

# How Reveal identified lending disparities in federal mortgage data

By **Emmanuel Martinez** and **Aaron Glantz**



from The Center for Investigative Reporting

Reveal from The Center for Investigative Reporting's analysis of lending disparities within racial and ethnic groups relied on publicly available records, released through the [Home Mortgage Disclosure Act](#) and maintained by the [Federal Financial Institutions Examination Council](#). HMDA, passed in 1975, requires mortgage lenders to report basic data about home loan applications to ensure fair lending practices.

The HMDA data set contains information about nearly every mortgage application in the country in a given year. HMDA exempts lending institutions that don't meet [certain requirements](#). The data set includes a wide range of information from the lending institution that received the application, including the race, ethnicity and income of the applicant. It also includes the type and size of loan that's being sought, and information about the neighborhood in which the property is located.

Reveal analyzed 31 million mortgage records from 2015 and 2016 and found that in certain areas of the country, people of color were more likely to be denied a conventional mortgage than white applicants, even after controlling for a wide array of economic and social factors. Reveal's analysis exposed a pattern of denials in major metropolitan areas such as Atlanta, St. Louis and San Antonio; and in smaller ones, such as Chico, California; Iowa City, Iowa; and Mobile, Alabama.

We focused our story in Philadelphia because it consistently proved statistically significant regardless of which variables were included, and because among the largest metro areas, it has one of the widest lending disparities: Prospective black borrowers were almost three times as likely to be denied a conventional home purchase loan there than white applicants.

Reveal's story focuses on uneven access to the conventional mortgage market for applicants of color when compared with white ones.

We concentrated on conventional loans for one- to four-unit properties where prospective borrowers said they intended to live in the home they were looking to buy, a similar subset of the loans that the Federal Reserve analyzes when it tracks lending trends.

We looked at every variable in the HMDA data set and added additional ones based on feedback from experts and research. A full list of variables used in our analysis is in the appendix.

## **How we analyzed the data**

To evaluate whether disparities in lending exist, we used a statistical technique called a binary logistic regression. This type of regression assesses the relationship between multiple independent variables against a single binary output, whether a specific event occurred or not. In this case, we looked at whether a mortgage was denied.

Lending institutions look at several factors to determine whether they will approve someone for a loan. A person's income is one example. Because income varies among people of different races and ethnicities, it's important to reduce the influence that race and ethnicity have on income and hold income constant among all applicants when looking at the disparities in lending. A logistic regression allows us to control for those factors.

We separately analyzed data from 2015 and 2016, looking at nine independent variables to predict loans that were denied. Those factors were:

- Race/Ethnicity
  - Native American
  - Asian
  - Black
  - Native Hawaiian
  - Hispanic/Latino
  - Race and ethnicity were not reported
- Sex
- Whether there was a co-applicant
- Applicant's annual income (includes co-applicant income)
- Loan amount
- Ratio between the loan amount and the applicant's income
- Ratio between the median income of the census tract and the median income of the metro area
- Racial and ethnic breakdown by percentage for each census tract
- Regulating agency of the lending institution

## Regulating agencies

There are six agencies that regulate lending institutions:

- Office of the Comptroller of the Currency (OCC)
- Federal Reserve System (FRS)
- Federal Deposit Insurance Corporation (FDIC)
- National Credit Union Administration (NCUA)
- Department of Housing and Urban Development (HUD)
- Consumer Financial Protection Bureau (CFPB)

We also felt it was important to distinguish among the regulating agencies and look at their behavior. We added these variables because regulations differ among regulatory agencies.

## Credit score and debt-to-income ratio

Credit scores are an important factor for lending institutions in deciding whether to approve or deny a loan application.

Credit scores are assigned to borrowers by a credit rating agency that assesses their trustworthiness to pay back a loan. A higher credit score means borrowers are making consistent and on-time payments and are a good credit risk. Despite its importance in guiding lending decisions, credit score data is not included in the HMDA data set, so we couldn't control for it as a variable. Banks consider that data proprietary, and while government regulators can obtain access to individual applicants' credit scores, the Freedom of Information Act specifically [exempts certain financial information](#) used by bank regulators from release to the public.

Credit scores carry their own set of problems. [Studies](#) have cited a relationship between lower credit scores and borrowers of color. The credit score doesn't reflect on-time rent or utility bill payments, only those that are delinquent, which disproportionately affects people of color who may have less access to other types of credit, such as mortgages and credit cards.

Lenders have also pointed out that an applicant's total debt-to-income ratio weighs heavily in lending decisions. That data point isn't currently available in the data set either.

HMDA regulations are changing, however, and applicants' debt-to-income ratios will be added as a field in the publicly available data for 2019 – but those ratios will be placed into categories rather than expressed as actual figures. Credit scores will be included only in data available to government regulators, not in the publicly available data set.

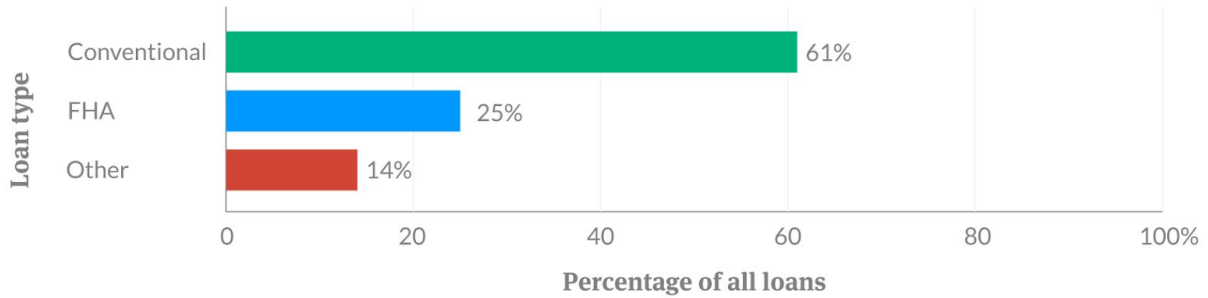
## **How we chose which loans to focus on**

### **Loan type**

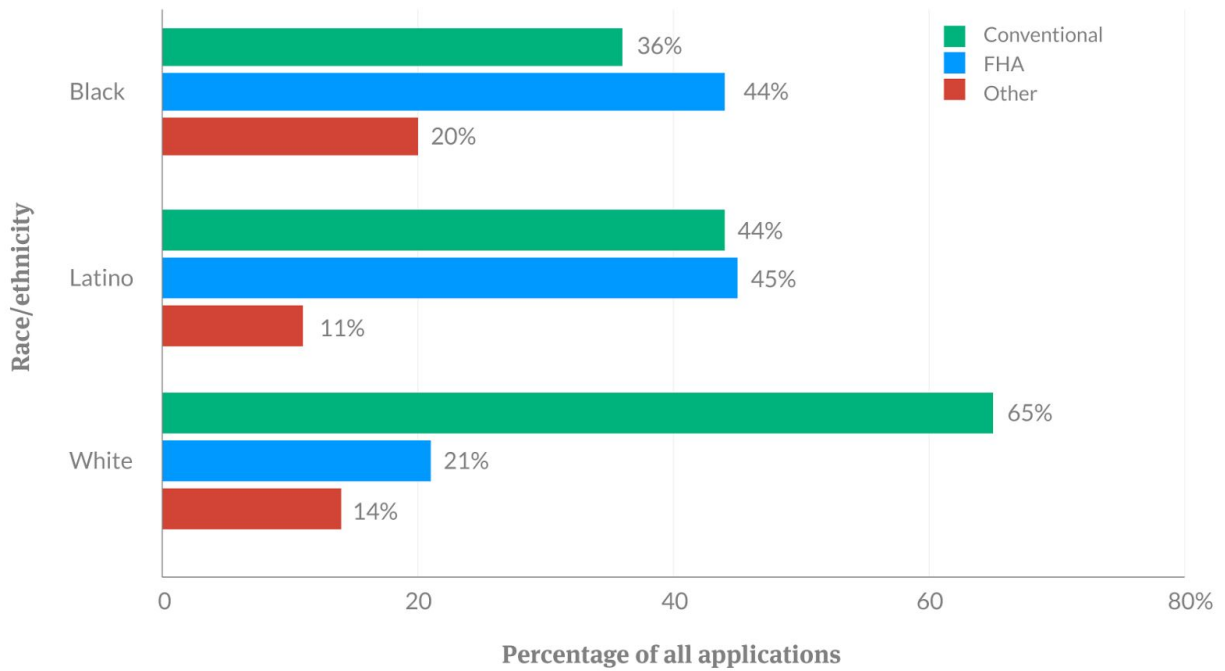
Conventional loans often provide the best terms for borrowers as they usually have lower interest rates and do not include extra fees and longer-term insurance premiums associated with government-backed loans. Applicants for conventional loans must frequently, depending on the lender and the applicant, put down a hefty down payment, which requires having more savings. The more money applicants put down the better it protects them from going underwater on their mortgage. Lenders consider those applicants less risky borrowers.

There are opportunities beyond conventional loans that get prospective homebuyers into a home. A Federal Housing Administration (FHA) mortgage is a type of loan that's backed by the federal government. That means the government assumes the risk if a borrower defaults, not the lending institution that made the loan. FHA loans provide an important bridge between applicants who would otherwise not qualify for a conventional loan and owning a home and building equity. They may be the only option for those who can't afford a large down payment or have a lower credit score. Other government-backed loans include those provided through the Veterans Administration (VA) and the Farm Service Agency or Rural Housing Service (FSA/RHS).

The 2016 HMDA data shows conventional loans made up a majority of the nationwide home purchase market, with FHA loans accounting for a quarter of all loans. The breakdown between conventional and FHA loans varies depending on the metro.



White applicants apply for conventional loans at higher rates than FHAs, while more black and Latino applicants apply for FHA loans.



Note: "Other" includes VA and Farm Service Agency/Rural Housing Service loans.

While we recognize the substantial presence of applicants of color in the FHA market, we focused on conventional home loans for several reasons.

We wanted to see whether applicants of color were being shut out of the conventional mortgage market and how lending institutions acted when the government wasn't directly involved. Conventional and FHA loans are two distinct products. Over the life of the loan, a conventional loan is usually a better deal. Most of the time, conventional loans have lower interest rates and borrowers have an easier time avoiding costly mortgage insurance.

Conventional loans are advantageous for lending institutions, too. They have an easier time getting them off their books because they are considered less risky. And many banks, such as [JPMorgan Chase & Co.](#), are retreating from the FHA mortgage market.

The decision between getting a conventional and FHA loan isn't always based on an applicant's merits. Our reporting has shown banks have [steered applicants who qualified for conventional mortgages into FHA loans](#). Philadelphia is [suing Wells Fargo](#), alleging such steering. We've also heard stories of banks sending only loan officers who specialize in FHA loans into neighborhoods of color. Because of these concerns, we decided to exclude FHA loans to avoid potential interference by other forms of discrimination in our analysis.

## Property type

Prospective borrowers submit loan applications for various types of structures: one- to four-unit properties, multifamily properties and manufactured homes. For this analysis, we focused on one- to four-unit properties.

## Occupancy

We included only borrowers who said they planned to live in the house they were looking to buy. We did this to exclude developers or individuals who were buying property as an investment or to subsequently flip it.

## Action Type

The HMDA data tracks several types of mortgage outcomes:

1. Originated, meaning the loan is made.
2. The loan is approved, but the applicant doesn't accept it.
3. The lending institution denies the applicant.
4. Applicant withdraws the application.
5. The application is closed because it's incomplete.
6. Lending institution purchases loan from another lending institution.
7. The mortgage is approved but not accepted at pre-approval.
8. The mortgage is denied at pre-approval.

We wanted to look at the reasons lending institutions deny people a mortgage. After conversations with former officials at HUD, we decided to include only those applications that resulted in originations (action type 1) or denials (action type 3). This is the methodology that the Justice Department uses to determine whether there is potential discrimination as defined in the Fair Housing Act of 1968. This is because these two are the most unequivocal outcomes. Our reporting has shown the mortgage application process may take a lot longer for applicants of color, which may cause them to withdraw their applications. Because we couldn't account for the reasons why applicants withdraw their applications or don't accept the terms of their loan, we excluded them from our analysis.

## **Income, loan and logarithm transformation**

An applicant's income isn't always reported in the data. In other cases, the data cuts off any incomes over \$9.9 million and any loan amounts over \$99.9 million, meaning there's a value in the database, but it's not precise. We focused only on those records where income and loan amount have an accurate estimation. This meant discarding about 1 percent of all conventional home loans in the country for 2016.

When we plotted the number of applicants, how much money they made and the size of the loan, we found that it skewed to the left, meaning the majority of applicants were clustered on the lower end of the income and loan amount scales. This was especially true for applicants of color. We took the logarithm transformation of income and loan amount to normalize the distribution of those variables and limit the effect of extreme outliers.

## **Lien status**

Lien identifies how the mortgage was secured. A first lien means it's a first mortgage and a subordinate lien refers to a second mortgage, for example. We included all cases in our analysis regardless of lien status. Although this accounts for a small fraction of loans, the data shows conventional, home purchase, owner-occupied loans aren't always secured through a first lien.

## **Race and ethnicity**

At first, we looked at race separate from ethnicity, but that approach introduced too many instances in which either the ethnicity or race was unknown. So we decided to combine race and ethnicity. Applicants who marked their ethnicity as Hispanic were grouped together as Hispanic/Latino regardless of race. Non-Hispanic applicants, as well as those who didn't provide an ethnicity, were grouped together by race: non-Hispanic white, non-Hispanic black, etc. This method is used by HUD investigators in determining whether there is potential discrimination against a particular racial or ethnic group as defined under the Fair Housing Act.

## **Metropolitan statistical areas/metro divisions**

HMDA data uses a combination of metropolitan statistical areas (MSA) and metro divisions (MD). The 11 largest MSAs are divided into 31 smaller divisions. We grouped our output by

metropolitan statistical areas or metro divisions, where applicable. Counties can have too small of a sample size, especially when looking at loan applications broken down by race and ethnicity, and grouping everything by state fails to capture the diversity and nuance within the state. Similarly, we chose not to group the metro divisions back into their respective MSAs because we wanted to preserve the differences among the populations within those divisions.

## **Loan purpose: Home purchase, home improvement or refinance**

We decided to look at home purchase, home improvement and refinance loans separately from each other. One reason we looked at them independently: Applicants may be better financially prepared when buying a home than when they are applying for a home improvement or refinance loan.

## **Results**

We found 61 metros out of 409 where applicants of color were more likely to be denied a conventional mortgage that's being used to purchase a home. We isolated those metros based on the following criteria:

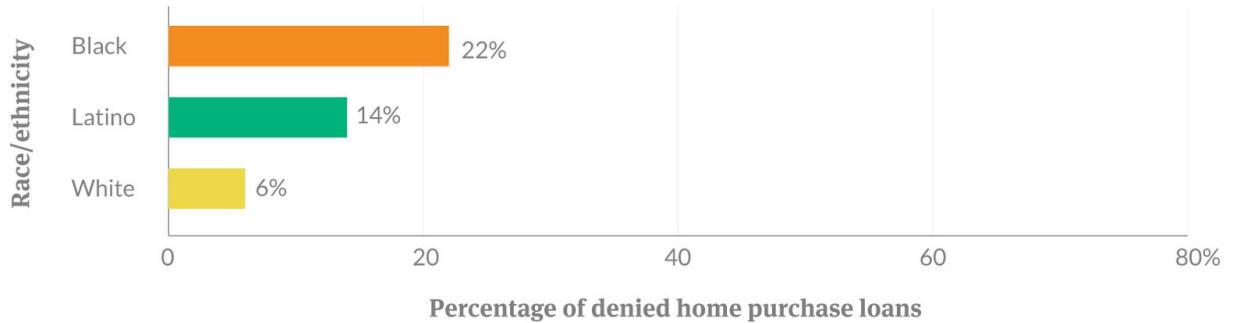
- Pseudo R-Square is greater than or equal to 0.1.
- P-value for applicants of color is less than or equal to 0.05.
- The odds multiplier for each variable is greater than or equal to 1.5.
- Each racial and ethnic group had to have 75 or more applicants.
- The metros had to individually meet these cutoffs for home purchase loans in both 2015 and 2016 to establish that the result was not anomalous. In our stories, we use numbers from 2016 to show the most recent data.

Below are the results for Philadelphia. We included a list of all metros that were statistically significant for both 2015 and 2016 in the appendix section.



## Conventional home purchase loan denial rate in the Philadelphia metro division, 2016

Less than a quarter of all black applicants in the Philadelphia metro division were denied a conventional home purchase loan, according to 2016 HMDA data. That's the highest percentage of any racial and ethnic group in Philadelphia with more than 100 applications that year.



## Regression analysis for conventional home purchase loans for the Philadelphia metro division, 2016

Pseudo R Square: 0.148

Hosmer and Lemeshow Goodness of Fit: 0.449

Number of cases in 2016: 11,367

Prospective black borrowers were nearly three times as likely to be denied a home purchase loan, and Latinos were 1.6 times as likely to be denied when compared with white, non-Hispanic applicants, even after controlling for income.

Variable	B	S.E.	Wald	Sig.	Exp(B)
Native American	1.961	0.588	11.113	0.001	7.104
Asian	0.261	0.111	5.517	0.019	1.298
Black	0.995	0.111	79.993	0.000	2.706
Native Hawaiian	0.955	0.529	3.262	0.071	2.598
Latino	0.513	0.145	12.482	0.000	1.671
Race N/A	0.350	0.128	7.480	0.006	1.419
Male	0.127	0.071	3.196	0.074	1.135
Sex N/A	0.562	0.173	10.626	0.001	1.755
Co-applicant	0.110	0.080	1.891	0.169	1.116
Co-applicant N/A	0.056	0.213	0.069	0.792	1.058
Log of loan	-0.554	0.072	59.272	0.000	0.574
Log of income	0.290	0.097	8.885	0.003	1.336
Loan-to-income ratio	0.335	0.038	77.468	0.000	1.399
Tract-MSA median income ratio	0.002	0.001	5.714	0.017	1.002
Percent black in tract	0.658	0.158	17.296	0.000	1.931
Percent Latino in tract	0.867	0.332	6.831	0.009	2.379
Percent Asian in tract	2.120	0.433	23.995	0.000	8.329

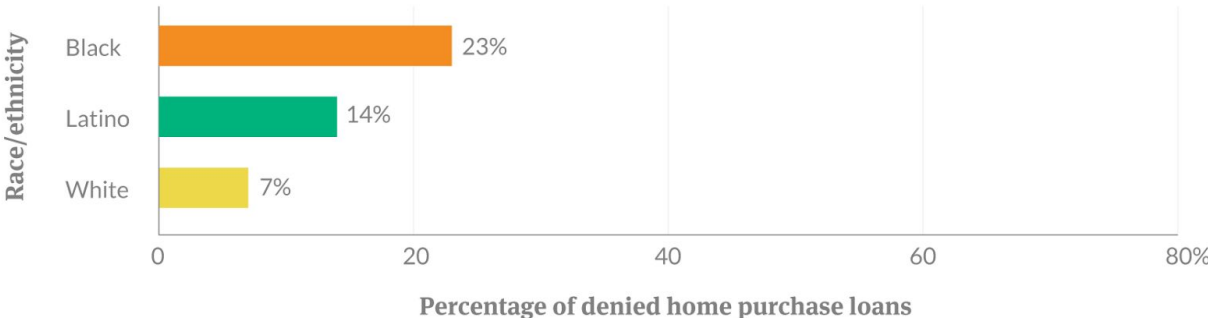
Percent Native American in tract	-10.716	9.025	1.410	0.235	0.000
Percent Native Hawaiian in tract	-2.920	19.921	0.021	0.883	0.054
OCC	0.824	0.248	11.010	0.001	2.279
FRS	0.169	0.194	0.760	0.383	1.184
NCUA	1.421	0.148	91.795	0.000	4.142
HUD	0.081	0.130	0.390	0.532	1.085
CFPB	0.966	0.128	57.209	0.000	2.627

Reference variables

- White non-Hispanic applicants
- Female applicants
- Federal Deposit Insurance Corporation (FDIC)

# Conventional home purchase loan denial rate in the Philadelphia metro division in 2015

The denial rate for black, Latino and white applicants was about the same in 2015 for these three racial/ethnic groups. Similarly to 2016, black applicants were denied nearly a quarter of the time they submitted a conventional mortgage application. White applicants were denied 7 percent of the time, the smallest percentage of any racial or ethnic group with more than 100 applications in Philadelphia.



## Regression analysis for conventional home purchase loans for the Philadelphia metro division, 2015

Pseudo R Square: 0.133

Hosmer and Lemeshow Goodness of Fit: 0.002

Number of cases in 2015: 10,126

Prospective black borrowers were about twice as likely to be denied a home purchase loan, even after controlling for income. We did not include Latinos in Philadelphia in our findings because in 2015 their likelihood of being denied was under our cutoff of 1.5 and we could not conclusively show a pattern of disparity.

Variable	B	S.E.	Wald	Sig.	Exp(B)
Native American	1.059	0.591	3.212	0.073	2.885
Asian	0.378	0.105	12.928	0.000	1.460
Black	0.751	0.115	42.484	0.000	2.119
Native Hawaiian	1.099	0.464	5.610	0.018	3.001
Latino	0.324	0.164	3.905	0.048	1.383
Race N/A	0.450	0.119	14.301	0.000	1.568
Male	0.121	0.070	2.928	0.087	1.128
Sex N/A	-0.090	0.176	0.262	0.609	0.914
Co-applicant	0.090	0.078	1.327	0.249	1.094
Co-applicant N/A	0.448	0.218	4.218	0.040	1.565
Log of loan	-0.228	0.056	16.399	0.000	0.796
Log of income	-0.356	0.070	26.052	0.000	0.701
Loan-to-income ratio	0.011	0.007	2.178	0.140	1.011
Tract-MSA median income ratio	0.002	0.001	6.900	0.009	1.002
Percent black in tract	0.622	0.160	15.042	0.000	1.863
Percent Latino in tract	0.719	0.351	4.212	0.040	2.053
Percent Asian in tract	1.065	0.440	5.854	0.016	2.902

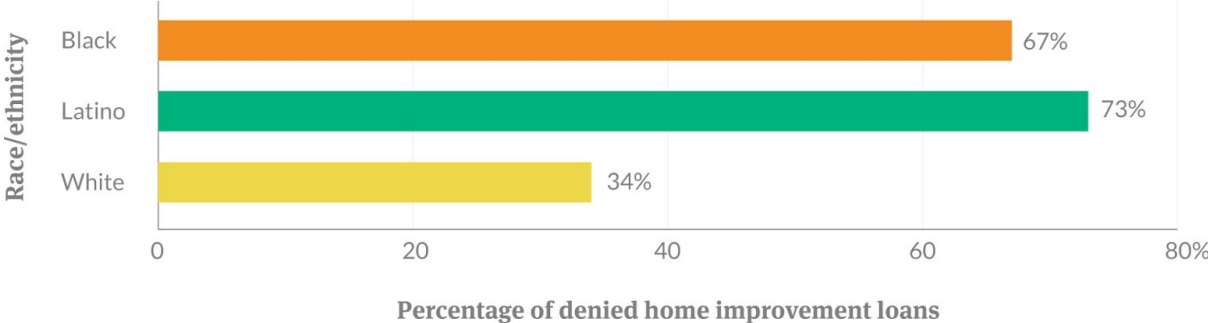
Percent Native American in tract	-25.637	9.697	6.990	0.008	0.000
Percent Native Hawaiian in tract	9.048	15.906	0.324	0.569	8501.434
OCC	0.317	0.201	2.490	0.115	1.372
FRS	-0.113	0.221	0.260	0.610	0.894
NCUA	1.302	0.153	72.479	0.000	3.676
HUD	0.118	0.140	0.712	0.399	1.125
CFPB	1.107	0.136	66.612	0.000	3.025

Reference variables

- White non-Hispanic applicants
- Female applicants
- Federal Deposit Insurance Corporation (FDIC)

# Conventional home improvement loan denial rate in the Philadelphia metro division, 2016

Every racial and ethnic group has a higher denial rate when looking at home improvement loans compared with home purchase. However, applicants of color were still denied at higher rates than white, non-Hispanic applicants.



## Regression analysis for conventional home improvement loans in the Philadelphia metro division, 2016:

Nagelkerke Pseudo R Square: 0.275

Hosmer and Lemeshow Goodness of Fit: 0.156

Number of cases in 2016: 5,991

When we looked at home improvement conventional loans separately from home purchase and refinance loans, we still saw disparities in lending for black applicants. Although much lower than home purchase loans, black applicants were still more likely – nearly twice – to be denied a home improvement loan than white, non-Hispanic applicants. The likelihood of denial for Latino applicants increased slightly when looking at home improvement loans.

Variable	B	S.E.	Wald	Sig.	Exp(B)
Native American	0.717	0.352	4.144	0.042	2.048
Asian	0.330	0.168	3.863	0.049	1.391
Black	0.511	0.101	25.405	0.000	1.666
Native Hawaiian	0.741	0.548	1.828	0.176	2.098
Latino	0.630	0.142	19.803	0.000	1.878
Race N/A	0.480	0.131	13.469	0.000	1.616
Male	-0.052	0.063	0.675	0.411	0.949
Sex N/A	0.180	0.160	1.257	0.262	1.197
Co-applicant	0.238	0.081	8.660	0.003	1.269
Co-applicant N/A	0.166	0.215	0.598	0.439	1.181
Log of loan	-0.114	0.040	8.020	0.005	0.893
Log of income	-0.570	0.063	82.907	0.000	0.566
Loan-to-income ratio	0.121	0.046	7.057	0.008	1.129
Tract-MSA median income ratio	-0.002	0.001	4.080	0.043	0.998
Percent black in tract	0.243	0.142	2.932	0.087	1.275
Percent Latino in tract	0.409	0.276	2.201	0.138	1.505



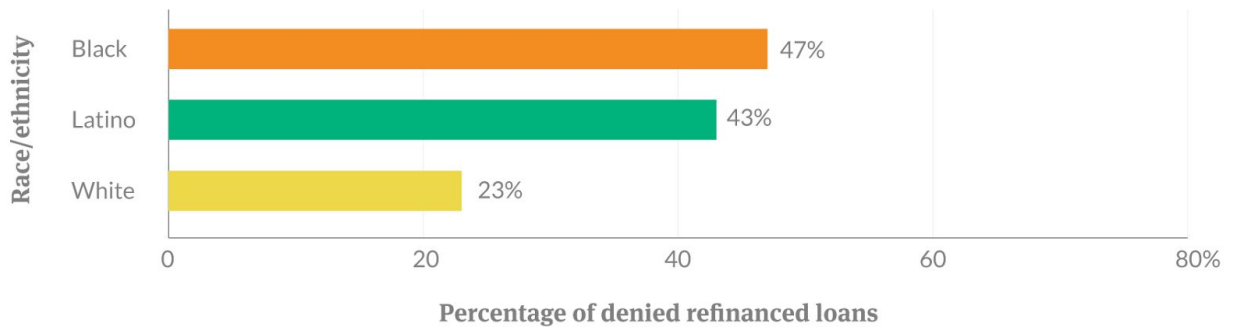
Percent Asian in tract	-0.426	0.453	0.885	0.347	0.653
Percent Native American in tract	16.068	9.113	3.109	0.078	9510041.445
Percent Native Hawaiian in tract	-29.146	23.216	1.576	0.209	0.000
OCC	0.123	0.289	0.182	0.670	1.131
FRS	0.145	0.242	0.357	0.550	1.156
NCUA	-0.050	0.130	0.150	0.698	0.951
HUD	-0.620	0.215	8.323	0.004	0.538
CFPB	0.807	0.129	39.197	0.000	2.241

Reference Variables

- White non-Hispanic applicants
- Female applicants
- Federal Deposit Insurance Corporation (FDIC)

## Conventional refinance loan denial rate in the Philadelphia metro division, 2016:

Refinance loans make up the bulk of all denied conventional loans. According to 2016 HMDA data, more than 40 percent of all Latinos who applied for a conventional refinance loan and almost half of all black applicants were denied a refinance conventional mortgage.



## Regression analysis for conventional refinance loans in the Philadelphia metro division, 2016:

Nagelkerke Pseudo R Square: 0.230

Hosmer and Lemeshow Goodness of Fit: 0.331

Number of cases in 2016: 16,297

We found that disparities in lending exists even when looking at refinance loans. Black and Latino applicants were both twice as likely to be denied a refinance conventional mortgage when compared with white, non-Hispanic applicants, even after we controlled for income.

Variable	B	S.E.	Wald	Sig.	Exp(B)
Native American	0.730	0.274	7.066	0.008	2.074
Asian	0.216	0.085	6.508	0.011	1.242
Black	0.699	0.064	120.765	0.000	2.013
Native Hawaiian	0.520	0.335	2.406	0.121	1.682
Latino	0.711	0.091	60.639	0.000	2.037
Race N/A	0.329	0.063	27.231	0.000	1.389
Male	0.105	0.039	7.147	0.008	1.111
Sex N/A	0.132	0.090	2.138	0.144	1.141
Co-applicant	0.242	0.043	31.799	0.000	1.274
Co-applicant N/A	-0.011	0.123	0.008	0.929	0.989
Log of loan	-0.407	0.050	67.028	0.000	0.665
Log of income	-0.056	0.060	0.879	0.348	0.945
Loan-to-income ratio	0.326	0.027	147.501	0.000	1.385
Tract-MSA median income ratio	-0.001	0.000	4.584	0.032	0.999
Percent black in tract	0.730	0.087	70.466	0.000	2.075
Percent Latino in tract	0.439	0.198	4.926	0.026	1.552
Percent Asian in tract	-0.368	0.284	1.684	0.194	0.692

Percent Native American in tract	-2.703	5.414	0.249	0.618	0.067
Percent Native Hawaiian in tract	-24.362	13.837	3.100	0.078	0.000
OCC	0.286	0.190	2.250	0.134	1.330
FRS	-0.527	0.240	4.817	0.028	0.590
NCUA	0.375	0.107	12.226	0.000	1.455
HUD	1.071	0.102	109.966	0.000	2.918
CFPB	0.517	0.105	24.456	0.000	1.678

Reference variables

- White non-Hispanic applicants
- Female applicants
- Federal Deposit Insurance Corporation (FDIC)

## **Conclusion:**

Our analysis shows that people of color – black, Latino, Asian and Native American applicants – don't access the conventional mortgage market at the same rate as their white counterparts. Prospective black borrowers have the toughest challenge: They are more likely to be denied a conventional mortgage in 48 metros across the country, even after holding income and the loan amount constant. Latinos are more likely to be denied in 25 metros, Asians in nine and Native American applicants in three.

## **Acknowledgements:**

- Angeliki Kastanis, The Associated Press
- Arthur Acolin, University of Washington
- Barbara van Kerkhove, Ph.D., Empire Justice Center
- Bruce C. Mitchell, Ph.D., National Community Reinvestment Coalition
- Elisabeth Risch, Metropolitan St. Louis Equal Housing and Opportunity Council
- Eric Sagara, Reveal from The Center for Investigative Reporting
- Kevin Stein, California Reinvestment Coalition
- Jennifer LaFleur, Investigative Reporting Workshop
- Peter Smith, Center for Responsible Lending
- Sinduja Rangarajan, Reveal from The Center for Investigative Reporting
- Stanton Glantz, Ph.D., University of California, San Francisco
- Vedika Ahuja, The Greenlining Institute
- William Axinn, University of Michigan

## Appendix:

List of metros where applicants of color are more likely to be denied a conventional mortgages (in alphabetical order):

- Atlanta-Sandy Springs-Roswell, GA
  - Black
  - Latino
  - Native American
- Augusta-Richmond County, GA-SC
  - Black
- Baton Rouge, LA
  - Black
- Bellingham, WA
  - Asian
- Camden, NJ
  - Asian
  - Black
- Chico, CA
  - Latino
- Columbia, SC
  - Black
  - Latino
- Corpus Christi, Texas
  - Latino
- Dayton, OH
  - Black
- Deltona-Daytona Beach-Ormond Beach, FL
  - Black
- Detroit-Dearborn-Livonia, MI
  - Black
- Dover, DE
  - Black
- Fayetteville-Springdale-Rogers, AR-MO
  - Latino
- Florence, SC
  - Black
- Fort Wayne, IN
  - Black
- Gainesville, FL
  - Black
- Greensboro-High Point, NC
  - Black
  - Latino
- Greenville, NC
  - Black
- Greenville-Anderson-Mauldin, SC
  - Asian
  - Black
  - Latino
- Gulfport-Biloxi-Pascagoula, MS
  - Black
- Huntsville, AL
  - Black
- Iowa City, Iowa
  - Latino
- Jackson, MS
  - Black
- Jacksonville, FL
  - Black
  - Latino

- Killeen-Temple, Texas
  - Black
- Knoxville, TN
  - Asian
  - Black
- Lafayette, LA
  - Black
- Lansing-East Lansing, MI
  - Black
- Little Rock-North Little Rock-Conway, AR
  - Black
  - Latino
- Longview, Texas
  - Latino
- Memphis, TN-MS-AR
  - Black
- Mobile, AL
  - Black
- Montgomery, AL
  - Asian
  - Black
- Myrtle Beach-Conway-North Myrtle Beach, SC-NC
  - Black
- New Orleans-Metairie, LA
  - Black
  - Latino
- Ocala, FL
  - Black
  - Latino
- Oklahoma City, OK
  - Black
- Orlando-Kissimmee-Sanford, FL
  - Black
  - Latino
- Philadelphia
  - Black
- Racine, WI
  - Latino
- Rockford, IL
  - Latino
- Rocky Mount, NC
  - Black
- Salinas, CA
  - Latino
- San Antonio-New Braunfels, Texas
  - Black
  - Latino
- Santa Fe, NM
  - Latino
- Savannah, GA
  - Asian
  - Black
- Shreveport-Bossier City, LA
  - Black
- St. Louis, MO-IL
  - Black
- Sumter, SC
  - Black
- Tacoma-Lakewood, WA
  - Black
  - Latino
- Tallahassee, FL
  - Black

- Tulsa, OK
  - Black
  - Native American
- Tuscaloosa, AL
  - Black
- Vallejo-Fairfield, CA
  - Black
- Virginia Beach-Norfolk-Newport News, VA-NC
  - Asian
  - Black
- Waco, Texas
  - Latino
- Warner Robins, GA
  - Black
- Washington-Arlington-Alexandria, DC-VA-MD-WV
  - Asian
  - Black
  - Latino
  - Native American
- Wenatchee, WA
  - Latino
- Wichita, KS
  - Latino
- Wilmington, DE-MD-NJ
  - Asian
  - Black
  - Latino

## HMDA variables

Respondent ID: The unique identifying number for each lending institution.

Agency code: Represents the regulating agency of the lending institution that received the mortgage application.

Loan type: Represents the type of loan requested – conventional, FHA, VA or FSA/RHS.

Property type: Identifies the type of property – structures that house one- to four-family units, manufactured housing or apartments.

Loan purpose: What the loan is being used for – home purchase, home improvement or refinance.

Occupancy: Represents if the owner is going to live in the home.

Loan amount: Size of the loan that's being requested.

Pre-approval: Whether the applicant requested pre-approval.

Action type: The outcome of the application: origination, denied, withdrawn, incomplete, etc.



MSA/MD: The numeric code for the metro statistical area or metro division.

State code: FIPS code for state.

County code: FIPS code for county.

Census tract number: Census tract name.

Applicant ethnicity: The ethnicity of the applicant.

Co-applicant ethnicity: The ethnicity of the co-applicant.

Applicant race: The race of the applicant.

Co-applicant race: The race of the co-applicant.

Applicant sex: The sex of main applicant.

Co-applicant sex: The sex of the co-applicant.

Applicant Income: Yearly income of applicant(s).

Purchase type: Whether the loan was purchased or not. If so, who purchased it.

Denial reason: Explains why someone was denied a loan.

Rate spread: The difference between the APR and the average prime offer rate.

HOEPA status: Whether a loan is subject to the provisions of Home Ownership and Equity Protection Act.

Lien status: Indicates whether it's a first or subordinate lien.

Population: Total population of census tract where the property is located.

Minority population percentage: The percentage of the minority population for the census tract where the property is located.

FFIEC median family income: The median family income of the metro statistical area/metro division where the property is located.

Tract to MSA/MD median family income percentage: The percent of median family income for the census tract where the property is located compared with the median family income of the MSA/MD where the census tract is located.

Number of owner-occupied units: Number of properties, including condominiums, that are lived in by the owner.

Number of one- to four-family unit properties: Properties that are built to house fewer than five families.

## **Variables added:**

Loan-to-income ratio: Looking at the raw numbers of an applicant's income and loan amount doesn't tell the whole picture. We needed to look at how much money applicants wanted to take out in relation to their income. This provides a proxy for whether or not the loan amount was manageable compared with the applicant's income. Experts agreed that this variable should be included.

Racial and ethnicity demographic data: Historical redlined neighborhoods – those that were defined “hazardous” to lend to – often coincided with neighborhoods of color. HMDA does have a minority population percentage column, but that lumps all people of color together. But that doesn't take into consideration how specific minority groups affect lending practices. Thus, we included racial and ethnic percentages for a neighborhood. To calculate this variable, we used tract-level American Community Survey data and divided the population of each racial and ethnic group by the total population of the census tract. Experts agreed on the inclusion of this variable.