the historically low levels of interest rates observed during the period. The wave of 2020 also shows the highest prepayment rate, with 8.05% of newly in-the money mortgages prepaid during the first few months of the year, compared to a 3.83% prepayment rate for the three previous waves of comparable interest rate reductions. In terms of observable characteristics, mortgages that became in the money in refinancing waves previous to 2020 have a FICO score of 728 points, compared to 737 points for mortgages that became in the money during 2020. We estimate income out of Debt-to-Income Ratios reported to McDash, and we find that borrowers in waves prior and during 2020 have similar estimated monthly income ($4,852 vs $4,736, respectively). Mortgages that became in the money prior to 2020 are about 1 year older than mortgages that became in the money in previous periods, and have slightly lower interest rates (40 bps) and loan to value ratio between 79% and 80%. The interest rate differential that captures incentives to refinance (mortgage rate - market rate) are very comparable across waves. Unpaid balances are slightly higher in 2020, and as a result, potential savings are also slightly higher for the 2020 wave, reaching a level of $11,714. Potential savings are defined as Agarwal, Driscoll and Laibson (2013) and Keys, Pope and Pope (2016) as the present value of the savings from refinancing at the market rate, adjusting for the probability of moving, tax incentives, up-front costs, and discounting over time. Overall, mortgages that became in-the-money before and after the pandemic are fairly comparable, with the differences in each of these variables representing less than 25% of a standard deviation. Nevertheless our analysis includes time and zip code fixed effects, along with a set of rich loan level controls to isolate the role of income on refinancing activities, from changes in the composition in the pool of newly in-the-money mortgages.

To study the relation between refinancing inequality and the COVID-19 pandemic, we complement the mortgage data with a rich set of variables tracking the impact of the pandemic on local economic conditions across different geographies. Specifically we look at mobility restrictions, initial unemployment insurance claims, percentage of mortgages under forbearance, and COVID-19 case rates. Except for the percentage of mortgages under forbearance, we download the data from the public repository created by Chetty et al. (2020) to track the impact of the pandemic across the United States.

Mobility Data In response to the COVID-19 pandemic, Google release data on GPS based mobility patterns (https://www.google.com/covid19/mobility/). The Community Mobility Reports provided estimated mobility for individuals aggregated different geographic levels, using GPS data. The baseline is the median value, for the corresponding day of the week, during the 5-week period Jan 3–Feb 6, 2020 and indices are reported as percentage differences from that baseline. Google released several indices including time spent at residential places. The data is reported at the county-by-day level, we aggregate it at the county-by-month level to match the frequency of the mortgage data.

Initial Unemployment Insurance Claims We use data on initial claims at the county level, reported to the Department of Labor. Weekly initial claims are averaged at the monthly level and expressed per 100 people in the 2019 labor force. Location is defined as the state liable for the benefits payment.

Forbearance To study the interaction between forbearance and refinancing inequality, we use data on financial hardship from TransUnion. Specifically we use data from the Monthly Industry Reports (TransUnion, 2020) showing the percentage of mortgages in hardship at the state level on a monthly basis between March 2020 and May 2020. This metric represents the percentage of accounts in a delinquency category that are hardship flagged (affected by natural/declared disaster, accounts reported as in forbearance, accounts reported as deferred or payment due amount removal, or freezing of account status and/or past due amount).

COVID-19 case rate The data comes as a seven-day moving average count of cases per capita at the county level, which we further average at the monthly level to match our mortgage data. The COVID-19 data is publicly available from the New York Times COVID-19 repository.

The coverage of these four variables is imperfect, and subject to availability by data providers in each case. COVID-19 case rates are available for 3,023 counties, which cover 99.9% of our mortgage data. We refer to these counties and mortgages as our base coverage for the pandemic analysis. Mobility measures are available only for 26% of those counties covering 86% of observations in our base coverage. Unemployment insurance claims at the county level are available for 52.2% of mortgages in our base coverage, and forbearance rates (at the state level) are available for 87% of observations in our base coverage. All four variables are available for a subset of 1.2 million observations at the mortgage-month level, representing 43% of our base coverage for the pandemic analysis. [2]

With this data, we perform two broad analyses. We measure the impact of the pandemic on refinancing inequality with our base coverage, to get the fullest possible impact of the pandemic with as much geographic coverage as possible. We then restrict the sample to counties for which all four variables of interest are available to explain the effect of the pandemic (COVID case rates) through mobility restrictions, unemployment and financial hardship. Figures A2 and A3 show the distribution of the intensity of the pandemic across geographies and over time.

Finally, we also use a proprietary dataset with refinancing applications submitted to Freddie Mac's Loan Product Advisor (LPA) tool. LPA collects information on applications submitted to Freddie Mac's underwriting system, which includes both loans funded by Freddie Mac, as well as loans not ultimately funded by Freddie Mac. These data broadly track trends seen in the Mortgage Bankers Association's Weekly Application Survey. About 18% of all new loans in the market are run trough Freddie Mac's LPA tool. We use it to investigate if the increases in refinancing inequality that we find are driven by changes in borrower or lender behavior. We note that we do not observe when an application is approved by a lender or not. Instead, we observe when an application is purchased by Freddie and when it is not. Our first part of the analysis will abstract away from approval status of the applications, to simply focus on the distribution of applications submitted to LPA across the income distribution, regardless of whether they were approved or not by lenders, and regardless of whether they were purchased by Freddie Mac or not. Our second analysis will be focused on loans that were purchased by Freddie Mac. For them, we will study changes in the time it takes to process applications over time, across the income distribution.

[2] For robustness, we also perform the analysis with unemployment insurance claims at the state level. This allows to increase our coverage to 79% of our original observations, however with a coarse measure of unemployment insurance claims. The results are qualitatively the same. The analysis with unemployment insurance at the state level is available upon request.

3 Refinancing inequality over time

3.1 Freddie Mac Matched-transactions data

We describe the evolution of savings from refinancing across the income distribution using our matched-transactions data set. This data set allow us to improve on existing characterizations of the distribution of savings from refinancing in several ways. First, we observe the interest rate of both the old and new existing loans. In contrast, most of the previous literature calculates savings from refinancing based on differences to the PMMS rate, thus missing out on potential differences in actual interest rates obtained by borrowers with different observable characteristics. Second, since we observe the purpose of the new loan, we can focus on actual rate refis, instead of the broader category of prepayments, which otherwise also includes cash refis, and sold properties and is a common proxy for refinancing in the literature. Finally, in addition to debt to income ratios calculated at the time of closing the original loan, we observe actual incomes at the time of closing both the original loan and the refinancing loan (i.e. the new loan). We can thus directly speak to the role of income as a predictor of savings from refinancing.

To do so, we study differences in the probability of refinancing across the income distribution, and differences in actual savings from refinancing conditional on refinancing. To do so, we estimate the following equation with different outcome variables:

$$\begin{aligned} y_{it} = \alpha + \sum_{j=2}^{5} \beta_j * Income\ quintile_{ji} + \gamma * Wave\ 2020_{it} \\ + \sum_{j=2}^{5} \phi_j * Income\ quintile_{ji} * Wave\ 2020_{it} + \delta * X_{it} + \epsilon_{it} \end{aligned} \tag{1}$$

Where y_{it} measures the outcome of interest for mortgage i in period t, $Wave$ 2020 is a dummy variable indicating if the observation corresponds to the 2020 window of analysis, and X_{it} is a vector of loan-level controls that will be added gradually across models. The omitted category is the bottom quintile of the income distribution during periods previous to 2020.To capture refinancing activity for the entire portfolio of Freddie Mac loans, we weight matched prepayments by the probability of being matched conditional on observable characteristics. Appendix A describes the matching process.

Depending on the outcome variable, this specification allows us to characterize refinancing activity or savings from refinancing across the income distribution, before and during the pandemic. This specification also provides direct estimates for differences across income quintiles and over time. For example, each coefficient β_j represents the difference in refinancing activity or savings from refinancing between the jth and the bottom quintiles of the income distribution, in periods previous to 2020. We refer to the difference between top and bottom quintiles of the income distribution as the *refinancing income gap*, and we use it as summary measure of inequality in refinancing activity and savings from refinancing, depending on the outcome variable. β_5 is our estimate of the refinancing income gap before pandemic. In turn, $\beta_5 + \phi_5$ is our estimate for the refinancing income gap during the 2020 refinancing wave, and ϕ_5 represents the change in the refinancing income gap before and during the pandemic. This specification also allows us to recover changes in refinancing activity and savings from refinancing within each quintile before and during the pandemic. For example, the coefficient γ represents the difference in refinancing activity or refinance savings for mortgages in the bottom quintile of the income distribution, and $\gamma + \phi_j$ represents our estimate of the change in refinancing activity or refinance savings in the jth quintile of the income distribution, holding everything else constant.

The first outcome variable we use, is a dummy variable that takes the value of one when a mortgage is refinanced, and zero otherwise. We study how these change with and without controlling for a set of off- the-shelve observable characteristics, namely zip code fixed effects, loan age, FICO score, LTV, original interest rates, and unpaid balance. The results are presented in Table 2.

Figure 2 shows our estimates of refinancing inequality change as we include different sets of control variables. Our first model doesn't use any control variable (Column 1 of Table 2) We can see that the refinancing income gap went from 6.73 percentage points (pp) before 2020, to 12.47 pp in 2020. We then add a first set of control variables: zip code FE, loan age, FICO score and LTV (Column 2 of Table 2). These control variables explain 66% of the refinancing income gap before 2020 (the estimate in column 2 for the refinancing income gap before 2020 decreases to 2.3 pp. This is a 4.43 pp reduction, from a basis of 6.73). In contrast the same set of controls explain only 22% of the refinancing income gap in 2020 (the estimate for the refinancing income gap during 2020 decreases to 9.7 pp. This is a 2.77 pp reduction from a basis of 12.47). We then show our estimates for the refinancing income gap before and during 2020, when we include our full set of controls (Column 3 of Table 2). In addition to the control variables used in column 2, we also include unpaid balances and original interest rates. We find that the our full set of controls explains 94% of the refinancing income gap before 2020 (column 3 shows an estimate of the refinancing income gap of only 0.42 pp. This is a reduction of 6.31 ppm, from a base of 6.73 pp without controls). In contrast, the same set of controls can only explain 52% of the refinancing income gap during 2020 (our estimate with the full set of controls is 5.98 pp, representing a 6.79 pp reduction from a base of 12.47 in column 1). Including our full set of control variables we can see that during 2020 the difference in refinancing activity between the top and bottom quintiles of the income distribution was 14 times higher than before 2020 (5.98/0.42).

We then restrict the analysis to mortgages that went through a refinancing transaction and are part of our matched transaction data. For them, we study the distribution of savings from refinancing conditional on refinancing. We do so estimating equation 1 with two additional outcome variables. The first one is the interest rate differential between the new refinancing loan and the original refinanced loan. The second one captures the value of savings from refinancing expressed in dollar terms, defined as present value of the difference in outflows under the old and new interest rates over the expected life of the loan. The expected life of the loan is parametrized by Agarwal, Driscoll and Laibson (2013). We note that we do not attempt to provide a causal interpretation to this analysis. Among other things, income clearly affects both average savings conditional on refinancing, as well as the probability of refinancing. Thus selection into refinancing is not random. Our goal is to describe average savings for individual across the income distribution who go through a refinancing transaction. We do so comparing average savings conditional on refinancing over a discrete set of (income) categories (Angrist, 2001).[3]

The results are presented in Table 3. In columns 1 to 3, we use interest rate differentials as the dependent variable, first without borrower level controls and then with our full set of borrower level controls. Here we can more clearly study variations in contract terms across the income distribution. In column 3 we can see that before 2020, conditional on refinancing, borrowers in the bottom quintile of the income distribution received a reduction of 166 basis points from their original interest rates (omitted category). Relative to them, comparable borrowers in the top quintile of the income distribution received slightly lower interest rate reductions of

[3]We choose to retain a linear model for this part of the analysis (instead of a two-step model or a conditional-on-positive Tobit estimate) to emphasize its descriptive nature. Nevertheless, in the next part of the analysis we consider a Tobit model for robustness purposes, and find that our results are very similar in all cases.

150 basis points (166 - 16). We note that this regression controls for original interest rates and unpaid balances, as well as zip code fixed effects and standard borrower level controls. During 2020, all borrowers received larges interest rate reductions (the coefficient for wave 2020, as well as its interaction with income quintiles are all positive and significant), but the improvement in contract terms for borrowers in the top quintile of the income distribution was larger than for borrowers in the bottom quintile of the income distribution (See Figure 3, panel A). Individuals in the bottom quintile of the income distribution improved their interest rate differentials by 16 basis points to reach a rate differential of 182 bps. Individuals in the top quintile of the income distribution improved their interest rate differentials by 36 bps to reach a rate differential of 186 bps. The slight edge of lower-income individuals who were refinancing before 2020 in terms of interest rate reductions disappeared in 2020.

In columns 4 to 6, we use dollar savings as the dependent variable, first without borrower level controls and then with our full set of borrower level controls. In column 6 we can see that before 2020, borrowers in the top quintile of the income distribution had $1,117.86 more in savings than comparable borrowers in the bottom quintile of the income distribution. This difference in savings increase to $3,532.16 in 2020 (See Figure 3, panel B).

We now describe average savings from refinancing on the entire portfolio of active mortgages, incorporating both the probability of refinancing, as well as savings from refinancing conditional on refinancing. We define savings from refinancing on the entire portfolio as a continuous variable that takes the value of zero for all mortgages which did no go through a refinancing transaction, or the corresponding value of savings from refinancing conditional on refinancing (as defined before) for mortgages that indeed went through a refinancing transaction.

Savings from refinancing, so defined over the entire portfolio, can be interpreted as a censored dependent variable. Therefore, we follow two approaches to estimate how they change across the income distribution. We estimate equation 1 as before and, in addition, we also estimate a Tobit model using the aforementioned equation as a latent linear index. The later approach imposes functional form assumptions to explicitly model savings as a variable censored at zero: a latent linear index feeds into normal distribution censored at zero which is then estimated by maximum likelihood. In contrast, the former takes a more agnostic approach describing changes in average savings over a set of discrete categories, namely income quintiles before and after 2020. (Angrist and Pischke, 2008; Angrist, 2001). Our results are robust to these different functional form assumptions.

Table 4 shows the results of estimating equation 1, as we gradually add control variables. Columns 1 and 4 show the coefficients of interests without loan-level controls. The difference in savings from refinancing between the top and bottom quintile of the income distribution before 2020 amounts to \$879 (or \$1,690 when measured with a Tobit model [4]). During 2020, the difference in refinancing activity between the bottom and top quintiles of the income distribution increases to \$2,288 (or \$5,009 when measured with a Tobit model). However, this change could be driven by changes in the composition of loans that became newly in-the money during the observation periods. To address this challenge, we gradually add a rich set of control variables to asses the sensitivity of our estimates. In columns 2 and 5 we add flexible controls for borrower and loan attributes, namely dummy variables for FICO score bins, loan to value bins, and bins of loan age. We find that this basic set of controls explains 35% (36%) of the difference in savings between the top and bottom quintile of the income distribution, which now accounts for \$453 (\$1,087) when estimated with the OLS model (Tobit model).

Finally, in Columns 3 and 6 we also include two additional controls which largely capture the potential savings from refinancing activity: baseline interest rate, and unpaid balance. Thus, Columns 3 and 6 estimate the role of income on refinancing activity for individuals with the same FICO, loan age, loan to value, and with the same unpaid balance and interest rate. In column 3, with OLS estimates, we find that the gap in refinancing activity between the top and bottom quintiles of the income distribution is fully explained, and even changes sign to reach a level of -\$131. In column 6, with a Tobit model, we find consistent results. Our full set of controls explains 77% of the difference in savings across the top and bottom quintiles of the income distribution, leading to a final difference of \$386. Nevertheless, even among comparable mortgages, the difference in savings from refinancing across the income distribution increased significantly. In Column 3, we can see that the difference in savings accounts to \$1,313 with our OLS estimates. This represents an 11 times increase from pre-2020 levels ((1,313 + 131)/131). Similarly, in Column 6 our Tobit estimates show that the difference in savings increased 9.8 times (3,798/386).

Panel A Figure 4 plots our estimates for savings from refinancing before and after the pandemic by income quintile, controlling for changes in the composition in the pool of newly in-the-money borrowers. We plot the coefficients β_j and $\beta_j + wave_{2020} + \phi_j$ from column

[4] For the Tobit models, the savings gap expressed in dollar terms is calculated as $(\frac{x1*\beta}{\sigma}) + \sigma * \phi(\frac{x1*\beta}{\sigma}) - (\frac{x1*\beta}{\sigma}) + \sigma * \phi(\frac{x1*\beta}{\sigma})$, where $, \phi$ are standard normal CDF/PDF, σ is the Tobit scale parameter and $x1 * \beta$ $(x0 * \beta)$ refers to regression coefficients evaluated at baseline.

3 of Table 4 for each quintile after adding up in both cases the prepayment rate in the omitted category (bottom quintile of the income distribution before 2020). We can see that before 2020, savings from refinancing were very similar across the income distribution, and even slightly lower for higher-income individuals. However, in 2020 savings from refinancing increased substantially, specially in the upper segments of the income distribution. In the second quintile of the income distribution, there was a $358 increase in refinancing savings from a basis of $121. In the third decile, there was a $661 increase from a basis of $111. The fourth quintile shows an increase of $1051, from a basis of $61. Finally, the top quintile experienced increase of $1520, from a basis of $5. Panel B shows the analogous results without controls.

Back of the envelope calculations using our estimates for refinancing savings across the income distribution, imply a gap in refinance savings of $5 billion between the top quintile of the income distribution and the rest of the market, i.e. if individuals in lower segments of the income distribution (without controlling for observable characteristics) were receiving the same savings from refinancing as individuals in the top quintile of the income distribution, they would capture an additional $5 billion in refinance savings. [5]

3.2 McDash Data

We now turn the analysis to the data from McDash Analytics. This is a very standard data set used in the literature which in principle represents around 60% of the market (although data quality issues leads us to drop a lot of observations). While this dataset suffers from certain limitations, we use it here for robustness purposes and to provide external validity to our analysis. In the McDash data, we only observe when a mortgage is prepapid, but cannot distinguish refinancing transactions from other types of prepayments. We thus proxy refinancing activity with prepayments and provide a variety of robustness test to argue that our results with McDash data are also driven by rate-refinances. Similarly, we do not observe income, but instead use an estimate of monthly income based on debt-to-income ratios reported at origination.

[5] To calculate this number, we start from a market size of 30.9 million mortgages with fixed rates at 30 years maturity (American Housing Survey, with data as of 2017), and extrapolate our estimates for average savings for mortgages that become newly in the money in each income quintile j ($In\ Money\ IQ_j$) and difference in refinance savings for each quintile j relative to the top quintile of the income distribution ($Gap\ Q_{j5}$). To do so, we use the results for 2020 in column 1 of Table 4, also depicted in Panel B of Figure 4. Specifically, we apply the following formula $\sum_{j=1}^{4} 30.4 * 0.2 * In\ Money\ IQ_j * Gap\ Q_{j5} =$ 4,964,017,020.

As before, to study the role of income on refinancing activity we regress a dummy variable indicating whether or not a mortgage was prepaid on a rich set of loan level covariates, income quintiles and their interaction with a binary variable identifying observations corresponding to the pandemic period.

Table 5 shows our main results from estimating Equation 1. Column 1 shows the coefficients of interests without any control variable. The difference in refinancing activity between the top and bottom quintile of the income distribution before 2020 amounts to 2.7 percentage points. After 2020, the difference in refinancing activity between the bottom and top quintiles of the income distribution increases to 7.4 percentage points (0.027 + 0.047). However, this change could be driven by changes in the composition of loans across zip codes that became newly in-the money during the observation periods. To address this challenge, we gradually add a rich set of control variables to asses the sensitivity of our estimates. In column 2 we add zip code fixed effects flexible controls for borrower and loan attributes, namely dummy variables for FICO score bins, loan to value bins, and bins of loan age.[6] Holding these characteristics constant, we find that the gap in refinancing activity between the bottom and top quintile of the income distribution went from 110 basis points (bps), to 550 bps. Finally, in Column 3 we also include original interest rates and unpaid balances. Column 3 thus estimates the role of income on refinancing activity for individuals in the same zip code, with the same FICO, loan age, loan to value, and with the same interest rates and unpaid balances (which determine potential savings from refinancing). We find that the gap in refinancing activity between the top and bottom quintiles of the income distribution increases from -40 bps before 2020, to a new level of 830 bps (-0.004+ 0.087).

Panel a) of Figure 5 plots our estimates for the levels of refinancing activity before and after pandemic by income quintile, after controlling for changes in the composition in the pool of newly in-the money borrowers. We estimate Equation 1 and plot the coefficients β_j and $\beta_j + wave_{2020} + \phi_j$ for each quintile after adding up in both cases the prepayment rate in the omitted category (bottom quintile of the income distribution before 2020). We can see that refinancing activity increased at all income levels, compared to previous waves of large drops in interest rates. However, refinancing activity in upper segments of the income distribution increased a lot more. In the bottom quintile, there was a 90 bps increase in refinancing activity. In the second quintile, the increase was of 260. In the third quintile,

[6] For credit score: 740+, [720,740), [680,720), [640,680),640-. For ltv: 95+, (90,95], (85,90], (80,85], (75,80],(70,75], (60,70], (0,60]. For age: 1 year - , 1-2 years, 2-3 years, 3-5 years, 5-7 years, 7+ years.

there was a 460 bps increase. The fourth quintile shows an increase of 730 bps. Finally, the top quintile experienced increase of 960 bps. Through the paper, we focus on the difference in refinancing activity between the top and bottom quintiles of the income distribution, as a summary measure of refinancing inequality.

To show that this increase in inequality is not the result of pre-pandemic trends, in Appendix B, we study the refinancing income gap over the 15 months periods previous to the pandemic, and compare its magnitude to the refinancing income gap during 2020. We find that while interest rates were consistently coming down from their 2018 peak, refinancing inequality was not following an upward trend. Instead, a dramatic increase takes place between February and June 2020, leading to inequality levels 7.3 times higher than in the 15 months immediately previous.

We investigate heterogeneities in the magnitude of increase inequality over time, by splitting our sample based on FICO score, and on potential savings from refinancing. In Table 6, we estimate our preferred specification (Equation 1) splitting the sample across three selected variables of interest. Columns 1 and 2 split the sample based on original interest rates. Column 1 shows that for individuals with the higher original interest rates, the difference in refinancing activity between the top and bottom quintile of the income distribution was 90 bps. But during the first few months of 2020, the refinancing income gap grew to 330 bps (0.009 + 0.024). Individuals with lower interest rates also experienced an important increase in the refinancing income gap from 90 bps to 510 bps, as can be seen in Column 2. This increase is larger than the increase in inequality experienced by individuals with the highest interest rate incentives. Similarly, in Columns 3 and 4, we split the sample based on unpaid balances. Column 3 shows that for individuals with balances above the median, the refinancing income gap increased from 120 bps to 540 bps. For individuals with balances below the median, Column 4 shows that the refinancing income gap also increased from 70 bps, to 270 bps. Finally, in Columns 5 and 6 we split the sample by FICO score. Borrowers with FICO scores greater than 740 are in Column 5. For them, we see that the refinancing income gap increased from 110 bps to 490 bps between 2020 and previous periods of similar interest rate drops. The increase is comparable from the change experienced by borrowers with FICO score below 740. For them, Column 6 shows an increase in the refinancing income gap from 70 bps before the pandemic to 440 bps in the first months of 2020.

4 Intensity of the COVID-19 Pandemic and Refinancing Activity

For the second component of our analysis, we estimate the impact changes in refinancing inequality as the severity of the pandemic increases. We measure the impact of the severity of the pandemic exploiting variation in COVID case rates. We explain the effect of the pandemic on inequality by adding a series of controls for local economic conditions: mobility restrictions, unemployment insurance claims, and financial hardship (mortgages in forbearance). For this part of the analysis, our dataset consists of a monthly panel that follows mortgage refinancing between February 2020 and July 2020. Specifically, we consider mortgages that were not in-the-money in February 2020 and became in-the-money in any of the subsequent periods until July 2020. For these mortgages we have monthly observations between the first month in which they turn in-the-money and up until the month in which they are prepaid, along with a vector of variables tracking the impact of COVID-19 at the county or state level.

To estimate the impact of the pandemic on the refinancing income gap, we estimate the following equation.

$$\begin{split} y_{izct} = \alpha_z + \alpha_t + \sum_{j=2}^{5} \beta_j * Income\ quintilej_{ji} + \sum_{k=2}^{5} \gamma_k * Severity\ Q_{kizct-1} \\ + \sum_{k=2}^{5} \sum_{j=2}^{5} \phi_{jk} * Income\ quintile_{ji} * Severity\ Q_{kizct-1} + \delta * X_{it} + \epsilon_{izcgt} \end{split} \tag{2}$$

Where y_{izct} indicates if mortgage i in zip code z and county or state c was refinancined in period t; Income quintile ji represents a set of dummy variables indicating if mortgage i belongs to income quintile j; $Severity\ Q_{kizct-1}$ is a dummy variable indicating if mortgage i in zip code z in county or state c belongs to the quintile k of the distribution of COVID severity in month $t-1$; and X_{it} is a vector of loan-level controls. To reflect that refinancing applications take between 1 and 1.5 months to be processed, we use a one month lag of the variables to measure the severity of the crisis. This way, the refinancing activity after households increased their time at home in month t-1, is measured in month t. This flexible specification allow us to identify non-linearities in the effect of the pandemic on refinancing inequality.

The coefficient β_5 represents the refinancing income gap in geography-months where the pandemic hit the least. The coefficient ϕ_{5k} represents increases in the refinancing income gap for mortgages in geography-months that lie on the jth quintile of the distribution of COVID severity, relative to those in the bottom quintile. Since we have zip code fixed effects, the coefficients measure changes in refinancing activity for mortgages with high or low income within a zip code over time, as the pandemic hit with different levels of severity. Since we have time fixed effects, we are also controlling for the effect of macro shocks affecting all geographies on a given month. The full set of coefficients is presented in Table A5. In the following we plot and interpret the main coefficients of interest for the analysis.

Figure 6 shows the evolution of the refinancing income gap across the distribution of COVID Severity using case rates as a direct measure of severity. Specifically, we plot the coefficients β_5 and $\beta_5 + \phi_{5j}$ for the bottom quintile of the distribution of COVID case rates, and the subsequent quintiles j=2 to 5 (black line). The refinancing income gap in the bottom quintile of the distribution of COVID case rates is -0.39 pp. The refinancing income gap in the second quintile of the distribution of COVID case rates is 1.59 pp. As the severity of the pandemic increases, refinancing inequality increases to 2.49 pp, 2.77 and 2.34 pp in the third, fourth and top quintiles of the distribution of COVID Severity. The impact of the pandemic has a slight inverse u-shape. The orange bars show the slope of black line at different points of the severity distribution, i.e. changes in the refinancing income gap as we move to county-months where the pandemic hit the hardest: all differences are statistically significant. We also plot the coefficients ϕ_{5j} to show more explicitly the changes in refinancing inequality relative to the bottom quintile of the distribution of COVID Severity and their statistical significance (blue bars). Overall, refinancing inequality increases by more than 700% as we more from the least affected to the more affected county-months (increase of 2.76 pp from a basis of -0.39 pp). This Figure is based on Column 1 of Table A5.

To explore the mechanisms behind the effect of COVID case rates on refinancing inequality, we restrict our sample to mortgages in areas for which there is coverage of GPS data (county-month data), unemployment insurance claims (county-month data), and mortgages on forbearance (state-month data). We reach a total of 1.2 million mortgage-months. With this restricted sample, we re-estimate equation 2 four times, each time using again COVID case rates, as well as our different measures of the severity of the pandemic on local economic conditions. [7] As before, the full set of coefficients is presented in Table A5. We plot and interpret the main coefficients of interest in Figure 7.

Figure 7 plots our estimates for the refinancing income gap across geography-months with different levels of COVID Severity. The black line represents projected levels of the refinancing income gap in county-months that fall in different quintiles of the COVID Severity distribution. These correspond to β_5 and $\beta_5 + \phi_{5j}$ for the bottom quintile of COVID severity and for quintiles j=2 to 5, respectively. As before, we can see a slight inverse U-shape when we measure severity with case rates (Panel a). The same shape is present when we measure the severity of the pandemic time spent at home (Panel d). Forbearance has a sustained positive correlation with refinancing inequality (panel b). Unemployment insurance claims on the other hand have very small effects across the board (panel c). The blue bars represent changes in the refinancing income gap relative to the bottom quintile of COVID severity, as we move to higher quintiles of COVID severity (i.e. ϕ_{5j} with j=2 to 5, in Equation 2). We note that the largest increases in inequality result from moving from the bottom quintile to any of the other quintiles. The effects from moving to one quintile to the other are smaller. The plots are based on columns 3 to 5 of Table A5.

We thus present a summary measure of the severity of the pandemic as follows: we define geography-months of low severity as those in the bottom quintile of the severity distribution and geography-months of medium to high severity as those in quintiles 2,3,4 or 5 of the severity distribution. Figure 8 shows the refinancing income gap in geography-months of low or medium to high severity, measured with four different variables. [8] There are several reasons why forbearance, unemployment insurance claims and time spent at home can affect refinancing inequality. We proceed to discuss them one by one.

Financial Hardship and Forbearance To mitigate the economic impact of the pandemic, several policies went into effect during the first few months of 2020. The CARES act offered up to 180 days of mortgage forbearance to individuals experiencing financial hardship due to the coronavirus pandemic. It is likely that lower-income individuals will be more likely to face financial hardship and request forbearance under the CARES Act. Under the rules of

[7] For robustness, we also replicate the analysis with unemployment insurance claims measured at the state-month level. This increases our observations to 2.3 million mortgage-months. The results are qualitatively the same and are available from the authors upon request.

[8] Specifically, we re-estimate equation 2 four times with our four measures of severity, using a binary variable for severity instead of quintiles as before. This binary variable takes the value of 0 when a mortgages in in a geography-month in the bottom quintile of the corresponding severity distribution and 1 otherwise.

the program, requesting forbearance does not preclude individuals from refinancing to take advantage of low interest rates. Furthermore, refinancing is still advantageous for those who requested forbearance, since refinancing not only reduce the debt service burden, but also reduces the overall long-term cost of debt. However, individuals may not be fully aware that forbearance does not preclude them to refinance. Furthermore, since forbearance removes a short-term need for liquidity, present bias and other cognitive biases may lead consumers to procrastinate and let go of the long-term savings from refinancing. The second set of bars in Figure 8, shows that increases forbearance rates (measured at the state-month frequency) are correlated with large increases in refinancing inequality. In state-months of low forbearance rates, refinancing inequality accounts for 0.71 pp. In state-months of medium to high forbearance, refinancing inequality reaches a level of 2.36 pp.

Initial Unemployment Insurance Claims It is well documented that the pandemic led to very large increases in unemployment across the US. We explore the role of unemployment as an additional mechanism behind the effect of the COVID-19 pandemic on refinancing inequality. The data is at the county-month level. We find no significant effects of increases in unemployment insurance claims (see the third set of bars in Figure 8). One potential explanation is that during the period of analysis unemployment insurance replacements rates were above 100%, meaning that they are eligible for benefits which exceed lost wages (Ganong, Noel and Vavra, 2020).

Time spent at home During the pandemic, mobility decreased significantly both as a result of stay at home orders (Alexander and Karger, 2020), and as a result of individuals taking precautions out of their own initiative (Chetty et al., 2020). During the COVID crisis, individuals with positions that allow telework have been more likely to keep their jobs (which had higher salaries to begin with (Adams-Prassl et al., 2020)). Furthermore, individuals with higher income have also hunkered down, reducing spending in activities that require contact, and reducing their mobility the most (Chen, Qian and Wen, 2020). This leaves the upper segment of the income distribution with lower commuting times, less activities outside home, and a constant stream of income. We argue that higher-income individuals, who are inherently better able to refinance their properties, will have more free time on their hands during the pandemic, and this could lead them to refinance their mortgages at higher rates. The last set of bars in Figure 8 shows that time spent at home is positively correlated with refinancing inequality. In county-months with small increases in time spent at home, refinancing inequality accounts for 1.13 pp. In county-months of medium to high increases in time spent at home, refinancing inequality reaches a level of 2.25 pp.

However, the measurement of time spent at home, unemployment insurance claims and forbearance is coarse. We have information aggregated at the county or state level. This lack of granularity introduces measurement error at the individual level since local economic conditions may affect households within the same geography differently. Furthermore, these three variables are highly correlated and it is not clear where the influence of one or the other ends. We thus interpret the effect of these three variables as reflecting one single shock to local economies and use them to measure how much of the increases in refinancing inequality tied to the pandemic can be explained through its effect on local economic conditions. We estimate equation 2 using COVID Case rates as our main measure of COVID Severity and include time spent at home, forbearance, and unemployment insurance claims as controls for local economic conditions. [9]

Figure 9 plots the results of estimating the effect of higher COVID case rates with and without controls for local economic conditions, using the restricted sample with full coverage of local variables in both cases. Specifically, the blue and orange bars represent coefficients $\phi_5 j$ with and without controls for local economic conditions, respectively (estimates presented in columns 2 and 5 of Table A5, respectively). We can see that our set of controls for local economic conditions explain about 40% of the increase in inequality for intermediate levels of case rates and up to 80% in county-months where the pandemic hit the hardest. Going from the first to the second quintile of the distribution of case rates leads to a 1.5 pp increase in the refinancing income gap, but this effect shrinks to 0.9 pp when we add controls for local economic conditions ((1.5-0.9)/1.6 = 0.39). Similarly, going from the first to the top quintile of the distribution of case rates increases the refinancing income gap by 1.5 pp, but this effect shrinks to a non-significant 0.4 pp when we add controls. This suggests that local economic conditions explain 74% of the increases in inequality attributable to the pandemic ((1.5-0.4)/1.5 = 0.74).

[9] Specifically we include quintiles of time spent at home, forbearance and unemployment insurance claims, as well as their interactions with income quintiles.

5 Mechanisms behind increases in refinancing inequality: borrower or lender behavior?

We have shown that increases in refinancing activity were concentrated on individuals with higher income, and that this is partly attributed to the pandemic, since the effect is stronger in areas and months where the COVID crisis hit the hardest. To understand the mechanisms behind this results, we explore whether changes in refinancing inequality were driven by lenders' behavior or borrowers' behavior. We note that both lender and borrower behavior are equilibrium outcomes, but we argue that splitting the analysis this way is nevertheless useful to describe frictions in this market during the pandemic that could potentially explain changes in refinancing inequality.

5.1 Lender behavior

During the period of analysis, most lenders saw increases in the number of refinancing applications received. If lenders faced binding capacity constraints and were not able to process all the applications they received, they could have prioritized applications of higher income individuals (which are arguably more profitable as they tend to carry larger balances). To explore this hypothesis, we first provide suggestive evidence for the presence of capacity constraints and then, we explore if this capacity constraints had a differential effect for applicants of different income.

Figure 12 shows the density distribution of lender-level growth rates in the number of applications submitted to Freddie Mac's LPA tool between the first six months of 2020, and the first six months of 2019. We can see that the mode is close to 100%, implying that, many lenders submitted twice as many applications to the LPA tool during the first six months of 2020, compared to the first six months of 2019. The growth in applications was unequal across lenders, with some experiencing increases of more than 700% and some others experiencing small decreases in the number of applications submitted.

Furthermore, lenders who experienced the largest growth rates in applications, show a decrease in the number of applications that were eventually funded by Freddie Mac. While we do not observe if an application is approved by a lender or not, we observe when an application is ultimately funded by Freddie Mac. Being funded by Freddie Mac is a sufficient condition to conclude that such application was approved by the lender. We argue that the unobserved changes in funding rates are uncorrelated with borrower income and with lender-level applications growth rates. We thus attribute changes in funding rates, conditional on borrower characteristics, to lender behavior.

Figure 13 shows the change in the percentage of applications eventually funded by Freddie Mac, across the distribution of lender-level growth rates in applications. Lenders in the bottom decile of the growth rate distribution experienced a 4.8 pp increase in the fraction of applications eventually funded by Freddie Mac. In contrast, lenders in the top decile of the distribution of application growth experienced a 6 pp decrease in the fraction of applications eventually funded.

If lenders are not able to process all applications, they may prioritize applications of higher-income individuals (which are likely to be more profitable). To explore this hypothesis, we aggregate the application-level data at the lender-by-income decile level. For each lender-by-income decile we calculate the change in funding rate between the first six months of 2019 and the first six months of 2020. In Column 1 of Table 7 we regress changes in funding rates at the lender-by-income quintile level, on lenders' application growth rate, controlling for differences in the income of applicants across lenders. Consistent with Figure 7 we find that growth in applications is negatively correlated with changes in funding rates. Column 2 interacts lender-level growth rates and deciles of applicants' income. We find that there are no clear differences in the effect of growth rates on the funding rate across the income distribution. Specifically, the interaction of growth rates with high income deciles is negative but not statistically significant. This constitutes evidence against the hypothesis that refinancing inequality is driven by capacity constraints leading lenders to prioritize high income borrowers in the extensive margin.

We now turn to the impact of lender behavior on the intensive margin of refinancing inequality. Specifically we describe the evolution on the time it takes lenders to process applications of borrowers with different levels of income. We focus on applications that were eventually funded by Freddie Mac. Figure 14 shows that, compared to previous waves of low interest rates, applications are being processed slightly faster than before. Furthermore, conditional on an application resulting in a loan funded by Freddie mac, income does not predict differences in processing times. In 01/2019, it took an average of 45 days to process applications of borrowers in the bottom decile of the income distribution, while it took 44.2 days to process applications of borrowers in the top decile (i.e. it took the same time in both cases). In 06/2020, applications of lowest income borrowers take about the same as before (44 days), and if anything only applications of borrowers with the highest income are being delayed by a few days (48.9 days).

5.2 Borrowers' profile and behavior

Income of applicants We use Freddie Mac application data to describe the income of applicants. We calculate income deciles of the mortgage portfolio, and plot the distribution of applications across those deciles. We find that 18.6% of mortgage refinancing applications come from borrowers with an income that places them in the top 10% of the income distribution of Freddie Mac's mortgage portfolio. If we stratify based on the income of borrowers newly in the money in 2020, we find that 17.5% of mortgage refinance applications come from borrowers in the top 10% of the income distribution. This suggests that lower-income borrowers for which interest rate differentials suggest that refinancing is financially beneficial, are nevertheless less likely to apply than their higher-income counterparts. If we extend the analysis to consider all outstanding loans (instead of just those that became newly in the money), we get a similar result.

Moving patterns across the income distribution The prospect of moving to a new home is one of the determinants of the decision to refinance a mortgage: individuals who expect to move sooner have less of an incentive to refinance, holding everything else constant. We explore the possibility that increases in refinancing inequality during the first half of 2020 were driven by differences in the probability of of moving to a new house across the income distribution. We test the hypothesis that low-income individuals would be more likely to move as a result of a negative shock to local economic conditions. Using the matched transactions data from Freddie Mac for newly in the money mortgages, we compare the probability of prepaying an existing mortgage and then buying a new home, across the income distribution and over time (for the five waves of low interest rates considered in the main analysis). Figure **??** shows that, on average, the probability of prepaying to buy a new home in periods of low interest rate before 2020 was 0.93% for borrowers in the bottom quintile of the income distribution. This probability shows a small decrease of 0.19 pp during the first half of 2020. Similarly, for borrowers in the top quintile of the income distribution, the probability of prepaying an existing mortgage to buy a new home before 2020 was 1.27%. During 2020, this group experienced a small decrease of 0.15 pp. Therefore, the difference over time in new purchases is about the same across the income distribution. If anything, high-income borrowers show slightly smaller decreases in the probability of prepaying an existing mortgage to buy a new home. We thus conclude that the increases in refinancing inequality are unlikely to be driven by a different evolution of moving patterns across the income distribution.

Delinquency patterns across the income distribution Another reason why low-income borrowers did not increase their refinancing activity at the same rate as high-income borrowers could be that low-income borrowers experienced a more than proportional increase in default rates during the period of analysis. For a refinancing transaction to make sense, borrowers need to retain their loans for a sufficiently long period of time. For individuals with a high probability of default, refinancing may not be optimal. To explore this possibility, we compare trends in delinquency for borrowers in the top and bottom quintiles of the income distribution. Figure 11 shows the evolution of the fraction of newly in the money mortgages that were ever delinquent during the five month windows of low interest rates considered in the analysis. We can see that, compared to periods of low interest rates previous to 2020, during the first months of 2020 individuals in the bottom quintile of the income distribution increased the probability of delinquency by 3.06 pp, from a basis of 4.34%. During the same period, individuals in the top quintile of the income distribution increased their probability of delinquency by 4.90 pp from a basis of 1.87%. Thus, while in general low-income borrowers have higher probabilities of delinquency, the evolution of delinquency cannot explain the change in refinancing patterns observed during the first months of the pandemic. If anything, high income borrowers are increasing their delinquency probabilities at a higher rate than low income borrowers.

6 Refinancing of low income borrowers and aggregate consumption

Motivated by our results we pose the hypothesis that, the effect of mortgage refinancing on spending would be higher if lower income borrowers were appropriating larger share of savings from mortgage refinancing.

Our main results on the distribution of savings from refinancing introduce a tension between two well documents forces affecting consumer spending. On the one hand, several papers have shown that refinancing affect consumer spending (Di Maggio et al., 2017; Boyce et al., 2012; Agarwal et al., Forthcoming; Berger et al., 2018, 2019; Eichenbaum, Rebelo and Wong, 2018;

Agarwal et al., 2017). On the other hand, it is also well known that low income individuals have higher marginal propensities to consume, compared to their higher income counterparts. We show that savings from refinancing are concentrated on high income individual. It is thus natural to hypothesize that the effectiveness of the mortgages refinancing channel for the transmission of monetary policy would be higher, if low income borrowers were capturing a higher share of savings from refinancing.

Formally testing this hypothesis is beyond the scope of this paper and we think is an interesting avenue for future research. We nevertheless provide graphical evidence of a positive relation between refinancing activity of low income borrowers and spending. Specifically, we correlate refinancing activity of borrowers in the bottom quintile of the income distribution, to credit and debit card spending. We measure refinancing activity between February and June 2020, and spending during the month of July 2020. February to June 2020 corresponds to the 2020 wave of refinancing activity that we have compared to previous periods, through the paper. To allow enough time for the savings from refinancing to influence consumption, we measure consumption at the end of this wave of refinancing activity. In that sense, we are capturing the cumulative effect of refinancing over a 5 month period, on spending. We define quintiles of the income distribution based on the pool of borrowers with active mortgage in the US as of 2020. We measure card spending with data from Affinity Solutions compiled by Chetty et al. (2020), also at the county level. Our spending data is in the form of an index reflecting changes in spending relative to January 2020. We calculate the refinancing rate of borrowers in the bottom quintile of the income distribution of mortgage borrowers at the national level. This measure is defined as the number of mortgages of low income borrowers in the county, divided by the number of active mortgages of low income borrowers in the same county.

Panel a) of Figure 17 shows a county-level heat map for the change in refinancing rates of low income borrowers between the period Feb-June 2020 and previous waves of low interest rates considered in the analysis. Panel b) shows a county-level heat map of the spending index of Chetty et al. (2020) as of July 2020. We can see that the counties with the largest increases in refinancing activity of low income individuals are also the counties with largest values for the spending index, which represent counties with the strongest spending recovery seven months into the pandemic.

These results are only correlations and there are a a number of omitted variables that could jointly affect refinancing activity of low income borrowers and consumer spending. For example, the severity of the pandemic itself negatively affects refinancing of low income borrowers and consumption. However, we argue that the hypothesis posed in this section is a potential avenue for future research.

7 Robustness and Alternative Explanations

Our McDash dataset comprises the portfolio of active mortgages handled by mortgage servicers that report to McDash. For all active mortgages in the portfolio, we observe whenever a mortgage is prepaid. However, we cannot distinguish if a mortgages was refinanced, or liquidated for good as part of a sell-and-buy transaction. We argue that our results on the McDash dataset are driven by refinancing transactions and not by sell-and-buy transactions. The fact that our results are even stronger in the matched-transaction dataset gives us confidence that our McDash results are capturing refinancing activity. Nevertheless, in this section we provide a battery of robustness tests to support our use of prepayments in the McDash analysis as a valid proxy for refinancing transactions.

We first note that, as can be seen in several industry reports, the spike in mortgage originations in early 2020 was driven by refinancing activity, and not by home purchases. Haughwout et al. (2020) for example, shows that refinances accounted for more than 600 billion dollars, out of a total mortgage origination volume of 846 billion in the second quarter of 2020. Similarly, the August 2020 report of Urban Institute (2020) shows that, "with rates at historic lows the refinance share [at issuance] is very high; GSE's are in the 71 to 75 percent range."

Nevertheless, one could argue that the relation between prepayment and income observed in our data comes from higher-income individuals buying houses in areas with lower exposure to the pandemic. To investigate this possibility, we perform three complementary tests.

For the first two, we study new originations during the observation period. We first plot the distribution of new originations by loan purpose across the income distribution and over time. Panel A of Figure 16 shows that during 2020 there is a stronger increase in refinancing activity by individuals in the top sections of the income distribution. Panel B of the same figure shows that, in contrast, new purchases in 2020 are fewer and are not more concentrated than before in the top sections of the income distribution. Appendix D shows a break down by type of refinance, including rate refinances, cash refinances, and other (unclassified) refinances.

Then we study the geographic distribution of high prepayment rates and high levels of mortgages originated for new purchases. Figure 18 shows that the counties with higher prepayment activity are not the counties with the higher purchases of new homes.

Finally, we pay special attention to cash-out refinances, by which borrowers extract equity from their houses while refinancing. Specifically, we estimate our main equation 1 with the matched transaction database, but using cash-refis, instead of rate-refis as our dependent variable. Table 8 shows the results. In general, before 2020, higher-income individuals had a slightly higher probability of taking out a cash-out refinance, compared to borrowers in the bottom quintile of the income distribution (the bottom quintile refinanced at a rate of 1.13%, whereas the top quintile refinanced at a rate of 1.16%). During 2020, this relationship remains mostly unchanged, with a slight *decrease* in the cash refinancing income gap, i.e., in 2020 individuals in the bottom quintile increase their refinancing activity by 0.9 pp, whereas individuals in the top quintile of the income distribution increase their refinancing rate only by 0.7 pp (0.009 - 0.002). We thus conclude that our McDash results based on prepayments are unlikely to be driven by cash-out refis.

8 Final Comments

In this paper we introduce the concept of refinancing inequality, by which we refer to differences in refinancing activity across the income distribution. We use the refinancing income gap, defined as the difference in refinancing activity between the top and bottom quintiles of the income distribution, as a summary measure to describe refinancing inequality over time, and across geographies.

We find that during the COVID-19 pandemic, refinancing inequality increased significant. Furthermore, increases in refinancing inequality track the severity of the pandemic: areas that were hit the hardest by the pandemic experience a two-fold increase in the refinancing income gap. To produce we result, we estimate the refinancing income gap within zip-code, exploiting county-by-month variation on the severity of the pandemic.

These results have implications for the evolution of wealth inequality in the US for several reasons. First, differences in refinancing activity and in interest rates conditional on refinancing lead to a $5 billion gap in savings between the top quintile of the income distribution and the rest of the market. Second, expansionary monetary policy itself tends to have redistributive consequences through its effect on inflation, which has different impacts on individuals depending on their asset holdings (net borrowers or net lenders): the wealth of higher-income individuals tend to be negatively affected by inflation. We find that during the pandemic, the stabilizing role of expansionary policy on inequality was counteracted by the ability of high-income individuals to appropriate significant savings at a much higher rate, compared to lower-income individuals.

Our results also have implications for the effectiveness of monetary policy, since mortgage refinancing is one of the channels to increase individual spending, but insofar as higher-income individuals have lower propensity to consume, the effectiveness of monetary policy will be hindered.

References

Adams-Prassl, Abi, Teodora Boneva, Marta Golin, and Christopher Rauh. 2020. "Inequality in the impact of the coronavirus shock: Evidence from real time surveys."

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, Tomasz Piskorski, and Amit Seru. 2017. "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. Forthcoming. "Mortgage refinancing, consumer spending, and competition: Evidence from the home affordable refinancing program." *Review of Economic Studies*.

Agarwal, Sumit, Itzhak Ben-David, and Vincent Yao. 2017. "Systematic mistakes of borrowers in the mortgage markets." *Journal of Financial Economics*, 123: 42–58.

Agarwal, Sumit, John C Driscoll, and David I Laibson. 2013. "Optimal Mortgage Refinancing: A Closed-Form Solution." *Journal of Money, Credit and Banking*, 45(4): 591–622.

Agarwal, Sumit, Richard J Rosen, and Vincent Yao. 2016. "Why do borrowers make mortgage refinancing mistakes?" *Management Science*, 62(12): 3494–3509.

Alexander, Diane, and Ezra Karger. 2020. "Do stay-at-home orders cause people to stay at home? Effects of stay-at-home orders on consumer behavior."

Andersen, Steffen, John Y. Campbell, Kasper Meisner Nielsen, and Tarun Ramadorai. 2020. "Sources of Inaction in Household Finance: Evidence from the Danish Mortgage Market." *American Economic Review*, 110(10): 3184–3230.

Angrist, Joshua D. 2001. "Estimation of limited dependent variable models with dummy endogenous regressors: simple strategies for empirical practice." *Journal of business & economic statistics*, 19(1): 2–28.

Angrist, Joshua D, and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.

Baker, Scott R, Robert A Farrokhnia, Steffen Meyer, Michaela Pagel, and Constantine Yannelis. 2020. "Income, liquidity, and the consumption response to the 2020 economic stimulus payments." National Bureau of Economic Research.

Bartik, Alexander, Marianne Bertrand, Feng Lin, Jesse Rothstein, and Matthew Unrath. 2020. "Measuring the Labor Market at the Onset of the COVID-19 Crisis."

University of Chicago, Becker Friedman Institute for Economics Working Paper, , (2020-83).

Bennett, Paul, Richard Peach, and Stavros Peristiani. 2000. "Implied mortgage refinancing thresholds." *Real estate economics*, 28(3): 405–434.

Beraja, Martin, Andreas Fuster, Erik Hurst, and Joseph Vavra. 2017. "Regional heterogeneity and monetary policy." National Bureau of Economic Research.

Berger, David, Veronica Guerrieri, Guido Lorenzoni, and Joseph Vavra. 2018. "House prices and consumer spending." *The Review of Economic Studies*, 85(3): 1502–1542.

Berger, David W, Konstantin Milbradt, Fabrice Tourre, and Joseph Vavra. 2019. "Mortgage prepayment and path-dependent effects of monetary policy." National Bureau of Economic Research.

Bhutta, Neil, and Benjamin J Keys. 2016. "Interest rates and equity extraction during the housing boom." *American Economic Review*, 106(7): 1742–74.

Boyce, Alan, Glenn Hubbard, Chris Mayer, and James Witkin. 2012. "Streamlined refinancings for up to 14 million borrowers." *Unpublished working paper. Columbia University, New York, NY.*

Campbell, John Y. 2006. "Household finance." *The journal of finance*, 61(4): 1553–1604.

Chen, Haiqiang, Wenlan Qian, and Qiang Wen. 2020. "The impact of the COVID-19 pandemic on consumption: Learning from high frequency transaction data." *Available at SSRN 3568574.*

Chen, Hui, Michael Michaux, and Nikolai Roussanov. 2020. "Houses as ATMs: mortgage refinancing and macroeconomic uncertainty." *The Journal of Finance*, 75(1): 323–375.

Chetty, Raj, John N Friedman, Nathaniel Hendren, and Michael Stepner. 2020. "Real-Time Economics: A New Platform to Track the Impacts of COVID-19 on People, Businesses, and Communities Using Private Sector Data." Mimeo.

Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber. 2020. "The cost of the covid-19 crisis: Lockdowns, macroeconomic expectations, and consumer spending." National Bureau of Economic Research.

Cox, Natalie, Peter Ganong, Pascal Noel, Joseph Vavra, Arlene Wong, Diana Farrell, and Fiona Greig. 2020. "Initial impacts of the pandemic on consumer behavior: Evidence from linked income, spending, and savings data." *University of Chicago, Becker Friedman Institute for Economics Working Paper*, , (2020-82).

D'Amico, Stefania, Vamsidhar Kurakula, and Stephen Lee. 2020. "Impacts of the Fed Corporate Credit Facilities through the Lenses of ETFs and CDX." *Available at SSRN 3604744*.

DeFusco, Anthony A, and John Mondragon. 2020. "No job, no money, no refi: Frictions to refinancing in a recession." *The Journal of Finance*.

Di Maggio, Marco, Amir Kermani, and Christopher J Palmer. 2020. "How quantitative easing works: Evidence on the refinancing channel." *The Review of Economic Studies*, 87(3): 1498–1528.

Di Maggio, Marco, Amir Kermani, Benjamin J Keys, Tomasz Piskorski, Rodney Ramcharan, Amit Seru, and Vincent Yao. 2017. "Interest rate pass-through: Mortgage rates, household consumption, and voluntary deleveraging." *American Economic Review*, 107(11): 3550–88.

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

Dunn, Abe, Kyle Hood, and Alexander Driessen. 2020. "Measuring the Effects of the COVID-19 Pandemic on Consumer Spending Using Card Transaction Data." *US Bureau of Economic Analysis Working Paper WP2020-5*.

Dynarski, Susan, Jonathan Gruber, Robert A Moffitt, and Gary Burtless. 1997. "Can families smooth variable earnings?" *Brookings papers on economic activity*, 1997(1): 229–303.

Eichenbaum, Martin, Sergio Rebelo, and Arlene Wong. 2018. "State dependent effects of monetary policy: The refinancing channel." National Bureau of Economic Research.

Ganong, Peter, Pascal J Noel, and Joseph S Vavra. 2020. "US Unemployment Insurance Replacement Rates During the Pandemic." National Bureau of Economic Research.

Goodstein, Ryan M. 2014. "Refinancing trends among lower income and minority homeowners during the housing boom and bust." *Real Estate Economics*, 42(3): 690–723.

Haughwout, Andew, Lee Donghoon, Joelle Scally, and Wilbert Klaauw. 2020. "A monthly peek into Americans' Credit during the COVID-19 pandemic." New York FED.

Johnson, Eric, Stephan Meier, and Olivier Toubia. 2015. "Money left on the kitchen table: Exploring sluggish mortgage refinancing using administrative data, surveys, and field experiments." *Unpublished working paper*, 2(2).

Karger, Ezra, and Aastha Rajan. 2020. "Heterogeneity in the Marginal Propensity to Consume: Evidence from Covid-19 Stimulus Payments."

Keys, Benjamin J, Devin G Pope, and Jaren C Pope. 2016. "Failure to refinance." *Journal of Financial Economics*, 122(3): 482–499.

Koijen, Ralph SJ, Otto Van Hemert, and Stijn Van Nieuwerburgh. 2009. "Mortgage timing." *Journal of Financial Economics*, 93(2): 292–324.

Kurmann, Andre, Etienne Lale, and Lien Ta. 2020. "The impact of covid-19 on us employment and hours: Real-time estimates with homebase data." *May). http://www. andrekurmann. com/hb_ covid.*

Laibson, David, Peter Maxted, and Benjamin Moll. 2020. "Present Bias Amplifies the Household Balance-Sheet Channels of Macroeconomic Policy."

Mathy, Gabriel. 2020. "The COVID-19 Epidemic will be the First Services Recession and it Could be a Bad One."

Molly, Kinder, and Ross Martha. 2020. "Reopening America: Low Income Workers have suffered badly from COVID-19 so policy makers should focus on equity." Brookings Institution.

Mongey, Simon, Laura Pilossoph, and Alex Weinberg. 2020. "Which workers bear the burden of social distancing policies?" National Bureau of Economic Research.

Nothaft, Frank E, and Yan Chang. 2005. "Refinance and the accumulation of home equity wealth." *Building Assets Building Credit: Creating Wealth in Low-Income Communities*, 71–102.

Scharfstein, David, and Adi Sunderam. 2016. "Market power in mortgage lending and the transmission of monetary policy." *Unpublished working paper. Harvard University*, 2.

Sterk, Vincent, and Silvana Tenreyro. 2013. "The transmission of monetary policy operations through redistributions and durable purchases."

TransUnion. 2020. "Monthly industry snapshot: financial services." TransUnion.

Urban Institute. 2020. "Housing finance at a glance: a monthly chartbook." Urban Institute.

Willen, Paul, and David Hao Zhang. 2020. "Do Lenders Still Discriminate? A Robust Approach for Assessing Differences in Menus."

Wong, Arlene. 2019. "Refinancing and the transmission of monetary policy to consumption." *Unpublished manuscript.*

Figures and Tables

Figure 1: Evolution of Interest Rates and Refinancing Activity (2013-2020)

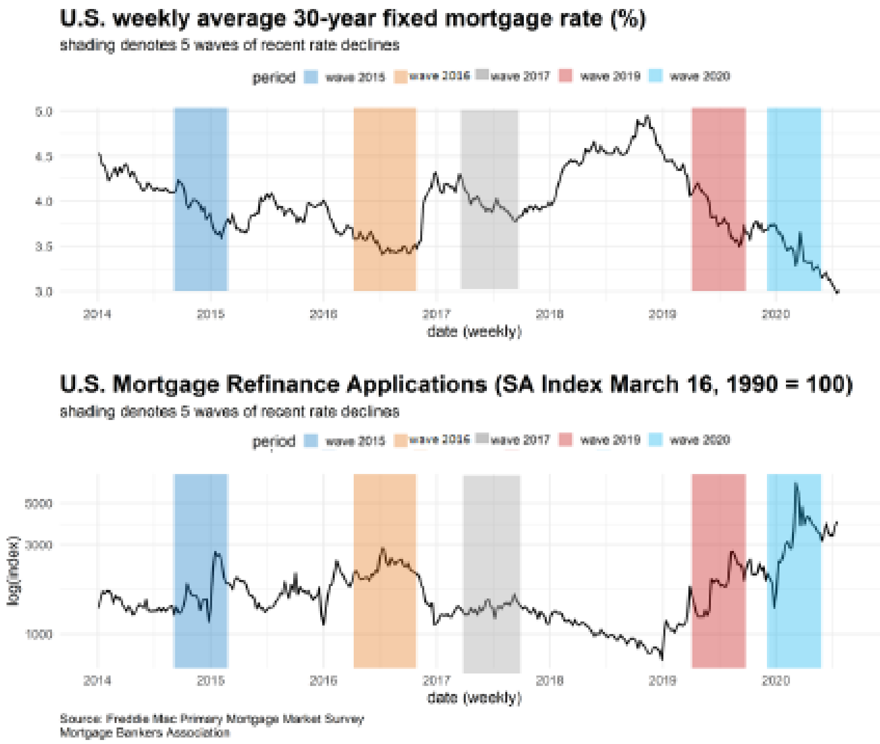
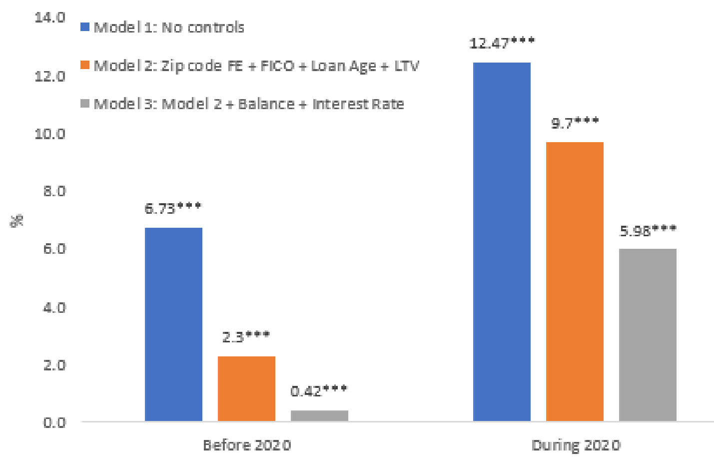


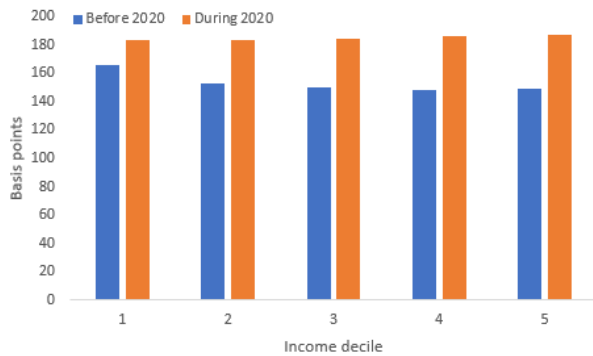
Figure 2: Refinancing Income Gap Before and During 2020, estimated with different sets of control variables (OLS model)



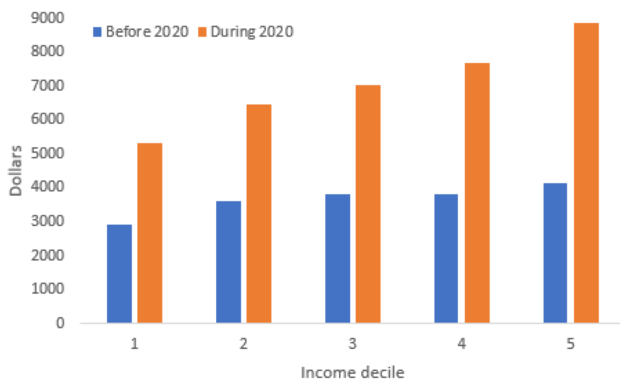
Notes: The refinancing income gap is defined as the difference in refinancing activity between the top and bottom quintiles of the income distribution. It is represented by β_5 before 2020, and by $\beta_5 + \phi_5$ during 2020, based on the coefficients that result from estimating equation 1 with different sets of control variables.

Figure 3: Savings from refinancing conditional on refinancing, by Income Quintile, Before and During the Pandemic (Full set of controls)

(a) Interest Rate reductions (bps)



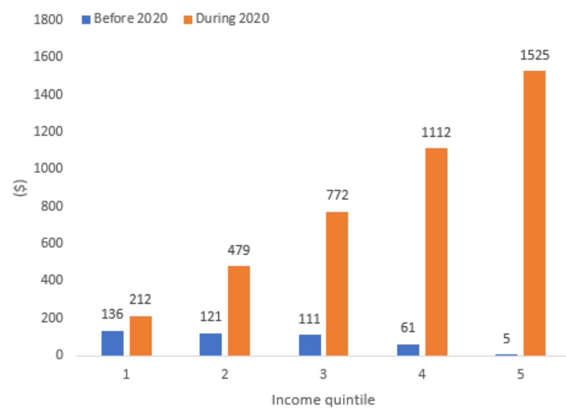
(b) Dollar Savings



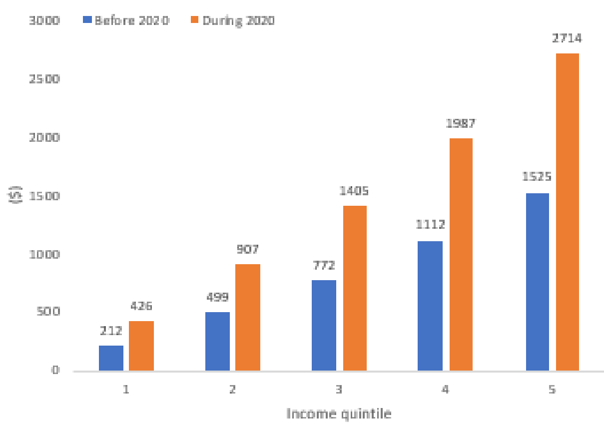
Notes: Savings are projected based on the coefficients in column 3 of Table 3 using refinancing levels in the bottom quintile of the income distribution before 2020 as the omitted category. These coefficients include the full set of controls, and as a result, the projection holds control characteristics fixed at the levels observed on individuals in the bottom quintile of the income distribution.

Figure 4: Savings from refinancing for the entire portafolio, by Income Quintile, Before and During the Pandemic

(a) Model with full set of controls

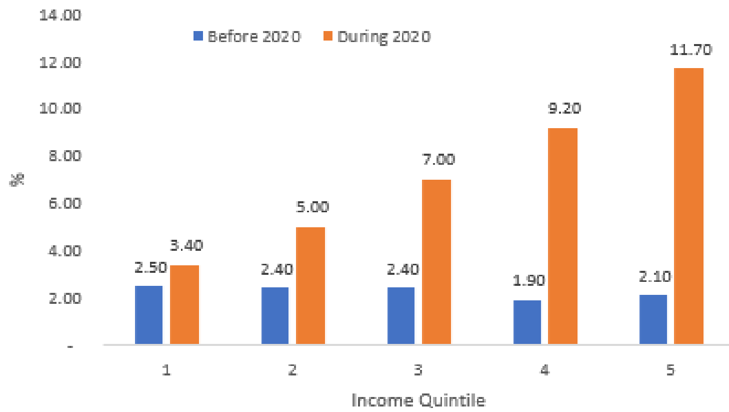


(b) Model without controls



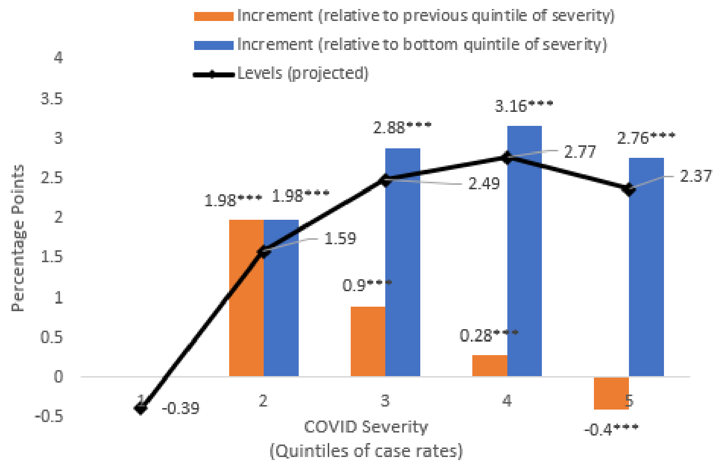
Notes: Savings are projected based on the coefficients of Table 4 using refinancing levels in the bottom quintile of the income distribution before 2020 as the omitted category. The projections in panel (a) are based on coefficients estimated with the full set of control, and as a result, the projection holds control characteristics fixed at the levels observed on individuals in the bottom quintile of the income distribution. The projections in panel (b) are based on coefficients without controls.

Figure 5: Refinancing Activity by Income Quintile: Estimates with full set of control variables



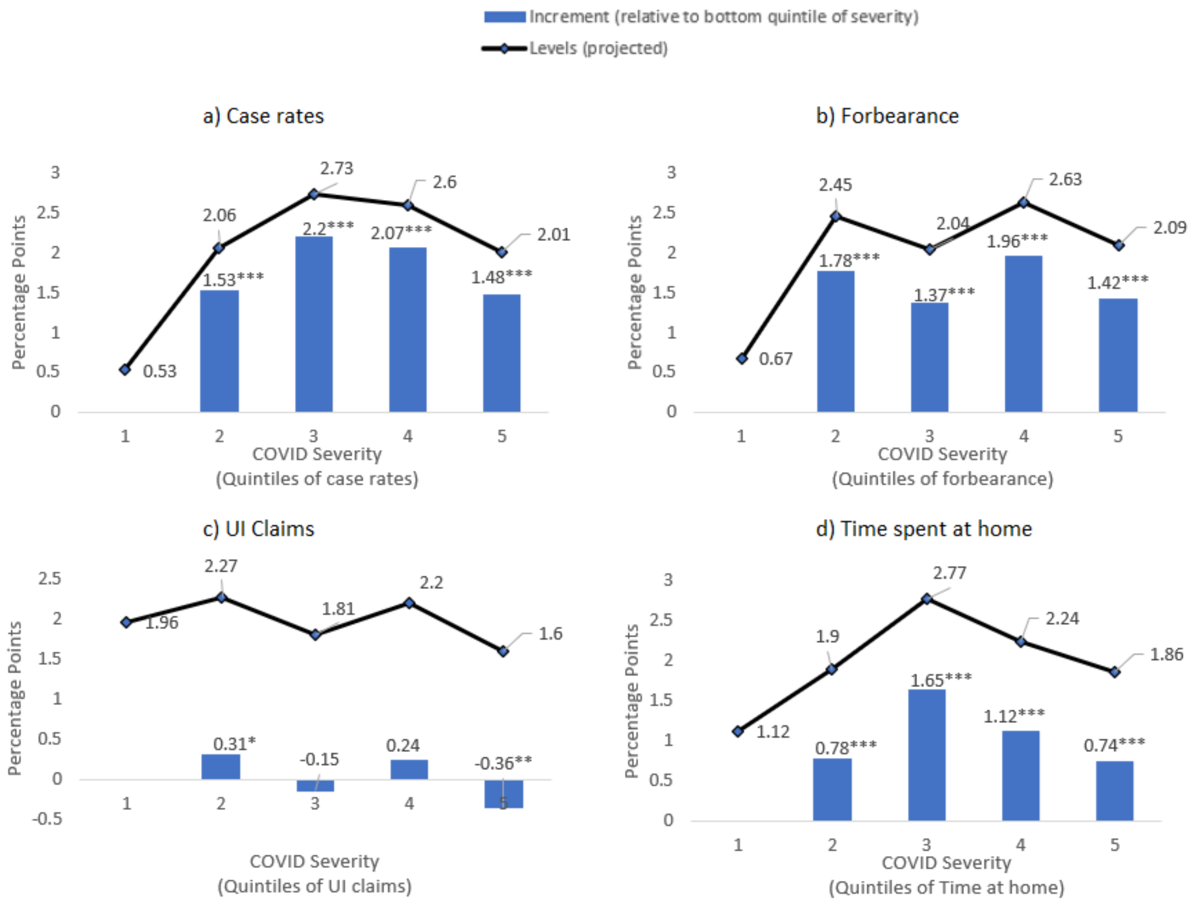
Notes: Refinancing activity is projected based on the coefficients in column 3 of Table 5 using refinancing levels in the bottom quintile of the income distribution before 2020 as the omitted category. The coefficients are estimated with the full set of control variables, and as a result, the projections hold control characteristics fixed at the levels observed on individuals in the bottom quintile of the income distribution.

Figure 6: Estimates of the refinancing income gap across the distribution of COVID-19 case rates (levels and changes relative to the bottom quintile of the distribution of case rates)



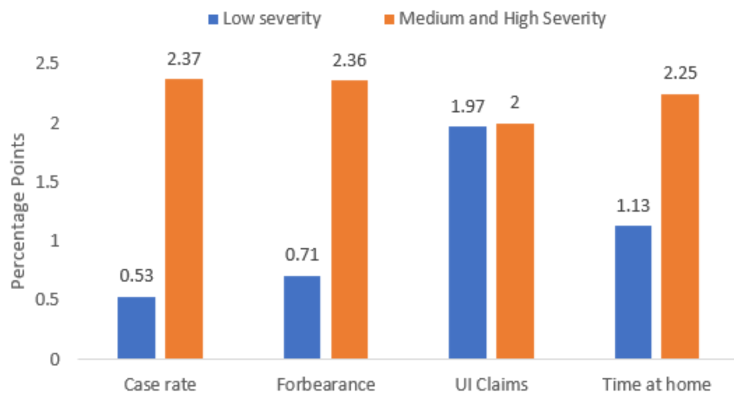
Notes: The black line represents (levels of) the refinancing income gap in county-months that fall in different quintiles of the COVID case rate distribution. The refinancing income gap is defined as the difference in refinancing activity between the top and bottom quintiles of the income distribution. For mortgages in county-months in the bottom quintile of the COVID case rate distribution, it is represented by β_5. For mortgages in county-months in quintiles $k = 2 - 5$ of the COVID case rate distribution, it is represented by $\beta_5 + \phi_{5k}$. Where β_5 and ϕ_{5k} result from estimating equation 2. The blue bars represent changes in the refinancing income gap relative to the bottom quintile of COVID case rates, captured by ϕ_{5k}. The orange bars represent increments in the refinancing income gap relative to the previous quintile of COVID Severity (i.e. the slope of the black line).

Figure 7: Refinancing income gap across county-months with different levels of COVID Severity



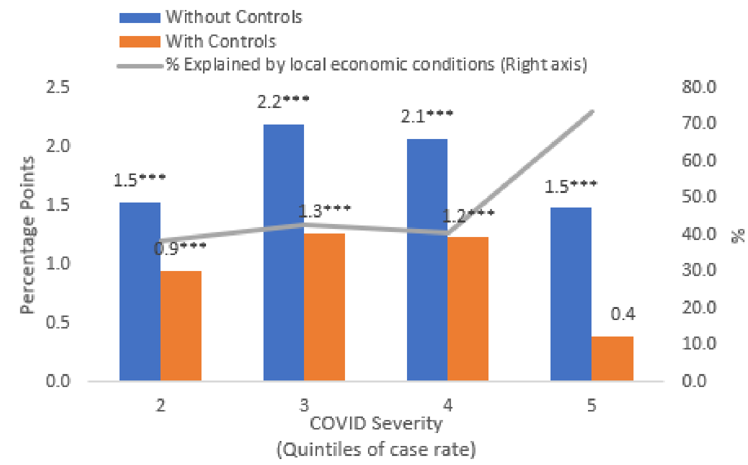
Notes: COVID Severity is measured separately in each panel with four different measures: Case rates, Forbearance, UI Claims and Time spent at home. The black line represents projected levels of the refinancing income gap in geography-months that fall in different quintiles of the COVID case rate distribution. The refinancing income gap is defined as the difference in refinancing activity between the top and bottom quintiles of the income distribution. For mortgages in county-months in the bottom quintile of the COVID Severity distribution, it is represented by β_5. For mortgages in geography-months in quintiles k =2 to 5 of the COVID Severity distribution, it is represented by $\beta_5 + \phi_{5k}$. Where β_5 and ϕ_{5k} result from estimating equation 2 with the corresponding measure of COVID Severity. The blue bars represent changes in the refinancing income gap relative to the bottom quintile of COVID Severity, captured by ϕ_{5k}. Sample is restricted to observations for which all severity measures have coverage.

Figure 8: Impact of local economic conditions on the refinancing income gap (summary)



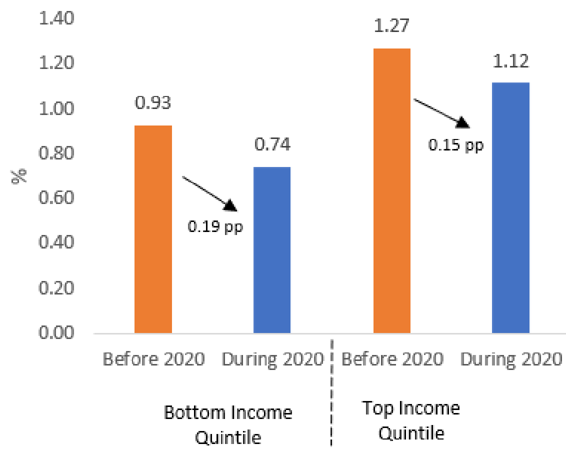
Notes: The refinancing income gap is defined as the difference in refinancing activity between the top and bottom quintiles of the income distribution. The impact of different measures of local economic conditions is estimated separately in four independent regressions, using one of four measures of COVID severity. Geography-months are split into quintiles of Severity. Low severity corresponds to the bottom quintile, Medium to High Severity correspond to quintiles 2 to 5 of the Severity distribution. Sample is restricted to observations for which all severity measures have coverage. All differences between low and medium to high severity estimates are statistically significant at standard levels, when severity is measured with unemployment insurance claims.

Figure 9: Changes in the refinancing income gap across the distribution of COVID-19 case rates (restricted sample with full coverage of controls for local economic conditions)



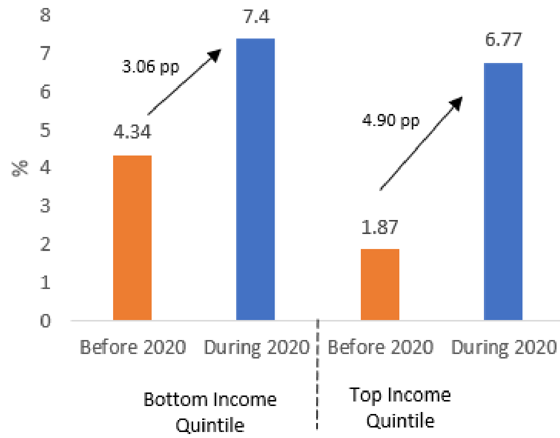
Notes: The refinancing income gap is defined as the difference in refinancing activity between the top and bottom quintiles of the income distribution. Changes in the refinancing income gap between the bottom and kth quintile of the COVID case-rate distribution is represented by ϕ_{5k}. The blue bars are based on the estimates in column 2 of Table A5. The orange bars are based on the estimates in column 3 of Table A5. The gray line plots the difference between the estimates with and without controls, as a percentage of the estimate without controls.

Figure 10: Probability of prepaying an existing mortgage and buying a new home, over time and across the income distribution



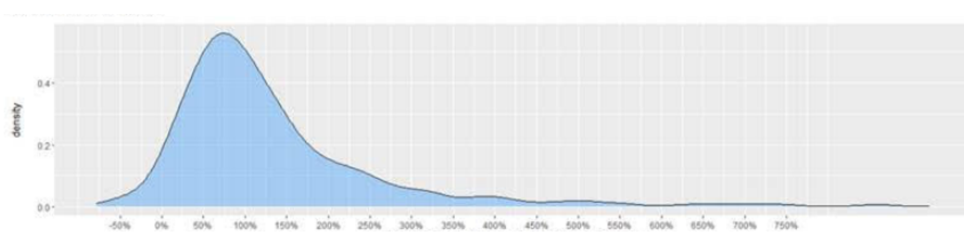
Note: This figure considers observations from mortgages in the matched-transactions database of Freddie Mac, that became newly in the money during the five waves of low interest rates considered in the analysis. The figure shows the probability that a given borrower prepays a mortgage and enters a new mortgage property for a new purchase within 45 days. Probabilities are weighted to match prepayment rates in every wave and income quintile. The orange bars consider observations corresponding to periods of low interest rates in 2014, 2015, 2016 and 2019. The blue bars consider observations corresponding to 2020.

Figure 11: Probability of delinquency, over time and across the income distribution



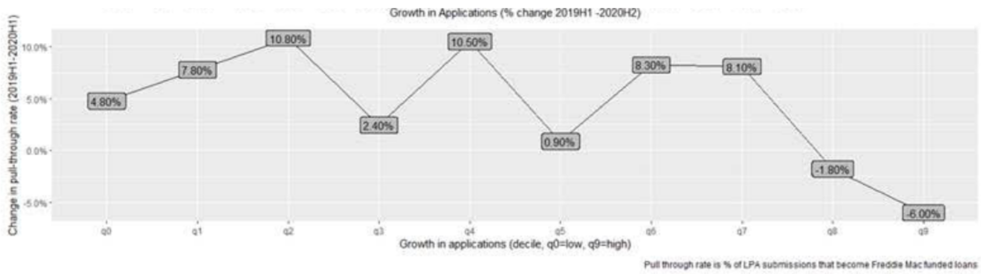
Note: This figure considers observations from mortgages that became newly in the money during the five waves of low interest rates considered in the analysis. The figure shows the probability of delinquency across the income distribution and over time. A mortgage is flagged as delinquent if it is ever delinquent during the five month window that defines each wave. The orange bars consider observations corresponding to periods of low interest rates in 2014, 2015, 2016 and 2019. The blue bars consider observations corresponding to 2020.

Figure 12: Lender-level growth rates in number of applications: First half of 2020 vs First half of 2019 (Density)



Note: Figure shows the distribution of the lender level growth rates in applications submitted to Freddie Mac's Loan Product Advisor tool, between the first six months of 2020 and the first six months of 2019.

Figure 13: Change in funding rates by decile of lender-level growth rate in applications: First half of 2020 vs First half of 2019 (Density)



Note: Figure shows the change in the fraction of applications funded by Freddie Mac for lenders with different growth rates in applications. Changes and growth rates are calculated between the first six months of 2020 and the first six months of 2019.

Figure 14: Time to process applications, across the income distribution

Mean time application to close for conventional loans by income decile

q0=low, q9=high, PU=purchase, RC= cashout refi, RN= rate/term refi

purpose	incq	2019_1	2019_2	2019_3	2019_4	2019_5	2019_6	2019_7	2019_8	2019_9	2019_10	2019_11	2019_12	2020_1	2020_2	2020_3	2020_4	2020_5	2020_6
Rate term Refinance																			
RN	q0	45.0	41.6	40.7	37.1	40.5	40.8	40.0	41.0	40.7	44.4	45.5	46.2	46.5	36.4	34.4	39.8	43.3	44.0
RN	q1	46.9	39.9	38.2	35.6	39.6	39.7	39.0	39.6	41.3	44.7	44.7	44.8	45.5	35.4	33.2	39.5	43.5	43.6
RN	q2	47.8	40.5	39.7	34.5	39.6	37.9	38.6	39.4	39.9	43.9	45.2	45.2	46.1	35.1	32.8	40.1	43.9	43.1
RN	q3	46.7	40.8	40.0	33.9	40.6	38.5	38.7	40.3	39.5	43.6	45.2	44.8	44.3	33.9	32.9	40.0	44.1	43.4
RN	q4	45.8	41.3	41.4	34.0	41.1	40.8	38.1	38.6	39.6	43.6	46.1	44.2	44.2	34.0	33.0	40.4	45.4	44.3
RN	q5	47.4	40.1	41.8	32.7	39.3	40.2	37.4	40.6	39.5	44.1	45.7	47.3	45.3	35.4	33.2	40.4	44.2	44.0
RN	q6	41.5	44.5	38.8	35.0	40.0	39.9	38.8	40.2	39.8	44.7	46.0	47.5	44.8	33.9	32.8	40.1	44.6	45.0
RN	q7	43.3	38.5	44.0	34.0	40.7	38.0	40.0	39.9	39.5	44.2	46.8	45.9	45.9	34.4	32.9	40.6	46.0	45.2
RN	q8	47.4	43.2	43.0	36.0	39.3	38.8	40.1	41.5	40.2	45.9	47.1	48.3	45.8	38.0	33.2	41.8	47.2	45.2
RN	q9	44.2	44.2	38.7	35.3	43.9	44.2	40.0	43.0	42.7	47.6	48.0	49.5	51.4	36.7	35.3	42.0	48.8	48.9

Note: This table considers observations that were ultimately funded by Freddie Mac, and therefore approved by the lender.

Figure 15: Refinance Applications By Income

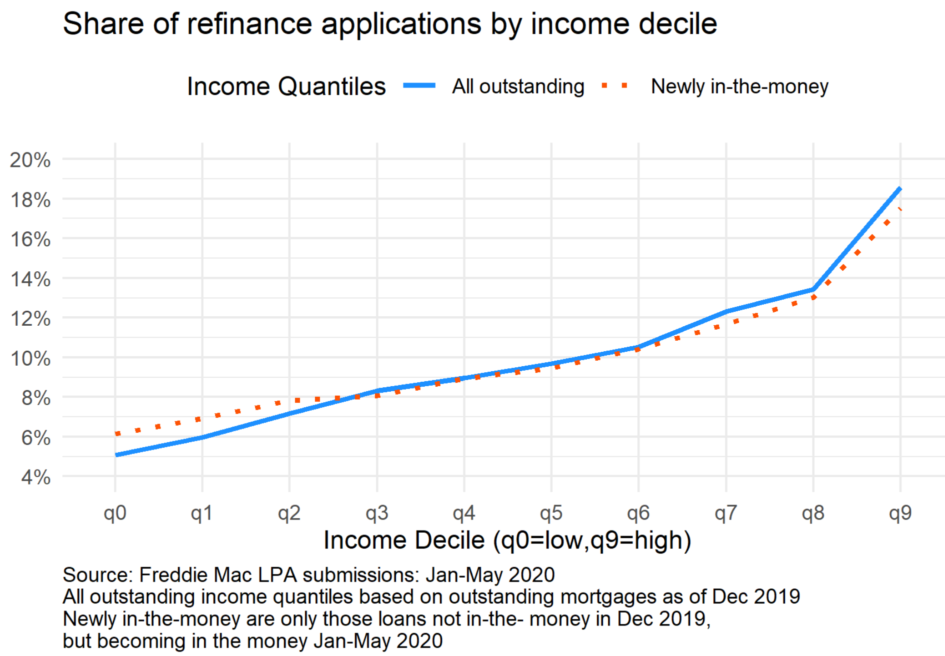
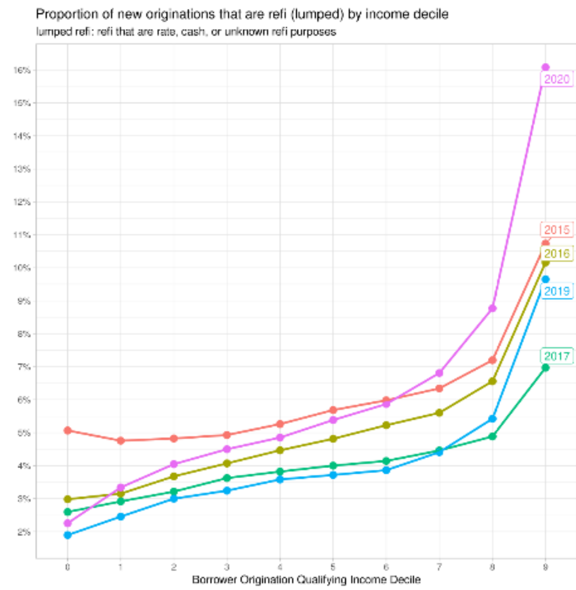
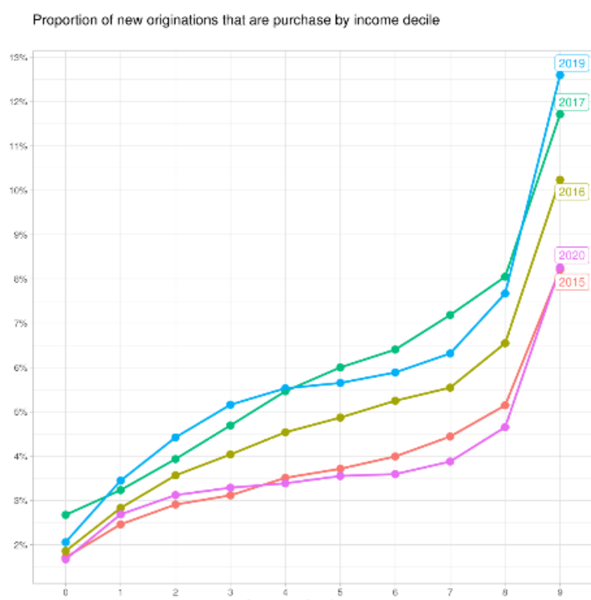


Figure 16: New originations by income decile and loan purpose, over time

(a) Fraction of new originations that are refinances

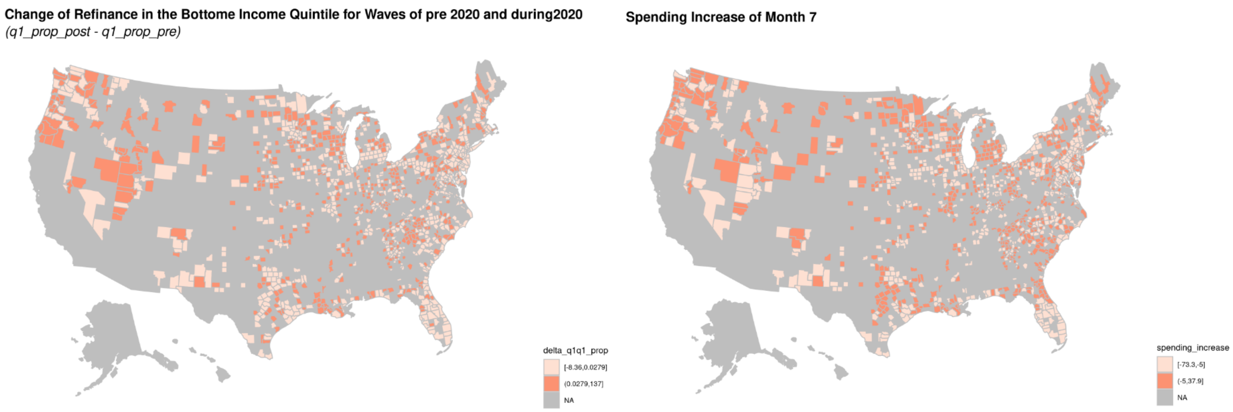


(b) Fraction of new originations that are new purchases



Notes: This figure is built with data reported to McDash. It plots the distribution of new originations by loan purpose across the income distribution and over time. The fractions presented in this figure add up to 100 across all income deciles, loan purposes and periods.

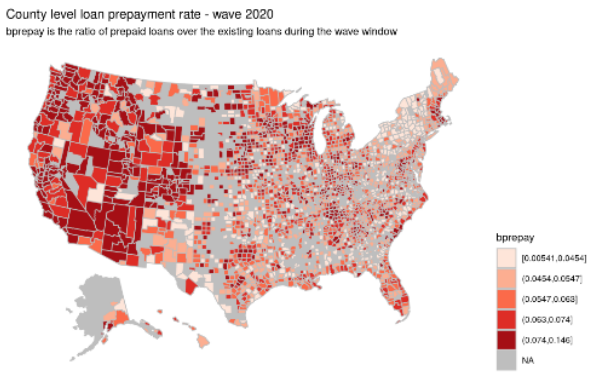
Figure 17: Refinancing of low income borrowers and aggregate spending



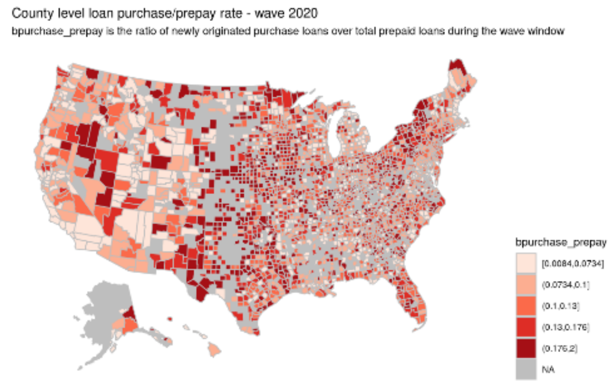
Notes: The left panel shows changes in refinancing activity for borrowers in the bottom quintile of the income distribution before and during the pandemic (Feb-June 2020). The right panel shows the value of the spending index of Chetty et al. (2020) as of July 2020.

Figure 18: Geographic distribution of prepayment and purchase rates, by county.

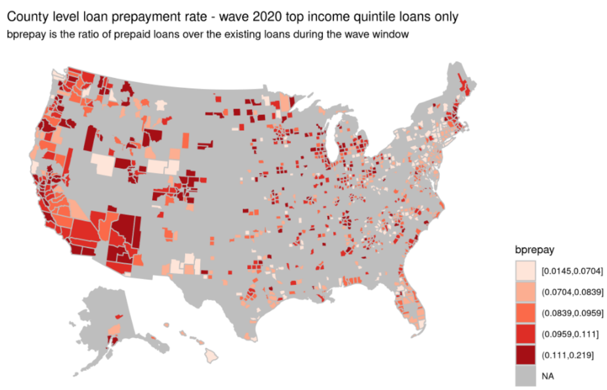
(a) Prepayment rate by county (all mortgages)



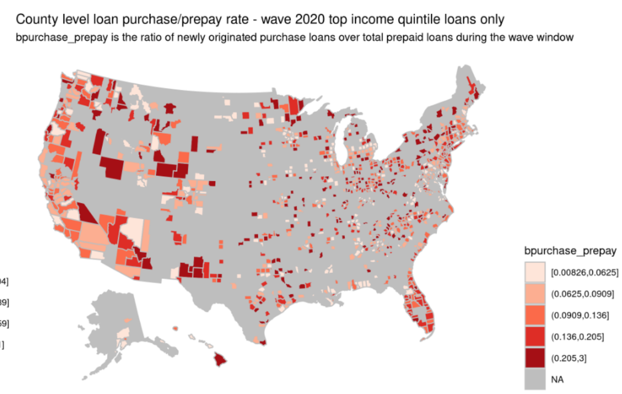
(b) Purchase rate by county (all mortgages)



(c) Prepayment rate by county (Top income quintile)



(d) Purchase rate by county (Top income quintile)



Notes: This figure is built with data reported to McDash corresponding to the refinancing wave of 2020. It plots the geographic distribution of prepayments and purchase rates at the county level. Panels (a) and (b) consider all mortgages across the income distribution. Panels (c) and (d) consider only mortgages in the top quintile of the income distribution.

Table 1: Descriptive statistics of newly in the money mortgages before and during 2020

	Freddie Mac Data				Mc Dash Data			
	Prior to 2020		During 2020		Prior to 2020		During 2020	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
FICO score	740	52	743	49	727	54	738	50
Income (monthly, thousands)	7.35	4.75	7.65	4.79	4.86	10.36	4.73	6.94
Loan Age (months)	55.41	49.89	58.55	52.20	68.49	45.94	56.86	46.05
Interest Rate (%)	4.58	0.93	4.39	0.77	5.14	0.30	4.68	0.22
LTV (%)	77.32	20.54	77.73	18.94	52.52	20.01	59.48	21.00
Unpaid Balance (thousands USD)	188.66	110.95	202.74	116.79	225.77	125.17	248.97	141.61
Potential Savings (thousands USD)	1.97	6.21	5.19	6.34	10.39	5.56	11.85	6.40
Rate Incentive (pp)	1.02	0.95	1.26	0.77	1.51	0.26	1.51	0.22
Number of newly in the money mortgages	2,041,992		1,351,845		1,095,038		680,882	
% of active mortgages that became newly in the money during each wave	9.03		20.14		7.30		15.16	
% of newly in the money mortgages that were prepaid during the period	8.68		16.59		3.92		7.31	

Notes: This table presents descriptive statistics for the main variables considered in the analysis. The left panel corresponds to mortgages in the portfolio of Freddie Mac. The right panel corresponds to mortgages reported to McDash. In both cases we include mortgages that were active at the beginning of each refinancing wave and were newly in-the-money. We say that a mortgage is newly in-the-money during a refinancing wave when it was not in-the-money at the beginning of the refinancing wave but becomes in-the-money during the corresponding refinancing wave. We say that a mortgage is in-the-money when it satisfies the conditions outlined in Agarwal, Driscoll and Laibson (2013). For Freddie Mac data, income corresponds to the income reported at origination. For McDash data, income is estimated out of debt-to-income ratios reported at origination. Potential savings are defined as the present value over the expected life of the loan of the difference in outflows calculated with the original interest rate of each loan and the PMMS rate at the end of the corresponding period. The expected life of the mortgage is parametrized as in Agarwal, Driscoll and Laibson (2013). Rate incentive is defined as the difference between the original interest rate of each loan and the PMMS Rate at the end of the corresponding refinancing wave.

Table 2: Refinancing inequality before and during the pandemic: Probability of refinancing (percentage point)

	(1)	(2)	(3)
incquintileq2	1.59***	-0.1	0.07
	(0.07)	(0.1)	(0.07)
incquintileq3	3.07***	0.2***	0.05
	(0.08)	(0.1)	(0.08)
incquintileq4	4.70***	0.9***	0.11
	(0.1)	(0.1)	(0.1)
incquintileq5	6.73***	2.3***	0.42***
	(0.13)	(0.1)	(0.12)
wave20201	4.80***	0.9***	1.23***
	(0.13)	(0.1)	(0.11)
incquintileq2:wave20201	2.61***	2.4***	1.88***
	(0.19)	(0.2)	(0.17)
incquintileq3:wave20201	4.04***	4.4***	3.36***
	(0.2)	(0.2)	(0.18)
incquintileq4:wave20201	5.35***	6.4***	4.94***
	(0.21)	(0.2)	(0.2)
incquintileq5:wave20201	5.74***	7.4***	5.56***
	(0.23)	(0.2)	(0.22)
Mean of the dep. var. (weighted)-bottom income quintile before 2020	1.14	1.14	1.14
Zip Fixed Effect	No	Yes	Yes
Borrower Controls in Regression	No	Yes	Yes
UPB and original interest rate	No	No	Yes
Observations	3002394	3,002,394	3,002,394
R2	0.04	0.141	0.14
Residual Std. Error	27.73 (df = 3002384)	26.4 (df = 2974421)	26.30 (df = 2974411)

Notes: This table presents presents the results of estimating equation 1 with data from Freddie Mac. We consider observations that were prepaid and matched to a new rate-refinancing loan during the period of analysis. The dependent variable is a dummy variable that takes the value of one when a mortgage goes through a rate refinancing transaction, and zero otherwise. Full list of control variables: zip code fixed effects, loan age, FICO score, LTV, original interest rate, unpaid balance (continuous variables are split into discrete categories and controlled for as dummies). Income quintile 1 is the lowest income quintile.

Table 3: Refinancing inequality before and during the pandemic: Savings conditional on refinancing

	(1) Rate difference (pp)	(2) Rate difference (pp)	(3) Rate difference (pp)	(4) Savings ($)	(5) Savings ($)	(6) Savings ($)
incquintileq2	-0.26***	-0.16***	-0.12***	1,621.93***	1,077***	604.26***
	(0.01)	(0.01)	(0.01)	(57.32)	(99)	(65.53)
incquintileq3	-0.40***	-0.25***	-0.14***	3,207.19***	2,302***	722.41***
	(0.01)	(0.01)	(0.01)	(62.44)	(96)	(67.87)
incquintileq4	-0.53***	-0.32***	-0.17***	4,551.42***	3,354***	762.02***
	(0.01)	(0.01)	(0.01)	(65.28)	(94)	(73.98)
incquintileq5	-0.62***	-0.37***	-0.16***	6,362.34***	4,906***	1,117.86***
	(0.01)	(0.01)	(0.01)	(72.56)	(93)	(83.6)
wave20201	-0.11***	0.07***	0.16***	2,326.31***	990***	2,485.52***
	(0.01)	(0.01)	(0.01)	(63.41)	(111)	(85.63)
incquintileq2:wave20201	0.12***	0.06***	0.13***	274.25***	598***	524.13***
	(0.01)	(0.01)	(0.01)	(90.58)	(137)	(99.97)
incquintileq3:wave20201	0.18***	0.07***	0.16***	386.59***	995***	989.87***
	(0.01)	(0.01)	(0.01)	(92.42)	(132)	(101.24)
incquintileq4:wave20201	0.25***	0.10***	0.19***	644.15***	1,515***	1,548.78***
	(0.01)	(0.01)	(0.01)	(94.58)	(129)	(105.47)
incquintileq5:wave20201	0.28***	0.10***	0.20***	1,061.15***	2,133***	2,414.30***
	(0.01)	(0.01)	(0.01)	(104.31)	(126)	(116.47)
Mean of the dep var (weighted)- bottom income quintile before 2020	1.66	1.66	1.66	2,898	2,898	2,898
Zip Fixed Effect	No	Yes	Yes	No	Yes	Yes
Borrower Controls in Regression	No	Yes	Yes	No	Yes	Yes
UPB and original interest rate	No	No	Yes	No	No	Yes
Observations	76,955	76,955	76,955	76,955	76,955	76,955
R2	0.12		0.55	0.25		0.54
Residual Std. Error	0.91 (df = 76945)	0.84 (df = 76928)	0.71 (df = 63340)	10,469.84 (df = 76945)	10(df = 76928)	9,044.82 (df = 63340)

Notes: This table presents presents the results of estimating equation 1 with data from Freddie Mac. We consider observations that were prepaid and matched to a new rate-refinancing loan during the period of analysis. For columns 1 to 3, the dependent variable is defined as the difference between interest rates of the old (refinanced) and new (refinancing) loans. For columns 4 to 6 the dependent variable takes the value of the dollar savings from refinancing, as defined in Agarwal, Driscoll and Laibson (2013). Full list of control variables: zip code fixed effects, loan age, FICO score, LTV, original interest rate, unpaid balance (continuous variables are split into discrete categories and controlled for as dummies). Income quintile 1 is the lowest income quintile.

Table 4: Savings from refinancing for the entire portfolio, across the Income Distribution, Before and During the Pandemic

	Dep. Var. Realized Refi Savings ($US)					
	(1)	(2)	(3)	(4)	(5)	(6)
incquintileq2	161***	14*	5	2,793***	1,362***	1,138***
	(7)	(7)	(7)	(71)	(69)	(69)
incquintileq3	350***	103***	-25***	4,637***	2,293***	1,360***
	(7)	(7)	(8)	(69)	(68)	(68)
incquintileq4	572***	231***	-75***	6,423***	3,283***	1,534***
	(7)	(8)	(8)	(67)	(67)	(70)
incquintileq5	879***	453***	-131***	8,135***	4,277***	1,494***
	(7)	(8)	(8)	(66)	(66)	(71)
wave20201	290***	-75***	76***	4,602***	1,159***	1,528***
	(9)	(9)	(9)	(79)	(79)	(80)
incquintileq2:wave20201	320***	298***	282***	1,428***	1,657***	1,260***
	(12)	(12)	(12)	(104)	(102)	(100)
incquintileq3:wave20201	629***	663***	585***	2,172***	2,961***	2,338***
	-12	-12	-12	-100	-98	-98
incquintileq4:wave20201	989***	1,089***	975***	2,713***	4,085***	3,377***
	(12)	(12)	(12)	(98)	(96)	(96)
incquintileq5:wave20201	1,409***	1,573***	1,444***	3,075***	4,969***	4,245***
	(12)	(12)	(12)	(96)	(94)	(96)
Model	OLS	OLS	OLS	Tobit	Tobit	Tobit
Tobit scale				15,334	14,712	14,505
Borrower Controls (FICO, LTV, Loan Age)?	No	Yes	Yes	No	Yes	Yes
Borrower Controls (Original Rate and UPB)?	No	No	Yes	No	No	Yes
Omitted group	Low income pre-2020	Low income pre-2020	Low income pre-2020	Low income pre-2020	Low income pre-2020	Low income pre-2020
Savings for omitted group (dollars)	136	136	136	3,114	3,114	3,114
Savings gap before 2020	879	453	-131	1,690	1,087	386
Savings gap in 2020	2,288	2,026	1,313	5,009	5,073	3,798
Observations	3,002,394	3,002,394	3,002,394	3,002,394	3,002,394	3,002,394
Log Likelihood				-2,237,309	-2,195,549	-2,188,227
Residual Std. Error	3 (df = 3002384)	3 (df = 3002367)	3 (df = 3002358)			
F Statistic	19,106*** (df = 9; 3002384)	9,963*** (df = 26; 3002367)	8,905*** (df = 35; 3002358)			

Notes: This table presents presents the results of estimating equation 1 or a Tobit model using equation 1 as the linear index, with data from Freddie Mac. We consider observations that were not prepaid during the period of analysis or were prepaid and matched to a new rate-refinancing loan. Matched prepayments are weighted by the inverse of the probability of a match. When a mortgage is refinanced, the dependent variable takes the value of the dollar savings from refinancing according to the formula of Agarwal, Driscoll and Laibson (2013). When a mortgage is not refinance, the dependet variable takes the value of 0. Full list of control variables: zip code fixed effects, loan age, FICO score, LTV, original interest rate, unpaid balance (continuous variables are split into discrete categories and controlled for as dummies). Income quintile 1 is the lowest income quintile.

Table 5: Refinancing inequality before and during the pandemic: Probability of Refinancing (pp)

	(1)	(2)	(3)
incquintile2	0.42***	−0.18***	−0.11*
	(0.06)	(0.06)	(0.06)
incquintile3	1.24***	0.08	−0.09
	(0.06)	(0.06)	(0.06)
incquintile4	1.97***	0.39***	−0.56***
	(0.06)	(0.07)	(0.07)
incquintile5	2.72***	1.14***	−0.44***
	(0.07)	(0.07)	(0.08)
wave20201	0.52***	−0.90***	0.86***
	(0.08)	(0.08)	(0.09)
incquintile2:wave20201	1.28***	1.02***	1.69***
	(0.10)	(0.11)	(0.11)
incquintile3:wave20201	2.63***	2.36***	3.75***
	(0.11)	(0.11)	(0.11)
incquintile4:wave20201	4.26***	4.09***	6.40***
	(0.11)	(0.12)	(0.13)
incquintile5:wave20201	4.66***	4.40***	8.66***
	(0.12)	(0.13)	(0.14)
Mean of the dependent variable - bottome quintile before 2020	2.46	2.46	2.46
Zip Fixed Effect	No	Yes	Yes
Borrower Controls	No	Yes	Yes
Controls for UPB and Original Interest Rate	No	No	Yes
Observations	1775920	1775920	1775920
R2	0.011	0.04	0.049
Residual Std. Error	0.221 (df = 1775910)	0.220 (df = 1750405)	0.219 (df = 1750395)

Notes: This table presents presents the results of estimating equation 1 with data from McDash. The dependent variable takes the value of 1 when a mortgage was prepaid, and 0 otherwise. Full list of control variables: zip code fixed effects, loan age, FICO score, LTV, original interest rate, investor type fixed effects (GSA, private label or portfolio), unpaid balance (continuous variables are split into discrete categories and controlled for as dummies). Income quintile 1 is the lowest income quintile.

Table 6: Heterogeneity Analysis: Sample Splits

	Dep. Var. Refinancing {0,1}					
	Interest Rate Split		Unpaid Balance Split		FICO Split	
	(1)	(2)	(3)	(4)	(5)	(6)
incquintile2	0.05	−0.95***	−0.24	0.09	0.05	0.02
	(0.06)	(0.29)	(0.26)	(0.06)	(0.11)	(0.08)
incquintile3	0.46***	−0.65**	−0.05	0.53***	0.33***	0.30***
	(0.07)	(0.27)	(0.24)	(0.07)	(0.11)	(0.08)
incquintile4	0.61***	−0.15	0.40*	0.56***	0.43***	0.33***
	(0.08)	(0.27)	(0.24)	(0.08)	(0.11)	(0.09)
incquintile5	0.87***	0.87***	1.24***	0.68***	1.12***	0.72***
	(0.09)	(0.27)	(0.23)	(0.12)	(0.12)	(0.10)
wave20201	−0.45***	−0.08	−0.68	−0.45***	−1.19***	−0.54***
	(0.08)	(0.85)	(0.64)	(0.09)	(0.14)	(0.11)
incquintile2:wave20201	0.90***	1.38	1.40**	0.95***	0.98***	1.08***
	(0.11)	(0.89)	(0.69)	(0.11)	(0.18)	(0.14)
incquintile3:wave20201	1.63***	2.99***	2.56***	1.56***	2.21***	2.05***
	(0.12)	(0.86)	(0.65)	(0.13)	(0.18)	(0.15)
incquintile4:wave20201	2.22***	5.06***	4.20***	2.19***	3.90***	3.12***
	(0.13)	(0.86)	(0.65)	(0.16)	(0.18)	(0.16)
incquintile5:wave20201	2.44***	4.20***	4.18***	1.96***	3.79***	3.74***
	(0.17)	(0.86)	(0.65)	(0.22)	(0.19)	(0.18)
Mean Prepay rate - omitted category		2.42	3.35	2.9	2.44	2.76
2.29						
Sample filter	Rate incentive high	Rate incentive low	UPB high	UPB low	FICO ge 740	FICO lt 740
Zip Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Controls for borrower attributes	Yes	Yes	Yes	Yes	Yes	Yes
Controls for UPB and Interest Rate	No	No	No	No	Yes	Yes
Observations	848,357	848,355	848,362	848,350	815,989	880,723
R2	0.063	0.036	0.051	0.040	0.054	0.039
Residual Std. Error	0.259 (df = 825596)	0.176 (df = 825122)	0.247 (df = 829152)	0.195 (df = 823989)	0.242 (df = 794243)	0.204 (df = 857407)

Notes: This table presents presents the results of estimating equation 1 with data from McDash. The dependent variable takes the value of 1 when a mortgage was prepaid, and 0 otherwise. Full list of control variables: zip code fixed effects, loan age, FICO score, LTV, original interest rate, investor type fixed effects (GSA, private label or portfolio), unpaid balance (continuous variables are split into discrete categories and controlled for as dummies). Income quintile 1 is the lowest income quintile.

Table 7: Funding rate, application growth and borrower income

	(1) Change in Funding Rate	(2) Change in Funding Rate
Application Growth Rate	-0.0325***	-0.0257***
	(0.0012)	(0.0056)
incqq1	0.0035	0.0063
	(0.0098)	(0.0133)
incqq2	0.0024	0.0053
	(0.0095)	(0.0129)
incqq3	0.0009	0.0007
	(0.0094)	(0.0127)
incqq4	0.0001	0.0039
	(0.0093)	(0.0127)
incqq5	0.000003	0.0052
	(0.0093)	(0.0126)
incqq6	0.0003	0.0054
	(0.0092)	(0.0125)
incqq7	-0.0003	0.0155
	(0.0090)	(0.0122)
incqq8	0.0012	0.0161
	(0.0088)	(0.0120)
incqq9	0.0059	0.0190*
	(0.0084)	(0.0114)
Application Growth Rate:incqq1		-0.0027
		(0.0074)
Application Growth Rate:incqq2		-0.0029
		(0.0071)
Application Growth Rate:incqq3		-0.0007
		(0.0070)
Application Growth Rate:incqq4		-0.0038
		(0.0069)
Application Growth Rate:incqq5		-0.0048
		(0.0068)
Application Growth Rate:incqq6		-0.0049
		(0.0067)
Application Growth Rate:incqq7		-0.0122*
		(0.0066)
Application Growth Rate:incqq8		-0.0114*
		(0.0065)
Application Growth Rate:incqq9		-0.0100
		(0.0062)
Constant	0.0311***	0.0231**
	(0.0074)	(0.0098)
height		
Observations	3,090	3,090
R2	0.1977	0.2009
Adjusted R2	0.1951	0.1960
Residual Std. Error	3.1469 (df = 3079)	3.1453 (df = 3070)
F Statistic	75.8822*** (df = 10; 3079)	40.6270*** (df = 19; 3070)

Notes: This table uses observations at the lender-by-borrower income decile level. The dependent variable is given by the change in the fraction of applications submited through Freddie Mac's LPA tool that were eventually funded by Freddie Mac. Application Growth Rate is calculated at the lender level. Changes and growth rates are calculated between the first half of 2020 and the first half of 2019. Income quintile 0 is the lowest income quintile. Robust standard errors in parenthesis.

Table 8: Cash out refis before and during the pandemic, across the income distribution

	(1)	(2)
incquintileq2	0.002***	0.002***
	(0.001)	(0.001)
incquintileq3	0.004***	0.004***
	(0.001)	(0.001)
incquintileq4	0.004***	0.005***
	(0.001)	(0.001)
incquintileq5	0.003***	0.003***
	(0.001)	(0.001)
wave20201	0.008***	0.009***
	(0.001)	(0.001)
incquintileq2:wave20201	-0.001	-0.001
	(0.001)	(0.001)
incquintileq3:wave20201	-0.001	-0.002
	(0.001)	(0.001)
incquintileq4:wave20201	-0.0002	-0.001
	(0.001)	(0.001)
incquintileq5:wave20201	-0.001	-0.002*
	(0.001)	(0.001)
Mean of the dependent variable (weighted)- bottom income quintile before 2020	0.0113	0.0113
Borrower Controls in Regression	None	FICO+LTV+AGE+log(UPB)+Rate Incentive
Controls for Savings (UPB and Rate Incentive as controls)	No	Yes
Observations	3,002,394	3,002,394
R2	0.04	0.043
Residual Std. Error	0.140 (df = 2974438)	0.139 (df = 2974419)

Notes: This table presents presents the results of estimating equation 1 with data from Freddie Mac. We consider observations that were not prepaid during the period of analysis or were prepaid and matched to a new cash-refinancing loan. Matched prepayments are weighted by the inverse of the probability of a match. The dependent variable takes the value of 1 when a mortgage was refinanced, and 0 otherwise. Full list of control variables: zip code fixed effects, loan age, FICO score, LTV, original interest rate, unpaid balance (continuous variables are split into discrete categories and controlled for as dummies). Income quintile 1 is the lowest income quintile.

A Data Description: Freddie Mac Matched Transactions

We use a unique administrative loan-level dataset for conventional single-family loans funded by Freddie Mac. This dataset includes all outstanding single-family 30-year fixed-rate mortgages that were funded by Freddie Mac and were active at the beginning of each refinance wave. We followed those loans through the entire duration of each wave and observed whether or not the loan was prepaid during the wave. Table A1 show descriptive statistics for the data.

Table A1: Descriptive Statistics Freddie Mac Loans (Averages)

wave	N	FICO	LTV	Loan Age (Months)	UPB	RATE
2015	715,363	726	0.77	79	$175,416	5.46%
2016	300,145	731	0.79	69	$198,219	5.04%
2017	232,306	714	0.78	121	$144,102	5.72%
2019	794,178	725	0.80	67	$221,727	5.24%
2020	1,351,845	733	0.80	57	$224,665	4.84%
All	3,393,837	728	0.79	69	$205,743	5.14%

Notes: This table presents descriptive statistics for the main variables considered in the analysis. The sample is restricted to all outstannding 30-year fixed rate mortgages on single-family properties which were not in-the-money for a refinance at the beginning of the wave, and became in-the-money during the wave.

In addition, for a subset of loans that were newly in-the-month and prepaid, we matched a new loan that was originated at the same property address within a 45-day window of the closure of the prepaid loan. For those matched transactions, we collected loan-level attributes of the newly originated loan at the same address. In cases where the loan was refinanced, we observed the new loan product and loan attributes, including the new interest rate. We also identified cases where the prepayment was not for a refinance, but rather a home purchase. Freddie Mac guarantees about 1 in 5 home loans in the United States. Consistent with that, we find that we have matches for approximately 20% of the prepaid loans. The match rate varies by loan attributes, with borrowers in the middle of the income distribution having slightly higher match rates than borrowers in the lowest and highest income quintiles. In the 2015 wave, we get a higher match rate of about 27% due to the inclusion of Home Affordable Refinance Program (HARP) loans. TABLE A2 contains summary match rate across our sample.

To assess the extent to which the matched loans broadly represent the full population of prepaid loans, we first compared the characteristics of matched loans to the unmatched

Table A2: Refinancing Rate by Income and Wave

			Income Quintile (q1=low, q5=high)				
Wave	N	All	q1	q2	q3	q4	q5
2015	52,205	27.1%	30.7%	30.3%	28.3%	27.1%	22.1%
2016	29,842	21.2%	20.1%	22.9%	23.1%	21.9%	18.9%
2017	16,716	19.5%	19.6%	21.8%	21.0%	19.9%	15.4%
2019	78,550	19.6%	14.6%	19.5%	20.9%	20.9%	19.5%
2020	224,235	19.3%	18.6%	19.9%	19.7%	19.7%	18.7%
All	401,548	20.5%	19.5%	21.5%	21.4%	21.1%	19.2%

Table A3: Comparison of matched and unmatched loans (averages)

Wave	Matched?	N	FICO	LTV	DTI	Loan Age	UPB	RATE
before 2020	No Match	138,158	735	0.789	0.360	60	$242,226	5.18%
before 2020	Match	39,155	734	0.786	0.360	61	$238,389	5.18%
2020	No Match	180,960	744	0.811	0.360	37	$273,533	4.70%
2020	Match	43,275	745	0.819	0.358	33	$278,064	4.68%

loans across waves. In TABLE A3 below, we compare origination FICO score, origination LTV, origination DTI, interest rate, and UPB (at the beginning of the wave) for matched and unmatched loans. On these observables, the matched and unmatched loans are very similar.

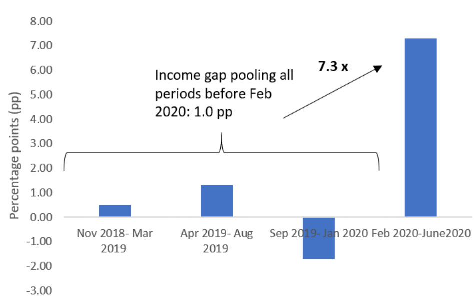
The matched loans could be different in ways that the univariate distributions do not capture. To help mitigate this possibility in our analysis of the matched loans, we use sampling weights. The sampling weights are derived from our estimate of the likelihood of a prepaid loan being matched using observable characteristics. For each prepaid loan in our sample, we code it as 1 if there is a match, 0 otherwise. We fit a linear probability model for the likelihood of a loan being matched using the wave, income quintile, FICO score, UPB, LTV, loan age, and potential savings from refinancing. We then use the inverse of the fitted probability from this regression as our sampling weight for matched loans (using a weight of 1 for all non-prepaid loans). The regression shows that higher FICO loans, higher LTV loans, and younger loans have a slightly lower chance of having a match. But the difference is relatively small. For example, going from a FICO score of 741 to 739 increases the match probability by only 0.8 percentage points. A loan that was prepaid in less than one year (loan age under 12 months) it is 1.7 percentage points less likely to be matched than a loan that was over 7 years old.

B Refinancing inequality over the last 20 months

We study the refinancing income gap over the 15 months periods previous to the pandemic, and compare its magnitude to the refinancing income gap during the pandemic. As before, we consider 5 month windows to allow a reasonable amount of time for refinancing. The results are presented in Figure A1. Panel a) of Figure A1 shows the refinancing income gap in each period ($\beta_5 + \phi_5$), and for reference, panel b) shows the evolution of mortgage rates during the period. We can see while interest rates were consistently coming down from their 2018 peak, refinancing inequality was not following an upward trend. Instead, a dramatic increase takes place between February and June 2020, leading to inequality levels 7.3 times higher than in the immediately previous 15 months.

Figure A1: Pre-pandemic Short Term Trends in Refinancing Inequality (last 20 months)

(a) Refinancing income gap



(b) Fixed-Term 30 year mortgage rate

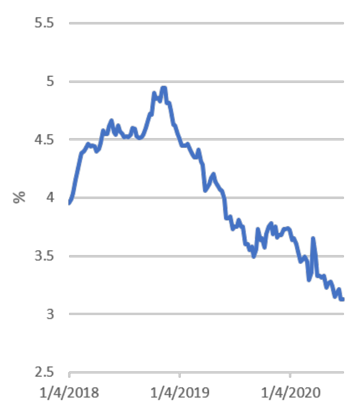


Table A4: Refinancing inequality over the last 20 months

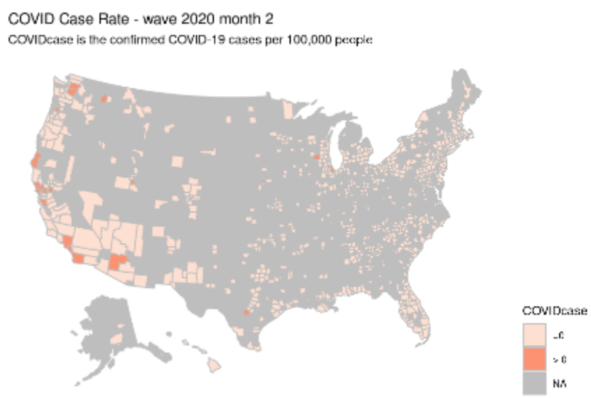
	(1)	(2)	(3)
incquintile2	0.005***	-0.003***	-0.002***
	(0.001)	(0.001)	(0.001)
incquintile3	0.010***	-0.003***	-0.011***
	(0.001)	(0.001)	(0.001)
incquintile4	0.016***	-0.002*	-0.023***
	(0.001)	(0.001)	(0.001)
incquintile5	0.022***	0.005***	-0.026***
	(0.001)	(0.001)	(0.001)
wave20201	0.015***	0.004***	0.039***
	(0.001)	(0.001)	(0.001)
incquintile2:wave20201	0.012***	0.013***	0.026***
	(0.001)	(0.001)	(0.001)
incquintile3:wave20201	0.028***	0.031***	0.053***
	(0.001)	(0.001)	(0.001)
incquintile4:wave20201	0.047***	0.051***	0.083***
	(0.001)	(0.001)	(0.001)
incquintile5:wave20201	0.052***	0.056***	0.107***
	(0.001)	(0.001)	(0.002)
Mean of the dependent variable - bottome quintile before 2020	0.015	0.015	0.015
Zip Fixed Effect	No	Yes	Yes
Borrower Controls	No	Yes	Yes
Controls for UPB and Interest Rate	No	No	Yes
Observations	1127525	1127525	1127525
R2	0.018	0.051	0.068
Residual Std. Error	0.225 (df = 1127515)	0.224 (df = 1103045)	0.222 (df = 1103035)

Notes: This table presents presents the results of estimating equation 1 with data from McDash. We consider observations corresponding to the last 20 months before June 2020. The dependent variable takes the value of 1 when a mortgage was prepaid, and 0 otherwise. Full list of control variables: zip code fixed effects, loan age, FICO score, LTV, original interest rate, investor type fixed effects (GSA, private label or portfolio), unpaid balance (continuous variables are split into discrete categories and controlled for as dummies). Income quintile 1 is the lowest income quintile.

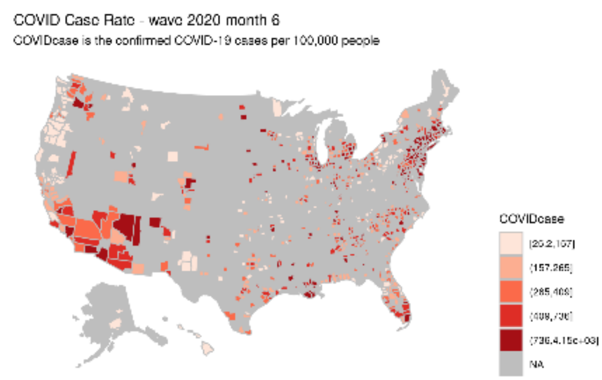
C Geography of the COVID-19 pandemic: severity over time by county and state

Figure A2: Geography of the COVID-19 pandemic, county level variables.

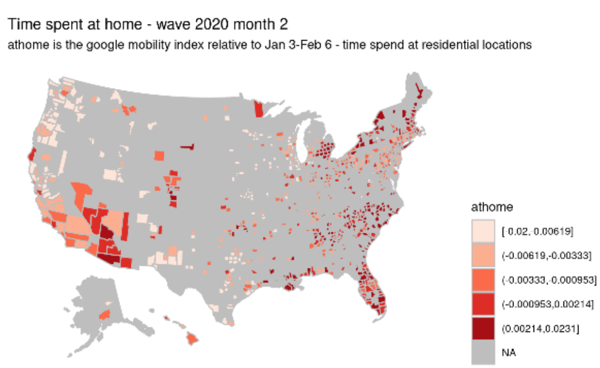
(a) COVID-19 cases per 100,000 people (Feb 2020)



(b) COVID-19 cases per 100,000 people (June 2020)



(c) Time spent at home (February 2020)



(d) Time spent at home (June 2020)

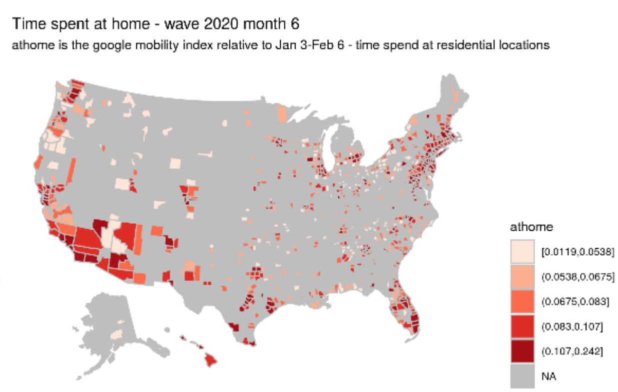
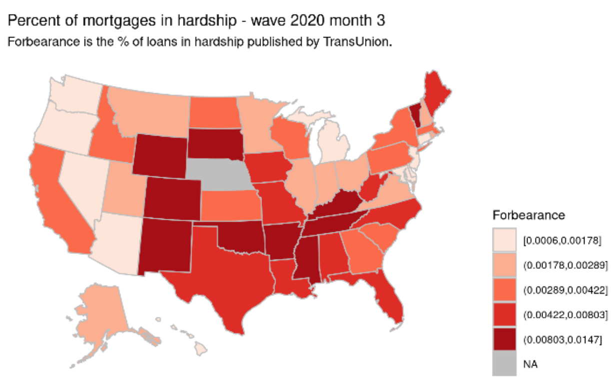
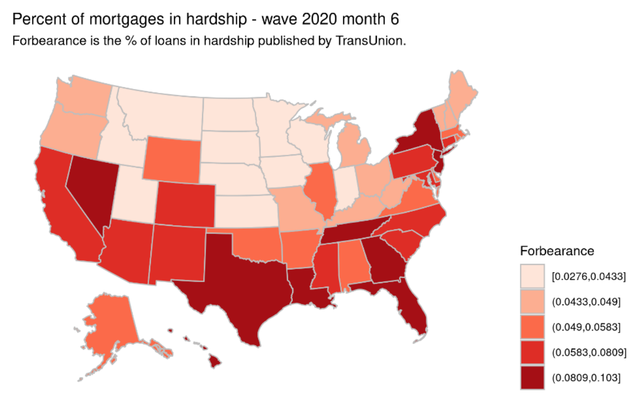


Figure A3: Geography of the COVID-19 pandemic, state level variables.

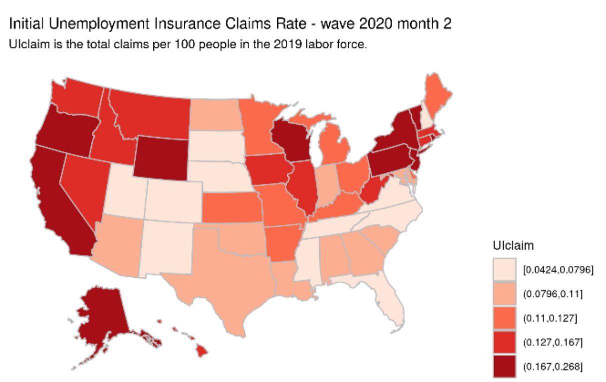
(a) Forbearance (March 2020)



(b) Forbearance (June 2020)



(c) UI weekly claim rate (February 2020)



(d) UI weekly claim rate (June 2020)

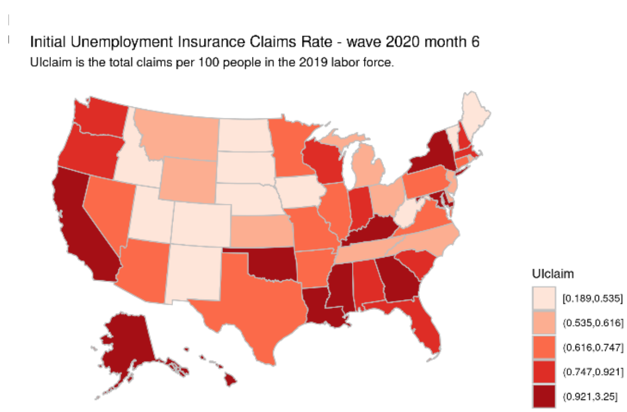


Table A5: Refinancing inequality across the distribution of COVID-19 case rates (different measures of severity)

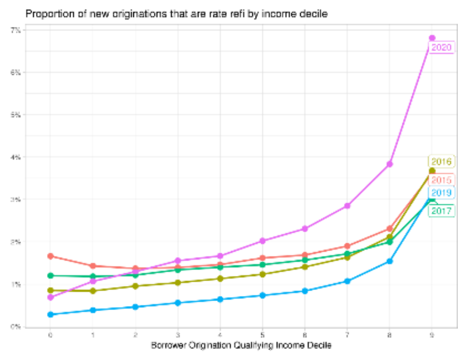
Severity Measure	(1) Case rate	(2) Case rate	(3) Forbearance	(4) UI Claims	(5) Time at Home	(6) Case rate
β_2 IncomeQuintile2	0.01	0.23***	0.40***	0.78***	0.58***	0.29**
	(0.05)	(0.09)	(0.08)	(0.09)	(0.08)	(0.14)
β_3 IncomeQuintile3	-0.12**	0.22**	0.46***	1.29***	0.78***	0.18
	(0.05)	(0.09)	(0.09)	(0.1)	(0.09)	(0.15)
β_4 IncomeQuintile4	-0.44***	0.16*	0.36***	1.77***	0.95***	0.16
	(0.06)	(0.1)	(0.1)	(0.11)	(0.1)	(0.16)
β_5 IncomeQuintile5	-0.39***	0.53***	0.67***	1.96***	1.12***	0.14
	(0.07)	(0.11)	(0.11)	(0.13)	(0.12)	(0.18)
γ_2 Severity2	-1.28***	-0.82***	-1.33***	0.06	-0.33***	-0.64***
	(0.07)	(0.12)	(0.13)	(0.11)	(0.11)	(0.16)
γ_3 Severity3	-1.85***	-1.13***	-1.48***	0.59***	-0.47***	-0.78***
	(0.09)	(0.13)	(0.14)	(0.11)	(0.12)	(0.17)
γ_4 Severity4	-1.97***	-1.16***	-1.25***	0.30***	-0.2	-0.77***
	(0.09)	(0.13)	(0.18)	(0.11)	(0.14)	(0.18)
γ_5 Severity5	-1.98***	-0.89***	-1.20***	0.23*	0.27	-0.45**
	(0.1)	(0.16)	(0.15)	(0.12)	(0.18)	(0.2)
$\phi_{2,2}$ IncomeQuintile2:Severity2	0.64***	0.44***	0.42***	-0.04	0.1	0.47***
	(0.07)	(0.13)	(0.12)	(0.13)	(0.12)	(0.18)
$\phi_{3,2}$ IncomeQuintile3:Severity2	1.09***	0.90***	0.84***	-0.08	0.39***	0.88***
	(0.08)	(0.14)	(0.13)	(0.14)	(0.13)	(0.19)
$\phi_{4,2}$ IncomeQuintile4:Severity2	1.65***	1.16***	1.49***	-0.15	0.55***	0.93***
	(0.08)	(0.14)	(0.14)	(0.15)	(0.14)	(0.2)
$\phi_{5,2}$ IncomeQuintile5:Severity2	1.98***	1.53***	1.78***	0.31*	0.78***	0.94***
	(0.09)	(0.15)	(0.16)	(0.16)	(0.16)	(0.22)
$\phi_{2,3}$ IncomeQuintile2:Severity3	0.84***	0.62***	0.32***	-0.33**	0.27**	0.61***
	(0.08)	(0.13)	(0.12)	(0.13)	(0.13)	(0.19)
$\phi_{3,3}$ IncomeQuintile3:Severity3	1.53***	1.18***	0.90***	-0.62***	0.70***	1.01***
	(0.08)	(0.13)	(0.13)	(0.14)	(0.13)	(0.2)
$\phi_{4,3}$ IncomeQuintile4:Severity3	2.33***	1.85***	1.46***	-0.95***	0.99***	1.32***
	(0.09)	(0.14)	(0.14)	(0.15)	(0.14)	(0.22)
$\phi_{5,3}$ IncomeQuintile5:Severity3	2.88***	2.20***	1.37***	-0.15	1.65***	1.26***
	(0.1)	(0.15)	(0.16)	(0.16)	(0.16)	(0.23)
$\phi_{2,4}$ IncomeQuintile2:Severity4	0.83***	0.59***	0.2	-0.14	0.08	0.63***
	(0.08)	(0.12)	(0.15)	(0.13)	(0.12)	(0.19)
$\phi_{3,4}$ IncomeQuintile3:Severity4	1.49***	1.08***	0.71***	-0.32**	0.31**	0.96***
	(0.08)	(0.13)	(0.16)	(0.14)	(0.13)	(0.21)
$\phi_{4,4}$ IncomeQuintile4:Severity4	2.47***	1.68***	1.21***	-0.38**	0.69***	1.13***
	(0.09)	(0.14)	(0.16)	(0.15)	(0.14)	(0.22)
$\phi_{5,4}$ IncomeQuintile5:Severity4	3.16***	2.07***	1.96***	0.24	1.12***	1.23***
	(0.1)	(0.15)	(0.17)	(0.17)	(0.16)	(0.24)
$\phi_{2,5}$ IncomeQuintile2:Severity5	0.70***	0.38***	0.19	-0.14	-0.02	0.40**
	(0.08)	(0.13)	(0.13)	(0.13)	(0.14)	(0.2)
$\phi_{3,5}$ IncomeQuintile3:Severity5	1.34***	0.88***	0.37***	-0.30**	-0.07	0.67***
	(0.08)	(0.13)	(0.13)	(0.14)	(0.14)	(0.21)
$\phi_{4,5}$ IncomeQuintile4:Severity5	2.08***	1.17***	0.81***	-0.67***	-0.1	0.52**
	(0.09)	(0.14)	(0.14)	(0.15)	(0.15)	(0.23)
$\phi_{5,5}$ IncomeQuintile5:Severity5	2.76***	1.48***	1.42***	-0.36**	0.74***	0.39
	(0.1)	(0.15)	(0.14)	(0.16)	(0.16)	(0.25)
Mean Prepay rate - omitted category	0.71	0.98	1.04	1.34	1.11	0.93
Zip Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Clustered Error	Zip	Zip	Zip	Zip	Zip	Zip
Controls for borrower attributes	Yes	Yes	Yes	Yes	Yes	Yes
Controls for rate and UPB	Yes	Yes	Yes	Yes	Yes	Yes
Controls for other severity measures	No	No	No	No	No	Yes
Observations	2,895,722	1,242,204	1,242,204	1,242,204	1,242,204	1,242,204
R2	0.02	0.02	0.02	0.02	0.02	0.02

Notes: This table presents presents the results of estimating equation 2. The dependent variable takes the value of 1 when a mortgage was prepaid , and 0 otherwise. We consider mortgages that were not in-the-money in February 2020 and became in-the-money in any of the subsequent periods until July 2020. For these mortgages we have monthly observations between the first month in which they turn in-the-money and up until the month in which they are prepaid. The reference category for calculating mean prepay rates in columns 1 to 5 is defined as bottom quintile of income and corresponding severity measure. The reference category for calculating mean prepay rates in column 6 is defined as bottom quintile of income and bottom quintile in all severity measures. The coefficient β_5 represents the refinancing income gap in geography-months where the pandemic hit the least. The coefficient ϕ_{5k} represents increases in the refinancing income gap for mortgages in geography-months that lie on the jth quintile of the distribution of COVID severity, relative to those in the bottom quintile. Standard errors clustered at the zip code level.

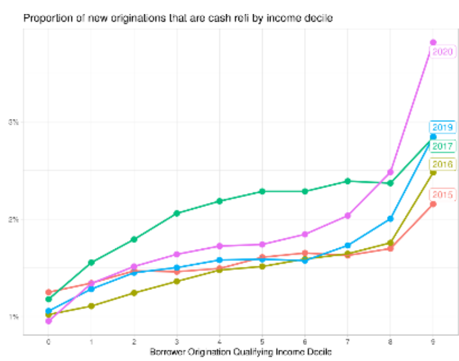
D New originations by purpose across the income distribution and over time

Figure A4: New originations by income decile and loan purpose, over time

(a) Fraction of new originations that are rate refis



(b) Fraction of new originations that are cash refis



(c) Fraction of new originations that are other refis (i.e. cannot be classified as rate refinance or cash refi)