CONSUMER FINANCIAL PROTECTION BUREAU | DECEMBER 2023

The Neighborhood Geography of Mortgage Lending

Office of Research Staff Report No. 2023-02

Giordano Palloni

NOTE: This report represents the views of the author and does not indicate concurrence either by the CFPB or other members of the CFPB staff.

Consumer Financial Protection Bureau

Table of contents

1. Introduction

Housing was the largest expenditure category for American consumers in 2020, on average accounting for 34.9 percent of total household spending.[1] For the nearly 50 million households in owner-occupied housing with a mortgage,[2] this spending is driven by mortgage-related costs, and changes in these expenses can importantly affect their ability to save and afford goods and services other than housing. For many renter households and others in non-owner-occupied housing, the financial and non-financial costs of a mortgage can be daunting; for renter households on the margin of homeownership, changes in mortgage pricing and availability can affect whether and where they are able to purchase a home. While perfectly competitive markets should result in similar loan prices and rejection rates for borrowers posing a similar credit risk,[3] there is growing evidence that price dispersion and differences in rejection rates exist even between observably similar applicants.[4] Understanding how the availability and pricing of home financing varies across consumers is therefore relevant to the welfare of millions of American consumers.

This report explores one potential correlate of variation in mortgage pricing and availability between observably similar borrowers: the number of mortgage originators in different neighborhoods. The analysis begins by describing how the number of mortgage originators per capita[5]—defined as the number of institutions originating mortgages per 1,000 residents

[1] See U.S. Bureau of Labor Statistics, "Consumer Expenditures in 2020," December, 2021, https://www.bls.gov/opub/reports/consumer-expenditures/2020/home.htm.

[2] U.S. Census Bureau. 2020. *2016-2020 American Community Survey 5-year Data Release, Table A10041*.

[3] Admittedly, there is evidence of some price dispersion even in markets with homogeneous goods and close to zero marginal search cost to consumers. See Michael R. Baye, John Morgan, and Patrick Scholten, "Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site." The Journal of Industrial Economics, Volume 52, No. 4 (December 2004), https://doi.org/10.1111/j.0022-1821.2004.00236.x.

[4] See Neil Bhutta and Aurel Hizmo, "Do Minorities Pay more for Mortgages?" The Review of Financial Studies, Volume 34, Issue 2, (February 2021), https://academic.oup.com/rfs/article/34/2/763/5827007; Neil Bhutta, Andreas Fuster, and Aurel Hizmo, "Paying Too Much? Borrower Sophistication and Overpayment in the US Mortgage Market." (July 19, 2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3422904; Greg Buchak and Adam Jorring, "Do Mortgage Lenders Compete Locally? Implications for Credit Access." (January 7, 2021), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3762250; and Alexei Alexandrov and Elizabeth Saunders "Mortgage Data Shows that Borrowers Could Save $100 a month (or more) by choosing cheaper lenders," CFPB Blog, May 24, 2023, https://www.consumerfinance.gov/about-us/blog/mortgage-data-shows-borrowers-could-save-100-month-choosing-cheaper-lenders/.

[5] I use originators *per capita* to refer to the number of originators per 1,000 residents. I scale by 1,000 residents instead of by resident—which would enable the correct use of *per capita*—to ensure that the results do not require more than two decimal places to be interpretable. I also use the terms originators per 1,000 residents, originating lenders per 1,000 residents, and originating institutions per 1,000 residents.

between 2018–2020—differed across neighborhoods in the United States.[6] This measure includes all financial institutions that originated mortgages in a census tract regardless of whether they had a physical presence in the neighborhood. A benefit of this measure is that it includes nondepository institutions that did not have physical branches. Variation in originators per capita across neighborhoods, however, undoubtedly reflects differences in both demand-side and supply-side factors that are not observable in the data. Such demand-side factors could include, for instance, consumer shopping behavior, peer information and referrals, and financial literacy; relevant supply-side factors could include the physical proximity of financial institution branches, marketing behavior, and the competitive landscapes for financial institutions and for service purveyors (for example, home appraisers or title insurance policy providers). Next, the report investigates how neighborhood demographic characteristics were related to the observed differences in originators per capita and whether the variation in originators per capita was associated with differences in neighborhood-level loan outcomes. The analysis then turns to the question of whether neighborhood originators per capita was associated with mortgage market outcomes for individual consumers, using regressions that focus on transactions that posed similar credit risk to lenders.

1.1 Key Findings

Key findings of this report include:

- The number of mortgage-originating institutions and the number of mortgage originations differed starkly across neighborhoods in the United States between 2018–2020:
  - Ranked by the number of institutions originating mortgages per 1,000 residents, census tracts ranged from 8.5 institutions originating mortgages per 1,000 residents in the 10^{th} percentile census tract to 35.7 institutions originating mortgages per 1,000 residents in the 90^{th} percentile census tract.
  - Ranked by the number of mortgages originated per 1,000 residents, census tracts ranged from 21.0 originations per 1,000 residents in the 10^{th} percentile census tract to 169.7 originations per 1,000 residents in the 90^{th} percentile census tract.

[6] Throughout the report neighborhoods are defined as 2010 census tracts. Census tracts are relatively stable county sub-divisions constructed to have between 1,200 and 8,000 residents with an ideal population of 4,000. Census tracts are constructed to encompass populations that have homogenous socioeconomic characteristics. They are also large enough that data on census tract characteristics are publicly available for a wide range of topics. In more rural areas, census tracts can be spatially large and may not describe what most readers consider to be a “neighborhood.” Nevertheless, census tract and neighborhood are used interchangeably in the remainder of the report.

- Observable demographic characteristics were strongly associated with the number of institutions that originated mortgages in a neighborhood. The difference between the 10th percentile census tract and the 90th percentile census tract with respect to neighborhood characteristics was:
  - By income: 15.1 additional originating institutions per 1,000 residents.
  - By poverty: 14.3 fewer originating institutions per 1,000 residents.
  - By internet access: 15.3 additional originating institutions per 1,000 residents.
- The number of originators in a neighborhood per 1,000 residents was correlated with loan and applicant outcomes in a neighborhood. The difference between the 90th percentile census tract and the 10th percentile census tract with respect to originators per capita was:
  - A 7.3 percentage point lower likelihood that applicants were rejected for a loan.
  - A 3.5 percentage point lower likelihood that applicants were rejected for a loan conditional on Automated Underwriting System (AUS) approval.
  - Borrower-paid origination charges that were lower by 0.8 percent of the total loan amount.
- Credit unions originated a similar share of mortgages across all percentiles of neighborhood originators per capita; nondepository institutions' share of originations in neighborhoods increased with originators per capita whereas the share of commercial bank or thrift originations fell with originators per capita.
- Even within groups of borrowers defined to pose a similar credit risk, the expected difference between transactions that occurred in a neighborhood with originators per capita equal to the 90th percentile value relative to the 10th percentile value was:
  - A 0.5 percentage point lower likelihood of rejection conditional on AUS approval overall, a 1.1 percentage point lower likelihood of rejection for Black non-Hispanic and Hispanic applicants, and a 0.7 lower likelihood of rejection for AAPI non-Hispanic applicants.
  - Origination charges that were $74 lower overall, $170 lower for Black non-Hispanic borrowers, $306 lower for Hispanic borrowers, and $316 lower for AAPI non-Hispanic borrowers.

2. Data

Home Mortgage Disclosure Act Data

The Home Mortgage Disclosure Act (HMDA) is a data collection, reporting, and disclosure statute enacted in 1975. HMDA data are used to assist in determining whether financial institutions are serving the housing credit needs of their local communities; to facilitate public entities' distribution of funds to local communities to attract private investment; and to help identify possible discriminatory lending patterns and enforce antidiscrimination statutes. Financial institutions covered by HMDA—both depository and nondepository—must report information about each covered mortgage application acted upon and covered mortgage purchased.[7] HMDA data are the most complete publicly available data on mortgage market activity, with estimates suggesting they included between 90 and 98 percent of annual, closed-end mortgage origination activity during the past decade.

HMDA data include the disposition of each application for mortgage credit (e.g., whether an application was accepted, rejected, withdrawn); the type, purpose, and characteristics of each home mortgage application or purchased loan; the census tract designations of the covered properties; loan pricing information; demographic and other information about loan applicants and co-applicants; and information about loan sales.[8] The analysis uses several data points that were first reported in 2018 including information on borrower paid origination charges, discount points, lender credits, loan interest rates, debt-to-income ratios, combined loan-to-value ratios, and Automated Underwriting System (AUS) results.[9] I identify lenders based on lender IDs that also map institutions to one of three institution types: bank and thrifts, credit unions, and nondepository institutions.[10] To characterize the mortgage market activity for

[7] Not all mortgage applications and purchase are reportable. See Home Mortgage Disclosure (Regulation C), 80 Fed. Reg. 66127 (2015) for more details, available at https://www.federalregister.gov/documents/2015/10/28/2015-26607/home-mortgage-disclosure-regulation-c.

[8] For a brief history of HMDA, see Federal Financial Institutions Examination Council, "History of HMDA," available at www.ffiec.gov/hmda/history2.htm (last modified September 6, 2018).

[9] For a complete list of HMDA data points collected in 2020, see "2020 Mortgage Market Activity and Trends," HMDA Data Point, August, 2021, https://files.consumerfinance.gov/f/documents/cfpb_2020-mortgage-market-activity-trends_report_2021-08.pdf.

[10] Institutions are classified as banks and thrifts if they are identified as being a national bank; a state member bank; a state non-member bank; a state chartered thrift; a federally chartered thrift; a federal branch or agency of a foreign banking organization; a mortgage banking subsidiary of a national bank, a state member bank, a state non-member bank, a bank holding company, a savings and loan holding company, a state chartered thrift, or a federally chartered thrift. Institutions are classified as credit unions if they are identified as being a credit union or a mortgage banking subsidiary of a credit union. Institutions are classified as nondepository institutions if they are identified as being an independent mortgage bank. The institution-type classification is done by the CFPB during the production of the HMDA panel dataset.

potential borrowers, I rely heavily on the census tract where properties are located. In each census tract, I use the count of originations, applications, and the count of financial institutions (as measured by unique lender IDs) originating mortgages[11] between 2018–2020 to describe mortgage activity.

Loan and application outcomes are based on application- and origination-level characteristics from the HMDA data. The outcomes I consider are whether an application was rejected, whether an application was rejected despite approval from an automated underwriting system (AUS),[12] the interest rate, the borrower-paid origination charges, the borrower-paid total loan costs,[13] and total points and fees.[14] For the outcomes used in the regression analysis, I construct origination-level outcomes using only originated, conforming[15] mortgages and application-level outcomes using only applications for conforming mortgages. I do not restrict the sample to conforming transactions for the data used to generate the binned scatter plots. Total points and fees, total loan costs, and origination charges are only reported for subsets of covered, originated loans.[16] The sample of originations with non-missing total points and fees is particularly thin: about 38 percent of all census tracts had no originations with information on total points and

[11] Throughout the report, originations refer to dwelling-secured mortgages for home purchase, home improvement, refinancing, or cash out refinancing.

[12] An automated underwriting system is defined in Regulation C as an electronic tool that provides guidance regarding the credit risk of the applicant and information about whether the loan is eligible to be originated, purchased, insured, or guaranteed by that organization. Up to five AUS results are reported in the data. I use only the first-reported AUS result in this report. Of the applications I analyze, 9.3 percent have more than one AUS result reported though just 1.2 percent of these involve AUS approval in the first-reported result and at least one AUS non-approval.

[13] Total loan costs include origination charges, charges for services that borrowers cannot shop for (appraisal fees, credit report fees), and charges for services borrowers can shop for (settlement agent, title insurance fees). For more details, see "An Updated Review of the New and Revised Data Points in HMDA" CFPB Data Point, August 2020, available at https://files.consumerfinance.gov/f/documents/cfpb_data-points_updated-review-hmda_report.pdf.

[14] Total points and fees allow borrowers to trade off monthly mortgage expenses and closing costs. In particular, points lead to paying more up front but reducing the interest rates/monthly payments, and lender credits result in paying less up front but accepting a higher interest rate/monthly payment. Total points and fees, total loan costs, and origination charges are all expressed as a percentage of the loan amount in binned scatter plots and in 2020 dollars in the regression analyses.

[15] Conforming loans are loan amounts below conforming loan limits set by the Federal Housing Finance Agency (FHFA) that determine whether the loan can be acquired by Fannie Mae and Freddie Mac. The conforming loan limits for 2018–2020 are available at https://www.fhfa.gov/DataTools/Downloads/Pages/Conforming-Loan-Limit.aspx. All conforming loans are conventional mortgages, implying they are not insured or guaranteed by the government.

[16] Origination charges, total points and fees, and total loan costs are only reported for subsets of originated loans. Total loan costs and origination charges are reported for originated loans covered by TILA-RESPA integrated disclosure (TRID) requirements; total points and fees are reported only for originated loans not covered by TRID but subject to ability-to-pay requirements in Reg Z. TRID covers most dwelling-secured consumer credit excluding HELOCs, reverse mortgages, or manufactured housing loans secured by the home but not the land ("chattel" loans). Institutions eligible for partial exemption under the Economic Growth, Regulatory Relief, and Consumer Protection Act (EGRRCPA) do not have to report these three fields.

fees, and, on average there were just 1.3 such originations with reported total points and fees per 1,000 residents.

For the transaction-level data restricted to conforming mortgages, I match loans to Government Sponsored Enterprise (GSE) securitization pricing information based on the borrower and—if relevant—co-borrower credit scores[17] as well as the loan-to-value, loan term, loan amount and debt-to-income ratio. The GSE pricing match enables me to identify transactions that pose a similar level of credit risk, which I use in the regression analysis described below. The regression analysis also converts loan amount values to a set of indicator variables for each percentile of the loan amount distribution and includes them as controls; in addition, the specifications include a set of 36 indicators for each month between January 2018–December 2020 based on the month when the financial institution originated, purchased, or made a decision about a loan application.[18] The loan amount percentile indicators ensure that comparisons are being made between similarly sized loans, and the month indicators help to adjust for macroeconomic shifts like changes in interest rates or housing prices. Finally, I use information on borrower race and ethnicity to group borrowers according to whether they were Black non-Hispanic, white non-Hispanic, Hispanic, or Asian American Pacific Islander (AAPI) non-Hispanic.

Census Data

To complement HMDA data on borrower and loan characteristics, I merge in publicly available census tract-level data from the 2010 Decennial Census and the 2014–2018 American Community Survey (ACS) 5-year estimates. The characteristics from the 2010 Decennial Census include the total 2010 population and the shares of the 2010 population that were Black non-Hispanic, Hispanic, white non-Hispanic, and AAPI non-Hispanic.[19] The characteristics from the ACS include median family income, the share of adults over the age of 25 with at least a high school degree, the unemployment rate, the poverty rate, and the share of households with internet access. Household internet access is defined based on the reported presence of internet subscriptions for any type of connection, including dial-up, broadband, cellular data plans, and satellite internet. I convert median family income from the ACS into a measure of census tract relative income by scaling median census tract family income by median family income in the Core-Based Statistical Area (CBSA) or county containing the census tract. Low-income census tracts are defined as those where the resulting ratio is below 0.5; moderate income census tracts

[17] For applications with both an applicant and a co-applicant credit score available in the data, I use the minimum observed credit score.

[18] As is necessary, one action month indicator and one loan amount percentile indicator are dropped from the regression specifications.

[19] 2020 Decennial Census data were not available at the census tract-level when this report was being written.

are those with a ratio between 0.5 and 0.8; middle income census tracts are those with a ratio between 0.8 and 1.2; and upper income census tracts are those with a ratio at or above 1.2.

I use population and race and ethnicity data from the 2010 Census rather than more recent ACS data to help mitigate issues related to the sampling uncertainty for small area (census tract) estimates.[20] The trade-off for doing so is relying on population estimates, and race/ethnicity-specific population estimates, that are based on data from up to 10 years prior. Using older data will result in population estimates that are too low for rapidly growing census tracts and too high for shrinking census tracts. Similarly, other demographic characteristics will be less accurate for rapidly changing areas. I opt to use 2010 Census estimates whenever possible and rely on ACS estimates only when the 2010 Census values are unavailable.

The census tract population numbers play an important role in the analysis below, providing a scaling factor for the number of originating institutions in my preferred measure of originator activity, the number of originating institutions per 1,000 residents between 2018–2020. The measure is imperfect. In census tracts where housing units are more likely to be renter-occupied, there will mechanically be fewer originating financial institutions. More broadly, census tracts with less demand for mortgage financing are likely to have fewer financial institutions originating mortgages; this does not necessarily imply that potential borrowers in these neighborhoods lacked access to mortgage financing. An advantage of the measure is that it recognizes that consumers can search for mortgages in different neighborhoods, that financial institutions may theoretically make mortgages available in any neighborhood, and that no level of geographic partition will perfectly align with the choices considered by consumers.

[20] Each address has approximately a 1-in-40 chance of being sampled in the ACS in a year. See https://www.census.gov/content/dam/Census/library/publications/2020/acs/acs_general_handbook_2020.pdf for more details.

3. Results

3.1 Sample and Summary Statistics

The main analysis uses combined HMDA and Census data. I first generate two census-tract-level datasets, one with means and medians of the origination-level HMDA data (e.g., median annual interest rate for originated mortgages) and one with application-level HMDA data (e.g., the share of applications that were rejected), both merged to census tract-level population and demographic characteristics. I also generate two transaction-level datasets, one at the application-level and one at the origination-level, by merging application- and origination-level HMDA data onto census-tract-level demographic information.

For the analysis datasets, I impose as few restrictions as possible. For the census-tract-level datasets used to produce the binned scatter plots, I require only that transactions be applications or originations for home purchase, home improvement, refinance, or cash-out refinance. Given the years of HMDA data used, this implies that a variety of transactions are included in the analysis dataset (e.g., open-ended lines of credit, multifamily homes, reverse mortgages, manufactured housing, and closed-end mortgages with non-standard loan terms). The regression analysis, however, relies on being able to identify transactions that pose a similar level of credit risk. I am only able to do this for transactions with certain characteristics including some related to loan size, borrower debt-to-income (DTI), and consumer credit scores. Therefore, for the regression analysis I additionally restrict the sample to conforming mortgages. I also confirm that the results do not change when further limiting the sample to a more relevant subset of originations, when adding flexible controls for loan term, when controlling for the number of originations and the total population in the census tract, when adding flexible controls for loan-to-value (LTV),[21] and when adding controls for loan purpose, property type, lien status, and occupancy.

Table 1 presents the 10th, 50th (median), and 90th percentiles for census-tract-level characteristics and mortgage outcomes. To facilitate exposition, I convert all census tract-level shares to percentages. As expected, ranking census tracts by population, the median population among the 74,099 included census tracts was approximately 4,000 (3,995); the 10th percentile and 90th percentile census tract contained 2,020 and 6,665 residents in 2010, respectively. Census tract relative income—calculated as the ratio of median family income in the census tract

[21] More precisely, I include a linear spline in LTV with knot points at key thresholds used in loan pricing by the Government Sponsored Enterprises (GSEs).

to median family income in the surrounding core-based statistical area (CBSA) or county multiplied by 100—ranged from 55.9 in the 10th percentile census tract (ranked by relative income) to 155.8 in the 90th percentile census tract. I also use relative income to classify census tracts as low-, moderate-, middle-, or upper-income based on whether the census tract relative income was below 50 (low-income), between 50 and 80 (moderate income), between 80 and 120 (middle income), or above 120 (upper income).[22] 6.8 percent of sample tracts were designated as low-income, 22.2 percent as moderate income, 42.9 percent as middle income, and 28.0 as upper income.

Based on ACS data from 2014–2018, in the 10th percentile census tract (ranked by adult education) 72.5 percent of adults aged 25 and older had at least a high school degree, while in the 90th percentile census tract this rose to 97.2 percent. When ranking tracts by poverty, the 10th percentile census tract had 3.6 percent poverty while the 90th percentile census tract had 32.5 percent poverty; there was substantially less variation when ranking tracts by unemployment, which varied from 2.3 percent in the 10th percentile tract to 12.9 percent in the 90th percentile tract. Ranked by household internet access, 66.9 percent of households in the 10th percentile census tract had some access to the internet as compared to 95.6 percent of households in the 90th percentile census tract.

All six mortgage outcomes exhibited important differences across census tracts. Ranking tracts by the average interest rate for originated loans, interest rates ranged from 3.70 in the 10th percentile census tract to 4.49 in the 90th percentile tract. Borrower-paid origination charges and total loan costs (both expressed in 2020 dollars), ranged from $1,138 and $3,035 in the 10th percentile up to $2,607 and $6,122 in the 90th percentile when tracts were ranked by origination charges and total loan costs, respectively. 14.9 percent of mortgage applications were rejected in the median census tract (ranked by rejection rate), while 7.0 percent were rejected after AUS approval (ranked by rejection rate conditional on AUS approval). In the 90th percentile census tracts, 26.1 percent of all applications were rejected, and 13.5 percent were rejected despite receiving AUS approval.

[22] This classification is based on Community Reinvestment Act (CRA) definitions of relative income. See https://www.federalreserve.gov/consumerscommunities/cra_resources.htm.

Table 1: Census Tract Summary Statistics

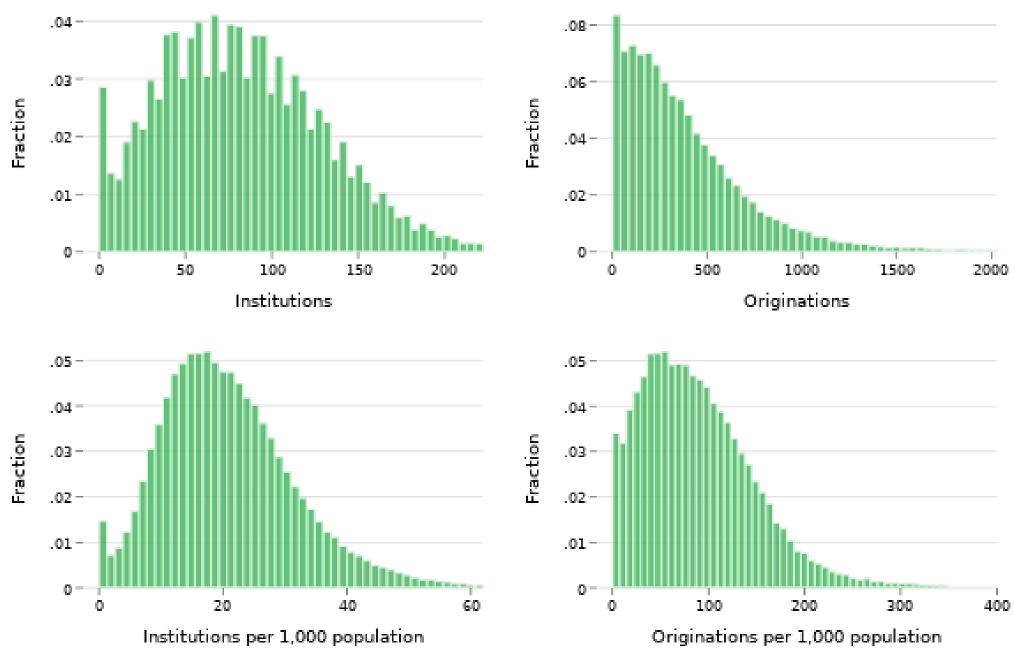
	10th Percentile	50th Percentile	90th Percentile
Demographic Characteristics			
Total population	2,020	3,995	6,665
Census tract relative income	55.9	99.0	155.8
Percent Adults (>25) with HS degree	72.5	89.8	97.2
Percent Poverty	3.6	12.2	32.5
Percent Unemployment	2.3	5.6	12.9
Percent Households with internet	66.9	84.6	95.6
Percent Black non-Hispanic	0.03	3.7	42.9
Hispanic	1.2	6.2	45.7
Percent AAPI non-Hispanic	0.2	1.5	11.3
Percent White non-Hispanic	11.7	73.9	95.3
Mortgage Characteristics			
Originations	53	304	843
Originators	24	79	148
Originations per 1,000 residents	21	81	170
Originators per 1,000 residents	9	20	36
Average interest rate	3.70	3.98	4.49
Average Borrower-paid origination charges	$1,138	$1,715	$2,607
Average total borrower-paid loan costs	$3,035	$4,456	$6,122
Average total points and fees	$140	$2,099	$5,898
Percent of Applications Rejected	9.3	14.9	26.1
Percent of Applications Rejected with AUS approval	3.7	7.0	13.5
Number of census tracts		74,099	

Note: Summary statistics from 2010 Census, 2014–2018 5-year American Community Survey (ACS) tables, and 2018–2020 HMDA data. Poverty, unemployment, households with internet, Black non-Hispanic, Hispanic, AAPI non-Hispanic, White non-Hispanic presented as the percent of the total tract population in the 2010 Census in that group. Borrower-paid origination charges, total loan costs, and total points and fees shown in 2020 dollars. Rejected and rejected with AUS approval displayed as the census tract-level percent of total applications.

To complement the summary statistics shown in Table 1, Figure 1 shows histograms of the census tract-level number of institutions originating mortgages and the number of mortgage originations between 2018–2020 with and without scaling by population (in 1,000's). All four histograms in Figure 1 have a long right tail, suggesting there were a small number of census tracts with an especially large number of mortgage originations and originating institutions. Similarly, all four have excess mass at zero, indicating that there were sharply more census tracts with no origination activity between 2018–2020 than there were with very small amounts.

Though the lenders available to borrowers seeking a mortgage are not observable, the census tract-level counts of originating institutions suggest borrowers in many neighborhoods *used* a limited number of institutions for their mortgage financing. While many financial institutions offered on-line applications for mortgage credit that were accessible from anywhere, differences in information, technology access, marketing activity, and beliefs could still have limited perceived options for mortgage credit. While these largely describe demand-side variation, many could be partially attributable to supply-side historical market features like restrictions on lending, price disparities, or broader differences in activity across neighborhoods.

Figure 1: Mortgage Lending Activity, 2018–2020



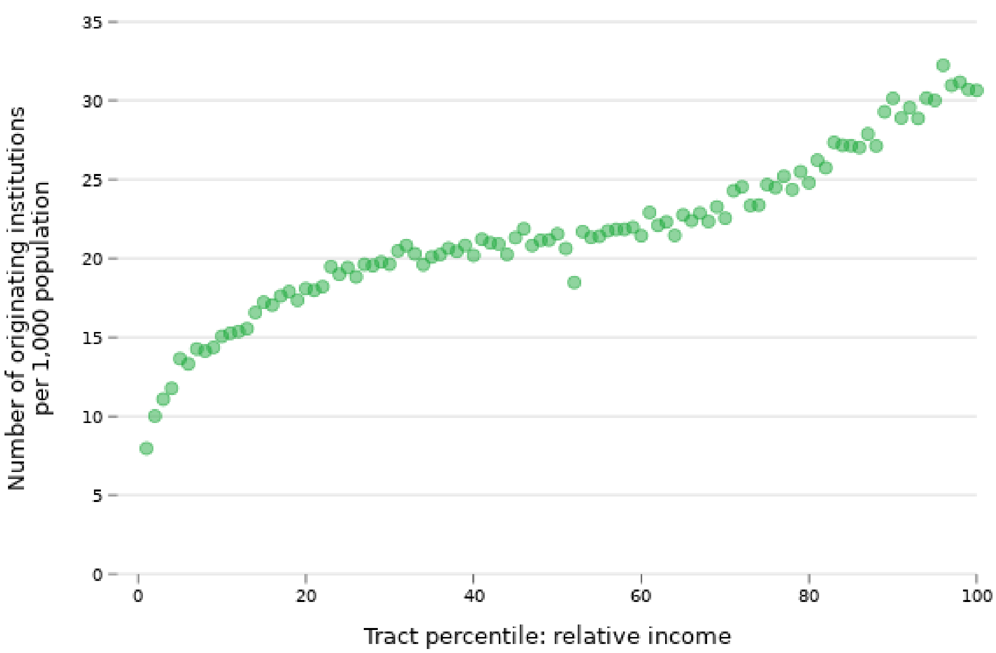
Note: Histogram based on HMDA census tract-level counts of mortgage originating institutions, mortgage originations, mortgage originating institutions per 1,000 residents, and mortgage originations per 1,000 residents between 2018–2020.

3.2 Originators Per Capita and Neighborhood Characteristics

I rely primarily on binned scatter plots to explore the relationships between demographic characteristics and originators per capita. Binned scatter plots collapse the census tract-level data by percentiles of different demographic characteristics from the 2010 Decennial Census or the 2014–2018 5-year ACS. The figures display the mean number of originators per 1,000 residents in each percentile of the characteristic to illustrate the relationship. For census tract relative income—a characteristic likely to be strongly associated with mortgage demand—I also plot the density of the number of originators per 1,000 residents separately for low-, middle-, moderate-, and high-income census tracts.

Figure 2 presents a binned scatter plot of the number of originators per 1,000 residents between 2018–2020 by census tract relative income.[23] The relationship was positive with originators per capita generally increasing as tract relative income increases. The steepest increases between percentiles were concentrated in the lowest relative income percentiles, suggesting consumers in these tracts were likely to have their mortgages originated by especially few lenders. In the 90th percentile tract there were 30.1 originating institutions per 1,000 residents; in the 10th percentile tract, this number was 15.1 institutions per 1,000 residents.

Figure 2: Originators Per Capita and Census Tract Relative Income



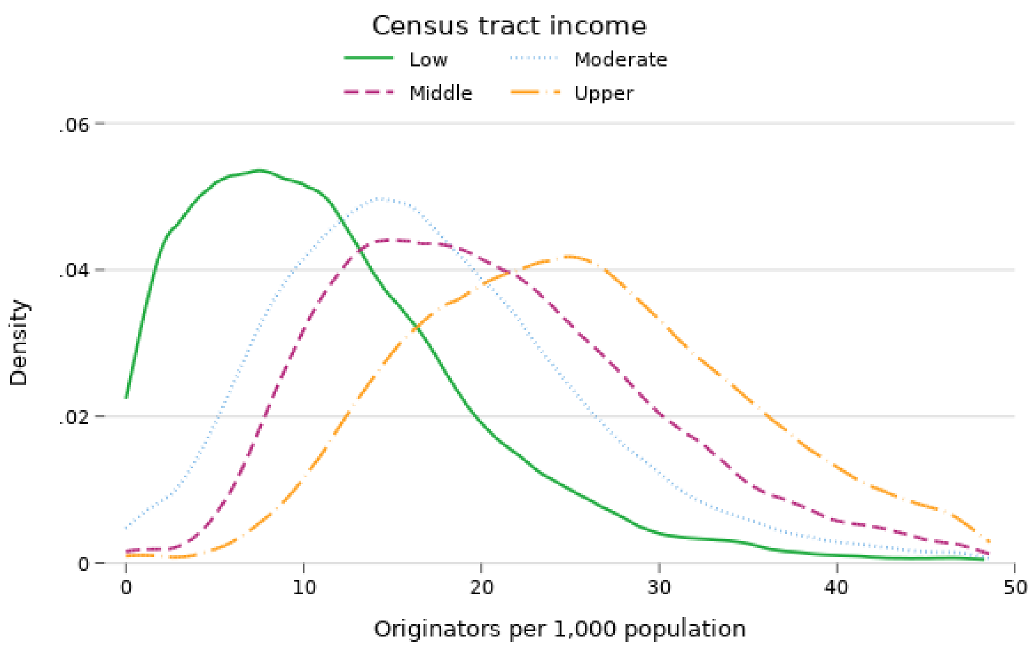
Note: Binned scatter plot displays census tract originators per capita (originating lenders per 1,000 residents between 2018–2020) by census tract percentile of relative income. Census tract relative income is the ratio of median family income in the census tract to median family income in the surrounding core-based statistical area (CBSA) or county.

The density plot of the number of originators per 1,000 residents partitioned by census tract relative income—shown in Figure 3—similarly illustrates a clear relationship between census tract relative income and lending institution activity. As relative income increased (from low to

[23] Because the number of originators per 1,000 residents is an equilibrium outcome reflective of both mortgage supply and demand, the relationships illustrated in the binned scatter plots are associations that should not be interpreted causally.

moderate to middle to high), the distribution of originators per 1,000 residents shifted to the right.[24]

Figure 3: Originators Per Capita by Census Tract Relative Income Group



Note: Kernel density plots of the number of originating institutions per 1,000 in population by census tract relative income group (low, moderate, middle, and upper).

The relationship between census tract relative income and the number of originating lenders is expected. Mortgage demand is tightly related to household income,[25] and originating institutions are likely to focus resources on neighborhoods with more potential borrowers. Despite this, Figure 3 also suggests that there remained considerable variation in originators per capita both within and across census tract-based income groups. That is, there are tracts within

[24] Low-income census tracts are those with census tract relative income—the ratio of median family income in the census tract to median family income in the surrounding core-based statistical area (CBSA) or county, multiplied by 100—below 50; moderate income census tracts are those with relative income between 50 and 80; middle income census tract are those with census tract relative income between 80 and 120; and upper income census tracts are those with relative income above 120.

[25] Data from the Consumer Expenditure Surveys (CES) suggest that in 2021 homeowners with a mortgage have an average household income 41 percent greater than the overall average household income and 120 percent greater than renter households. See https://www.bls.gov/cex/tables/calendar-year/mean-item-share-average-standard-error/cu-housing-tenure-2021.pdf.

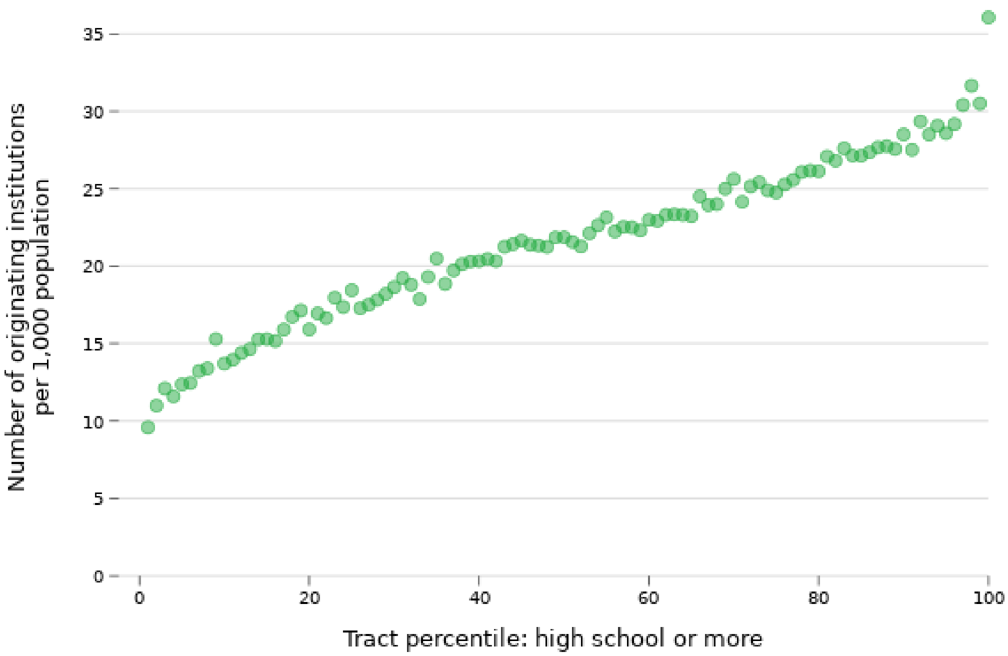
each of the four relative income categories with, for example, fewer than ten originators per 1,000 residents and more than twenty-five originators per 1,000 residents.

Binned scatter plots for demographic characteristics associated with household income[26] show similarly close relationships with originators per capita. Figure 4 displays a binned scatter plot of originators per capita based on percentiles of the share of adults older than 25 with at least a high school degree; Figure 5 does the same for percentiles of the census tract poverty rate; Figure 6 and Figure 7 do so for percentiles of the census tract unemployment rate and share of households with internet access.

The number of lenders originating mortgages was largely increasing in the share of adults with at least a high school education and decreasing in the poverty rate. Contrasting neighborhoods at the 10th percentile to neighborhoods at the 90th percentile with respect to the share of adults with a high school degree, neighborhoods at the 90th percentile had 14.8 additional originators per 1,000 residents; the same comparison with respect to the poverty rate shows there were 14.3 fewer originators per 1,000 residents in neighborhoods at the 90th percentile. The binned scatter plot for unemployment was comparatively flat: the 10th–90th percentile difference in the number of originators per 1,000 residents was just 6.0, less than half of the analogous differences for adult education and poverty.

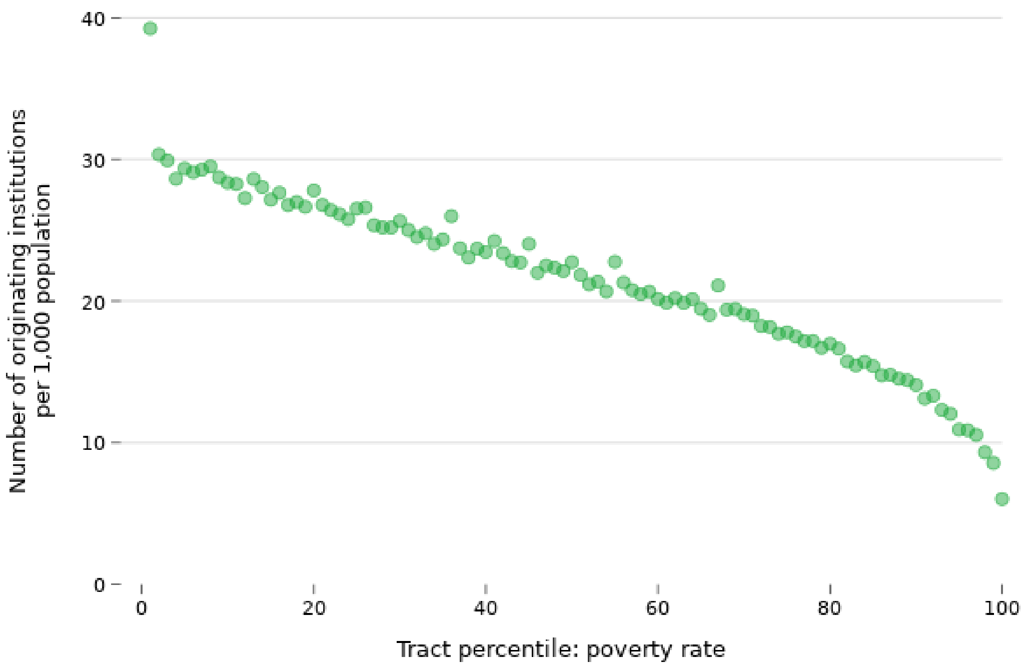
[26] The census tract-level correlation with relative income was -0.7 for poverty, 0.6 for the share of adults with at least a high school degree, -0.3 for unemployment, and 0.6 for the share of households with any internet access.

Figure 4: Originators Per Capita and Adult Educational Attainment



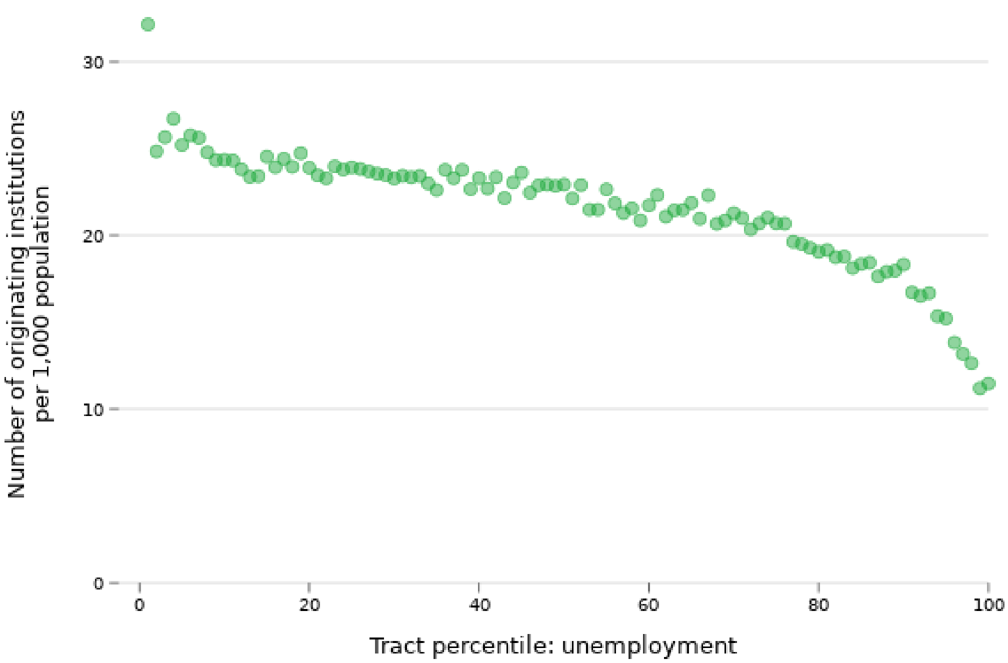
Note: Binned scatter plot displays census tract originators per capita (originating lenders per 1,000 residents between 2018–2020) by census tract percentile of the share of adults over age 25 with at least a high school degree.

Figure 5: Originators Per Capita and Census Tract Poverty



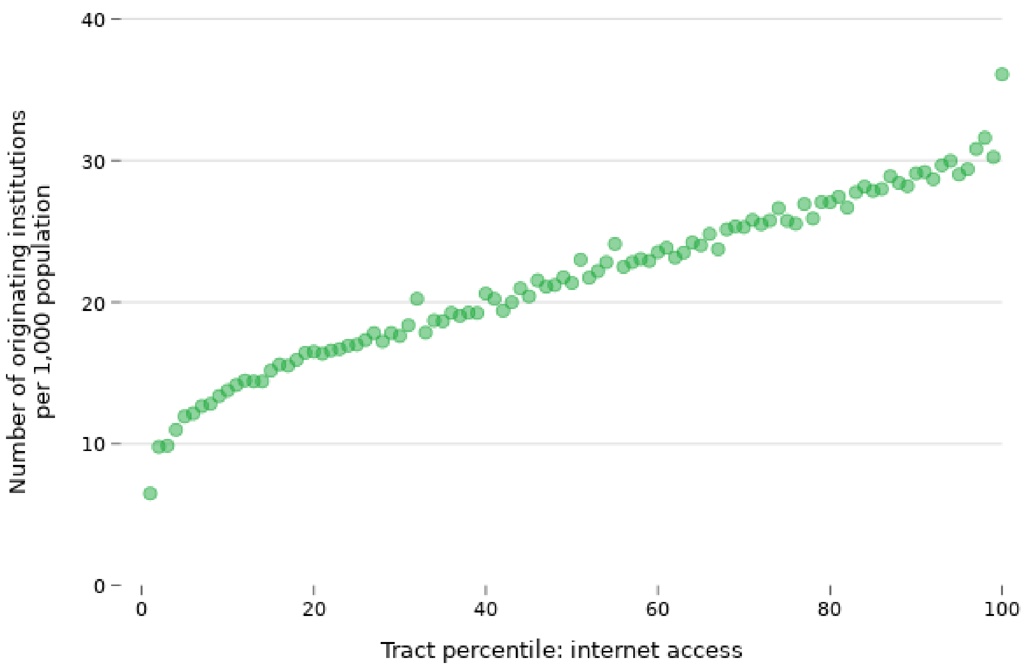
Note: Binned scatter plot displays census tract originators per capita (originating lenders per 1,000 residents between 2018–2020) by census tract percentile of the share of residents below the poverty line.

Figure 6: Originators Per Capita and Unemployment



Note: Binned scatter plot displays census tract originators per capita (originating lenders per 1,000 residents between 2018–2020) by census tract percentile of the unemployment rate.

Figure 7: Originators Per Capita and Household Internet Access



Note: Binned scatter plot displays census tract originators per capita (originating lenders per 1,000 residents between 2018–2020) by census tract percentile of the share of households with internet access.

Given the expansion of nondepository mortgage lenders[27] and their reliance on digital marketing, applications, and underwriting, the binned scatter plot of the share of households with internet access is particularly germane for understanding the use of lenders. The census tract-level association between household internet access and income was clear, but not perfect: a one percentile increase in relative income was associated with a 0.6 percentile increase in the share of households with internet access. The binned scatter for neighborhood internet access is, however, even steeper than the binned scatter for neighborhood relative income. Relative to neighborhoods at the 10th percentile of the internet access distribution, neighborhoods at the 90th percentile had 15.3 more originating institutions per 1,000 residents. The same difference

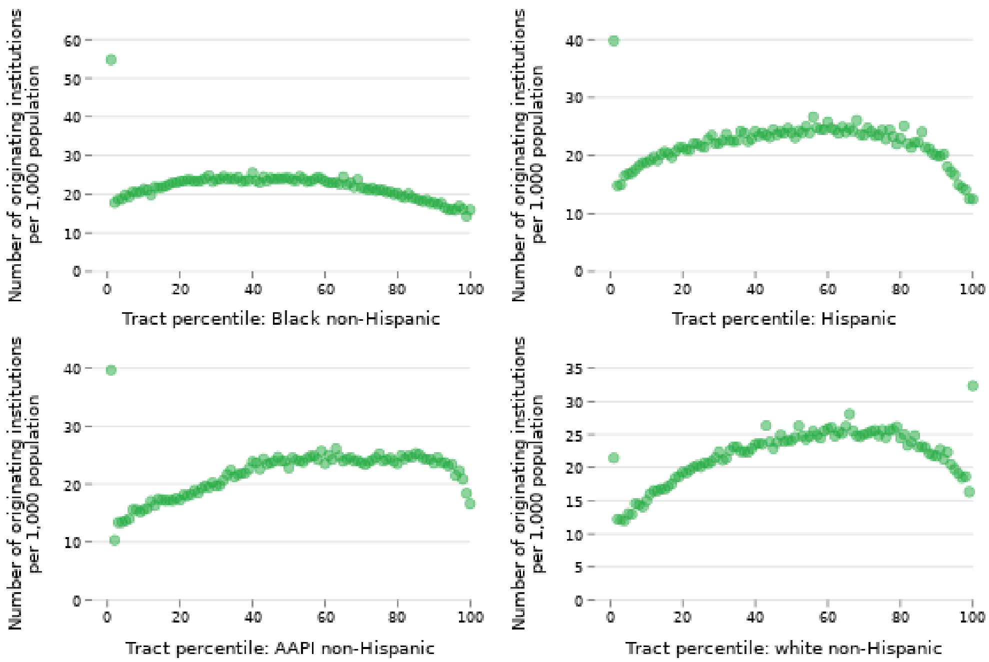
[27] See "2020 Mortgage Market Activity and Trends" CFPB Data Point, August 2021, available at https://files.consumerfinance.gov/f/documents/cfpb_2020-mortgage-market-activity-trends_report_2021-08.pdf and "2017 Mortgage Market Activity and Trends" CFPB Data Point, May 2018, available at https://files.consumerfinance.gov/f/documents/bcfp_hmda_2017-mortgage-market-activity-trends_report.pdf. Between 2016 and 2021 the share of first-lien, owner-occupied, site-built home-purchase loans originated by nondepository independent mortgage companies increased from 53.3 percent to 63.9 percent.

based on percentiles of the relative income distribution showed there were 15.1 more originating institutions per 1,000 residents in neighborhoods at the 90th percentile. Figure 7 therefore underscores the potential role played by internet access in connecting borrowers to mortgage suppliers.

Census tract-level measures of race/ethnicity were also associated with neighborhood originators per capita. Figure 8 shows binned scatter plots of originators per capita by census tract percentiles for each of the race/ethnicity groups. All four binned scatters display an "inverted-U" shaped relationship; originators per capita generally increased most steeply at census tracts in the lowest percentiles, began to flatten between the 20th and 40th percentiles, and began declining between the 60th and the 80th percentiles.

The percentile where originators per capita peaks differed somewhat across the figures: around the 20th percentile for Black non-Hispanic neighborhood shares, around the 40th percentile for AAPI non-Hispanic and Hispanic neighborhood shares, and near the 60th percentile for white non-Hispanic neighborhood shares. The difference in originators per capita between the most concentrated neighborhoods in terms of race and ethnicity and similar, but slightly less concentrated neighborhoods was also apparent. This was likely driven, in part, by the existence of neighborhoods that were extreme outliers with respect to some demographic characteristics—that is, census tracts with sharply more (or fewer) members of a particular demographic group than other neighborhoods. These neighborhoods also tended to differ with respect to other demographic characteristics (e.g., population, poverty, relative income) and had different levels of mortgage lending activity.

Figure 8: Originators Per Capita and the Share of Residents in Different Race/Ethnicity Groups



Note: Binned scatter plot displays census tract originators per capita (originating lenders per 1,000 residents between 2018–2020) by census tract percentile of the share of the population that is Black non-Hispanic, Hispanic, Asian Pacific Islander (AAPI) non-Hispanic, and white non-Hispanic.

The data show variation across neighborhoods in mortgage originators per capita and highlight characteristics that were predictive of this variation. It is not, however, clear that the observed differences in this measure translated into differences in welfare for borrowers. Some recent work even suggests that, all else equal, fewer lenders originating mortgages in a market may, in some instances, help borrowers.[28] The combined 2018–2020 HMDA data and the expanded data points contained therein, offer a means to explore this further.

3.3 Loan Outcomes and Originators Per Capita

I continue using binned scatter plots to investigate whether a greater number of institutions originating mortgages in a neighborhood was associated with differences in the likelihood of a

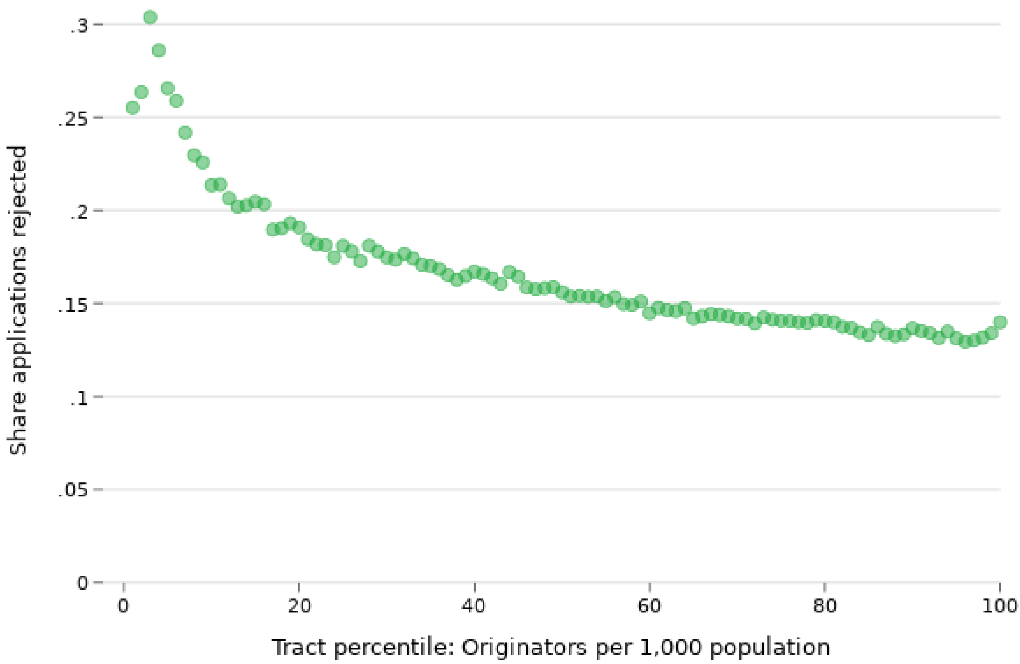
[28] See Constantine Yannelis and Anthony Lee Zhang, "Competition and Selection in Credit Markets" NBER Working Paper No. 29169, available at https://www.nber.org/papers/w29169.

mortgage application being rejected or in the terms that borrowers received. The binned scatter plots are constructed based on percentiles of the census tract number of originators between 2018–2020 per 1,000 residents. Borrower-paid origination charges, borrower-paid total loan costs, and total points and fees are all expressed as a percentage of the total loan amount and the figures for these outcomes as well as for interest rates show percentile-specific medians rather than means to reduce the influence of outliers in the data.

Loan pricing and origination decisions were the outcome of complicated processes that relied on more extensive data than were available in even the extended HMDA data points and more complex modelling than was feasible for this analysis. Caution is therefore warranted to avoid over-interpreting associations in all the subsequent analysis, particularly for application and loan characteristics related to pricing (interest rates and total points and fees) and origination decisions. Nevertheless, explorations of how these outcomes varied across neighborhoods with differing numbers of originating institutions can still be a useful starting point for more rigorous research. Origination fees and total loan costs, which include charges for services that consumers "can shop around for," may be less likely to be associated with the expected credit risk posed by a potential loan after conditioning on available data, though care is still necessary.

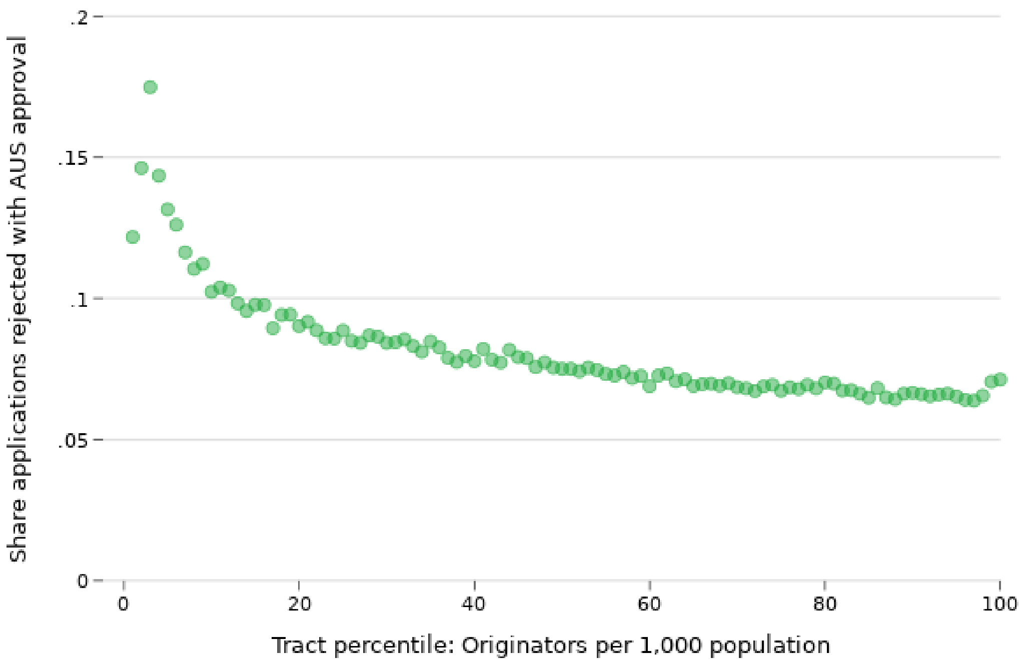
Figure 9 shows a binned scatter plot of the share of applications in a census tract that were rejected and Figure 10 does the same for the share of applications that were rejected after being approved by an AUS. Both figures suggest rejection/denial rates that were decreasing in the number of originators per 1,000 residents. For neighborhoods in the 10th percentile of originators per capita, 21 percent of covered applications were rejected; in the 90th percentile this dropped to 13.7 percent, a 56 percent reduction relative to the 10th percentile value. Application rejections after AUS approval were less common, but the pattern across neighborhood-level originators per capita was the same. In the 10th percentile census tract 10.2 percent of covered applications were rejected after AUS approval; in the 90th percentile census tract this fell to 6.7 percent.

Figure 9: Mortgage Application Rejections and Originators Per Capita



Note: Binned scatter plot displays the census tract-level share of mortgage applications rejected by originators per capita (originating lenders per 1,000 residents between 2018–2020) percentile.

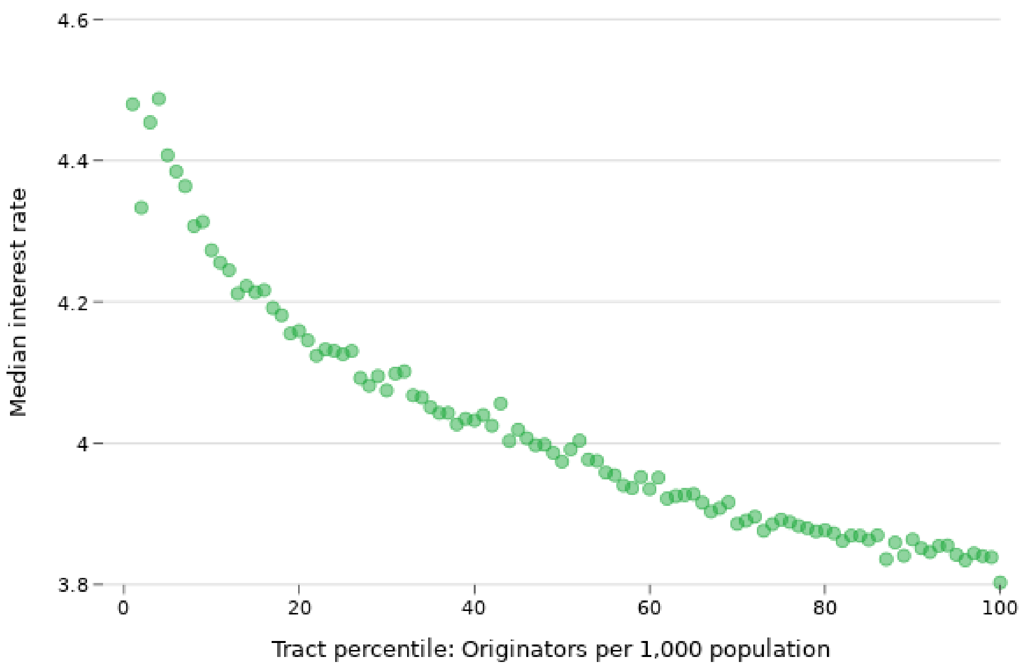
Figure 10: Mortgage Application Rejections with AUS Approval and Originators Per Capita



Note: Binned scatter plot displays the census tract-level share of mortgage applications rejected with automated underwriting system (AUS) approval by originators per capita (originating lenders per 1,000 residents between 2018–2020) percentile.

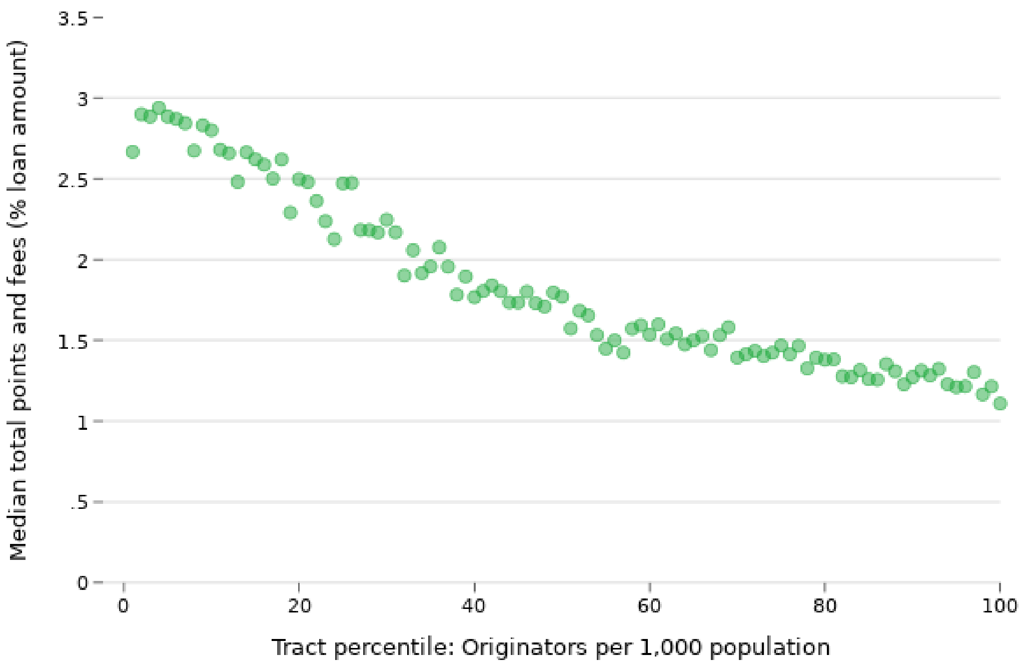
Figures 11 and 12 display binned scatter plots for the median interest rate and median total points and fees of originated loans (expressed as a percentage of the loan amount). Median interest rates were generally decreasing in the number of originators per 1,000 residents, with the steepest differences observed between the 1st and 10th percentiles. In the 10th percentile census tract, the median interest rate for covered, originated loans was 4.3 percent; in the 90th percentile neighborhood the median interest rate fell to 3.9 percent, a 41 basis point decrease. Total points and fees showed the same, decreasing pattern overall. Borrowers in neighborhoods at the 90th percentile of originators per capita paid 1.5 percent of the loan amount less, relative to borrowers in neighborhoods at the 10th percentile. This difference was equivalent to 54.6 percent of the median total points and fees amount in neighborhoods at the 10th percentile. This could simply reflect an increased willingness of borrowers in neighborhoods with fewer originating lenders to trade-off lower interest rates for increased up-front costs. Aside from being explained by differences in originators per capita, other potential drivers of the relationship include differences in the likelihood that borrowers faced binding liquidity constraints at the time of origination, time preferences that weighted future consumption more (or less), and differences in financial literacy. It is important to caveat that total points and fees were reported for only a small number of loans and that these loans were more likely to be HELOCs and manufactured housing loans than HMDA transactions overall.

Figure 11: Median Interest Rate and Originators Per Capita



Note: Binned scatter plot displays the census tract median interest rate for originated mortgages by originators per capita (originating lenders per 1,000 residents between 2018–2020) percentile.

Figure 12: Median Total Points and Fees and Originators Per Capita

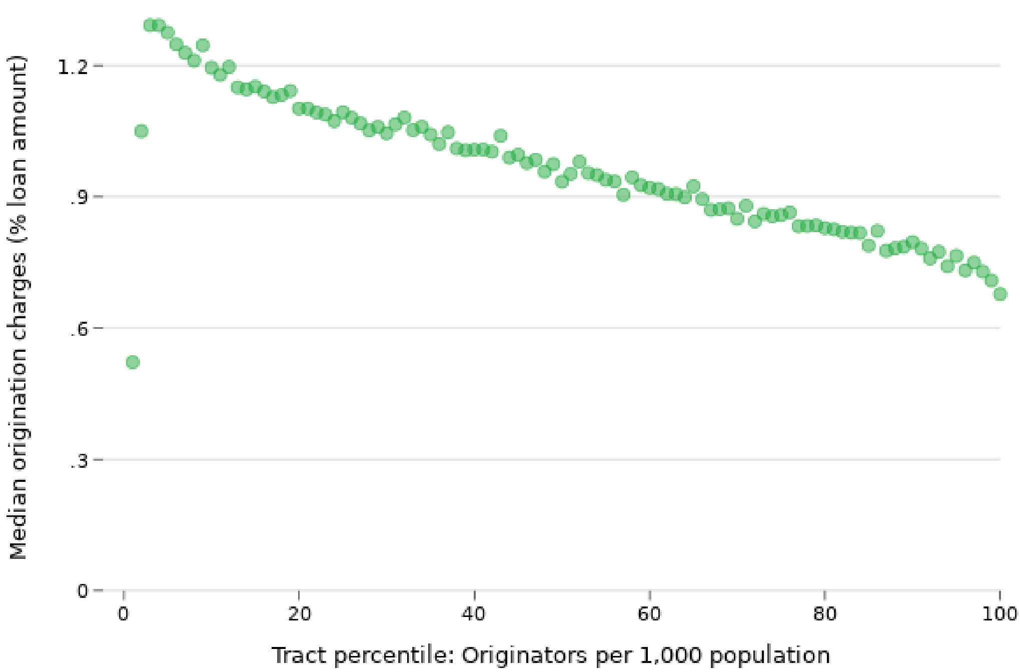


Note: Binned scatter plot displays the census tract median total points and fees (as a percent of the total loan amount) for originated mortgages by originators per capita (originating lenders per 1,000 residents between 2018–2020) percentile.

While total borrower-paid points and fees was reported for few loans, the relationships between points and fees and originators per capita and between interest rates and originators per capita were similar for the sub-sample with non-missing values. This suggests, if anything, that the relationship between the interest rates and originators per capita for these loans may be understated; if borrowers in neighborhoods with relatively low originators per capita opted for the (lower) borrower-paid points and fees observed in neighborhoods with relatively higher originators per capita, they would likely have needed to pay *higher* interest rates. This would have steepened the slope of the relationship between interest rate and originators per capita in the binned scatter plot.

Figures 13 and 14 show binned scatter plots of median origination costs and median total loan costs as a percentage of the total loan amount. Both costs tended to decrease with the number of originating lenders per 1,000 residents. In the 10th percentile of originators per capita median origination charges were 1.2 percent and total loan costs were 3.2 percent of the loan amount; in the 90th percentile these were 0.8 percent (for origination charges) and 2.0 percent (for total loan costs). Origination fees typically cover the cost of processing the loan application, underwriting, and funding the loan, and administration services incurred in the course of originating the loan. Some of the items included in origination costs scale directly with the size of the loan, while many common non-scaling origination charges are often charged at the discretion of the lender. It is therefore surprising to see such a clear pattern of declining origination costs with the originators per capita in a neighborhood. Borrowers can shop for alternative providers for many closing costs, and lenders have scope to reduce borrower-paid costs for other services. The variation displayed in Figure 13 and Figure 14 suggests that originators could have faced and passed on lower costs of originating mortgages in neighborhoods where there were more originating institutions per capita, reduced pass through rates for investors, or competed with one another through origination and total loan costs in these areas.[29] However, the data are not sufficiently rich to rule out other potential mechanisms that could explain the observed relationships.

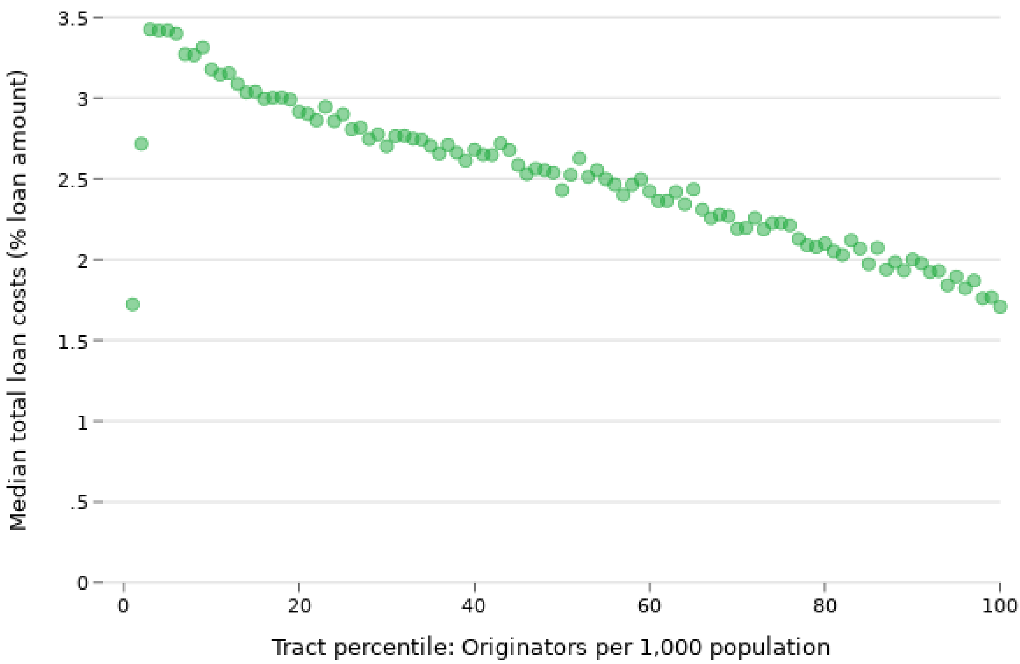
Figure 13: Median Origination Charges and Originators Per Capita



Note: Binned scatter plot displays the census tract median total borrower paid origination charges (as a percent of the total loan amount) for originated mortgages by originators per capita (originating lenders per 1,000 residents between 2018–2020) percentile.

[29] Loans in neighborhoods with fewer originating institutions per 1,000 residents were smaller, on average. Given financial institutions' need to cover their fixed costs of originating loans, the smaller loan sizes likely also contribute to the relationship between origination charges as a percent of the loan amount and originators per capita. As with all the binned scatter plots other loan- or borrower-characteristics that are correlated with originators per capita could also be affecting the observed relationships.

Figure 14: Median Total Loan Costs and Originators Per Capita



Note: Binned scatter plot displays the census tract median total borrower paid loan costs (as a percent of the total loan amount) for originated mortgages by originators per capita (originating lenders per 1,000 residents between 2018–2020) percentile.

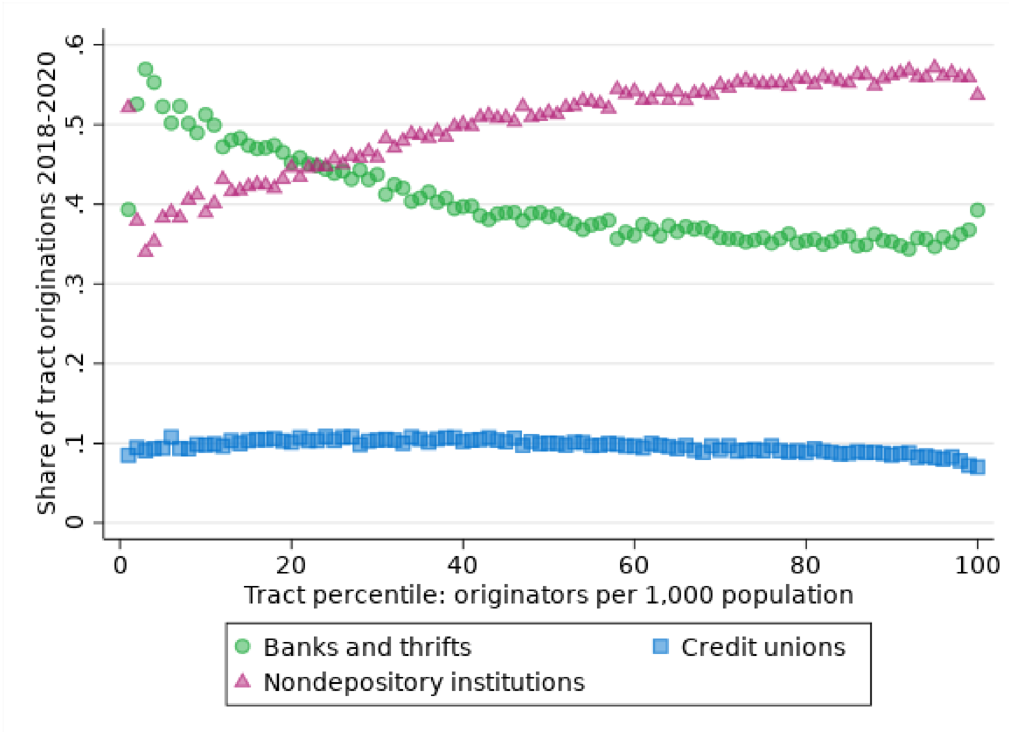
Three types of financial institution originate mortgages: banks and thrifts (banks), credit unions (CUs), or nondepository institutions (non-DIs). In recent years, non-DIs, including fintech firms specializing in mortgage originations, have rapidly increased their share of the mortgage market. Many non-DIs relied more heavily than banks and CUs on automated decision-making and therefore may have offered different financial and time costs to borrowers. Some non-DIs used customer-interfaces that were mostly or entirely virtual and less dependent than banks and CUs on physical proximity to borrowers, potentially increasing their capacity to reach remote neighborhoods. More automation could also have made it easier for non-DIs to find and fund borrowers even in neighborhoods where there were already many financial institutions originating mortgages. Some recent work suggests that CUs, as non-profit institutions with the explicit goal of helping CU members, have different incentives than banks and non-DIs.[30] These alternative incentives could have constrained some CUs from expanding beyond pre-specified geographic boundaries or away from areas with a high concentration of existing members.

[30] See Andrés Shahidinejad, "Are (Nonprofit) Banks Special? The Economic Effects of Banking with Credit Unions." Working Paper, November 2021, available at https://cpb-us-w2.wpmucdn.com/sites.northeastern.edu/dist/6/2602/files/2022/08/JMP.pdf.

To explore the relationship between originators per capita and institution type, Figure 15 combines three binned scatter plots by percentiles of the census tract number of originating institutions per 1,000 residents, one each for the share of originations made by banks, CUs, and non-DIs. CUs originated roughly 5–10 percent of mortgages across all percentiles. Non-DIs and banks originated a larger share of mortgages than CUs at all percentiles—each was responsible for between 40 and 60 percent of originations—but banks originated a higher share in neighborhoods with fewer originators per capita while non-DIs originated relatively more mortgages in neighborhoods with more originators per capita.[31] The pattern of origination shares by institution type is consistent with the idea that non-DIs may have lower fixed costs of entry, enabling them to originate mortgages even in neighborhoods that already have many institutions originating mortgages. Conversely, banks and CUs may be more dependent on having proximate physical branches, and the costs of opening and maintaining branches to serve new neighborhoods could slow or preclude entry into already-crowded areas. However, the associations do not provide definitive evidence to support these explanations; other unobserved factors could also have driven the observed pattern of origination shares by institution type.

[31] The figure plotting shares of originations by institution type against household internet access has a similar shape, though the bank and thrift and nondepository curves are somewhat flatter. The average bank and thrift share in the bottom ten percentiles of household internet access was 0.43; in the top ten percentiles the average share fell to 0.39. The nondepository share of originations was 0.48 in the bottom ten percentiles and 0.52 in the top ten percentiles of household internet access.

Figure 15: Institution Type and Originators Per Capita



Note: Binned scatter plot displays the census tract share of originations by institution type by originators per capita (originating lenders per 1,000 residents between 2018–2020) percentile.

3.4 Borrower Outcomes and Originators Per Capita

The associations between demographic, application, and loan characteristics, and the number of originators per 1,000 residents rely on neighborhood-level aggregates. While useful, neighborhood-level aggregates could mask important differences across transactions that occur within the same neighborhoods. And, ultimately, neighborhood-level loan and application characteristics are not the primary outcome of interest; rather, I am focused on understanding whether *consumers* in neighborhoods with more or fewer originating institutions per capita end up with different mortgage outcomes.

An ideal experiment for answering this question would be assigning a new group of financial institutions to receive mortgage applications, make origination decisions, and originate mortgages in randomly selected neighborhoods, and to avoid engaging in the same activity in non-selected neighborhoods. Comparing the loan rejection rates and loan terms across the two neighborhood groups would yield an estimate of the causal impact of potential access to the new lenders.[32] Because the selected and non-selected neighborhoods were chosen at random, with a large enough sample size researchers can be confident that the non-selected neighborhoods are a valid *counterfactual*—or approximation of what would have happened in the absence of the additional lenders—for the selected neighborhoods.

Generating random variation or finding "as good as random" variation that is not related to outcomes for reasons other than changes in originators per capita, is beyond the scope of this report. But the HMDA expanded datapoints enable more rigorous analyses of the relationship between originators per capita and loan- or application-level outcomes. To do so, I borrow a strategy from earlier work[33] and rely on a central feature of the residential lending market: Government Sponsored Enterprise (GSE) securitization. The GSEs (Fannie Mae, Freddie Mac, and the Federal Home Loan banks) help ensure liquidity and stability in the mortgage market by guaranteeing interest and principal payments to investors. Originators can securitize mortgages if they meet pre-specified term, amount, and credit risk criteria.[34] Securitizing these "conforming" loans requires a fixed, monthly payment (the g-fee) that is a proportion of total loan size and can vary across lenders but not borrowers or loans, and an up-front loan-level price adjustment (LLPA) that depends on loan and borrower characteristics. Because lenders always have the option to securitize conforming loans through the GSEs at prices determined by the LLPA, differences in rejection rates, interest rates, origination charges, and total loan costs within LLPA-based groups are more likely to reflect borrower and lender choices than credit risk.

The observable GSE pricing for each HMDA loan is used to further the analysis of originators per capita, by grouping each application and originated conforming loan into cells based on credit score and loan-to-value.[35] I then run ordinary least squares (OLS) regressions that condition on GSE cell, loan action date (the decision or loan origination month), and loan amount percentile fixed effects.[36] By adjusting for these characteristics, differences in outcomes associated with the number of originating institutions per 1,000 residents in the census tract are

---

32 Even this hypothetical context would not guarantee that the new financial institutions were included in any consumer's choice set.

33 See Robert Bartlett, Adair Morse, Richard Stanton, and Nancy Wallace, "Consumer-lending discrimination in the FinTech Era." Journal of Financial Economics, Volume 143, Issue 1, (January 2022), available at https://doi.org/10.1016/j.jfineco.2021.05.047.

34 In 2020, 59.2 percent of first-lien residential mortgages were GSE-securitized. See, "Housing Finance at a Glance: A monthly Chartbook", Urban Institute Research Report, February, 2021, available at https://www.urban.org/research/publication/housing-finance-glance-monthly-chartbook-february-2021.

35 These are the characteristics that determine up-front LLPA pricing through the GSEs.

36 Fixed effects refer to indicator variables that account for the differences in the regression outcome across sub-groups. For example, month fixed effects would allow for monthly variation in the typical number of originations.

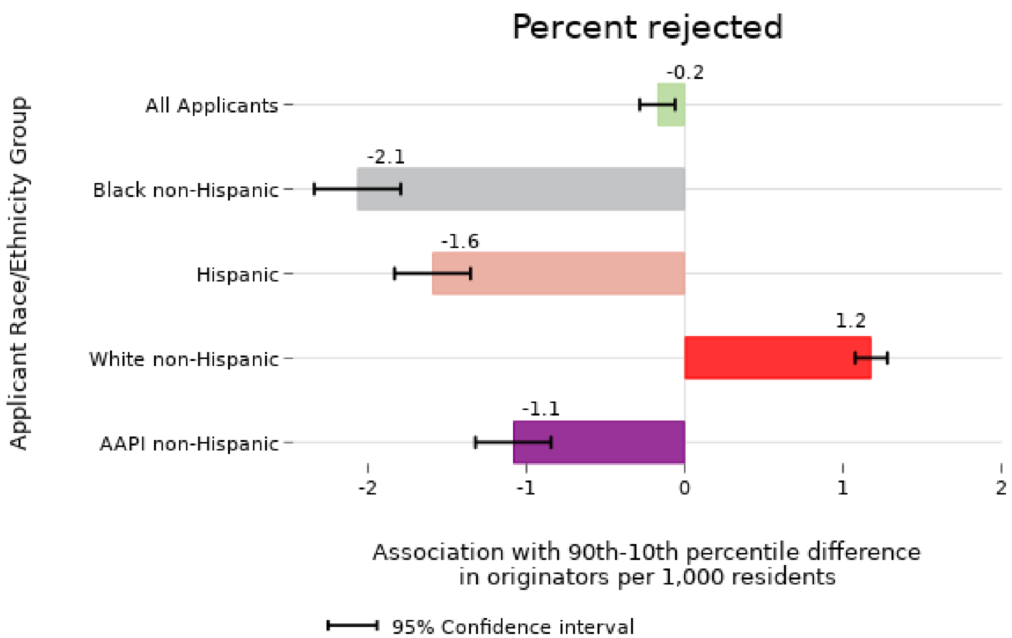
based on variation in originators per capita within more comparable groups of loans and borrowers. I estimate regressions for all sample applications or originations as well as separate regressions for subgroups based on coarse categories of applicant race and ethnicity (Black non-Hispanic, Hispanic, Asian American Pacific Islander (AAPI) non-Hispanic, and white non-Hispanic).

Figures 16–21 present estimates of the relationships between the number of originating lenders in a census tract per 1,000 residents and the outcomes. Because a difference of one originator per 1,000 residents is small (less than 3 percent of a standard deviation), the figures display the expected difference in the outcome for a borrower in a neighborhood that had the 90th percentile value of originators per capita (from the overall neighborhood ranking) relative to a borrower in a neighborhood that had the 10th percentile value of originators per capita, holding other observable characteristics constant.[37]

Both rejection rates (Figure 16) and rejection conditional on AUS approval (Figure 17) were higher in neighborhoods with fewer originators per capita. This was true for the model for all applicants and for each of the separate models by race/ethnicity groups except the model for white non-Hispanics. Applicants in neighborhoods at the 90th percentile of originators per capita were 0.2 percentage points (1.4 percent of the overall mean) less likely to have their application rejected than applicants in neighborhoods at the 10th percentile; the corresponding difference for the likelihood of a rejected application after AUS approval suggests applicants in the 90th percentile neighborhood were 0.5 percentage points (7.8 percent of the combined sample mean) less likely to be rejected conditional on AUS approval. The race/ethnicity group-specific estimates were largest for Black non-Hispanic and Hispanic applicants: a 2.1 percentage point difference in the likelihood of rejection for Black non-Hispanic applicants and a 1.6 percentage point difference in the likelihood of rejection for Hispanic applicants. The estimates for rejection conditional on AUS approval were a 1.1 percentage point difference for both Black non-Hispanic and Hispanic applicants. For AAPI non-Hispanic applicants, estimates were always between the pooled estimates and those for Black non-Hispanic and Hispanic applicants; the coefficients for white non-Hispanics were positive for both outcomes.

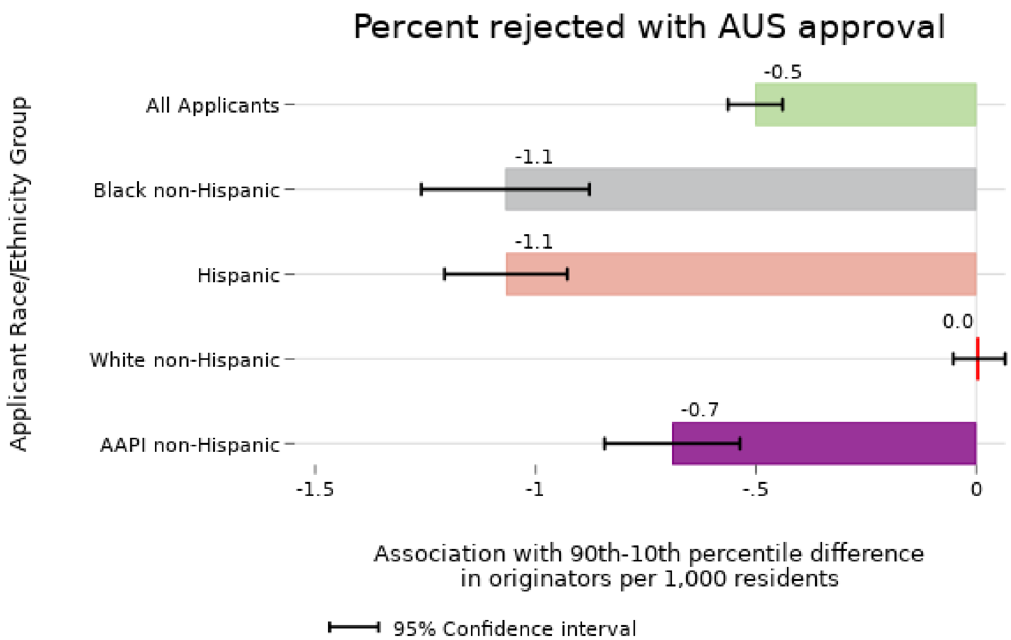
[37] Under the assumption that the relationship of interest is linear, this is simply the coefficient from the GSE fixed effects regression multiplied by the difference in originating institutions per 1,000 residents in the 90th percentile census tract and in the 10th percentile census tract: 27.2 originators per 1,000 residents. 95 percent confidence intervals are based on Wald tests of linear combinations of the regression estimates.

Figure 16: Percent of Applications Rejected and Originators Per Capita



Note: Figure presents ordinary least squares estimates and 95% confidence intervals of the associations between census tract-level originators per capita and the percent of conforming mortgage applications that were rejected by applicant race/ethnicity. All specifications include loan percentile fixed effects, action month fixed effects, and Government Sponsored Enterprise (GSE) loan-level price adjustment fixed effects. 95% confidence intervals are based on standard errors clustered at the census tract level.

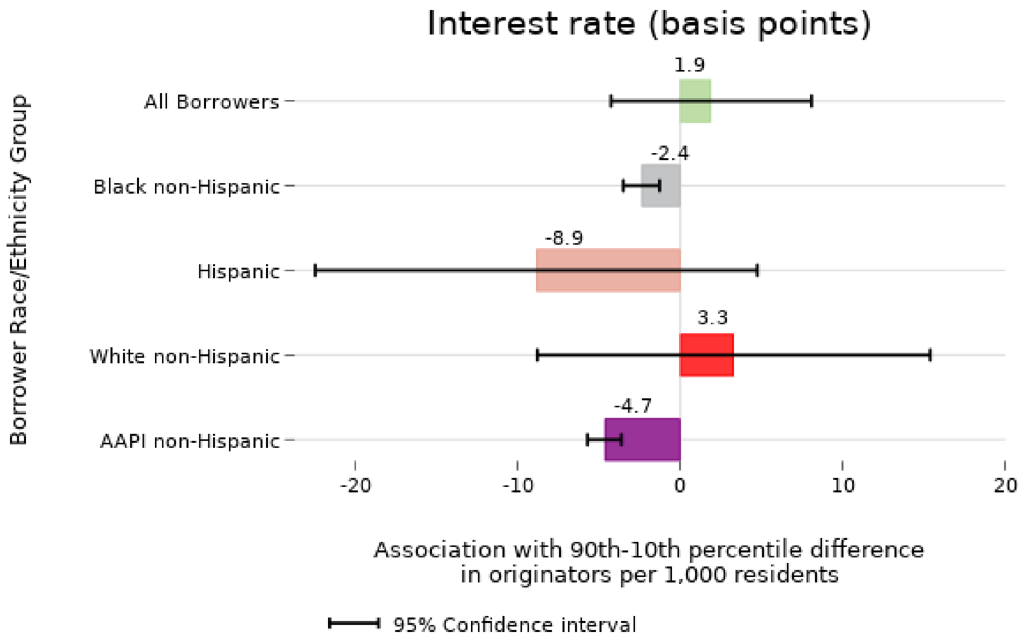
Figure 17: Percent of Applications Rejected with AUS Approval and Originators Per Capita



Note: Figure presents ordinary least squares estimates and 95% confidence intervals of the associations between census tract-level originators per capita and the percent of conforming mortgage applications that were rejected with Automated Underwriting System approval by applicant race/ethnicity. All specifications include loan percentile fixed effects, action month fixed effects, and Government Sponsored Enterprise (GSE) loan-level price adjustment fixed effects. 95% confidence intervals are based on standard errors clustered at the census tract level.

The relationship between the number of originating lenders per 1,000 residents and interest rates for conforming, originated loans is displayed in Figure 18. Ninety-five percent confidence intervals for the interest rate outcome include zero for all borrowers as well as for Hispanic and white non-Hispanic borrowers. For these groups, originated, conforming loans were similarly priced in census tracts with more and fewer originators per capita for loans identified as posing similar credit risk. Associations were negative, but small in magnitude for Black non-Hispanic and AAPI non-Hispanic borrowers. In expectation, interest rates were 2.4 basis points higher for originations in a 10th percentile neighborhood than they were in 90th percentile neighborhoods for Black non-Hispanic borrowers, and 4.7 basis points higher for AAPI non-Hispanic borrowers. At the median loan amount for all conforming loans, an interest rate that is 4.7 basis points lower for a 30-year fixed mortgage would imply just $6 less in monthly housing costs; a 2.4 basis point decrease in the interest rate would reduce monthly costs by just $3. These average differences are not likely to be economically meaningful for most consumers.

Figure 18: Interest Rates for Originated Loans and Originators Per Capita



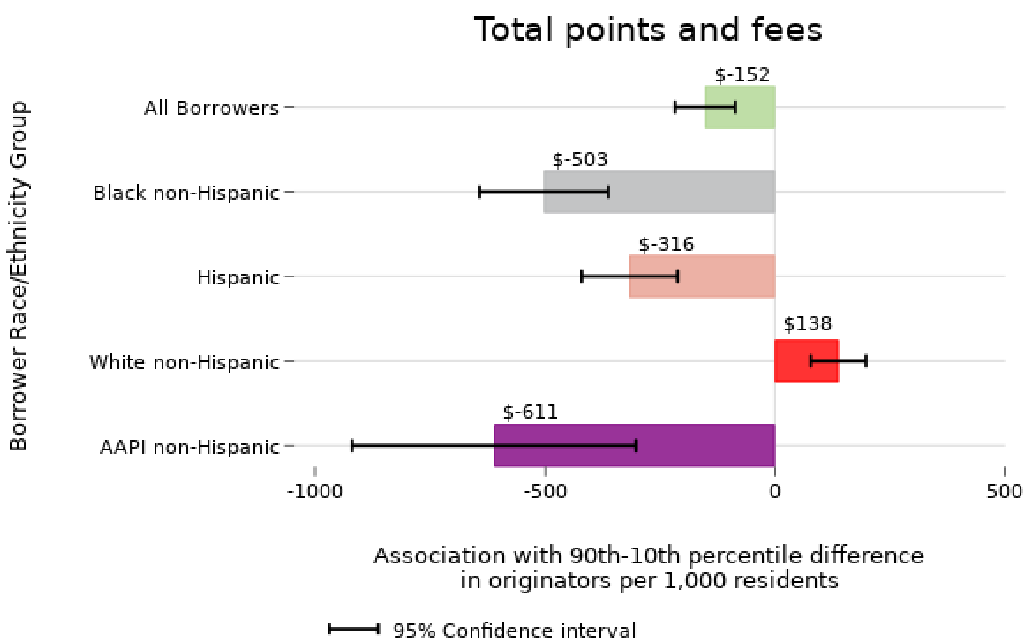
Note: Figure presents ordinary least squares estimates and 95% confidence intervals of the associations between census tract-level originators per capita and interest rates (in basis points) for originated conforming mortgage loans by applicant race/ethnicity. All specifications include loan percentile fixed effects, action month fixed effects, and Government Sponsored Enterprise (GSE) loan-level price adjustment fixed effects. 95% confidence intervals are based on standard errors clustered at the census tract level.

Figure 19 displays the associations between total points and fees and originators per capita. For the full sample, total points and fees are $152 lower, on average, in neighborhoods at the 90^{th} percentile value of originators per capita relative to neighborhoods at the 10^{th} percentile. The same comparison was associated with larger differences for Black non-Hispanic ($503 lower), Hispanic ($316 lower), and AAPI non-Hispanic ($611 lower) borrowers. White non-Hispanic borrowers were the only group for which neighborhoods with more originators per capita had higher borrower-paid total points and fees: $138 higher, on average. For all groups except white non-Hispanics, more originators per capita in a neighborhood was associated with both lower interest rates and lower total points and fees. Although—as mentioned above—total points and fees are reported for relatively few originations,[38] the point estimates for that sub-sample suggest that borrowers paid smaller up-front costs (total borrower paid points and fees) *and* made smaller monthly payments (interest rates) if there were a greater number of originators per capita in their census tract. At least for these transactions, the point estimates were not

[38] Around 1 percent of originated, conforming loans with an interest rate also report total points and fees.

consistent with the results being driven solely by differences in borrower preferences that result in different intertemporal choices (e.g., paying more up-front to make smaller subsequent monthly payments).

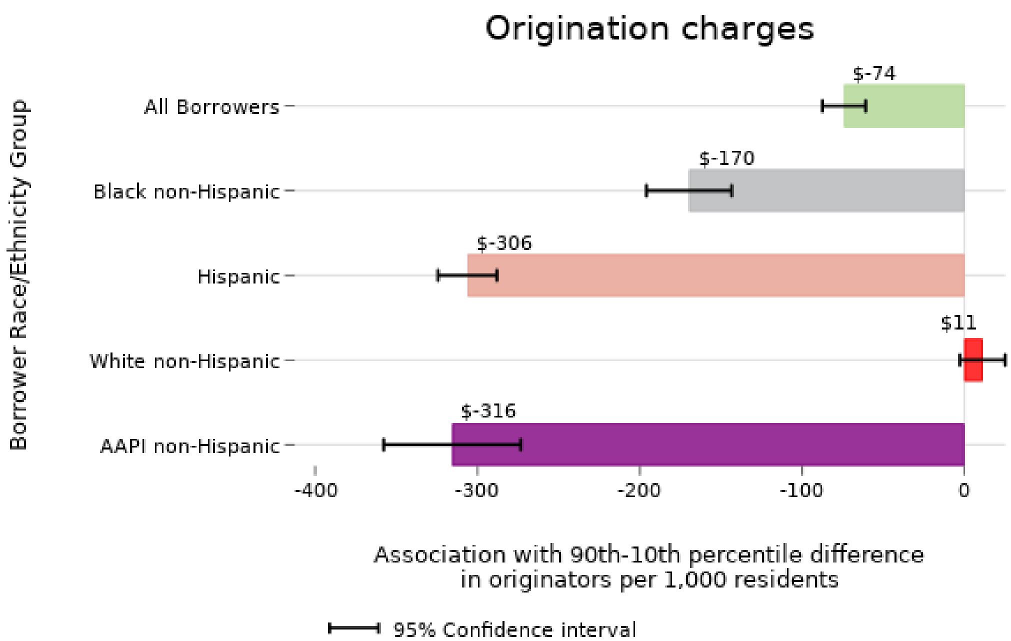
Figure 19: Total Points and Fees for Originated Loans and Originators Per Capita



Note: Figure presents ordinary least squares estimates and 95% confidence intervals of the associations between census tract-level originators per capita and total points and fees (in dollars) for originated conforming mortgage loans by borrower race/ethnicity. All specifications include loan percentile fixed effects, action month fixed effects, and Government Sponsored Enterprise (GSE) loan-level price adjustment fixed effects. 95% confidence intervals are based on standard errors clustered at the census tract level.

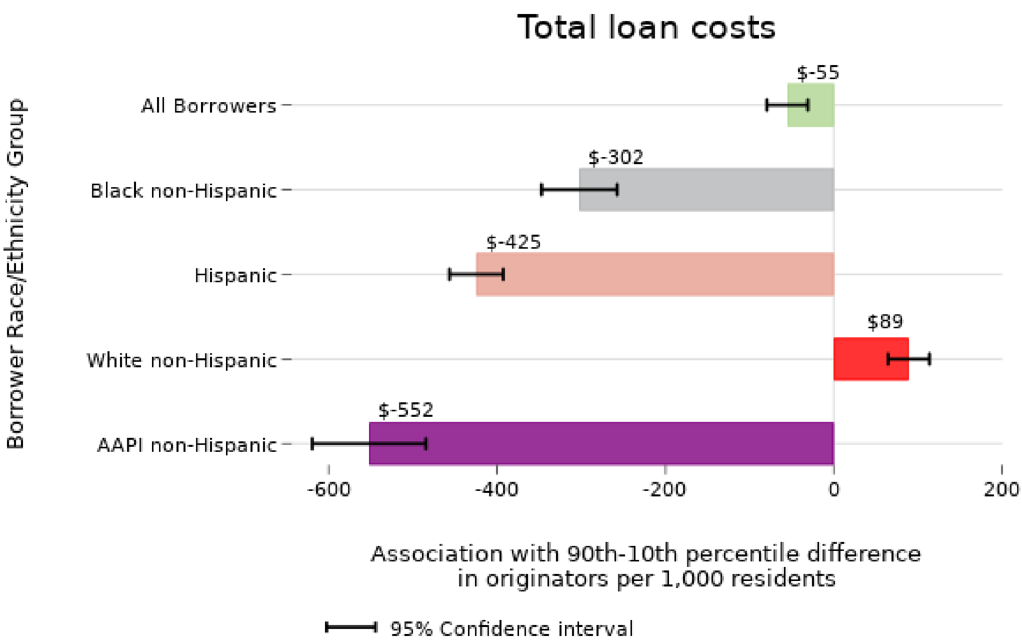
Figures 20 and 21 present the estimated relationships between originators per capita, origination charges, and total loan costs. The associations were similar in sign and magnitude for both outcomes across most demographic groups. For the pooled sample, origination charges in the 90th percentile neighborhood with respect to originators per capita were $74 lower than in the 10th percentile neighborhood, while total loan costs were $55 lower. For Black non-Hispanic and Hispanic borrowers, origination charges were $170 and $306 lower, while total loan costs were $302 and $425 lower. AAPI non-Hispanic borrowers had origination charges that were $316 lower and total loan costs that were $552 lower in neighborhoods with originators per capita equal to the 90th percentile value relative to those in neighborhoods with originators per capita equal to the 10th percentile value. White non-Hispanic borrowers are the only group for whom neighborhoods with more originators per capita did not have lower origination charges—though the 95 percent confidence interval included zero and the point estimate suggests that origination charges were just $11 higher in neighborhoods with originators per capita equivalent to the 90th percentile value relative to those with originators per capita equal to the 10th percentile value. Total loan costs for white non-Hispanic borrowers were also higher in the 90th percentile neighborhood (by $89), but the difference is modest.

Figure 20: Borrower-paid Origination Charges for Originated Loans and Originators Per Capita



Note: Figure presents ordinary least squares estimates and 95% confidence intervals of the associations between census tract-level originators per capita and total borrower-paid origination charges for conforming mortgage originations by borrower race/ethnicity. All specifications include loan percentile fixed effects, action month fixed effects, and Government Sponsored Enterprise (GSE) loan-level price adjustment fixed effects. 95% confidence intervals are based on standard errors clustered at the census tract level.

Figure 21: Borrower-paid Total Loan Costs for Originated Loans and Originators Per Capita



Note: Figure presents ordinary least squares estimates and 95% confidence intervals of the associations between census tract-level originators per capita and total borrower-paid total loan costs for conforming mortgage originations by borrower race/ethnicity. All specifications include loan percentile fixed effects, action month fixed effects, and Government Sponsored Enterprise (GSE) loan-level price adjustment fixed effects. 95% confidence intervals are based on standard errors clustered at the census tract level.

The median reported origination charge for a conforming loan in the sample was $1,311. The results above suggest that—for loans and borrowers with similar credit risk to lenders, of a similar size, and originated in the same calendar month—origination charges in neighborhoods at the 90th percentile of originators per capita were lower by nearly 25 percent of the median origination charge, relative to origination charges in neighborhoods at the 10th percentile of originators per capita. For Black non-Hispanic borrowers, the difference was 13.0 percent; for Hispanic borrowers the difference was 23.1 percent; and for AAPI borrowers the difference was 24.1 percent.

The results shown in Figures 16–21 indicate that applicants were less likely to be rejected for a mortgage—both conditional on AUS approval and unconditional on AUS approval—and paid lower origination costs and total loan costs when they applied for a mortgage in neighborhoods where more financial institutions per capita originated mortgages. An immediate follow-up question is whether the results are driven by the types of institutions that accept applications and originate loans in neighborhoods with higher originators per capita or by variation in lenders' decisions about applications and loan pricing across neighborhoods. To investigate, controls for the lending institution[39] are added to the specification used to produce the results in Figures 16–21. This isolates variation in originators per capita that occurs *within* the same lending institution, comparing outcomes for consumers applying to an institution in a neighborhood with more originators per capita to outcomes for similar consumers applying to the same institution but in a neighborhood with fewer originators per capita.

The point estimates from the specifications that use variation within institutions hewed closely to the results shown in Figures 16–21. For the pooled sample, the 90th percentile–10th percentile difference in originators per capita (27.2 originators per 1,000 residents) was associated with: a 1 percentage point lower likelihood of rejection; a 0.7 percentage point lower likelihood of rejection conditional on AUS approval; $135 lower origination charges for originated, conforming loans; and $166 lower total loan costs. The differences in association sizes across race and ethnicity groups seen in Figures 16–21 are somewhat muted when adding controls for the lending or application receiving institution, suggesting some of those differences in outcomes can be explained by variation in the institutions that these groups utilize for mortgage financing.

[39] More precisely, lender fixed effects, or a series of indicator variables that take on the value of one if each possible lending institution was the originating institution or application receiving institution, and zero otherwise, are added to the regression specifications.

4. Conclusion

This report uses the expanded HMDA data points to describe how mortgage activity varied across neighborhoods between 2018–2020, with a focus on differences associated with the number of originators per 1,000 residents in a neighborhood. Neighborhoods with more originators per capita had, on average, higher incomes, higher levels of adult educational attainment, and more household internet access; these neighborhoods also had lower rejection rates, lower rejection rates conditional on automated underwriting system approval, and paid lower origination charges and total loan costs as a percent of the total loan amount. The report confirms that the associations with mortgage outcomes persist when considering transaction-level outcomes and only relying on variation in originators per capita that occurred between transactions that posed a similar level of credit risk, that were of a similar loan size, and that occurred in the same calendar month.

The observed differences in the neighborhood-level number of originating institutions per capita and the corresponding variation in borrower outcomes are undoubtedly shaped by a variety of consumer and firm-level choices. These potentially include demand-side factors such as consumer shopping behavior, peer information and referrals, and financial literacy, as well as supply-side factors such as the physical proximity of financial institution branches, marketing behavior, and the feasible choices for consumers among originating financial institutions and origination service providers. While distinguishing between these factors is beyond the scope of this report, the links between institution and consumer choices and mortgage outcomes often depend, at least in part, on consumers having imperfect information about the mortgage options available to them. In mortgage markets, it can be difficult for consumers to get information on the prices and features of the mortgage products available to them.[40] And nearly half of consumers report not getting more than one mortgage offer, despite there being substantial price dispersion and large potential benefits to engaging in search behavior.[41] While existing work has focused on dispersion in interest rates, differences in borrower paid origination charges and loan costs are also likely to be affected by search and can meaningfully impact consumer welfare.

---

[40] Even within markets for consumer financial products this is not unique. See Lawrence M. Ausubel, "The Failure of Competition in the Credit Card Market," *The American Economic Review,* 81(1): 50-81, available at https://econ.umd.edu/sites/www.econ.umd.edu/files/pubs/aerhigh.pdf.

[41] See Alexei Alexandrov and Sergei Koulayev, "No Shopping in the U.S. Mortgage Market: Direct and Strategic Effects of Providing Information," available at https://dx.doi.org/10.2139/ssrn.2948491 , and "Does shopping for a mortgage make consumers better off?" Bureau of Consumer Financial Protection Research Brief, May 2018, available at https://files.consumerfinance.gov/f/documents/bcfp_mortgages_shopping-study_brief-2-experimental-results.pdf.