Ditching the Middle Class with Consumer Protection Regulation

Francesco D'Acunto*
University of Maryland

Alberto G. Rossi†
University of Maryland

January 8, 2017

Abstract

We analyze the effects of a recent piece of consumer-protection regulation – Dodd-Frank – on mortgage originations. Dodd-Frank aimed at reducing mortgage fees and abuses against vulnerable borrowers, but increased the costs of originating mortgages. We find it triggered a substantial redistribution of credit from middle-class households to wealthy households. Lenders reduced credit to middle-class households by 15%, and increased credit to wealthy households by 21%, after controlling for drivers of the demand for housing, local house prices, and foreclosures. Large lenders found reacting to Dodd-Frank to be less costly. We thus instrument households' exposure to Dodd-Frank with the pre-crisis share of mortgages originated by large lenders in each county. The redistribution of credit from the middle class to the wealthy was higher in counties more exposed to large lenders, which are similar to other counties. Results hold at the individual-loan level and zip-code level, at the intensive margin (amount lent) and extensive margin (number of loans originated), and for accepted and rejected loans. The collapse of the private-label securitization market, banks' risk-management concerns, and wealth polarization after the crisis do not explain the results.

Keywords: Mortgage Market, Financial Crisis, Dodd-Frank, Household Finance, Banking.

*Smith School of Business, University of Maryland, 4422 Van Munching Hall, College Park, MD 20742. Email: fdacunto@rhsmith.umd.edu.

†Smith School of Business, University of Maryland, 4457 Van Munching Hall, College Park, MD 20742. Email: arossi@rhsmith.umd.edu. For very helpful comments, we thank Michael Barr, Michael Faulkender, Adam Levitin, and Michael Weber. All errors are our own.

1 Introduction

Financial regulation aims at guaranteeing the stability, transparency, and fairness of the financial system to the ultimate benefit of consumers. But changes to the regulatory environment are often implemented in the aftermath of crises, and in times of pressure from public opinion and the media. Despite their aim, new rules change financial institutions' incentives in ways that legislators might not envision, and that could ultimately harm vulnerable consumers.

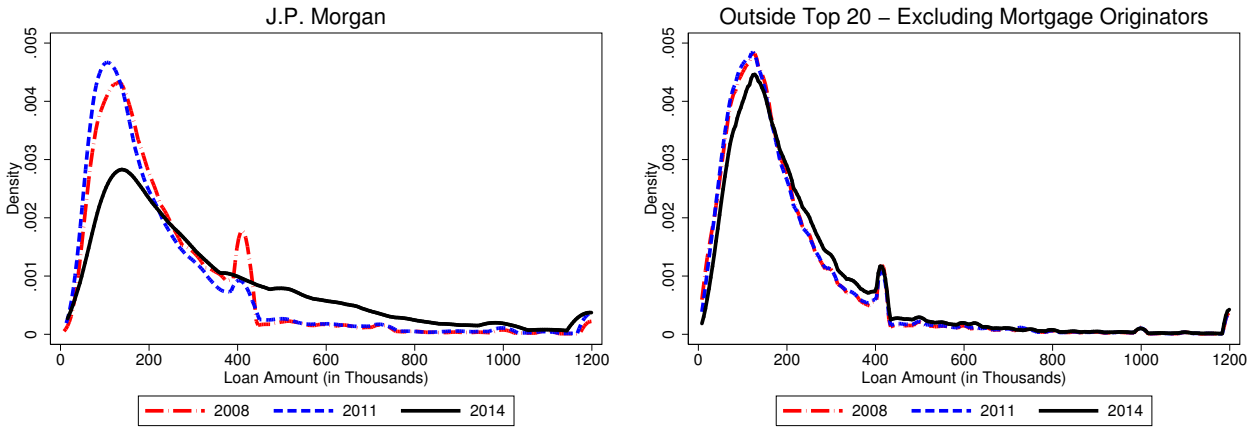
In this paper, we document the unexpected consequences of one of the most controversial reforms of financial regulation in the United States, the *Dodd-Frank Wall Street Reform and Consumer Protection Act* (Dodd-Frank). Two of the main objectives of Dodd-Frank were to eliminate abusive practices in the residential mortgage market, and to improve borrowing conditions for middle-class households. But Dodd-Frank also increased the fixed and per-loan costs of originating mortgages. This increase in the costs of origination made larger loans more profitable. Consistently, we document that financial institutions – in particular, large banks – started supplying more credit to households demanding large mortgages, and decreasing lending to middle-class households. Lenders could barely cut credit to the low-end of the distribution, which includes borrowers that are more likely to belong to demographics protected by the Fair Lending Act.

The redistribution of credit from the middle class to wealthier households is substantial. A back-of-the-envelope calculation that keeps constant the mortgage-demand characteristics of 2010 shows financial institutions reduced their lending to medium-sized loans by 15% in 2014, and increased lending to large loans by 21%. We find this redistribution is robust to controlling for individual- and county-level determinants of the demand for mortgages, such as applicants' race and income, the local racial composition and average household income, local house prices, and the local share of foreclosed properties. The results are similar if we control for time-varying local economic shocks related to the demand or supply of mortgages, and if we restrict the variation within counties. The results are also similar if we run the analysis at the individual-loan level or the zip-code level.

Financial institutions reacted immediately after the approval of Dodd-Frank, even though several provisions were not self executing. This behavior is expected, because lenders had to invest substantial

resources, such as building new training infrastructures and hiring specialized employees, well before the time of execution of the new provisions in order to be compliant at the time of execution.[1]

We find the reaction to this change in incentives increased monotonically with the size of financial institutions. The plot below shows this point, which we discuss further in Section 4. The left panel shows the distribution of mortgages originated by JP Morgan Chase, one of the top 20 lenders in our sample period. Although we focus on JP Morgan Chase as a representative example, the effects described below are similar for other large banks, as we show in Panel A of Figure 1. The right panel plots the distribution across all lenders outside the top 20, excluding mortgage originators.[2] The figure shows the bunching at the jumbo-loan limit was very pronounced for JP Morgan Chase in 2008. The bunching dropped in 2011, and remained low until 2014. On the other hand, the bunching is almost unaltered for smaller lenders. For the lending below the conforming limit, both distributions moved towards larger loans, but the effect was more evident for JP Morgan Chase compared to the smaller banks.



Our results show larger lenders reacted more aggressively to the newly imposed regulation than small lenders. Our interpretation is that reacting would have been profitable for all lenders, but it

[1]Note that several provisions in Dodd-Frank were not self-executing. The Federal Reserve (Fed), and subsequently the Consumer Financial Protection Bureau, had to produce the regulations needed to make the provisions executable. Banks had to be compliant to the new provisions at the execution date, and hence they faced the costs of compliance before that date. Because the execution date was uncertain at the time of approval of Dodd-Frank, banks' expected costs and revenues of different types of mortgages changed immediately. In the Online Appendix, we show direct evidence that large banks paid these costs as early as Januray 2011.

[2]The results are qualitatively similar if we include mortgage originators (see the right plot in Panel B of Figure 1), but, as we explain in section 4, mortgage originators have different incentives for originating jumbo loans.

was less costly for larger lenders for three reasons. First, larger lenders could set lower rates for jumbo loans they did not sell to government-sponsored enterprises (GSEs), because they did not need to engage in costly private securitization. They could keep large loans on their balance sheets because of their high amount of deposits.[3] Second, larger banks could offer a wider menu of services to wealthy customers. Acquiring wealthy customers was more profitable for them. Third, larger lenders operated in several counties with varied demographic characteristics. They could more easily move their origination activities from middle-class counties to wealthier counties.

Our baseline analysis builds on the different extent of reaction by large and small lenders to changing incentives, and the fact that households can barely shop for loans outside their county of residence.[4] Households in counties with a higher exposure to large lenders were thus also more exposed to the effects of Dodd-Frank on mortgage-origination behavior. We focus on individual-loan applications, and compare the loan amounts before and after Dodd-Frank, for households in counties with high or low shares of mortgages originated by large lenders. The share of mortgages originated by large lenders varies substantially across counties - the 5th percentile is 16%, and the 95th percentile is 62%. Variation is large even within states that are homogeneous in terms of economic activity and demographic characteristics.

We consider the difference in the distribution of loan amounts for accepted applications across counties with varying shares of activity by the top 20 lenders – which account for about 40% of mortgage originations from 2007 to 2014. All the effects are stronger if we look at exposure to the top 5 and top 10 lenders, and weaker if we look at exposure to the top 50 and top 100 lenders, as implied by our interpretation of the reaction to Dodd-Frank based on lenders' size. After Dodd-Frank, relative to small lenders, large lenders reduced the amounts lent between $100K and the conforming-loan limit of $417K, and increased the amounts lent above the limit. Systematic differences in the demand for mortgages across counties served by large or small lenders, such as applicants' income, race, and local house prices, do not explain this differential reaction. Our baseline results are robust to using linear or non-linear estimators, to controlling for proxies of the full distribution of house prices

[3] Using the HMDA data, we confirm that after Dodd-Frank, top 20 lenders by size are more likely to keep mortgages on their balance sheets, compared to before Dodd-Frank and to smaller banks.

[4] We measure the share of mortgage originated by lenders at the county level, because households can easily obtain mortgages in zip codes or census tracts different from the ones where they reside.

in counties, and to allowing for all our covariates to vary systematically before and after Dodd-Frank. The results are similar if we control for dimensions we can observe only for a subset of loans, such as the share of foreclosed properties in the zip-code in which the loan was originated. Results are also similar if we exclude non-bank originators. If we restrict the analysis to sand states (California, Arizona, Florida, Nevada), which reacted less than other states to unconventional monetary policy (see DiMaggio, Kermani, and Palmer (2016)), the effects are similar for loans below $417K, and larger for loans above $417K.

We assess three natural alternatives to our interpretation of the results. First, the dynamics of the residential-mortgage private-label securitization market after the crisis might be a concern.[5] It is a well documented fact that the market for private securitization of mortgages has collapsed during the recent financial crisis. Suppose this market had been revived after 2010, and large banks were more active in it than small banks. It would then be easier for larger banks to originate jumbo loans than for other banks, which is consistent with our baseline results. We show that this explanation cannot explain our results, because the residential-mortgage private-label securitization market remained stagnant throughout our analysis period for banks of all sizes. Furthermore, we show that large banks were not more likely than small banks to securitize their loans privately after 2010, which is a necessary condition for the dynamics of the private-label securitization market to explain our results.

The second explanation involves risk management considerations. One might think that large banks moved to originating jumbo loans to decrease the riskiness of their assets and hence comply with capital requirements. This explanation requires that jumbo loans are less likely to default than non-jumbo loans, and entails that riskier banks drive our results. We rule out both of these necessary conditions.

The third explanation relates to the anecdotal evidence that the US wealth distribution has polarized since the financial crisis. If counties that have polarized more are also served mainly by large banks, our results might be driven by wealth polarization, and hence demand, as opposed to the supply-side channel we describe in Section 4.2. Using additional data from the American Community Survey. We show that, if anything, wealth polarization reduces the size of our estimated effects.

[5] We thank Jonathan Parker for proposing this alternative explanation.

Our primary data source, HMDA, includes information on denied mortgage applications. We can thus compare the characteristics of denied mortgages across counties with a high or low share of mortgages originated by large lenders. We consider the share of income to loan amount for denied applicants, and find that in counties with a higher presence of large lenders, the average income per loan amount increases between $100K and the conforming-loan limit of $417K, and decreases above the limit. This result suggests that, after Dodd-Frank, large lenders might have applied stricter standards than small lenders to approve mortgages in the middle of the size of the distribution, and more lenient standards to approve large mortgages.

Two issues hinder a causal interpretation of the baseline results. First, Dodd-Frank is not a policy shock exogenous to the dynamics of the mortgage market during and after the financial crisis. It is not even the only policy intervention that targeted the mortgage market after the financial crisis. Other policy shocks include unconventional monetary policy (e.g., see DiMaggio, Kermani, and Palmer (2016) and Rodnyansky and Darmouni (2016)) and fiscal policy (e.g., see Agarwal et al. (forthcoming) and Berger, Turner, and Zwick (2016)). The ideal experiment to test for the causal effect of Dodd-Frank on the distribution of loan amounts would implement the reform randomly across counties. Furthermore, unobserved shocks might have caused the differential change in the lending behavior of large and small lenders. Ideally, counties should be assigned randomly to a high share or a low share of large lenders.

We cannot obtain exogenous variation in the implementation of Dodd-Frank across space or over time. Instead, we propose an instrumental-variable approach to obtain quasi-exogenous variation to the share of large lenders across counties. We instrument the share of large lenders active in each county-year from 2008 to 2014 with the share of large lenders active in the county in 2007, before the collapse of Lehman Brothers and the financial crisis.[6] The rationale is that the variation in the share of large lenders across counties before the financial crisis cannot have been determined by the wave of bankruptcies during the financial crisis, or by the fiscal and monetary policy measures implemented after the financial crisis. Because of the costs of moving branches and relocating workers, inertia

[6]Strategies based on supply-side pre-program variation to instrument for the extent exposure of consumers to a program have been used recently in economic research (e.g., see Mian and Sufi (2012), Chodorow-Reich (2014), and Agarwal et al. (forthcoming)).

occurs in the local supply structure of financial services. The variation in the share of large lenders in 2007 is highly correlated with the variation in the share of large lenders in each of the following years, and hence our instrument is relevant.

The crucial identifying assumption is that the variation in the share of large lenders in 2007 does not affect the amounts lent after Dodd-Frank through channels different from the share of large lenders after Dodd-Frank. This exclusion restriction cannot be tested, and hence we propose evidence to assess the extent to which the exclusion restriction might be plausible. First, we show our outcome variable and covariates are balanced across counties, based on the shares of mortgages originated by large lenders in 2007. Also, no differential growth in loan amounts occurs before 2011. Second, we use the reduced-form specification to assess the extent to which our instrument might be associated with the outcome variable through channels different from the share of large lenders after 2007. We find the instrument explains the outcome variable in the reduced-form specification in a similar way as the endogenous variable does in the baseline results, but once both the instrument and the endogenous regressor enter the same specification, we can never reject the null that the coefficient is equal to zero for both coefficients at the same time. This result suggests that alternative channels through which our instrument is associated with the outcome variable are unlikely to be economically relevant.

The instrumental-variable results confirm our baseline analysis both in terms of statistical significance and magnitude of the effects. In the second-stage regressions, we find that in counties with a higher share of large lenders, the average amount of approved loans between $100K and $417K decreased, whereas the average amount of approved loans above $417K increased. For denials, we find the income over loans of denied applications increased between $100K and $417K, and decreased above $417K.

Our analysis focuses mainly on individual-loan applications. This analysis allows us to study the effect of Dodd-Frank on the intensive margin of mortgage origination – the size of loans originated before and after Dodd-Frank across the size distribution – but it cannot inform us on the extensive margin – the number of loans originated before and after Dodd-Frank across the size distribution. To investigate the effect of Dodd-Frank on the intensive margin, we perform an alternative analysis in which the unit of observation is a zip code. We regress the change in the average amount and the

number of mortgages for each size group from 2010 to 2014 on the zip-code-level change in demographic characteristics capturing the demand for mortgages from 2010 to 2014, and the share of large lenders in the zip code's county in 2007. The zip-code-level analysis confirms the redistribution of loan amounts from below the median to above the median. We also find the number of loans decreases below the median, and increases above the median, although the latter effect is not statistically different from zero.

2 Related Literature

This paper contributes to three strands of the economics and finance literature. First, we fit in the upcoming literature that studies the mortgage-origination behavior of financial institutions. So far, this literature has focused on origination before the 2008-2009 financial crisis. The literature is debating whether subprime borrowers and the credit-supply channel (Mian and Sufi (2009), Agarwal et al. (2014), Palmer (2015), Mian and Sufi (2016)) or middle-class borrowers and the demand channel (Adelino, Schoar, and Severino (2016), Albanesi, DeGiorgi, and Nosal (2016), Foote, Loewenstein, and Willen (2016)) were ultimately responsible for the collapse of the real estate market after 2007.[7] Our paper focuses on lenders' originating behavior *after* the crisis. Our analysis suggests Dodd-Frank might have permanently changed well-known facts regarding the mortgage-originating behavior of lenders. For instance, the bunching of the largest lenders at the conforming loan limit, a robust fact that was used to estimate the elasticity of house prices to interest rates before the crisis (DeFusco and Paciorek (2016)), might have disappeared for good.

Second, we contribute to the line of research that studies the effects of fiscal and monetary policy on households' consumption and saving choices, such as fiscal stimulus and unconventional monetary policy (e.g., see Mian and Sufi (2012), Green et al. (2014), Broda and Parker (2014)). DiMaggio, Kermani, and Palmer (2016), and Rodnyansky and Darmouni (2016) study the effects of unconventional monetary policy – the long-duration large-scale asset purchase programs (LSAPs) – on lending behavior in the period 2008-2013. A concern with our results is we might capture the reaction of

[7] See also Guiso, Sapienza, and Zingales (2013), who find that social network connections increase the likelihood of strategic defaults, hence proposing an alternative channel for demand-driven defaults.

lenders to LSAPs. Our results have characteristics that do not seem consistent with this concern. First, we exploit the different incentives of large and small lenders after Dodd-Frank. Large lenders react more to Dodd-Frank, whereas Rodnyansky and Darmouni (2016) find small lenders react more to LSAPs, and in particular to QE3 and QE1. If our analysis captured the effect of unconventional monetary policy, we should find the opposite of our baseline result. We also find our results for loans above the conforming loan limit originated in sand states (California, Nevada, Arizona, and Florida) are stronger than for similar loans originated in other states. Instead, in DiMaggio, Kermani, and Palmer (2016), loans originated in sand states – the states more hit by the credit freeze before QE1 – did not react to LSAPs. Moreover, if unconventional monetary policy explained large lenders' move toward jumbo loans, the effect should arise after 2013. After 2013, tapering started and the rates of conforming mortgages increased, making jumbo loans comparatively less costly. Instead, we find the bunching at the conforming loan limit dropped for large banks in 2011. Finally, our analysis is at the level of the new-purchase mortgages, whereas DiMaggio, Kermani, and Palmer (2016) focus on refinancing loans to study within-borrower deleveraging.

On the fiscal policy side, Agarwal et al. (forthcoming) study the effects of the Home Affordable Modification program (HAMP), which provided monetary incentives for both lenders and borrowers to renegotiate the terms of loans at risk of default. Higher renegotiation rates may suggest the banks affected would engage in less originations, because their customers would not be in the position to sign new loans. Agarwal et al. (forthcoming) find large banks are much less likely to engage in renegotiation after the program,[8] so our baseline estimates are, if anything, a lower bound of the true effect of bank size on origination. Because Agarwal et al. (forthcoming) also find higher exposure to HAMP resulted in lower house price declines and lower foreclosure rates, we control for these quantities directly in our baseline or robustness specifications. Similar to Agarwal et al. (forthcoming), we instrument for household exposure to a policy using lender characteristics before the policy was implemented. They exploit the inertia in lenders' ability to engage in renegotiation, whereas we exploit the inertia in household exposure to large banks across counties.

Another measure of fiscal policy that targeted the mortgage market after the crisis was the First-

[8] Although Agarwal et al. (forthcoming) analyze servicers, they show the institutions originating the loan service the majority of loans.

Time Homebuyer Credit (FTHC). Berger, Turner, and Zwick (2016) study the effect of FTHC on home sales and house prices.[9] They find FTHC accelerated home purchases from several years in the future, and its effects were concentrated in active home-sale markets, which suggests the overall stimulative effect of the policy might not have been large. As noted above, we find the higher increase in lending by large lenders above the limit is stronger for depressed home-sale markets, which suggests the effects of the FTHC cannot explain our results.[10]

Third, the paper falls into the literature that studies the effects of consumer protection regulation on demand- and supply-side incentives. The view that regulation is necessary to eliminate abuses against vulnerable agents dates back at least to President Woodrow Wilson (see Glaeser and Shleifer (2003)).[11] Despite its aims, regulation changes the incentives of agents in ways that might not be envisioned at the time the rules are implemented. For instance, the *Freedom of Information Act* (FOIA) was enacted to allow all US citizens, especially those with limited financial resources and social networks, to access the truth about otherwise obscure procedures and actions in the governmental sector. But Gargano, Rossi, and Wermers (forthcoming) find sophisticated institutional investors commonly gather information through FOIA requests to make money on their trades, whose counterparties are uninformed investors.

Together with financial stability, consumer protection is the main aim of financial regulation. The need for consumer protection in finance is especially compelling in times of increasing consumer autonomy in the financial realm (Campbell et al. (2011)). Some authors argue ill-designed regulation created incentives that were prodromic to the recent financial crisis (e.g., see Barth, Caprio, and Levine (2012)), whereas others blame the lack of more stringent regulation (e.g., see Stiglitz (2009)).

This debate was particularly heated during the discussion of Dodd-Frank, and the establishment of the Consumer Financial Protection Bureau (CFPB). Proponents of the CFPB highlighted the need of an agency that protected vulnerable categories from abuses in the financial realm (Warren (2007), Krugman (2013)). Its opponents argued the agency imposed costly and unnecessary one-size-fits-all

[9] For the effects fo FTHC, see also Brogaard and Roshak (2011) and Hembre (2015).

[10] Throughout our analysis, we also control for median house prices in the counties in which the applicant resides.

[11] President Wilson thought regulation was the only shield of consumers against large corporations, because courts were unable to prosecute large corporations (see Wilson (1913)).

regulation, favoring big financial institutions compared to small financial institutions, and ultimately damaging consumers. Other scholars suggest if the institutions that had the same powers as the CFPB were unable to defend vulnerable categories, the CFPB would also possibly be unable to do so (see Barth, Caprio, and Levine (2015)).

In this paper, we highlight one channel through which a piece of consumer financial regulation has reduced the availability of credit to middle-class households, and redistributed credit toward wealthy households. Although we emphasize this channel, we are agnostic about the overall welfare effects of Dodd-Frank.

3 Data Description

Our main source of data is the Home Mortgage Disclosure Act (HMDA) data set for the years 2007-2014, which we obtain through the CFPB's website. The data set contains the universe of mortgage applications over the sample period. For each mortgage application, it includes information regarding the characteristics of the loan, the applicant, and the lender. The covariates on which we focus are the loan amount, the loan status (approved/rejected), the lien of the loan, the purpose of the loan (home purchase/refinancing/home improvement), the owner occupancy, the lender identifier, the applicant's income, the race and ethnicity of the applicant, and the location of the applicant (county/census tract).

Our second source of data is Zillow, through which we obtain the time-series of house prices and the number of houses foreclosed for every county-year. Because individual zip codes may contain more than one census tract, and because individual census tracts may belong to more than one zip code, we follow Adelino, Schoar, and Severino (2016) and use the Missouri Census Data Center bridge – and the population weights therein – to aggregate our individual-loan data at the zip-code level. Because the census tracts definitions vary over time, we use the two alternative bridges available from the Missouri Census Data Center for the periods 2007–2011 and 2012–2014.[12]

[12]The bridges are freely available at http://mcdc2.missouri.edu/websas/geocorr2k.html and http://mcdc.missouri.edu/websas/geocorr12.html.

For all individual-loan results, the working sample uses only loan applications for home-occupied new purchases secured by a first lien. We run the loan-level analysis separately for approved and rejected loans. Our preferred definition of rejected loans includes two types of action taken in the HMDA form, that is, denied applications (code 3) and approved but not accepted applications (code 2). We use this broad definition, because the applications approved but not accepted are effectively also rejected loans. All our results on rejected loan applications are similar if we only define denied applications as rejected applications.

Table 1 reports summary statistics for the main covariates and outcomes. Panel A refers to the sample of approved loans. Our broader sample includes 13,532,723 individual-loan applications that were approved from 2007 to 2014. The median amount lent is $186K, but the loan amounts vary widely from $65K at the 5th percentile to $518K at the 95th percentile. On average, in the counties in which applicants obtain loans, 39% of the mortgages are originated by the top 20 US lenders, and large variation exists in the share of the top 20 US lenders, ranging from 16% (5th percentile) to 62% (95th percentile). The median reported gross income of approved applicants is $71K, and the variation in income is similar to the variation in requested loan amounts. On average, 7% of the applicants are Black, 7% are Asian, and 9.5% are Latino. The average median house price in the counties of approved loans is $184K. The average share of foreclosed properties in the zip code in which the approved borrowers reside is only available for a subset of zip codes in our sample, and hence for a subset of individual-loan applications. It is, on average, 0.08%.

Panel B of Table 1 reports the corresponding descriptive statistics for the subsample of rejected loan applications, which includes 5,983,994 applications. The median loan amount requested by rejected applicants ($176K) is similar to the one requested by approved applicants, but this median masks a distribution with fatter tails. Rejected applicants are more likely to be found among those requesting small amounts or large amounts, compared to approved applicants. A similar comparison holds for the distribution of rejected applicants' income and of average median house prices in the counties in which applicants are rejected. Rejections seem slightly more likely than approvals in counties served by the top 20 US lenders. Moreover, rejected applicants are more likely than approved applicants to be Black (12%) and Latino (14%), but similarly likely to be Asian (7.8%). Finally, the share of

foreclosed properties in the zip code in which the rejected borrowers reside (8.6%) is two orders of magnitude higher than the one for approved lenders. This fact compels us to ensure we check the robustness of our baseline results to controlling for the share of foreclosed properties in the applicants' zip codes, even if this variable is only available for a subset of the applications.

To analyze the extensive margin of mortgage origination, and for comparability with the rest of the literature,[13] we also report results computed at the zip-code level.

Panel C of Table 1 reports summary statistics for the zip-code-level sample. The variables we show are similar to the ones in the individual-level loan analysis, but are computed as averages for each zip-code-year observation from 2007 to 2014. The median loan amount is $170K, which is in line with the loan-level samples. The top 20 US lenders originated on average 47% of the loans at the zip-code level. The statistics for the median income and racial groups of applicants are also similar to the loan-level sample.

A measurement issue that has garnered considerable attention in the mortgage origination literature is the fact that households' gross income is self-reported in HMDA, and hence overstated. We believe overstatement of income is not a relevant issue for our paper – and if anything, it could help us – for two reasons. First, we do not test for the effect of income growth on credit growth, as earlier research did. In our case, income is an observable we use to capture the determinants of demand for loans at the household level. Income overstatement is motivated by the household's demand for a larger loan than what their gross income would allow. If we used the true gross income of the household, we would not capture this determinant of households' demand for loans. Suppose we have two households with the same gross income, but one household head overstates her income, because the household wants a larger loan than what her income suggests. Overstated income captures this unobserved component of demand. If we were controlling for true income, we would assume the two households have the same demand for loans. The second reason is that one of the provisions Dodd-Frank mandated, and that increased the costs of mortgage origination, was a more detailed and costly inquiry on households' income and characteristics before approving a loan. This provision aimed at hindering households from obtaining loans that were excessively large based on their actual gross in-

[13] e.g., see Mian and Sufi (2009) and Adelino, Schoar, and Severino (2016)

come. We cannot test whether the income-overstatement problem in HMDA became less relevant after 2010, but one could believe the issue has at least weakly improved because of this explicit provision in Dodd-Frank.

4 Dodd-Frank and Mortgage-Originating Behavior

Dodd-Frank is a comprehensive set of rules that reformed the US financial system in the aftermath of the 2008-2009 financial crisis. It was enacted in July 2010, and consists of 2,300 pages. It covers numerous aspects of the functioning of financial markets and financial services, ranging from financial stability to investor protection and the transparency of service providers. Below we detail our preferred economic explanation for the redistribution of mortgage lending after Dodd-Frank, whereas in section 5, we rule out alternative explanations after presenting the baseline results.

4.1 Origination after Dodd-Frank

Figure 1 depicts the change in mortgage-originating behavior by larger and smaller financial institutions after the approval of Dodd-Frank, that is, after 2010. Figure 1 plots raw data, and describes facts that hold in the data before we perform any multivariate analysis.

In Panel A of Figure 1, we plot the densities of the mortgage amounts originated by the top 3 financial institutions throughout our sample period (2007-2014), based on the share of mortgages originated each year. In each graph, the long-dashed red line refers to the density of loan amounts in 2008, before the approval of Dodd-Frank; the short-dashed blue line refers to 2011, just after the approval of Dodd-Frank; and the solid black line refer to 2014, the last year for which the HMDA data are available. Changes in originating behavior are similar for these three institutions. First, before the approval of Dodd-Frank, all financial institutions had an incentive to originate loans below the conforming loan limit of $417k. Bunching of originated mortgages just below the value of the limit in 2008 emphasizes this incentive. DeFusco and Paciorek (2016) document this bunching in US mortgage data from 1997 to 2007, and use it to estimate the interest rate elasticity of mortgage demand.[14]

[14]The smaller bunching at the top of the distribution is due to our winsorization of loan amounts at the 0.5% level. It

The novel fact we document is that the incentives to bunch loans below the conforming loan limit has decreased dramatically for large financial institutions since 2011, just after the approval of Dodd-Frank. The incentive was even lower in 2014, when the bunching was quite limited for Wells Fargo to completely eliminated for J.P.Morgan. This fact holds for all large financial institutions, as can be seen in the left plot of Panel B of Figure 1. Note the top 20 financial institutions by size account for about 40% of all mortgage originations in the period 2007-2014. The drop in bunching is associated with a large increase in the mass of loans above the conforming loan limit, which stresses even more the fact that large financial institutions lost the incentives to originate mortgages just below the limit. The distribution of originated loan amounts had a fat right tail in 2014 that was virtually non-existent in 2008.

The drop in bunching is not evident for smaller financial institutions and non-bank mortgage lenders (see the middle and right plots in Panel B of Figure 1). Consistent with this fact, the density of loan sizes had a thinner right tail in 2014 than the density for the largest financial institutions. At the same time, smaller lenders and non-bank lenders also reduced the share of loans originated below the mean of the distribution, and increased the share of loans just below the conforming loan limit.

In Figure 2, we present the raw percentage change in loans originated by bank size between 2008 and 2014 to show that our results are not an artifact of the choice of financial institutions included in the various groups reported in Figure 1. The left panel considers loans between \$100K and \$417K. The right panel considers loans above \$417K. We group institutions in 15 equal-size groups based on total lending, and report the value-weighted change in lending for each group. To avoid including very small banks, we limit the sample to the 1,000 top lending institutions. The plots show larger banks have decreased their lending in the \$100K – \$417K category and have increased their lending above the conforming loan limit of \$417K. Below, we interpret the different reactions to the changing incentives to bunch mortgages below the conforming loan limit by types of institutions.

Figure 3 emphasizes the patterns described above. The figure plots the value-weighted residual percentage change in the number of loans for financial institutions within and outside the top 20 by share of activity from 2010 to 2014. Value-weighted residuals are computed after regressing the

does not reflect a choice on the part of the originators.

percentage changes on a set of observables, averaged at the lender level in 2007, before the approval of Dodd-Frank. The observables are defined as follows: *ApplicantIncome* is the applicant income from HMDA. *Black*, *Asian*, and *Latino* are dummy variables that equal 1 if the applicant belongs to the respective demographic group. *MedianHousePrice* comes from Zillow and is the median price of properties in the county in which the loan was originated. Lenders within or outside the top 20 by share of activity differ in the change of loans originated below and above the conforming loan limit. Large institutions decreased the share of loans below the limit more than smaller institutions. At the same time, large institutions increased the share of loans above the limit more than smaller institutions. Lenders could barely cut loans below $100K, because these loans are most common among demographics that are protected by the Fair Lending Act.

4.2 The Reaction of Large vs. Small Financial Institutions to Dodd-Frank

The data show small and large financial institutions modified their mortgage-originating behavior differently after the approval of Dodd-Frank. We argue Dodd-Frank changed the incentives of all financial institutions in a similar way, but large financial institutions could react to this change more than smaller institutions.

Dodd-Frank modified both the per-loan costs and the overall costs of mortgage origination. A first set of provisions reduced the per-loan profitability of small loans, compared to large loans. For instance, Dodd-Frank imposed a 3% cap on all mortgage-related service fees, such as title searches, home inspections, and closing costs.[15] The rationale for the cap was to avoid abusive practices against consumers, and to force institutions to reduce the fees they would charge on loans. But this cap reduced the profitability of originating loans, and especially smaller loans. For this reason, the financial industry considered the cap a major deterrent to the origination of smaller loans. According to Sue Johnson, the director of Real Estate Services Providers Council, "[smaller loans] would not be made at all or lenders would charge higher interest rates to make up for the higher liability risk" *National Mortgage News* (2013). Moreover, Dodd-Frank imposed a thorough yet costly verification

[15]The CFPB implemented the final fee structure in 2014. It allowed originators to charge fees above 3% for loans below $100k, but it imposed the 3% cap on all loans above $100k.

of the income customers reported at the time of the application. This procedure, which required a thorough screening of the applicants' ability to repay, also increased the per-loan originating costs. Because the procedure has the same fixed cost, irrespective of the size of the loan, it decreased the per-loan returns of smaller loans more than those of larger loans.

A second set of provisions dealt with the overall costs of originating mortgages for financial institutions. For instance, lenders had to establish an internal training system and to provide special training to all loan officers at their branches. The costs of such training decreased the institutions' incentives to originate any mortgages. According to a 2016 survey by the American Bankers Association, the average cost per transaction increased between $300 and $1,000 due to the implementation of the disclosure rules imposed by the Truth in Lending Act in 2010. Moreover, survey respondents estimated an increase in the average length of processing transactions between 8 and 20 additional days (ABA (2016)).

Dodd-Frank affected all financial institutions' incentives to originate mortgages, but we argue that large financial institutions had the ability to react more to this change in incentives than small institutions for at least three reasons.

First, jumbo loans cannot be sold to GSEs, and hence institutions must either keep them on their balance sheets or sell them to private counterparties, which impose worse conditions than GSEs. Larger banks have a comparative advantage in originating jumbo loans, because they can keep more of them on their balance sheets, due to their larger amount of liabilities in the form of retail customers' deposits. The same is not true for non-bank mortgage originators and smaller credit unions and local banks. Accordingly, we find that, on average, the top 20 lenders keep 66.4% of the jumbo loans they originate on their balance sheet. Lenders outside the top 20 keep 61.3% of the jumbo loans they originate on their balance sheet. After Dodd-Frank, the top 20 lenders were 44% more likely to keep the loans they originated on their balance sheets, and 45% more likely to securitize them privately.

Large financial institutions could therefore be more aggressive in setting lower rates for jumbo loans. Consistently, starting in 2010, the gap between the interest rates charged on jumbo loans and those charged on smaller loans started to close. On average, rates were the same for all type of loans

as of 2013, and the gap was even negative for larger banks.[16]

Second, larger banks can offer a much broader set of financial services to their customers than mortgage originators, small banks, and credit unions. Such services include wealth management, brokerage accounts, and credit cards. These services will be especially appealing to wealthier customers, who demand larger loans. Therefore, large financial institutions may be willing to entice new customers with lower interest rates on large mortgages, even if they only break even. This argument is quite salient to the financial industry. For example, according to Keith Gumbinger, vice president at HSH, "There is a potentially significant longer time frame to offer wealthier customers additional products and services. Banks can offer investment services, other loan products or other kinds of services" (Morrison, 2013).

Third, large lenders have more geographically diversified operations than small lenders, and their customers are distributed across more counties. Large lenders can cater to different groups of customers and larger loans by moving their operations across counties, even if the within-county demand for mortgages does not change. On the contrary, small financial institutions normally operate in only a few counties, and their origination behavior is constrained by the characteristics of local demand.

5 Dodd-Frank and Mortgages: Approved Loans

Our baseline analysis builds on the different reaction of large and small lenders to originating loans after Dodd-Frank. We compute the shares of mortgages originated by the top 20 lenders in the United States in each county and compare the within-county lending behavior across counties with higher or lower shares of top 20 lenders, before and after Dodd-Frank.

[16] "One indication of banks' eagerness to woo jumbo borrowers is that average interest rates on 30-year fixed-rate jumbos in 2014 dropped below those on smaller mortgages for the first time in decades" *Wall Street Journal* (2016).

5.1 Baseline Specification

The baseline specification is as follows:

$$\begin{aligned} Log(LoanAmount)_{i,k,t} = \alpha + \beta\ Top20\_Share_{k,t} + \gamma\ Top20\_Share_{k,t} \times DoddFrank_t \\ + X'_{i,k,t}\delta + D'_{k,t}\phi + \eta_k + \eta_t + \epsilon_{i,k,t}, \end{aligned} \tag{1}$$

where $Log(LoanAmount)_{i,k,t}$ is the log amount of the mortgage obtained by applicant i in county k in year t; $Top20\_Share_{k,t}$ is the percentage of loans generated within a county by the top 20 mortgage originators by lending activity; $DoddFrank_t$ is a dummy variable equal to zero for the years before the Dodd-Frank reform was implemented (2008-2010) and 1 thereafter (2011-2014); $X'_{i,k,t}$ is a vector of individual-specific covariates that includes *Income*, the log income of the applicant, as well as dummy variables indicating whether the applicant is *Black*, *Asian*, and/or *Latino*; $D'_{k,t}$ is a county-specific vector of covariates that includes the average number of applicants – for a given year – that are black, *Avg-Black (county)*, asian, *Avg-Asian (county)*, and/or latino, *Avg-Latino (county)*; and *Median House Price* is the log median house price in a given county for a given year;[17] η_k and η_t denote county and year fixed effects.

The coefficients of interest are β and γ. The first captures the effect of the changes in the share of large lenders in a given county on the size of the loans in that county. The second captures the additional effect these changes in share have in the post-Dodd-Frank reform period.

As argued in section 4 and shown in Figure 1, Figure 2, and Figure 3, we expect the Dodd-Frank to have a differential impact on the lending behavior of large banks for different quantiles of the loan-size distributions.

We do not expect banks to reduce their lending differentially for loans below \$100K, because this segment is more likely to include demographics protected by the Fair Lending Act. We expect large lenders to reduce their lending in the \$100K-\$417K range, and move instead toward loans above the \$417K conforming limit.

[17] It is important to control for house prices, because Favara and Imbs (2015) find changes in local credit supply affect house prices.

To capture these effects across different loan sizes, we estimate separately equation 1 for loans in five size categories: loans between zero and \$100K, between \$100K and \$200K, between \$200K and \$417K, between \$417K and \$700K, and greater than \$700K. The results for the various groups are reported in the five columns of Table 2.

Starting from the first column (the smallest loans), we find an insignificant effect of the share of top 20 lenders on loan sizes post Dodd-Frank. Moving to the second column, we find instead a negative and significant coefficient of -0.022 for the loans in the \$100K-\$200K category. Economically, the coefficient implies that for a hypothetical county that has a 100% share, the average loan reduction post Dodd-Frank equals -2.2%. Likewise, for a county that has a 50% share, the effect equals 1.1%. Assuming the average loan in the category equals \$150K, the per-loan economic effect for a county with a 50% share equals: $-\$150K \cdot 2.2\% \cdot 50\% = -\$1,650$. Multiplying this quantity by the 3,912,441 loans generated post Dodd-Frank in this category, we obtain a total economic effect of $-\$1,650 \cdot 3,912,441 = -\$6.5$ billion dollars.[18]

Turning to the remaining three columns, the β coefficient for the third group is also negative – even though significant only at the 10% level – whereas the results for the fourth and fifth groups are positive, economically large, and statistically significant. If we repeat a procedure similar to the one described above for group 2, we obtain the following economic magnitudes for the three groups: group 3= $-\$300K \cdot 1.2\% \cdot 50\% \cdot 3,371,515 = -\$6,068,727$; group 4= $\$550K \cdot 4.3\% \cdot 50\% \cdot 690,453 = \$8,164,606$ and group 5= $\$800K \cdot 24.1\% \cdot 50\% \cdot 205,147 = \$19,776,171$. Note these economic magnitudes do not represent the overall effect of Dodd-Frank on lending, but the differential lending behavior of large and small banks.

These results show that, the higher the share of large banks in given county, the more – in the post-Dodd-Frank period – the approved loans we observe in that county shift from the medium-sized range (\$100K-\$417K) to the non-conforming size range (\$417K+).

[18] The regression estimates reported in Table 2 are based on the subset of loans for which all regressors do not contain missing values, that is, they are based on 4,463,568 loans rather than the 6,959,924 loans originated in the period. The computations reported above, however, use the number 3,912,441 (the full set of loans generated post 2010) rather than 2,513,290 (the set of loans generated post 2010 for which regressor), because our interest lies in providing aggregate economic magnitudes for the whole economy. The same applies for the results associated with the other groups we discuss below.

5.2 Robustness

We now proceed to assess the robustness of the baseline results to alternative specifications. We first verify the results are qualitatively similar if we change the number of financial institutions that we define as large or small. According to our interpretation, the larger a bank is, the higher the redistribution of lending from below the middle of the distribution to above it. Therefore, the smaller the set of large institutions we consider, the larger we expect the size of this redistribution to be. In Panel A and Panel B of Table 3, we find the results are qualitatively similar to our baseline analysis when we compare the originating behavior of the top 5 largest institutions, or the top 100 largest institutions, with the behavior of others. As expected, the size of the estimated coefficients decreases as the size of the top group increases.

A concern with the baseline analysis is that the linear specifications cannot keep constant the characteristics of the demand of mortgages if such characteristics have a nonlinear relationship with the outcome variable. For instance, in the baseline results, we control for the median house price in the county where the applicant resides in the year of the application. But if the originating behavior of financial institutions depends nonlinearly on the overall distribution of house prices, keeping constant the median price would not be enough. Thus, in Panel C of Table 3, we control for changes in other parts of the house-price distribution, by including bottom, middle and top tier house prices from Zillow. The results do not differ from our baseline analysis.

In Panel D of Table 3, we include the full set of interactions between our control variables and the Dodd-Frank dummy. These specifications allow for each of our covariates, including the variables that capture the demand of mortgages, to vary systematically before and after Dodd-Frank. The results are similar when we allow for this full set of interactions.

To control for time-varying local demand and supply shocks, in Panel E of Table 3, we add a set of state$\times$year fixed effects to the baseline specification, and we also find the baseline results are replicated.

In Panel F, we find no differential effect in sand states (California, Nevada, Arizona, and Florida) for the drop in average loan amount for loans below the non-conforming limit. If anything, the

positive effect on the amount for loans above the limit is higher in sand states than in other states. Note the specification includes the full set of interactions needed to interpret the coefficients on the triple interaction, which, because of space constraints, are the only ones we report.

So far, we have studied the reaction of all types of lenders after Dodd-Frank. One might wonder whether banks and non-bank mortgage originators reacted differently to the change in incentives. For instance, anecdotal evidence suggests that non-bank mortgage originators proliferated after the wave of bank bankruptcies during the financial crisis. In Panel G of Table 3, we exclude all loans originated by a non-bank lender, and confirm our baseline results. Thus, the nature of the lender does not seem to explain our baseline results.

In Panel H of Table 3, we add to our baseline specification the share of properties foreclosed in the zip code of the applicant as a covariate. We do so, because Favara and Giannetti (Forthcoming) find that lenders with a higher share of outstanding mortgages on their balance sheet are less likely to foreclose. The share of foreclosures is available at the zip-code level from Zillow, but only for a subset of zip codes. Adding this control thus reduces the size of the sample. We find the baseline results are replicated if we also control for the share of foreclosures at the time of the application.

5.3 Alternative Explanations

We conclude this section by discussing whether alternative explanations can explain our findings. We consider three natural alternatives to the framework discussed in section 4.2. The first one refers to the dynamics of the residential-mortgage private-label securitization market. It is a well documented fact that during and after the recent financial crisis the market for private securitization of mortgages has collapsed. If this market had been revived after 2010, and large banks were more active in it than smaller banks, we could expect that it would be easier for large banks to originate jumbo loans than for other banks, therefore explaining our baseline results. We show that this explanation is not consistent with the data, because the residential-mortgage private-label securitization market remained stagnant throughout our analysis period for banks of all sizes. Furthermore, we show that large banks were not more likely than small banks to securitize their loans privately after 2010, which is a necessary condition to explain our results.

The second explanation involves risk management considerations. One might think that large banks moved to originating jumbo loans to decrease the riskiness of their assets and hence comply with capital requirements. This explanation requires that jumbo loans are less likely to default than non-jumbo loans, and entails that riskier banks drive our results. We rule out both of these necessary conditions.

The third explanation relates to the anecdotal evidence that the US wealth distribution has polarized since the financial crisis. If counties that have polarized more are also served mainly by large banks, our results might be driven by wealth polarization, and hence demand, as opposed to the supply-side channel we describe in section 4.2. Using additional data from the American Community Survey. We show that, if anything, wealth polarization reduces the size of our estimated effects.

5.3.1 Changes in the Private-Label Securitization Market

During the recent financial crisis, private-label securitization came to a halt (Goodman (2015)). If banks wanted to originate jumbo loans, which cannot be sold to GSEs, they could have hardly securitized them privately. As documented in (Goodman, 2015, Figure 2), the aggregate amount of issuance of private-label residential mortgage-backed securities has dropped from approximately $700 billion in 2007 to approximately $60 billion in 2008. After the crisis, the private-label residential mortgage-backed securitization has remained stagnant or contracted even more in both our control and treatment periods, even if private-label securitization in other asset classes – such as credit cards, automobile loans, and student loans – started to recover after 2011. The fact that private-label residential mortgage securitization did not recover between 2008 and 2014, that is, before and after the approval of Dodd-Frank, is an indication that this channel cannot explain our results.

A remaining concern is that, even though the private-label residential mortgage-backed securitization market has remained frozen after the approval of Dodd-Frank, small banks might have stopped the private securitization of their loans completely after 2010, and large banks might have substituted them completely. To rule out this concern, we estimate a variation of equation 1 in which the dependent variable is a dummy that equals 1 if the originated loan is securitized privately. We find that large banks did not securitize privately more than small banks either below or above the conforming

loan limit after Dodd-Frank.

5.3.2 Risk Management

Dodd-Frank has intervened on several dimensions of bank-level activity, including risk management. An alternative interpretation to ours is that large banks have moved towards jumbo loans more than small banks to comply with capital requirements or reduce the risk of their pool of assets. To assess the extent to which this alternative explanation might explain our baseline results, we propose two arguments. First, there is no evidence that jumbo loans are at lower risk of default than smaller loans, even if the applicants are on average wealthier. If anything, evidence reported by the New York Times and based on customized CoreLogic data shows that in the second quarter of 2010, just before the approval of Dodd-Frank, the delinquency rate on mortgages for investment homes above $1 million was twice as high as the delinquency rate on mortgages below $1 million. Delinquency rates for large mortgages were higher than those for smaller mortgages for both owner-occupied and investment real estate.[19]

Second, we propose a test in line with the one in Figure 2. We rank the top 100 institutions in our sample based on the riskiness of their assets and match them with the list of institutions in our sample. We test whether the riskiest institutions increased the origination of jumbo loans after 2010, and decreased the origination of conforming loans after 2010. Our preferred measure of bank risk is the share of reserves over the total amount of non-performing loans held by the bank. We report the results in Figure A.1 in the Online Appendix, which also replicates Figure 2 for the subsample of institutions for which we observe our risk measure. As is apparent from the pictures, there is no relationship between riskiness and the change in lending behavior, whereas our baseline result is replicated for the subsample of banks for which we observe the risk measure.

5.3.3 Wealth Polarization

A third set of competing interpretations refers to wealth polarization within counties. First, anecdotal evidence suggests the US wealth distribution has polarized since the financial crisis. Households that

[19] *"Biggest Defaulters on Mortgages are the Rich"*, New York Times, 2010.

belonged to the middle class before the crisis have moved toward the left or the right tail of the wealth distribution. Wealth polarization might explain our results if large lenders had higher incentives to cater to wealthier borrowers, so that the higher demand for large loans in more polarized counties was more likely to be satisfied by large lenders than small lenders. To disentangle our supply-side-only interpretation from wealth polarization, we estimate the baseline specification across counties with a high or low share of middle-class households before the crisis. Wealth polarization predicts that in counties with a larger share of middle-class households before the crisis, more households moved to the right tail of the wealth distribution after the crisis, and hence the demand for large loans would be higher and the demand for middle-range loans lower. To the contrary, our interpretation predicts the effect should be larger in counties with a lower share of middle-class households prior to the crisis. The rationale is that in those counties, after the crisis, large banks would have more wealthy households to whom they could start to provide lending. Consistent with our interpretation, the results are stronger for counties that had a lower share of middle-class households before the crisis.

Another concern is that middle-class households might have accumulated higher mortgage debt before the crisis. They would thus be less likely to demand mortgages after the crisis. We therefore estimate our results in counties with a higher and lower share of middle-class households that had a mortgage outstanding before the crisis. The results are similar across these counties.

We also repeat the exercise using proxies for the ex-ante exposure of counties to the demand-side effects of the financial crisis. We find our results are similar for counties that were more or less exposed to the demand-side effects of the financial crisis, which corroborates our supply-side interpretation of the baseline results.[20]

Because of space constraints, we report the results described above as well as the description of the additional data sources we use for this analysis in Table A.1 of Online Appendix A.

[20] We consider the share of overall county income from stock dividends and interests in 2007, and the share of workforce in the public administration, because workers in the public administration are more likely to hold defined-contribution pension schemes. Table A.1 in the Online Appendix contains the results.

6 Dodd-Frank and Mortgages: Rejected Loans

The results reported so far consider only the loans approved. To provide further evidence on the differential behavior of lenders for small and large loans, we now present results for rejected mortgage applications. In this case, the outcome variable is the income-to-loan ratio. The idea is to have a proxy for how strict the lending standards of institutions are before and after Dodd-Frank. Higher income-to-loan ratios among denied loans suggest lending standards are stricter, because the lenders are rejecting wealthier households, for the same loan amount. We estimate:

$$\begin{aligned} Log(Income/LoanAmount)_{i,k,t} = \alpha + \beta\ Top20\_Share_{k,t} + \gamma\ Top20\_Share_{k,t} \times DoddFrank_t \\ + X'_{i,k,t}\delta + D'_{k,t}\phi + \eta_k + \eta_t + \epsilon_{i,k,t}, \end{aligned} \tag{2}$$

where $Log(Income/LoanAmount)_{i,k,t}$ is the log income-to-loan ratio for the mortgage denied to applicant i in county k in year t. The rest of the variables are defined as in equation (1) and the results are reported in Table 4.

Interestingly, we find that, across the five specifications reported, the γ coefficients mirror the ones in Table 2, but – as intuition suggests – have opposite signs. The effect for the loans below $100K is not significantly different from zero, and the coefficients for medium-sized loans are positive and insignificant (columns 2 and 3), whereas the coefficients for large loans (columns 4 and 5) are negative and significant.

Note, however, that the economic interpretation of the coefficients is slightly different from Table 2. For example, the γ coefficient estimate for the second specification is 0.03. Economically, it implies that for a hypothetical county that has a 100% share of large lenders, the income-to-loan ratio of the denied mortgage applications is 3% higher post Dodd-Frank. In other words, it implies that applicants need to have a 3% higher income-to-loan ratio in order for their application to not be denied. Likewise, for a county that has a 50% share, the effect equals 1.5%. The coefficient estimates for the third, fourth, and fifth specifications equal 3%, -4.6%, and -22.9% and are, therefore economically significant as well.

Overall, the results reported in sections 5 and 6 show the change in behavior by large institutions has had a large economic impact on the type of loans approved and denied across the United States

as a whole.

7 Identification Strategy

The ideal experiment to test for the causal effect of Dodd-Frank on the distribution of loan amounts would assign US counties randomly to a treated condition, in which Dodd-Frank is in effect after 2010, or a control condition, in which Dodd-Frank is never implemented. This ideal experiment is not feasible, because Dodd-Frank applied to all banks in the United States at the same time.

We tackle the identification problem by exploiting variation in the intensity of the treatment across similar counties. In section 2, we argue Dodd-Frank affected the origination incentives of large financial institutions more than it affected the incentives of small financial institutions. The penetration of large financial institutions in a county measures the extent to which households in the county are exposed to the effects of Dodd-Frank.

Our baseline analysis would document a causal effect only if at each point in time, counties with different levels of penetration of large financial institutions were similar in all other respects. In particular, households in these counties should have the same demand for mortgages, and should react the same way to changes in the supply of mortgages if they were not facing different levels of penetration over time.

Two issues plague the causal interpretation of the baseline analysis. First, we face a reverse-causality problem because Dodd-Frank is not a policy shock exogenous to the dynamics of the mortgage market. As a consequence of the crisis, counties with a higher penetration of large banks might have started to lend more to wealthy households. This tendency, or some unobservables related to this tendency, might be behind the approval of Dodd-Frank. For instance, suppose households in the middle of the distribution by loan amount were more likely to default than other households during the crisis, and large banks were serving these households more than smaller banks. This fact would have pushed large banks out of lending to the middle of the distribution, as well as inspired the provisions in Dodd-Frank. Alternatively, suppose small banks and credit unions went bankrupt more often in counties with more middle-class households than in other counties. The bankruptcy of

small institutions would have mechanically increased the penetration of large banks, and middle-class households would have gotten fewer mortgages. At the same time, Congress approved Dodd-Frank to deal with the bankruptcy waves of small institutions, among other issues.

Second, we face an endogeneity problem. Dodd-Frank, or other unobserved shocks that happened concurrently, might have affected unobservable characteristics at the household and county level. Such characteristics, and not the incentives Dodd-Frank created, might explain why large financial institutions changed their mortgage-originating behavior. Unobservable characteristics might have also affected households' mortgage demand across the income distribution, and hence the observed changes in the distribution of originated loans over time. For instance, Dodd-Frank or other pieces of legislation might have increased house prices more in counties with a higher penetration of large banks, and middle-class households might have asked for larger loans. In our baseline analysis, we control for the median level of house prices in the households' zip codes to address this issue, but other unobservables might have affected households' incentives similarly.

Concerns about reverse causality and endogeneity are relevant, because our baseline analysis exploits variation in the penetration of large financial institutions within counties over time. If penetration varied because of underlying unobservables that also affected the demand and supply of loans, our baseline results would document a spurious correlation.

The ideal source of exogenous variation in the penetration of large financial institutions should not be affected by the financial crisis and the approval of Dodd-Frank. This source of variation should also be unrelated to underlying unobserved characteristics of counties, which Dodd-Frank might have affected and that determine the supply and demand of mortgage credit across the income distribution. Such a source of exogenous variation would allow us to test for the causal effect of the penetration of large financial institutions on the distribution of mortgages before and after Dodd-Frank.

7.1 Instrument and Two-Stage Least-Squares Specification

To get close to such an ideal source of variation, we propose an instrumental-variable identification strategy. We instrument the yearly penetration of large financial institutions in a county in the period

from 2008 to 2014 with the penetration of large financial institutions in the county as of 2007. The rationale is that Dodd-Frank or the financial crisis could not have determined the penetration of large financial institutions in 2007. The first signs of distress in the US financial markets were in late 2007, and the financial crisis did not hit until October 2008. Moreover, unobservables that changed after 2010 could not have affected the penetration of large financial institutions in 2007.

This instrument is likely to be relevant, because inertia is present in the spatial penetration of bank branches. Changing a bank's penetration across counties is costly and takes time. Households exposed to a certain structure of the local banking system in 2007 were likely to be exposed to a similar structure after 2007, both before and after Dodd-Frank. We document the relevance of the instrument below.

Figure 4 and Figure 5 describe graphically the variation in the penetration of large financial institutions in 2007. Figure 4 plots the probability density function for the penetration of large financial institutions. The variable obtains values throughout its range, between 0 and 1. For the vast majority of counties, large financial institutions cover between 20% and 80% of the overall mortgage activity in 2007, and the modal value is about 50%. The variation in penetration across counties is therefore substantial.

Similar to our baseline analysis, we will focus on the variation in loan amounts within counties before and after Dodd-Frank. The systematic variation in penetration across counties will not identify the coefficients in the specifications described below. At the same time, the intensity of the treatment will be higher for counties with higher penetration than for other counties. We note substantial variation in penetration even across similar bordering counties. Panel A of Figure 5 plots the spatial variation in the penetration of large financial institutions in 2007 across all US counties. Panel B of Figure 5 plots the corresponding spatial variation for counties in Iowa. Counties in Iowa are homogeneous in terms of observable characteristics, including their racial composition, the median house prices, and average household income. Both panels document substantial spatial variation in the penetration of large financial institutions, including across areas that are otherwise similar.

To implement our instrumental-variable strategy, we estimate a set of two-stage least-squares regressions. The endogenous covariate is the interaction between the county-level penetration of large

institutions in each year from 2008 to 2014 with the dummy variable for the years after the approval of Dodd-Frank, that is, from 2011 to 2014. In the first stage, we predict the endogenous variable using the interaction between the county-level penetration of large institutions in 2007 with the dummy for the years after Dodd-Frank as the instrument. Specifically, we estimate the following specification:

$$\begin{aligned} LargePenetration_{k,t} * DoddFrank_t = \alpha + \gamma LargePenetration_{k,2007} \times DoddFrank_t \\ + X'_{i,k,t}\delta + D'_{k,t}\phi + \eta_k + \eta_t + \epsilon_{i,k,t}, \end{aligned} \tag{3}$$

where $LargePenetration_{k,t}$ is the percentage of large-institution activity in a county in year t and county k, $LargePenetration_{k,2007}$ is the percentage in year 2007 and county k, and X and D are a set of observables at the individual and county level in year t; η_k and η_t are sets of county and year fixed effects. Note the variable $LargePenetration_{k,2007}$ does not vary over time. Note that restricting the variation within years using year fixed effects absorbs completely the variation in the level of $LargePenetration_{k,2007}$ across counties, which is why the level of the variable does not appear in the RHS of equation 3.

Moreover, the specification in equation 3 mirrors our baseline analysis in restricting the variation within counties with the addition of county fixed effects. Therefore, the variation we exploit for identification excludes any systematic differences in the share of large lenders across counties. We only exploit the component of exposure to large lenders that has not changed within counties, and look at the differential effect of this component on origination before and after Dodd-Frank.

In the second stage, we use the instrumented interaction on the LHS of equation 3 as the main covariate of the following specification, which is otherwise the same as our baseline regression in equation 1:

$$\begin{aligned} Log(LoanAmount)_{i,k,t} = \alpha + \beta\ \widehat{LargePenetration_{k,t} \times DoddFrank_t} \\ + X'_{i,k,t}\delta + D'_{k,t}\phi + \eta_k + \eta_t + \epsilon_{i,k,t}, \end{aligned} \tag{4}$$

where $Log(LoanAmount)_{i,k,t}$ is the log amount of the mortgage obtained by applicant i in county k in year t.

In the rest of this section, we discuss the relevance of the instrument, the exclusion restriction we need to assume if we want to interpret the IV results in a causal fashion, and the threats to identification that our strategy implies.

7.2 Validity of the Instrument

To assess the validity of our instrument, we need to verify its relevance and discuss the plausibility that the exclusion restriction we assume for causal interpretation holds.

In terms of relevance, the instrument needs to be correlated with the endogenous variable we want to instrument, because a weak instrument would lead to inconsistent estimates and would invalidate statistical inference in the second stage. Our analysis does not appear to be prone to a weak-instrument problem. For each column of Table 7 and Table 8, we report the first-stage F-statistic associated with our instrument. The first-stage Kleibergen-Paap F-statistics are larger than 101 across all specifications.

If we want to interpret the estimated coefficient $\hat{\beta}$ in equation 4 causally, we need to assume an exclusion restriction. The penetration of large financial institutions in a county in 2007 should only affect the amounts lent in the county in the following years through the penetration of large financial institutions in the following years, and not through unobservable characteristics at the county or individual-borrower level. This exclusion restriction cannot be tested directly, but we propose two sets of results to assess its plausibility.

First, we test whether observable characteristics that are important determinants of mortgage demand are systematically related to the penetration of large financial institutions. In Table 5, we describe the balancing of the main outcome variables and the main covariates in the analysis, across levels of the penetration of large financial institutions in counties. The first four columns of Table 5 report the average value of each variable for the observations in all years, split into four equal-size groups based on penetration. The first column refers to observations in the bottom quarter of penetration by county, whereas the fourth column refers to observations in the top quarter of penetration by county. The fifth column reports the standard deviation for each of the listed variables,

which helps in assessing the size of the differences in the averages across quantiles. The outcome variables include the average change in the mortgage amounts for loans below $100k, between $100k and $200k, between $200k and $417k, between $417k and $700k, and above $700k, between 2007 and 2010. The covariates include dummy variables for the following: whether the applicant is Black, Asian, or Latino; the share of Black, Asian, and Latino population in each county-year; the log of the median house price in each county-year; and the share of foreclosed houses in each county-year.

Both the dependent and independent variables appear to be balanced for varying shares of large lenders. The average loan amounts for the groups above $417k show virtually no difference between the bottom quarter and the top quarter of observations based on share of large lenders. The two groups between $100k and $417k also appear to be substantially balanced, because the differences across group averages are small compared to the standard deviation in the underlying variables. The loans below $100k have a higher average for the top quarter by share of large lenders, but the difference with the other groups is less than one tenth of a standard deviation in the underlying variable, and this group is the one in which we find no effect of Dodd-Frank by penetration of large financial institutions. The covariates also appear to be balanced, with no monotonic relationships between group averages and penetration of large financial institutions, except for the median house price in county-years, which increases with the share of large lenders.

Second, we report the reduced-form specification, and we run our baseline specification adding the interaction between the penetration of large financial institutions in 2007 and the Dodd-Frank dummy. In these specifications, we add loan observations for 2007, because the instrument we add to the specifications is measured in 2007. The reduced-form specification in Panel A of Table 6 shows the baseline results are very similar when we use the instrument as a covariate on its own. At the same time, if we had substantial violations of the exclusion restriction, we should find that both the instrument and the endogenous variable have an autonomous correlation with the outcome variables once they enter in the same reduced-form specification. On the contrary, Panel B of Table 6 shows we cannot reject both null hypotheses that the coefficients associated with the two interactions are different from zero for any loan size. This result suggests that even if violations of the exclusion restriction existed, they could not be economically relevant enough to explain our baseline and reduced-form results.

7.3 Instrumental-Variable Results

We report the results for estimating equation 4 in Table 7. Table 7 refers to specifications using approved loans, and whose dependent variable is the average loan amount within each size group. The sign and magnitude of the coefficients are similar to the baseline OLS results. A 10% increase in the instrumented interaction between the penetration of large institutions in the county and the Dodd-Frank dummy variable decreases the average amount of approved mortgages between $100k and the conforming loan limit, whereas it increases the average amount of jumbo loans.

The results are very similar if we add the percent of foreclosed houses in the county by year. We do not use this specification as our main specification, because the percent of foreclosed houses is only available from Zillow for about two thirds of counties.

Table 8 refers to the specifications using denied loans, and whose dependent variable is the share of applicant income over the loan amount requested. Even in this case, the two-stage least-square results are similar to the baseline OLS regression.

8 Zip-Code-Level Analysis

In this section, we run our analysis at the zip-code level instead of at the individual-loan level. We have two motivations for a zip-code-level analysis. On the one hand, using a level of aggregation allows us to investigate the effect of Dodd-Frank not only on the amounts of the originated loans (intensive margin of lending), but also on the number of loans originated within each mortgage-size group (extensive margin of lending). Financial institutions are likely to adapt their originating behavior along both margins, because they can control both. On the other hand, most existing papers that study the mortgage-origination behavior of financial institutions before and during the financial crisis run their analysis at the zip-code level. We run most of our analysis at the individual level, because this level allows us to control for a set of important characteristics of households, which are likely to be correlated with loan-level characteristics and might not be captured by aggregate averages.[21] At the same time,

[21] For instance, Adelino, Schoar, and Severino (2016) write, "it is households, and not zip codes, which take on mortgage loans."

we want to ensure our results can be easily compared to the results in earlier literature.

To run the analysis at the zip-code-level, we follow Adelino, Schoar, and Severino (2016) and Mian and Sufi (2009) and compute the zip-code level growth of loan amounts and loan counts within each loan-size group from 2010 to 2014. These variables capture the growth in the intensive margin and extensive margin of originated loans from just before Dodd-Frank was approved to the latest year for which the HMDA data are available. Similarly, we compute the zip-code-level growth of the covariates we used in the individual-level analysis from 2010 to 2014. We then compare the growth of the loan amount and number of loans across zip codes based on the peneration of large financial institutions in 2007 in the counties in which the zip codes lie. We use the penetration of large financial institutions in 2007, before the financial crisis, because we use this year in the instrumental-variable analysis. The results do not change if we use the penetration in 2010.

A contentious point in the interpretation of mortgage-origination dynamics before the financial crisis is whether the analysis should use geographically unrestricted variation in the outcome and controls, or should only exploit variation within counties. In our case, the variation cannot be restricted at the county level, because our main covariate – the penetration of large financial institutions in counties in 2007 – does not vary at the county level. We will therefore propose our results when using unrestricted geographic variation and within-state variation.

We estimate the following specification with OLS:

$$\begin{aligned}\Delta Outcome_{2010\rightarrow 2014,z} = \alpha + \beta\ LargePenetration_{k,2007} \times DoddFrank_t \\ + \Delta X'_{2010\rightarrow 2014,z}\delta + \Delta D'_{2010\rightarrow 2014,z}\phi + \eta_s + \epsilon_{z,k},\end{aligned} \tag{5}$$

where $\Delta Outcome_{2010\rightarrow 2014,z}$ is the growth of the outcome variable – average loan amount or loan number – from 2010 to 2014 at the zip-code-level z, $LargePenetration_{k,2007}$ is the share of mortgage activity by the top 20 national financial institutions in year 2007 and county k, ΔX and ΔD are the change in the set of observables at the individual and zip-code level; η_s is a set of state fixed effects. Note the variable $LargePenetration_{k,2007}$ does not vary over time. Restricting the variation within years using year fixed effects absorbs completely the variation in the level of $LargePenetration_{k,2007}$ across counties, which is why the level of the variable does not appear in the RHS of equation 5.

Table 9 reports the results for estimating equation 5. In all specifications, we cluster the standard errors at the county level to allow for correlation of the residuals of unknown form within counties and over time. In Panel A and Panel B, the outcome variable is the growth of the average loan amount at the zip-code level from 2010 to 2014. We omit reporting the coefficients associated with each of the control variables. In both panels, zip codes in counties with a higher penetration of large financial institutions in 2007 experience less growth in loan amounts for loans between $100k and $200k, and more growth in loan amounts above the non-conforming limit of $417k. The results are similar if we restrict the variation within states, although statistical significance is more sparse.

In Panel C and Panel D of Table 9, the outcome variable is the growth in the number of loans at the zip-code level from 2010 to 2014. In Panel C, we find zip codes in counties with a higher penetration of large financial institutions have less growth in the number of loans both below and above the conforming loan limit. This result is consistent with the fact that financial institutions cut the loans below the conforming loan limit not only along the intensive margin, but also along the extensive margin. At the same time, the number of loans just above the conforming loan limit also decreased, and because the average amount for the same group of loans increased, banks might have cut on the loans closer to the limit, compared to the loans well above the limit. When we restrict the variation within states, the results are qualitatively similar, although statistical significance is more sparse. In particular, the negative growth in the number of loans just above the conforming loan limit is only robust to allowing for variation across zip codes of any state.

Overall, the zip-code-level analysis confirms our baseline results at the individual-loan level for the intensive margin of lending. They suggest loans decreased also along the extensive margin, both below and above the conforming loan limit.

9 Conclusions

The mortgage provisions in Dodd-Frank incentivized lenders to redistribute credit from mid-size loans to large loans. Large banks reacted the most to the change, which affected households exposed to large banks more than other households. Financial institutions reduced their lending to medium-sized

loans by approximately 15% from 2010 to 2014, and increased lending to large loans by 21%.

This redistribution is robust to controlling for individual- and county-level observable determinants of the demand for mortgages, and holds at the individual- and zip-code level.

Economists are debating why middle-class households have not resumed consuming at the pre-Great Recession levels, which contributes to the slow recovery. The US administration has implemented several costly policies in an attempt to stimulate consumption by the middle class. Whether these policies had any substantial aggregate effects is unclear, but the redistribution of mortgage credit from the middle class to the wealthy goes against the aims of such policies.

Our results speak to the debate about the costs and benefits of regulating economic activity. Proponents of regulation aim to help vulnerable consumers. But regulators often underestimate the fact that lenders are private organizations competing in a free market, and hence they react to the incentives regulation creates based on their own objective function. In the case of Dodd-Frank, middle-class households did not obtain cheaper mortgages, but were cut out of the mortgage market altogether. Future research at the intersection of Finance, Accounting, and Law should delve deeper into the welfare implications of Dodd-Frank for consumers, lenders, and society as a whole.

References

ABA. 2016. 2016 ABA TRID Survey.

Adelino, M., A. Schoar, and F. Severino. 2016. Loan originations and defaults in the mortgage crisis: The role of the middle class. *Review of Financial Studies* .

Agarwal, S., G. Amromin, I. Ben-David, S. Chomsisengphet, and D. D. Evanoff. 2014. Predatory lending and the subprime crisis. *Journal of Financial Economics* 113:29–52.

Agarwal, S., G. Amromin, I. Ben-David, S. Chomsisengphet, T. Piskorski, and A. Seru. forthcoming. Policy intervention in debt renegotiation: Evidence from the home affordable modification program. *Journal of Political Economy* .

Albanesi, S., G. DeGiorgi, and J. Nosal. 2016. Credit growth and the financial crisis: A new narrative.

BankofAmerica. 2011. Plan to enhance enterprise-wide compliance program submission.

Barth, J., G. Caprio, and R. Levine. 2012. Guardians of finance - making regulators work for us.

———. 2015. Misdiagnoses and incomplete cures of financial regulatory failures.

Berger, D., N. Turner, and E. Zwick. 2016. Stimulating housing markets.

Braunstein, S. 2011. Statement before the subcommittee on insurance, housing, and community opportunity committee on financial services.

Broda, C., and J. A. Parker. 2014. The economic stimulus payments of 2008 and the aggregate demand for consumption. *Journal of Monetary Economics* 68:S20–36.

Brogaard, J., and K. Roshak. 2011. The effectiveness of the 2008-2010 housing tax credit.

Campbell, J., H. Jackson, B. Madrian, and P. Tufano. 2011. Consumer financial protection. *Journal of Economic Perspectives* 25(1).

Chodorow-Reich, G. 2014. The employment effects of credit market disruptions: Firm-level evidence from the 2008-2009 financial crisis. *The Quarterly Journal of Economics* 129:1–59.

Citibank. 2011. In the matter of Citigroup inc. and CitiFinancial Credit Company.

DeFusco, A., and A. Paciorek. 2016. The interest rate elasticity of mortgage demand: Evidence from bunching at the conforming loan limit. *American Economic Journal: Economic Policy* .

DiMaggio, M., A. Kermani, and C. Palmer. 2016. Unconventional monetary policy and the allocation of credit.

Favara, G., and M. Giannetti. Forthcoming. Forced asset sales and the concentration of outstanding debt: Evidence from the mortgage market. *Journal of Finance* .

Favara, G., and J. Imbs. 2015. Credit supply and the price of housing. *The American Economic Review* 105:958–92.

Foote, C., L. Loewenstein, and P. Willen. 2016. Cross-sectional patterns of mortgage debt during the housing boom: Stocks and flows.

FRB. 2011. Interagency review of foreclosure policies and practices .

Gargano, A., A. G. Rossi, and R. Wermers. forthcoming. The freedom of information act and the race towards information acquisition. *Review of Financial Studies* .

Glaeser, E., and A. Shleifer. 2003. The rise of the regulatory state. *Journal of Economic Literature* 41:401–25.

Goodman, L. 2015. The rebirth of securitization. Urban Institute White Paper Series.

Green, D., B. Melzer, J. A. Parker, and R. Pfirrmann-Powell. 2014. Accelerator or brake? microeconomic estimates of the cash for clunkers and aggregate demand.

Guiso, L., P. Sapienza, and L. Zingales. 2013. The determinants of attitudes toward strategic default on mortgages. *Journal of Finance* 68:1473–515.

Hembre, E. 2015. The price of homeowners: An examination of the first-time homebuyer tax credit.

JPMorganChase. 2011. Federal Reserve Consent Order - JPMorgan Chase Response.

Krugman, P. 2013. Friends of fraud. *New York Times* .

Mian, A., and A. Sufi. 2016. Household debt and defaults from 2000 to 2010: The credit supply view.

———. 2009. The consequences of mortgage credit expansion: Evidence from the u.s. mortgage default crisis. *The Quarterly Journal of Economics* 124:1449–96.

———. 2012. The effects of fiscal stimulus: Evidence from the 2009 'cash for clunkers' program. *Quarterly Journal of Economics* 127.

Morrison, D. 2013. Low jumbo rates moving members to banks. *Union Times Magazine* .

Palmer, C. 2015. Why did so many subprime borrowers default during the crisis: loose credit or plummeting prices?

Rodnyansky, A., and O. Darmouni. 2016. The effects of quantitative easing on bank lending behavior.

Stiglitz, J. 2009. A crisis of confidence. *The Guardian* .

SunTrust. 2011. Consent order reponse (training) .

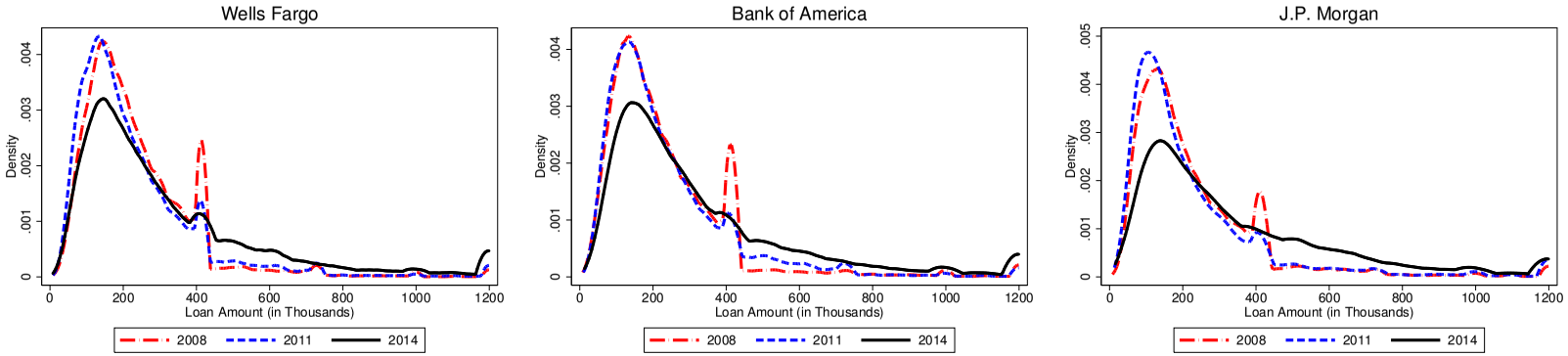
USBancorp. 2011. Federal reserve consent order action plan .

Warren, E. 2007. Unsafe at any rate. *Democracy* .

WellsFargo. 2011. FRB consent order implementation report .

Wilson, W. 1913. *The new freedom.* NY: Doubleday, Page Co.

Panel A. Lending Behavior by Top Originators



Panel B. Lending Behavior for Large Institutions, Small Institutions, and Non-bank Originators

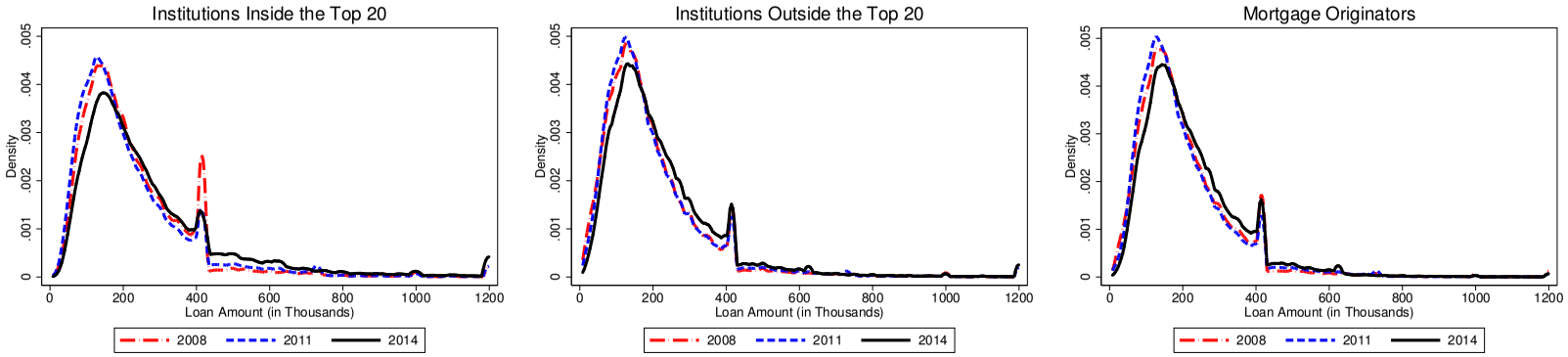


Figure 1: Panel A reports the loan size distributions for the top three institutions by mortgage lending activity over the sample: i.e. Wells Fargo, Bank of America and J.P. Morgan. Panel B reports the loan size distributions for the institutions that rank within (left plot) and outside (middle plot) the Top 20 by mortgage lending activity, as well as the loan size distributions for all the mortgage originators in our sample (right plot). Each plot reports densities for the years 2008, 2011, and 2014. Loan amounts have been winsorized at the 0.5% level.

Change in Lending by Bank size

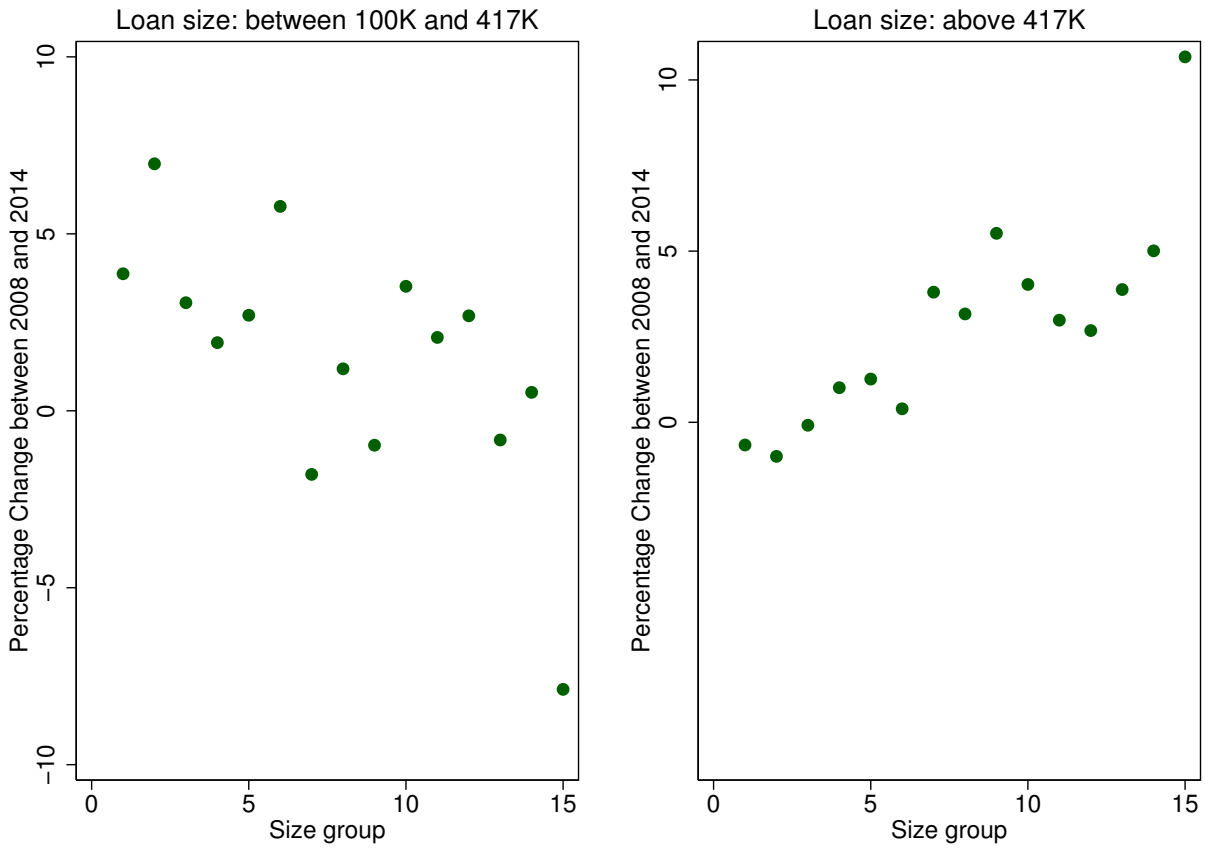


Figure 2: This figure reports the raw percentage change in loans originated by bank size between 2008 and 2014. The left panel considers loans between $100K and $417K. The right panel considers loans above $417K. We group institutions in 15 equal size groups based on total lending, and report the value-weighted change in lending for each group. We limit the sample to the 1,000 top lending institutions.

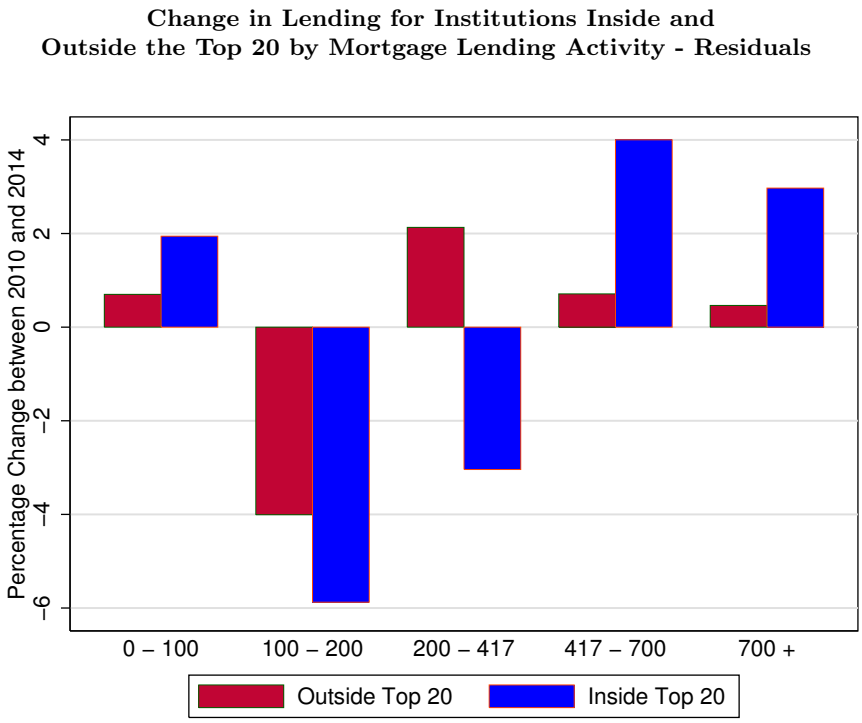


Figure 3: This figure reports residuals for regressing the value-weighted percentage changes between 2010 and 2014 in the fraction of loans generated within five size categories on a set of controls computed as averages at the lender level in 2007. Weights are given by each institution's mortgage lending activity. The controls are defined as follows: *ApplicantIncome* is the applicant income from HMDA. *Black*, *Asian*, and *Latino* are dummy variables that equal 1 if the applicant belong to the respective demographic group. *MedianHousePrice* is the median price of properties in the county in which the loan was originated from Zillow. The first size category comprises loans between zero and $100,000; the second loans between $100,000 and $200,000; the third loans between $200,000 and $417,000; the fourth loans between $417,000 and $700,000; and the fifth loans greater than $700,000.

Distribution of Percent Originations by Top 20 Institutions across Counties

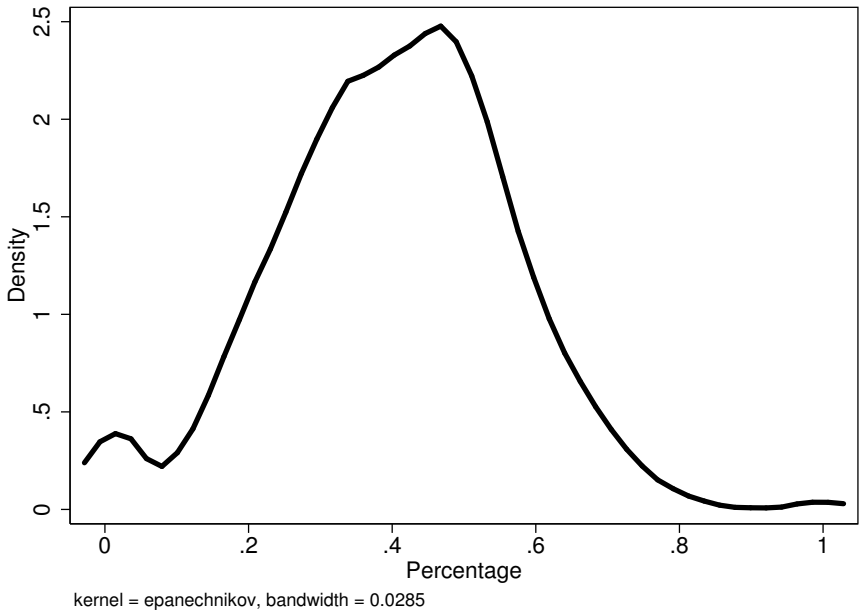
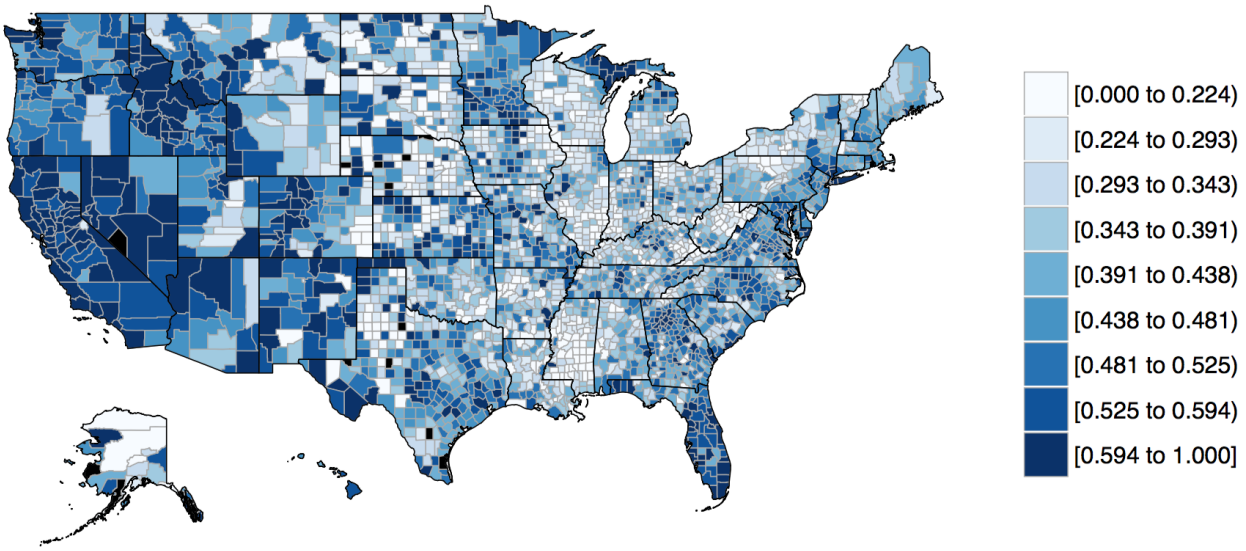


Figure 4: This figure reports kernel density estimates of the percentage of lending – across counties – generated by the Top 20 institutions for the year 2007.

Panel A. Percent Originations by Top 20 Institutions across US Counties



Panel B. Percent Originations by Top 20 Institutions across Iowan Counties

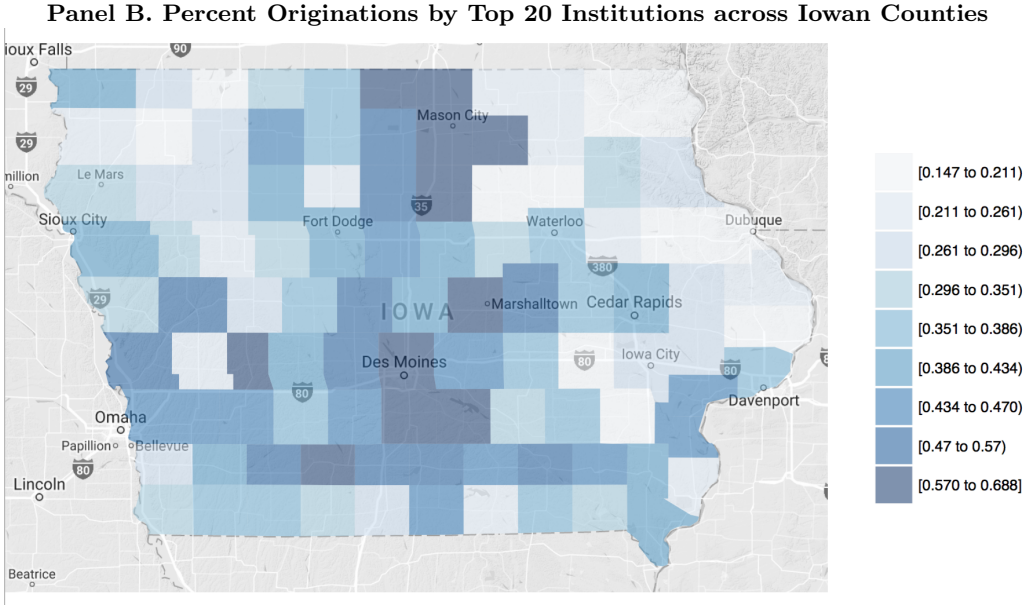


Figure 5: This figure reports choropleth maps of the percentage of lending – across counties – generated by the Top 20 institutions for the year 2007. Panel A focuses on the whole US. Panel B focuses on the state of Iowa.

Table 1. Summary Statistics

Panel A. Approved Loans

				Percentiles				
	Obs.	**Mean**	**St.dev.**	**5th**	**25th**	**50th**	**75th**	**95th**
Loan Amount ($000)	13,532,723	228.2	162.2	65.0	124.0	186.0	285.0	518.0
Top20 Share	13,532,723	.387	.137	.161	.286	.384	.485	.616
Applicant Income ($000)	13,532,723	95.0	121.8	28.0	47.0	71.0	109.0	225.0
Black	13,532,723	.070	.255	0	0	0	0	1
Asian	13,532,723	.070	.253	0	0	0	0	1
Latino	13,532,723	.095	.293	0	0	0	0	1
Median House Price ($000)	13,532,723	219.1	199.6	101.4	132.6	184.1	264.0	459.5
Share Foreclosed (perc. points)	10,516,574	.079	.085	.004	.023	.052	.103	.268

Panel B. Rejected Loans

				Percentiles				
	Obs.	**Mean**	**St.dev.**	**5th**	**25th**	**50th**	**75th**	**95th**
Loan Amount ($000)	5,983,994	225.3	182.1	45.0	106.0	176.0	290.0	563.0
Top20 Share	5,983,994	.409	.143	.167	.304	.411	.519	.648
Applicant Income ($000)	5,983,994	92.8	151.8	23.0	42.0	65.0	104.0	231.0
Black	5,983,994	.121	.326	0	0	0	0	1
Asian	5,983,994	.078	.269	0	0	0	0	1
Latino	5,983,994	.141	.348	0	0	0	0	1
Median House Price ($000)	5,983,994	229.1	127.0	99.1	134.2	187.6	297.9	497.6
Share Foreclosed (perc. points)	4,475,810	8.583	9.035	.409	2.302	5.738	11.362	29.119

Panel C. Zip code-Year Level

				Percentiles				
	Obs.	**Mean**	**St.dev.**	**5th**	**25th**	**50th**	**75th**	**95th**
Loan Amount ($000)	15,452	208.0	128.7	88.0	127.0	169.6	245.0	458.4
Top20 Share	15,452	.472	.124	.259	.388	.477	.558	.673
Applicant Income ($000)	15,452	93.4	58.0	49.0	63.3	77.5	102.4	189.7
Avg. Black	15,452	.057	.129	0	0	.012	.048	.280
Avg. Asian	15,452	.043	.098	0	0	.009	.037	.208
Avg. Latino	15,452	.072	.138	0	.004	.026	.067	.345
Avg. Median House Price ($000)	15,452	205.5	126.0	85.1	123.4	163.3	229.5	488.0
Number Loans (000)	15,452	55.2	113.9	.030	.603	5.7	51.8	290.2

This table reports descriptive statistics for the main variables in the analysis, observed at the individual loan level (Panel A and Panel B) or at the zip code-year level (Panel C) for the period 2008-2014. $LoanAmount$ and $ApplicantIncome$ are the loan amount requested and applicant income from HMDA. $Top20Share$ is the share of mortgages originated by top 20 financial institutions in the zip code in which each individual mortgage is originated. $Black$, $Asian$, and $Latino$ are dummy variables that equal 1 if the applicant belong to the respective demographic group. $MedianHousePrice$ is the median price of properties in the county in which the loan was originated from Zillow. $ShareForeclosed$ is the percentage of properties foreclosed in the zip-code in which the mortgage was originated, at the time of origination. Value in Panel C refer are the averages of the individual-loan level variables for approved loans at the zip code-year level.

Table 2. The Effect of Dodd-Frank on Approved Loan Amounts – Large vs. Small Institutions

	Below $100K	**$100K-$200K**	**$200K-$417K**	**$417K-$700K**	**Above $700K**
Top20×Dodd-Frank	-0.022	-0.022***	-0.012*	0.043***	0.241***
	(0.144)	(0.000)	(0.094)	(0.000)	(0.000)
Top20	-0.021	0.012	-0.003	-0.032**	-0.094***
	(0.199)	(0.117)	(0.762)	(0.012)	(0.009)
Log(Income)	0.098***	0.151***	0.214***	0.110***	0.159***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Black	0.011**	0.005**	0.001	0.001	-0.018***
	(0.036)	(0.033)	(0.661)	(0.765)	(0.000)
Asian	0.025***	0.009***	0.016***	-0.004	-0.009***
	(0.000)	(0.000)	(0.000)	(0.129)	(0.000)
Latino	0.011**	-0.002	-0.015***	-0.014***	-0.015***
	(0.033)	(0.426)	(0.000)	(0.000)	(0.000)
Avg-Black (county)	0.123*	-0.040	-0.049	-0.132***	0.204**
	(0.069)	(0.171)	(0.222)	(0.007)	(0.048)
Avg-Asian (county)	0.186*	-0.146***	0.011	0.112**	-0.061
	(0.086)	(0.007)	(0.849)	(0.020)	(0.366)
Avg-Latino (county)	0.266*	0.048	-0.220***	0.086	-0.144
	(0.065)	(0.374)	(0.000)	(0.455)	(0.207)
Median House Price	0.029	0.162***	0.121***	0.027**	0.137***
	(0.211)	(0.000)	(0.000)	(0.041)	(0.000)
Year Fixed Effects	✓	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓	✓
Observations	1,729,513	4,463,568	4,094,783	845,591	222,922
Adjusted R^2	0.061	0.170	0.286	0.190	0.286

This table reports regression coefficient estimates, and associated $p-$values, for the regression of the log loan size of the approved mortgages on the percentage of loans originated by the Top 20 mortgage lenders interacted with a Dodd-Frank dummy, the percentage of loans originated by the Top 20 mortgage lenders, and a number of control variables. The control variables are: *Log(Income)*, the log income of the applicant; dummy variables indicating whether the applicant is *Black*, *Asian*, and/or *Latino*; the average number of applicants for a given year that are black, *Avg-Black (county)*, asian, *Avg-Asian (county)*, and/or latino, *Avg-Latino (county)*; and *Median House Price*, the log median house price in a given county for a given year. The results are computed separately for loans in five size categories. The first set of results are reported in the first column and are associated with loans between zero and $100,000; the second column reports results for loans between $100,000 and $200,000; the third column for loans between $200,000 and $417,000; the fourth column for loans between $417,000 and $700,000; and the fifth column for loans greater than $700,000. All specifications include year fixed effects and county fixed effects. Standard errors are clustered at the county level.

Table 3. The Effect of Dodd-Frank on Approved Loan Amounts - Robustness

Panel A. Top 5	**Below $100K**	**$100K-$200K**	**$200K-$417K**	**$417K-$700K**	**Above $700K**
Top 5×Dodd-Frank	0.019	-0.042***	-0.041***	0.054***	0.316***
	(0.460)	(0.000)	(0.000)	(0.000)	(0.000)
Year and County F. E.	✓	✓	✓	✓	✓
Observations	1,729,513	4,463,568	4,094,783	845,591	222,922
Adjusted R^2	0.061	0.170	0.286	0.190	0.288
Panel B. Top 100	**Below $100K**	**$100K-$200K**	**$200K-$417K**	**$417K-$700K**	**Above $700K**
Top 100×Dodd-Frank	-0.012	-0.025***	0.000	0.026**	0.218***
	(0.312)	(0.000)	(0.940)	(0.029)	(0.000)
Year and County F. E.	✓	✓	✓	✓	✓
Observations	1,729,513	4,463,568	4,094,783	845,591	222,922
Adjusted R^2	0.061	0.170	0.286	0.190	0.285
Panel C. Quantiles House Prices	**Below $100K**	**$100K-$200K**	**$200K-$417K**	**$417K-$700K**	**Above $700K**
Top 20×Dodd-Frank	-0.027*	-0.017***	-0.015**	0.031***	0.214***
	(0.065)	(0.006)	(0.045)	(0.007)	(0.000)
Year and County F. E.	✓	✓	✓	✓	✓
Observations	1,729,513	4,463,568	4,094,783	845,591	222,922
Adjusted R^2	0.061	0.170	0.286	0.191	0.287
Panel D. All Interactions	**Below $100K**	**$100K-$200K**	**$200K-$417K**	**$417K-$700K**	**Above $700K**
Top 20×Dodd-Frank	-0.021	-0.013**	-0.004	0.048***	0.095***
	(0.171)	(0.037)	(0.635)	(0.000)	(0.000)
Year and County F. E.	✓	✓	✓	✓	✓
Observations	1,729,513	4,463,568	4,094,783	845,591	222,922
Adjusted R^2	0.061	0.170	0.286	0.192	0.290
Panel E. Local shocks	**Below $100K**	**$100K-$200K**	**$200K-$417K**	**$417K-$700K**	**Above $700K**
Top 20×Dodd-Frank	0.014	-0.022***	-0.002	0.115***	0.261***
	(0.445)	(0.001)	(0.788)	(0.000)	(0.000)
State*Year and County F. E.	✓	✓	✓	✓	✓
Observations	1,729,513	4,463,568	4,094,783	845,591	222,922
Adjusted R^2	0.063	0.171	0.287	0.193	0.291
Panel F. Sand States	**Below $100K**	**$100K-$200K**	**$200K-$417K**	**$417K-$700K**	**Above $700K**
Top 20×Dodd-Frank ×Sand State	0.191***	0.001	-0.018	0.142***	0.173***
	(0.000)	(0.953)	(0.414)	(0.000)	(0.009)
Year and County F. E.	✓	✓	✓	✓	✓
Observations	1,729,513	4,463,568	4,094,783	845,591	222,922
Adjusted R^2	0.061	0.170	0.286	0.191	0.287
Panel G. Exclude non-bank lenders	**Below $100K**	**$100K-$200K**	**$200K-$417K**	**$417K-$700K**	**Above $700K**
Top 20×Dodd-Frank	-0.029*	-0.033***	-0.023***	0.048***	0.255***
	(0.087)	(0.000)	(0.005)	(0.000)	(0.000)
Year and County F. E.	✓	✓	✓	✓	✓
Observations	976,130	2,214,717	1,959,778	471,795	175,286
Adjusted R^2	0.065	0.160	0.271	0.171	0.275
Panel H. Foreclosures	**Below $100K**	**$100K-$200K**	**$200K-$417K**	**$417K-$700K**	**Above $700K**
Top 20×Dodd-Frank	-0.038**	-0.021***	-0.029***	0.040***	0.263***
	(0.023)	(0.006)	(0.000)	(0.000)	(0.000)
Year and County F.E.	✓	✓	✓	✓	✓
Observations	1,169,782	3,368,168	3,398,971	758,441	196,416
Adjusted R^2	0.055	0.169	0.292	0.197	0.298

This table reports regression coefficient estimates, and associated $p-$values, for the regression of the log loan size of the approved mortgages on the percentage of loans originated by the Top 20 mortgage lenders interacted with a Dodd-Frank dummy, the percentage of loans originated by the Top 20 mortgage lenders, and a number of control variables. The control variables are as in Table 2. The results are computed separately for loans in five size categories, defined as in Table 2. Standard errors are clustered at the county level. Panel A and Panel B use the share of top 5 and top 100 large lenders serving each county-year. Panel C controls for 3 quantiles of county-year house prices. Panel D allows for a full set of interactions of the controls with the Dodd-Frank dummy. Panel E adds state-year fixed effects. Panel F compares the effect for counties in sand states (CA, NV, FL, AZ) and other states. Panel G excludes loans originated by non-bank lenders. Panel H controls for the share of foreclosed properties in the zip code in which the loan is originated.

Table 4. The Effect of Dodd-Frank on Denied Applicants' Income over Loan Amounts - Large vs. Small Institutions

	Below $100K	**$100K-$200K**	**$200K-$417K**	**$417K-$700K**	**Above $700K**
Top20×Dodd-Frank	0.010	0.030***	0.030***	-0.046***	-0.229***
	(0.647)	(0.000)	(0.006)	(0.000)	(0.000)
Top20	0.000	-0.011	-0.016	0.009	0.093***
	(0.992)	(0.207)	(0.194)	(0.555)	(0.003)
Log(Income)	0.875***	0.892***	0.839***	0.919***	0.879***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Black	-0.022***	-0.003	0.003	0.001	0.022***
	(0.000)	(0.227)	(0.364)	(0.589)	(0.001)
Asian	-0.072***	-0.010***	-0.015***	0.007**	0.015***
	(0.000)	(0.000)	(0.000)	(0.043)	(0.000)
Latino	-0.028***	0.002	0.020***	0.018***	0.012*
	(0.000)	(0.527)	(0.000)	(0.001)	(0.058)
Avg-Black (county)	-0.278***	0.088***	0.066**	-0.039	-0.072
	(0.001)	(0.002)	(0.014)	(0.266)	(0.341)
Avg-Asian (county)	-0.234*	0.085*	0.050	-0.146***	0.007
	(0.073)	(0.071)	(0.408)	(0.000)	(0.925)
Avg-Latino (county)	0.127	-0.093*	0.184***	-0.074	0.022
	(0.360)	(0.064)	(0.000)	(0.198)	(0.826)
Median House Price	0.120***	-0.141***	-0.105***	-0.042***	-0.119***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Year Fixed Effects	✓	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓	✓
Observations	1,133,521	1,676,520	1,521,463	359,659	116,519
Adjusted R^2	0.540	0.859	0.844	0.918	0.926

This table reports regression coefficient estimates, and associated $p-$values, for the regression of the log income-over-loan amounts of the denied mortgages on the percentage of loans originated by the Top 20 mortgage lenders interacted with a Dodd-Frank dummy, the percentage of loans originated by the Top 20 mortgage lenders, and a number of control variables. The control variables are: *Log(Income)*, the log income of the applicant; dummy variables indicating whether the applicant is *Black*, *Asian*, and/or *Latino*; the average number of applicants for a given year that are black, *Avg-Black (county)*, asian, *Avg-Asian (county)*, and/or latino, *Avg-Latino (county)*; and *Median House Price*, the log median house price in a given county for a given year. The results are computed separately for loans in five size categories. The first set of results are reported in the first column and are associated with loans between zero and $100,000; the second column reports results for loans between $100,000 and $200,000; the third column for loans between $200,000 and $417,000; the fourth column for loans between $417,000 and $700,000; and the fifth column for loans greater than $700,000. All specifications include year fixed effects and county fixed effects. Standard errors are clustered at the county level.

Table 5. Validity of the Instrument: Balancing of Variables

	Quantile Large Banks in 2007				
	1	**2**	**3**	**4**	**St. Dev.**
Growth Loan Amount 2007-2010 (<$100k)	0.018	0.014	0.029	0.046	0.175
Growth Loan Amount 2007-2010 ($100k-$200k)	0.009	0.003	0.006	-0.007	0.074
Growth Loan Amount 2007-2010 ($200k-$417k)	-0.011	-0.008	-0.009	-0.025	0.096
Growth Loan Amount 2007-2010 ($417k-$700k)	-0.044	-0.039	-0.052	-0.045	0.109
Growth Loan Amount 2007-2010 (>$700k)	-0.015	-0.014	-0.006	-0.015	0.140
Avg. Black county	0.044	0.047	0.048	0.052	0.139
Avg. Asian county	0.005	0.007	0.014	0.022	0.076
Avg. Latino county	0.122	0.036	0.054	0.069	0.154
Share Foreclosed Properties	0.004	0.005	0.004	0.007	0.010
Share Middle-Class Households 2007	0.388	0.378	0.370	0.369	0.041
Share Middle-Class Households with a Mortgage 2007	0.535	0.520	0.500	0.475	0.086
Share County Income from Stock Dividends 2007	0.060	0.066	0.067	0.072	0.026
Share Workforce in Public Administration 2007	0.085	0.085	0.083	0.084	0.033

This table reports the sample mean of a set of variables within four quantiles of US counties sorted by the percent of county-level mortgage activity by large financial institutions. The last column reports the standard deviation of each variable to allow the assessment of the magnitude of the differences of mean point estimates across quantiles. Growth Loan Amount 2007-2010 is the sample mean county growth of the average loan amounts between 2007 and 2010, when Dodd-Frank was not in effect. All other sample means are computed for observations throughout the sample period, that is, 2007-2014, because we need to verify the balancing of covariates both before and after Dodd-Frank was in effect. Avg. Black county, Avg. Asian county, and Avg. Latino county are the mean share of Black, Asian, and Latino population in the county in the period 2007-2014. Share Foreclosed Houses is the ratio of units subjects to foreclosure throughout the time period 2007-2014. Share Middle-Class Households 2007 is the share of households with an annual gross income between $25K and $75K in 2007. Share Middle-Class Households with a Mortgage 2007 is the share of households that had a mortgage outstanding in 2007 among the households with an annual gross income between $25K and $75K. Share County Income from Stock Dividends 2007 is the ratio between the gross income from stock dividends and interests over the overall gross income in 2007. Share Workforce in Public Administration 2007 is the share of households in the workforce that were employed in public adminsitration jobs in 2007.

Table 6. Validity of the Instrument: Reduced Form Regressions

Panel A. Reduced Form

	Below $100K	$100K-$200K	$200K-$417K	$417K-$700K	Above $700K
Top20 in 2007×Dodd-Frank	0.049**	-0.025***	-0.027***	0.056***	0.242***
	(0.011)	(0.001)	(0.007)	(0.000)	(0.000)
Controls	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓
County fixed effects	✓	✓	✓	✓	✓
Observations	2,031,971	5,277,768	4,902,513	1,035,592	284,879
Adjusted R^2	0.061	0.170	0.286	0.190	0.287

Panel B. Reduced Form with Endogenous Regressor

	Below $100K	$100K-$200K	$200K-$417K	$417K-$700K	Above $700K
Top20×Dodd-Frank	-0.028	0.001	0.004	0.088***	0.092
	(0.125)	(0.890)	(0.691)	(0.000)	(0.120)
Top20 in 2007×Dodd-Frank	0.070***	-0.026***	-0.030***	-0.022	0.157***
	(0.001)	(0.004)	(0.008)	(0.258)	(0.004)
Controls	✓	✓	✓	✓	✓
Year Fixed Effects	✓	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓	✓
Observations	2,031,971	5,277,768	4,902,513	1,035,592	284,879
Adjusted R^2	0.058	0.171	0.287	0.177	0.285

Panel A of this table reports regression coefficient estimates, and associated $p-$values, for the regression of the log loan amounts of the applied mortgages on the percentage of loans originated by the Top 20 mortgage lenders in 2007 interacted with a Dodd-Frank dummy and a number of control variables (reduced form specification). Panel B adds the interaction between the percentage of loans originated by the Top 20 mortgage lenders each year interacted with a Dodd-Frank dummy as a control variable. The control variables are: *Income*, the log income of the applicant; dummy variables indicating whether the applicant is *Black*, *Asian*, and/or *Latino*; the average number of applicants for a given year that are black, *Avg-Black (county)*, asian, *Avg-Asian (county)*, and/or latino, *Avg-Latino (county)*; and *Median House Price*, the log median house price in a given county for a given year. The results are computed separately for loans in five size categories. The first set of results are reported in the first column and are associated with loans between zero and $100,000; the second column reports results for loans between $100,000 and $200,000; the third column for loans between $200,000 and $417,000; the fourth column for loans between $417,000 and $700,000; and the fifth column for loans greater than $700,000. All specifications include year fixed effects and county fixed effects. Standard errors are clustered at the county level.

Table 7. Instrumental-Variable Results: Approved Loan Amounts - Large vs. Small Institutions

	Below $100K	**$100K-$200K**	**$200K-$417K**	**$417K-$700K**	**Above $700K**
Top 20×Dodd-Frank	0.008	-0.044***	-0.026**	0.036***	0.338***
	(0.749)	(0.000)	(0.041)	(0.002)	(0.000)
Log(Income)	0.098***	0.151***	0.214***	0.110***	0.159***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Black	0.011**	0.005**	0.001	0.001	-0.018***
	(0.036)	(0.033)	(0.660)	(0.766)	(0.000)
Asian	0.025***	0.009***	0.016***	-0.004	-0.009***
	(0.000)	(0.000)	(0.000)	(0.129)	(0.000)
Latino	0.011**	-0.002	-0.015***	-0.014***	-0.015***
	(0.032)	(0.432)	(0.000)	(0.000)	(0.000)
Avg. Black county	0.143**	-0.055*	-0.054	-0.127**	0.207*
	(0.036)	(0.061)	(0.186)	(0.012)	(0.088)
Avg. Asian county	0.175*	-0.145***	0.009	0.113**	-0.048
	(0.099)	(0.006)	(0.870)	(0.013)	(0.625)
Avg. Latino county	0.261*	0.052	-0.222***	0.074	-0.163
	(0.070)	(0.330)	(0.000)	(0.526)	(0.174)
Log(Median House Price)	0.030	0.161***	0.121***	0.031**	0.149***
	(0.198)	(0.000)	(0.000)	(0.017)	(0.000)
Year Fixed Effects	✓	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓	✓
Kleibergen-Paap F-Statistic	475.7	453.0	417.9	149.5	101.6
Observations	1,729,513	4,463,568	4,094,783	845,591	222,922

This table reports second-stage regression coefficient estimates, and associated $p-$values, for the regression of the log loan amounts of the applied mortgages on the percentage of loans originated by the Top 20 mortgage lenders interacted with a Dodd-Frank dummy and a number of control variables. The interaction is the endogenous variable, which is instrumented in the first stage with the percentage of loans originated by the Top 20 mortgage lenders in 2007 with a Dodd-Frank dummy. The level of the percentage of loans originated is not included, because the variation in its instrument is absorbed by the county fixed effects. The control variables are: *Log(Income)*, the log income of the applicant; dummy variables indicating whether the applicant is *Black*, *Asian*, and/or *Latino*; the average number of applicants for a given year that are black, *Avg-Black (county)*, asian, *Avg-Asian (county)*, and/or latino, *Avg-Latino (county)*; and *Median House Price*, the log median house price in a given county for a given year. The results are computed separately for loans in five size categories. The first set of results are reported in the first column and are associated with loans between zero and $100,000; the second column reports results for loans between $100,000 and $200,000; the third column for loans between $200,000 and $417,000; the fourth column for loans between $417,000 and $700,000; and the fifth column for loans greater than $700,000. All specifications include year fixed effects and county fixed effects. Standard errors are clustered at the county level.

Table 8. Instrumental-Variable Results: Denied Applicants' Income over Loan Amount - Large vs. Small Institutions

	Below $100K	**$100K-$200K**	**$200K-$417K**	**$417K-$700K**	**Above $700K**
Top 20×Dodd-Frank	0.041	0.040***	0.051***	-0.045***	-0.327***
	(0.038)	(0.013)	(0.018)	(0.016)	(0.037)
Log(Income)	0.875***	0.892***	0.839***	0.919***	0.879***
	(0.004)	(0.001)	(0.002)	(0.003)	(0.004)
Black	-0.022***	-0.003	0.003	0.002	0.022***
	(0.006)	(0.002)	(0.003)	(0.003)	(0.007)
Asian	-0.072***	-0.010***	-0.015***	0.007**	0.016***
	(0.006)	(0.001)	(0.003)	(0.003)	(0.003)
Latino	-0.028***	0.002	0.020***	0.018***	0.013**
	(0.006)	(0.003)	(0.002)	(0.005)	(0.006)
Avg. Black county	-0.273***	0.092***	0.072**	-0.039	-0.057
	(0.083)	(0.028)	(0.029)	(0.035)	(0.083)
Avg. Asian county	-0.247*	0.082*	0.045	-0.146***	0.031
	(0.132)	(0.046)	(0.058)	(0.039)	(0.080)
Avg. Latino county	0.133	-0.094*	0.182***	-0.070	0.085
	(0.137)	(0.051)	(0.052)	(0.057)	(0.101)
Log(Median House Price)	0.121***	-0.140***	-0.104***	-0.043***	-0.129***
	(0.033)	(0.010)	(0.010)	(0.012)	(0.027)
Year Fixed Effects	✓	✓	✓	✓	✓
County Fixed Effects	✓	✓	✓	✓	✓
Kleibergen-Paap F-Statistic	496.9	420.8	341.2	145.9	114.39
Observations	1,133,521	1,676,520	1,521,463	359,659	116,519

This table reports second-stage regression coefficient estimates, and associated $p-$values, for the regression of the log income-over-loan amounts of the denied mortgages on the percentage of loans originated by the Top 20 mortgage lenders interacted with a Dodd-Frank dummy and a number of control variables. The interaction is the endogenous variable, which is instrumented in the first stage with the percentage of loans originated by the Top 20 mortgage lenders in 2007 with a Dodd-Frank dummy. The level of the percentage of loans originated is not included, because the variation in its instrument is absorbed by the county fixed effects. The control variables are: *Log(Income)*, the log income of the applicant; dummy variables indicating whether the applicant is *Black*, *Asian*, and/or *Latino*; the average number of applicants for a given year that are black, *Avg-Black (county)*, asian, *Avg-Asian (county)*, and/or latino, *Avg-Latino (county)*; and *Median House Price*, the log median house price in a given county for a given year. The results are computed separately for loans in five size categories. The first set of results are reported in the first column and are associated with loans between zero and $100,000; the second column reports results for loans between $100,000 and $200,000; the third column for loans between $200,000 and $417,000; the fourth column for loans between $417,000 and $700,000; and the fifth column for loans greater than $700,000. All specifications include year fixed effects and county fixed effects. Standard errors are clustered at the county level.

Table 9. The Effect of Dodd-Frank on Approved Loan Amounts
Zip-Code Level Analysis

Panel A. Loan Amount Without State Fixed Effects

	Below $100K	$100K-$200K	$200K-$417K	$417K-$700K	Above $700K
Share Activity Top 20 in 2007	-0.013	-0.032***	0.009	0.074***	0.209***
	(0.704)	(0.007)	(0.657)	(0.003)	(0.000)
State Fixed Effects	✗	✗	✗	✗	✗
Growth Controls 2007-2010	✓	✓	✓	✓	✓
Observations	7,045	7,469	7,374	4,890	2,711
Adjusted R^2	0.001	0.086	0.061	0.012	0.030

Panel B. Loan Amount With State Fixed Effects

	Below $100K	$100K-$200K	$200K-$417K	$417K-$700K	Above $700K
Share Activity Top 20 in 2007	0.026	-0.042***	-0.030	0.056	0.315***
	(0.479)	(0.001)	(0.108)	(0.169)	(0.000)
State Fixed Effects	✓	✓	✓	✓	✓
Growth Controls 2007-2010	✓	✓	✓	✓	✓
Observations	7,045	7,469	7,374	4,890	2,711
Adjusted R^2	0.010	0.096	0.086	0.040	0.090

Panel C. Number of Loans Without State Fixed Effects

	Below $100K	$100K-$200K	$200K-$417K	$417K-$700K	Above $700K
Share Activity Top 20 in 2007	0.277*	-0.853***	-1.728***	-1.454***	0.335
	(0.079)	(0.000)	(0.000)	(0.001)	(0.435)
State Fixed Effects	✗	✗	✗	✗	✗
Growth Controls 2007-2010	✓	✓	✓	✓	✓
Observations	7,045	7,469	7,374	4,890	2,711
Adjusted R^2	0.001	0.086	0.061	0.012	0.030

Panel D. Number of Loans With State Fixed Effects

	Below $100K	$100K-$200K	$200K-$417K	$417K-$700K	Above $700K
Share Activity Top 20 in 2007	0.447***	-0.157	-1.247***	0.501	0.533
	(0.001)	(0.164)	(0.000)	(0.280)	(0.358)
State Fixed Effects	✓	✓	✓	✓	✓
Growth controls 2007-2010	✓	✓	✓	✓	✓
Observations	7,045	7,469	7,374	4,890	2,711
Adjusted R^2	0.104	0.139	0.083	0.112	0.053

This table reports regression coefficient estimates, and associated $p-$values, for the regression of the growth of average loan amount in zip codes from 2010 to 2014 (Panel A and Panel B), or the growth of the number of loans originated in zip codes from 2010 to 2014 (Panel C and Panel D), on the percentage of loans originated by the Top 20 mortgage lenders in 2007 in the counties where the zip codes lie, plus the growth of a number of control variables. The control variables, whose associated coefficients are not reported due to space constraints, are: *Income*, the log income of the applicant; the average number of applicants for a given year that are black, *Avg-Black (county)*, asian, *Avg-Asian (county)*, and/or latino, *Avg-Latino (county)*; and *Median House Price*, the log median house price in a given county for a given year. The results are computed separately for loans in five size categories. The first set of results are reported in the first column and are associated with loans between zero and $100,000; the second column reports results for loans between $100,000 and $200,000; the third column for loans between $200,000 and $417,000; the fourth column for loans between $417,000 and $700,000; and the fifth column for loans greater than $700,000. All specifications include year fixed effects and county fixed effects. Standard errors are clustered at the county level.

Online Appendix:
Ditching the Middle Class with Consumer Protection Regulation

Francesco D'Acunto and Alberto Rossi

Not for Publication

A.1 Additional Data Description and Alternative Explanations

In this section, we describe the construction of the variables that proxy for counties' exposure to wealth polarization and other demand shocks after the financial crisis.

Our source of data for information on household-level stock market participation, mortgage holding, and occupation is the American Community Survey (ACS). The ACS is a yearly survey-based repeated cross section that obtains demographic and economic information on 1% of the US population. The stratification is performed in two stages. First, the US Census Bureau selects a set of counties among which they then select the individual households. All members of the selected households have to respond to the survey.

The ACS collects detailed information on household-members' age, race, ethnicity, income, workforce status and occupation, as well as a set of economic dimensions, including the overall gross income and its components based on IRS classification. These components include: (i) income from wages and salaries; (ii) income from business and farms; (iii) income from interest, dividend, and rents; (iv) income from retirement plans.

The micro-data files are available for download from the Integrated Public Use Microdata Series (IPUMS) from the following website: https://usa.ipums.org/usa-action/variables/group. We downloaded the ACS three-year cross section for the year 2005-2007, which we denote as the 2007 sample. This sample period includes the three years before the collapse of Lehman Brothers and the start of the 2008-2009 financial crisis. The raw sample includes 8,842,783 individual observations. To run our analysis, we perform a set of steps to select the population we are interested in. In the first step, we exclude observations which only report their state of residence, but not the county of residence. In a second step, we exclude all observations the ACS categorize as not applicable in terms of employment status, and for which individual income is missing. In a third step, we exclude individuals for which the total family income is missing.

We construct the county-level measures as follows: (i) *share of middle-class households in 2007*, is the share of household that have an overall gross income between $25K and $75K in 2007; (ii) *share of middle class households with mortgages in 2007*, is the share of middle-class households that own

their primary residence and hold at least one mortgage; (iii) *exposure to the stock market in 2007*, is the ratio between the sum of income from interest, dividends, and rent over the sum of the gross income of households computed at the county level; (iv) *share of workers in the public administration*, is the share of individuals that declares to be part of the workforce, and whose occupation is in the 1-digit industry classification "9" in the US Census industry definition.

In Table A.1, we split the counties in terciles, and report estimates of our baseline specification for the bottom and the top terciles.

A.2 Execution and Timing of the Economic Effects of Dodd-Frank's Mortgage Provisions

President Obama signed Dodd-Frank into law on July 21, 2010. Most provisions in Title XIV (Mortgage Reform and Anti-Predatory Lending Act) were not self-executing. According to Section 1400, the provisions should have taken effect on the earliest date between the "date on which the final regulations implementing such section, or provision, take effect," or "18 months after the designated transfer date," in which the transfer date refers to the date on which the Federal Reserve was set to transfer its supervisory and regulatory powers based on the *Truth in Lending Act* to the Bureau of Consumer Financial Protection. This date was set to July 21, 2011, and hence – at the time of signing into law of Dodd-Frank – the latest possible date by which financial institutions had to comply with the provisions was January 21, 2013.

Whereas we know the last date by which financial institutions had to comply with the Dodd-Frank provisions, we do not know when each institution started to invest resources and effort towards this goal. On the one hand, it would take months or years to set up the training systems and hire the personnel the banks would have needed to comply with Dodd-Frank, especially for the larger banks. On the other hand, there was uncertainty about the actual date on which the provisions of Dodd-Frank would be executable, because the CFPB could have produced the required regulations before the final date of January 21, 2013. For these reasons, the regulation affected the fixed and marginal costs of originating loans well before it was executed. Because banks faced this increase in actual and projected costs well before 2013, their optimization implies a shift of their lending towards loan with a higher expected return, that is, larger loans.

Even though we do not observe directly the date at which all banks started their compliance process, we do know the exact timing of the compliance process for 14 servicers and the parent companies of 12 of them, which were deemed to be systemically important by the Federal Reserve System in November 2010. In April 2011, the Federal Reserve System disclosed a set of enforcement actions that imposed the immediate compliance to new standards for mortgage originations that were "substantially similar" to the provisions approved in Dodd-Frank (Braunstein (2011)). These banks,

therefore, had no choice but to start complying with the new provisions immediately after April 2011. The set of 12 parent companies includes the following institutions: Ally Bank/GMAC, Bank of America, Citibank, Everbank, HSBC, JPMorgan Chase, MetLife, OneWest, PNC, SunTrust, U.S. Bank, and Wells Fargo (see FRB (2011), footnote 1).

One might wonder whether financial institutions indeed started to face the costs of preparing their compliance to the provisions in Dodd-Frank starting in 2011. The enforcement actions are key in this respect, because the confidential filings of the institutions to the Federal Reserve Bank of New York describe these costs explicitly.

On July 12, 2011, Citibank's CEO Michael Corbat stated that "Citi [...] has dedicated significant management and financial resources to these efforts and ongoing business needs. Incremental expenses associated with these actions are estimated to be in excess of $(confidential figure) for 2011." (see Citibank (2011), page 3).

On the same date, Bank of America stated that "We recognize the significant effort, time, and resources necessary to implement the Plan and to verify its consistency and rigor once implemented." (see BankofAmerica (2011)). On December 8, 2011, JPMorgan Chase stated that "Home Lending Compliance has been actively adding resources and upgrading talent and, in mid-2010, a dedicated position of Head of Compliance for Home Lending, reporting to a newly-created Head of Compliance Retail Financial Services, was established and the role filled." (see JPMorganChase (2011)).

On December 9, 2011, US Bancorp stated that "The Bank is fully committed to securing all necessary resources to respond to the Orders in an effective and timely manner." (see USBancorp (2011)).

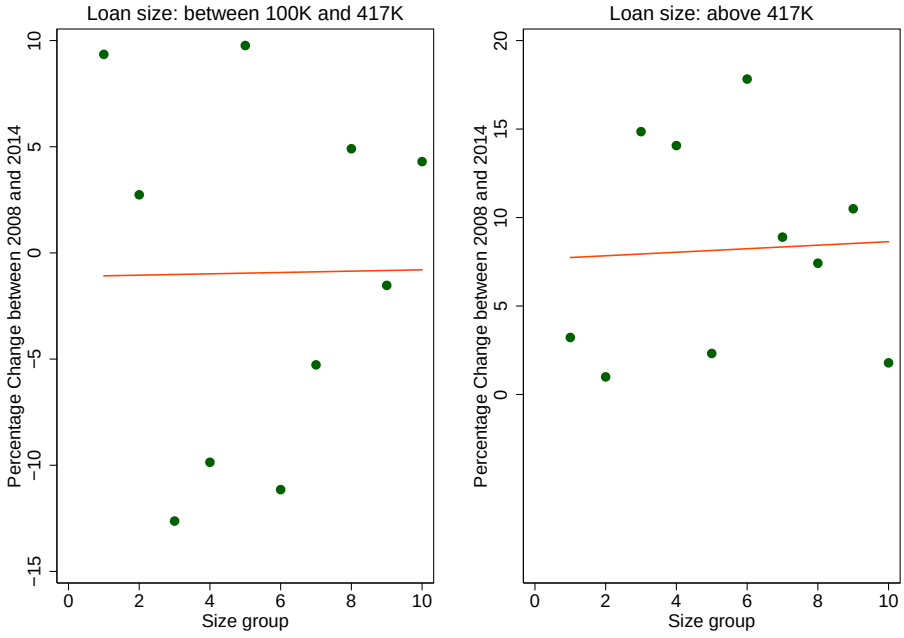
On December 16, 2011, SunTrust stated that "SunTrust's consent order became effective on April 13, 2011. In anticipation of the Order, SunTrust created a formal program organization, comprised of individual working teams for each section of the Consent Order." (see SunTrust (2011)).

On December 23, 2011, Wells Fargo stated that "we have expanded our Dodd-Frank Program Office [...] Although the nominal date for this office to be operating is 1/1/2012, in fact personnel are in place and managing efforts [...]" (see WellsFargo (2011)).

All these documents provide *prima facie* evidence that financial institutions started to invest resources, and hence faced additional fixed and marginal costs to originate loans, since early 2011. Please access the original versions of the documents at the following website: https://www.federalreserve.gov/newsevents/press/enforcement/20120227aletters.htm.

Change in Lending by Bank Risk and Bank Size

Panel A. Bank Risk



Panel B. Bank Size

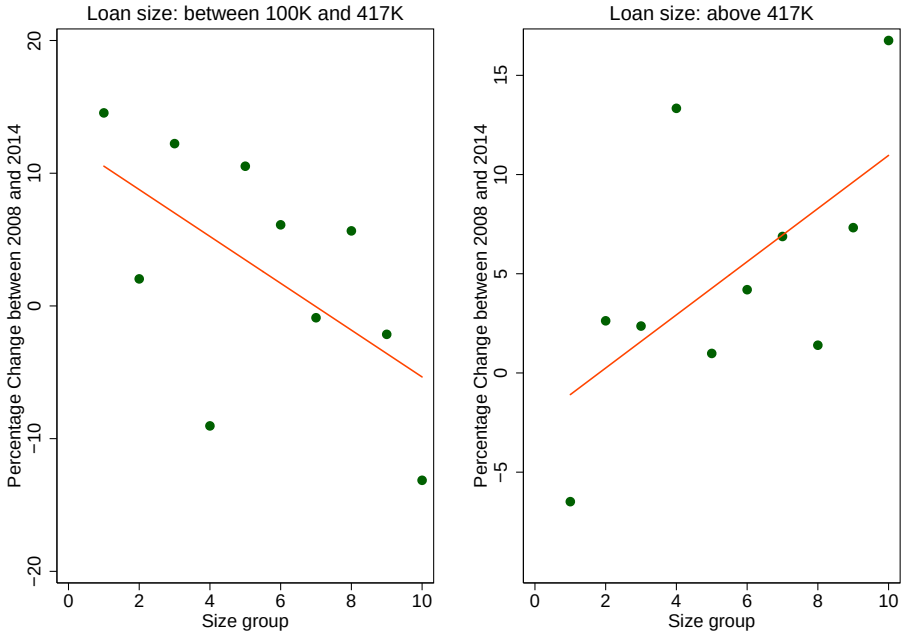


Figure A.1: This figure reports the raw percentage change in loans originated by bank risk between 2008 and 2014. The left panel considers loans between \$100K and \$417K. The right panel considers loans above \$417K. The measure of bank risk is the share of reserves over the total amount of non-performing loans held by the bank. Because we do not observe this measure of bank risk for all the institutions in our sample, we focus on the 100 riskiest institutions, and group them in 15 equal size groups based on bank risk. We report the value-weighted change in lending for each group.

Table A.1. Post-Crisis Wealth Polarization and Exposure to Financial Crisis

	Below $100K	$100K-$200K	$200K-$417K	$417K-$700K	Above $700K
	Panel A. Share Middle-Class Households in 2007				
	Bottom Third Counties				
Top20×Dodd-Frank	-0.019	-0.013	-0.044**	0.048***	0.231***
	(0.733)	(0.500)	(0.010)	(0.005)	(0.000)
	Top Third Counties				
Top20×Dodd-Frank	-0.014	-0.018	-0.010	0.022	-0.102**
	(0.567)	(0.277)	(0.388)	(0.336)	(0.015)
	Panel B. Share Middle-Class Households with Mortgages in 2007				
	Bottom Third Counties				
Top20×Dodd-Frank	-0.044	-0.058**	-0.034**	0.045**	0.208***
	(0.237)	(0.039)	(0.021)	(0.023)	(0.000)
	Top Third Counties				
Top20×Dodd-Frank	-0.034	-0.013	-0.016	0.046	0.043
	(0.153)	(0.444)	(0.530)	(0.186)	(0.606)
	Panel C. Exposure to Stock Market in 2007				
	Bottom Third Counties				
Top20×Dodd-Frank	-0.057*	-0.016	-0.003	0.004	0.205**
	(0.066)	(0.284)	(0.832)	(0.848)	(0.015)
	Top Third Counties				
Top20×Dodd-Frank	0.038	-0.003	-0.005	0.062***	0.218***
	(0.264)	(0.846)	(0.807)	(0.001)	(0.000)
	Panel D. Share Workforce in Public Administration in 2007				
	Bottom Third Counties				
Top20×Dodd-Frank	-0.026	-0.040*	-0.014	0.030**	0.310***
	(0.395)	(0.064)	(0.315)	(0.035)	(0.000)
	Top Third Counties				
Top20×Dodd-Frank	-0.048	-0.025	0.022	0.005	0.142**
	(0.204)	(0.188)	(0.197)	(0.799)	(0.035)

This table reports regression coefficient estimates, and associated $p-$values, for the regression of the log loan size of the approved mortgages on the percentage of loans originated by the Top 20 mortgage lenders interacted with a Dodd-Frank dummy, the percentage of loans originated by the Top 20 mortgage lenders, and a number of control variables. The control variables are as in Table 2. The results are computed separately for loans in five size categories, defined as in Table 2. Standard errors are clustered at the county level. In each Panel, results are reported separately for estimating the specification using approved loans in counties below the bottom tercile and above the top tercile of counties sorted by different variables. Panel A sorts counties based on the share of households with a gross income between $25K and $75K in 2007. Panel B sorts counties based on the share of households that had a mortgage outstanding in 2007 among the households with an annual gross income between $25K and $75K. Panel C sorts counties by the ratio between the gross income from stock dividends and interests over the overall gross income in 2007. Panel D sorts counties based on the share of households in the workforce that were employed in public adminsitration jobs in 2007.