An Unequal Dream
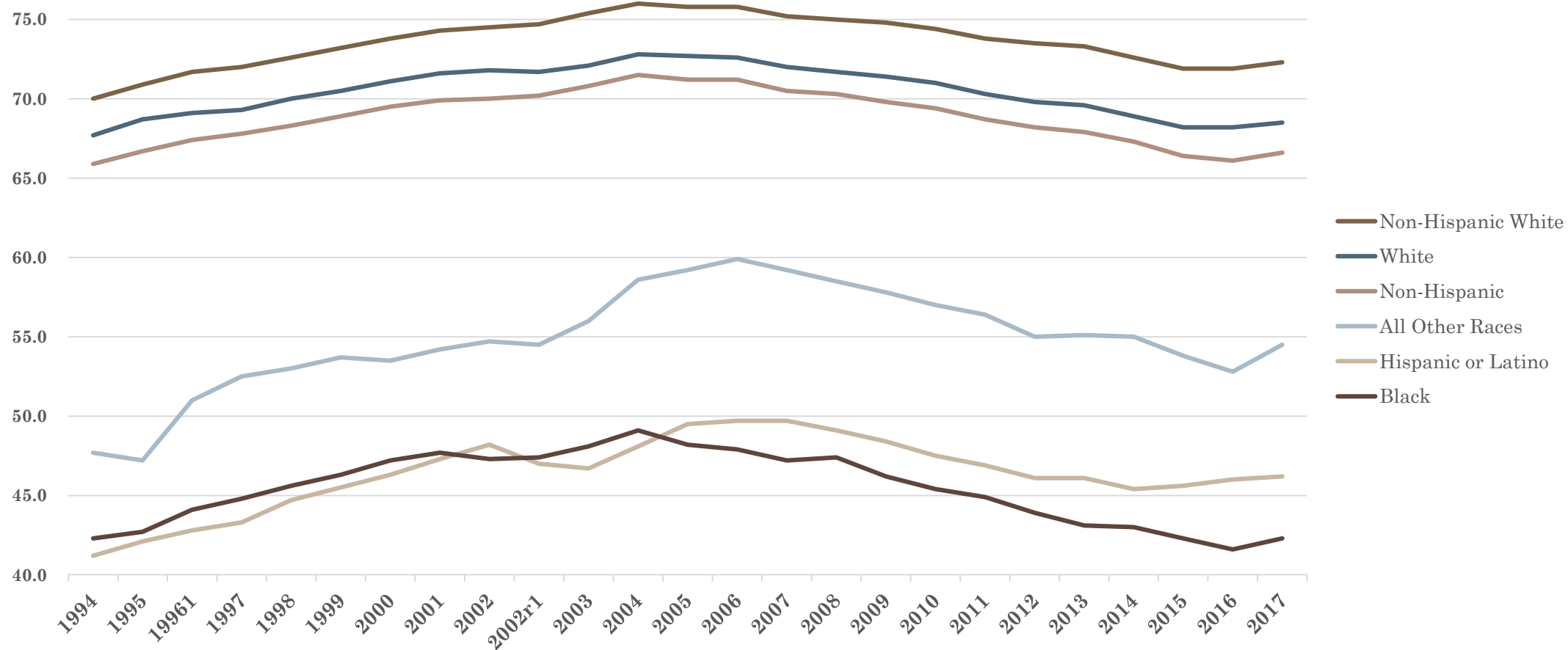
The Gap in Homeownership for White and Black Americans

By Michael Nicholson

1

The Gap in Homeownership

Homeownership by Race

Legend:
- Non-Hispanic White
- White
- Non-Hispanic
- All Other Races
- Hispanic or Latino
- Black

2018 Denial Rates

	% Denied	Number of Applications
White	8.53	4,191,319
Black	16.75	466,856
Asian	7.66	334,691
American Indian	18.93	30,246
Native Hawaiian	11.00	13,229
2+ Minority Races	14.22	9,878

Denial Reason	Other	Insurance	Incomplete	Unverified	Cash	Collateral	Employment	DTI	Credit
White	7.25	0.13	8.79	6.77	4.37	12.90	2.92	28.15	28.60
Black	6.05	0.12	6.49	6.97	3.84	8.34	2.31	29.65	36.05
Asian	8.20	0.12	13.99	8.99	6.02	11.69	3.70	36.16	11.12
American Indian	5.01	0.02	5.38	8.38	3.35	7.11	2.39	27.96	40.32
Native Hawaiian	7.97	0.34	9.27	6.39	5.29	10.16	2.61	33.59	24.38
2+ Minority Races	6.83	NA	7.33	6.98	3.13	8.33	1.92	31.32	34.16

Data

HDMA LAR Data

- Home Mortgage Disclosure Act
  - Enacted in 1975
  - Requires financial institutions to provide mortgage data to the public

- Loan Application Register
  - Loan level data released annually
  - Includes data on race, ethnicity, income, type (conventional loan, FHA loan, VA loan, etc.), amount, action taken, reason denied (if applicable), etc.

- Dodd-Frank
  - Beginning in 2018, the required characteristics in LAR data was expanded
  - These new data points include credit score, DTI, LTV, age, fees, interest rate and expanded information on commercial and business loans
    - Credit scores are stripped from publicly released versions of the data to protect the autonomy of the borrower

Excluded Data

- The HDMA data includes applications for refinancing, commercial loans, pre-approval requests, investment properties, second home, vacation properties and other non-primary home loan applications

- I focused my analysis on primary residence mortgage applications to normalize the data and to avoid any idiosyncrasies of these other loan types

- I also intend to focus on discrimination in credit availability that impacts the homeownership rate gap
  - These other loan types provide the potential for interesting analysis but do not directly contribute to the disparity in homeownership since you are a homeowner whether you own 1 or 10 homes

- A few data points contain obvious data entry errors, e.g. a loan with a maturity of "360360"
  - I removed these points from the dataset as they arose

State Level Dataset

- Using the HDMA loan level-data and a SQL database, I constructed a state-level dataset
  - This required several transformations depending on the datatype
    - E.g. converting factor and binned variables to percent
  - 87 variables for all 50 states and Washington DC
    - Variables are aggregated over white, black and total loan applicants for each state
      - Applicant demographic information: race, gender, age, income, etc.
      - MSA and tract-level demographics: income, population, percent minority
      - Purchaser: Fannie Mae, Ginnie Mae, Private securitizer, etc.
      - Denial rates and reason for denial
      - Mortgage Statistics: DTI, loan amount, property value, etc.
  - All data is from 2018
    - The HDMA data also includes information on the Virgin Islands, Guam and Puerto Rico, but these were outliers for several fields so I removed them to remove idiosyncrasies

County Level Data

- In order to conduct analysis, I re-aggregated the loan-level data at the county level
  - All mortgage and applicant characteristics from the state level are available in this dataset

- Additionally, I added demographic variables from the 2017 American Community Census Survey and FRED datasets

- The dataset now includes county-level demographic variables including:
  - Poverty statistics
  - Income, employment and education data
  - Violent crime statistics
  - Proportion of county population with subprime credit
    - Equifax scores under 650
  - Homeownership rates and home price indices
  - Inequality and segregation measures
  - Etc.

State Level Analysis

African American Denials

Percent Denied
0.34
0.00

Powered by Bing
© GeoNames
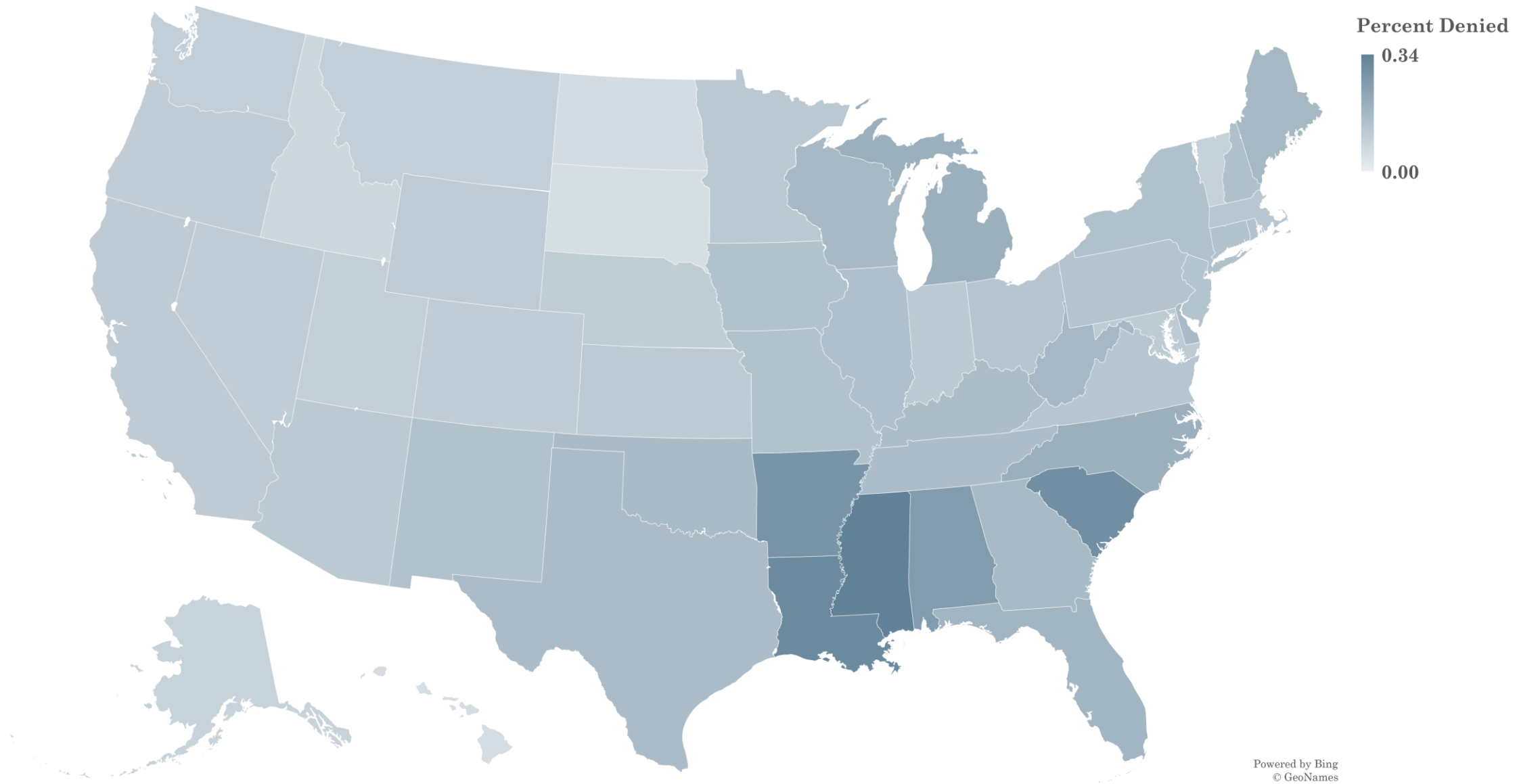
Black Denials

|	Beta	Std. Error	t-stat	p-value	
(Intercept)	1.43E+00	1.52E-01	9.415	1.77E-11	***
Average Income Black	5.35E-05	6.35E-05	0.843	0.40461	
Average Income White	-5.64E-05	6.01E-05	-0.939	0.35341	
Percent Denied White	4.95E-01	1.65E-01	3.006	0.00467	**
Hoepa Status Black	2.52E+00	3.60E+00	0.701	0.48773	
Total Loan Costs Black	2.43E-06	3.18E-06	0.764	0.44974	
Loan Term Black	-3.51E-03	4.22E-04	-8.323	4.32E-10	***
Percent Female Black	1.88E-01	4.06E-02	4.637	4.10E-05	***
Ginnie Purchased Black	5.67E-02	7.14E-02	0.794	0.43222	
Private Purchaser Black	-3.77E-01	2.72E-01	-1.387	0.17367	
Low Debt to Income Black	-2.92E-01	1.55E-01	-1.884	0.06718	.
Median MSA Income Black	-8.28E-07	3.72E-07	-2.226	0.03199	*
Tract Income as Percent of MSA	-1.01E-03	3.96E-04	-2.549	0.01495	*

Signifiance codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1

Multiple R-squared: 0.9251
Adjusted R-squared: 0.9014

State Residuals

- With an R^2 of .925, it appears the included mortgage and applicant demographic variables explain the majority of the variation of black denial rates between states
  - This would indicate there are not individual states that are significantly more discriminatory than others

- The graphed residuals show how individual states' denial rates compare to other states' after accounting for the included demographical information

- At first glance, it would seem Montana, Maine and New Hampshire have higher denial rates than their demographics would predict
  - Looking deeper, all three of these states have extremely low black populations
  - It is likely a simple lack of black applicants separating them from other states
    - Their raw black denial rates are also centered around the national average

State Residuals

residuals

0.04980501

−0.063334495
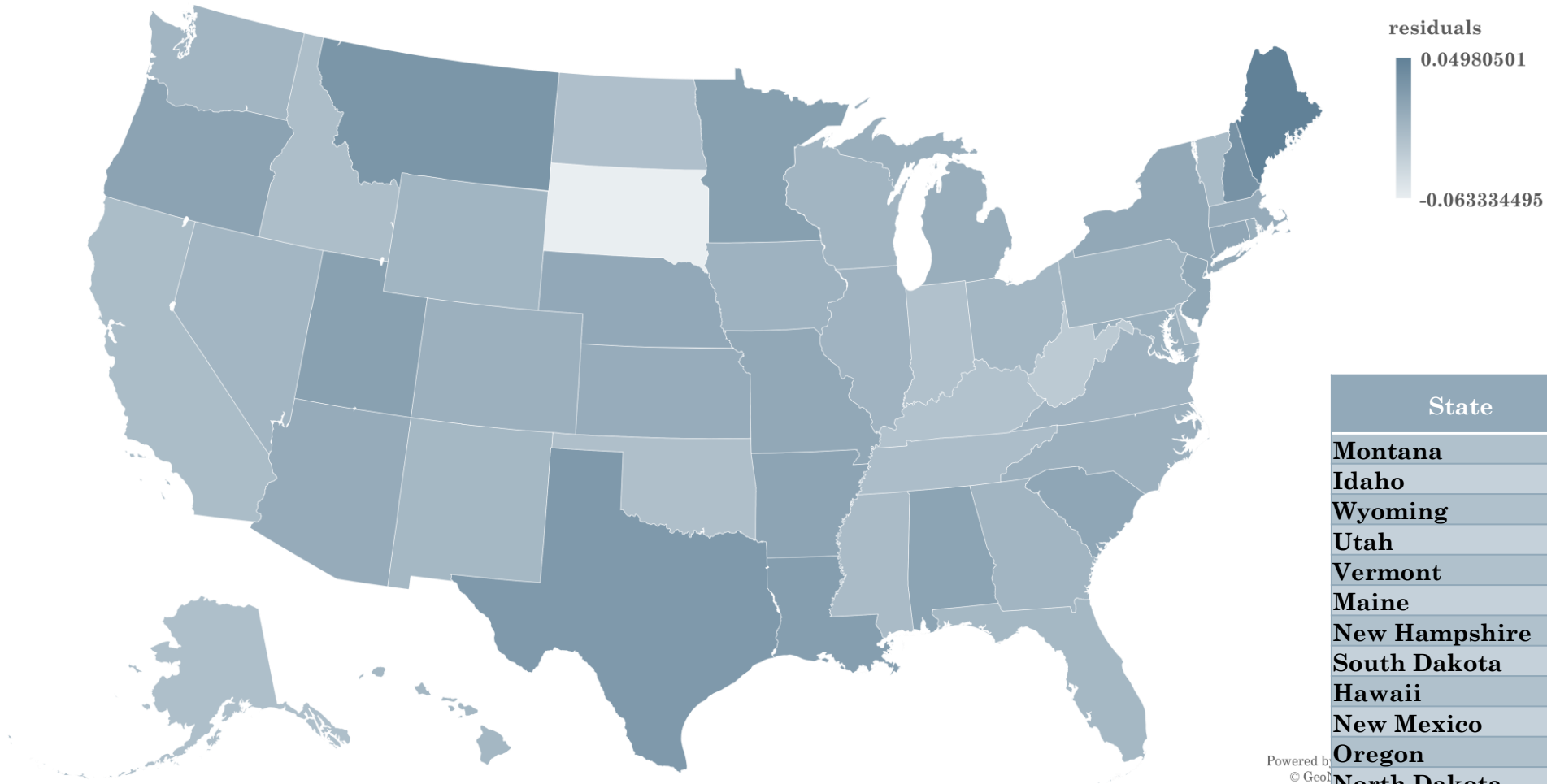
State	Black Population
Montana	0.40%
Idaho	0.64%
Wyoming	0.99%
Utah	1.06%
Vermont	1.21%
Maine	1.21%
New Hampshire	1.21%
South Dakota	1.67%
Hawaii	1.69%
New Mexico	1.79%
Oregon	1.80%
North Dakota	2.28%

Powered by
© Geo

Black Female Applicants

- One of the most significant variables in the regression was the share of black applicants that were female in each state

- Analyzing the national share of white and black applicants that are female, it becomes apparent that there is a divergence between races
  - Black applicants are 13 percentage points more likely to be female
  - Historically, female applicants have had lower loan acceptance rates than male for all races and ethnicities

- Together, these factors appear to be driving a significant portion of the variance in state denial rates for black applicants

Black Female Applicants

percent_female_black

0.509814613

0.149171271
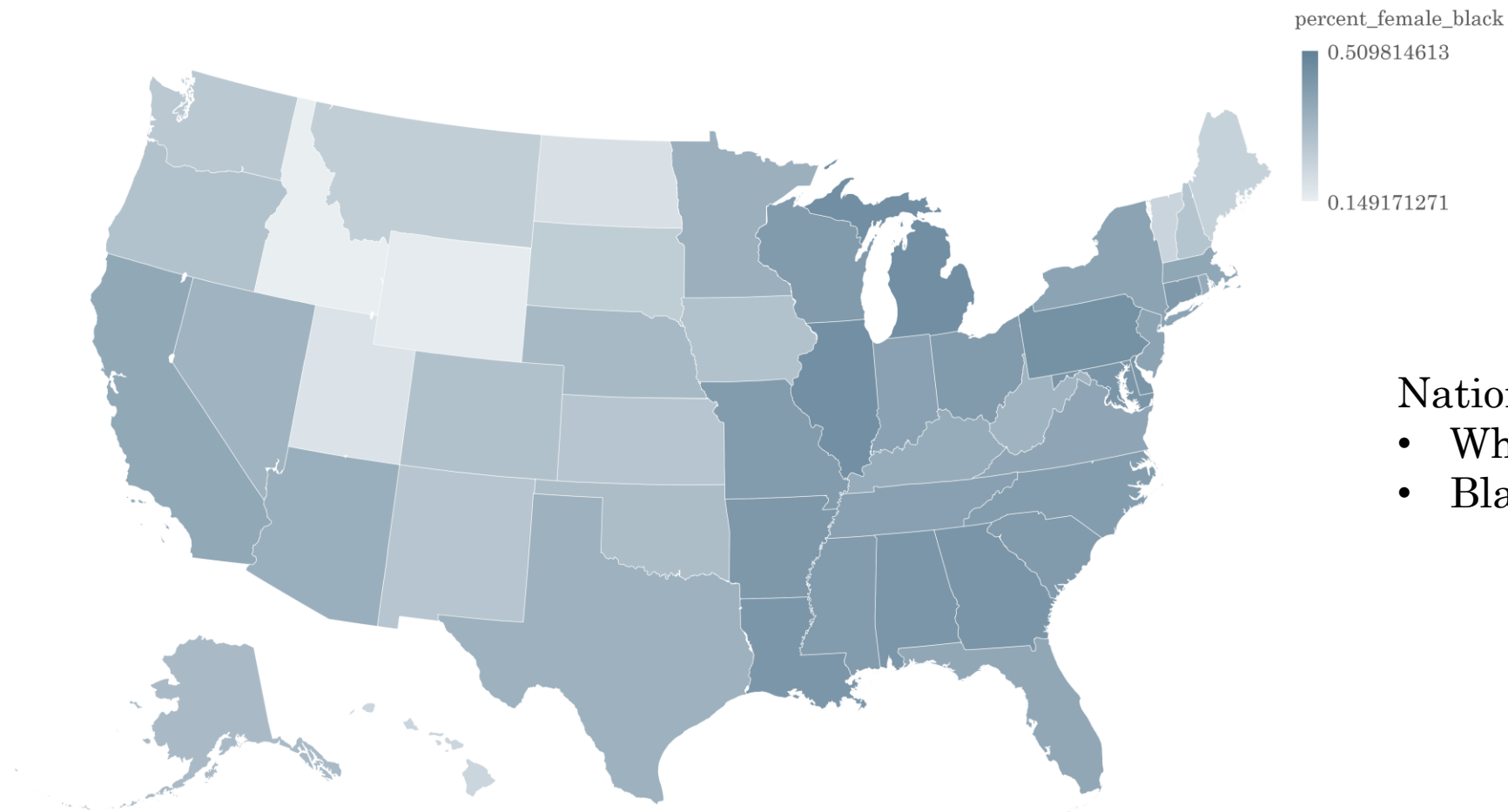
National Average:
- White: 22.5%
- Black: 35.4%

Powered by Bing
© GeoNames

Loan Term

- The other surprising outcome of the regression analysis was the significance of loan term
  - Loan term was the single most significant variable in the model

- Looking at the state map, it also appears to have an obvious negative correlation with black denial rates

- What would cause this relationship?
  - Several factors could be at play:
    - It is possible that originators purposely put more black applicants in shorter-term mortgages
    - Weaker applicants may systematically take out shorter mortgages, either at the behest of the originator or to increase the odds of approval
    - Certain low income or minority programs may utilize shorter-term mortgages
  - This will be analyzed further on the county level dataset

Loan Term

loan_term_black

360

307.8987069

Total Denials

- In contrast to black denials, a study on the effect of race on total denials allows for a direct measure of the effect of race

- Including applicant demographic variables and mortgage characteristics, race is not significant at the state level
  - This would indicate that adjusted for applicant quality and loan type, there is not discrimination on the basis of race at the state level
  - Though there may be discrimination on the basis of gender

	Beta	Std. Error	t-stat	p-value	
(Intercept)	1.58E+00	1.73E-01	9.108	3.37E-11	***
Average Income	-3.75E-06	1.73E-05	-0.217	0.829675	
Hoepa Status	3.33E-01	5.13E+00	0.065	0.948626	
Loan Costs	6.76E-06	2.39E-06	2.833	0.007264	**
Loan Term	-4.23E-03	4.90E-04	-8.618	1.45E-10	***
Percent Female	4.98E-01	1.14E-01	4.377	0.0000874	***
Ginnie Purchased	1.22E-02	5.28E-02	0.231	8.18E-01	
Private Purchased	2.72E-01	4.44E-01	0.612	5.44E-01	
Low Debt to Income	-2.47E-01	9.25E-02	-2.665	0.01113	*
MSA Median Income	-7.36E-07	1.95E-07	-3.769	0.000543	***
Tract Income as Percent of MSA	-6.53E-04	3.31E-04	-1.974	0.055515	.
Percent Black	-4.68E-02	5.19E-02	-0.902	0.37266	

Signifiance codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1

Multiple R-squared: 0.9081
Adjusted R-squared: 0.8822

County Level Analysis

Denials by County

Total Denials

- In contrast to the state-level analysis, race is significant
  - This is after adjustments for income, aggregate credit measures and other mortgage and demographic variables
  - This would indicate that there may be discrimination at the county/local level that is averaged out in more aggregated measures
- As in the state-level analysis, loan term is still among the most significant variables
- Purchased loans have a lower denial rate indicating the GSEs have a positive impact on credit availability

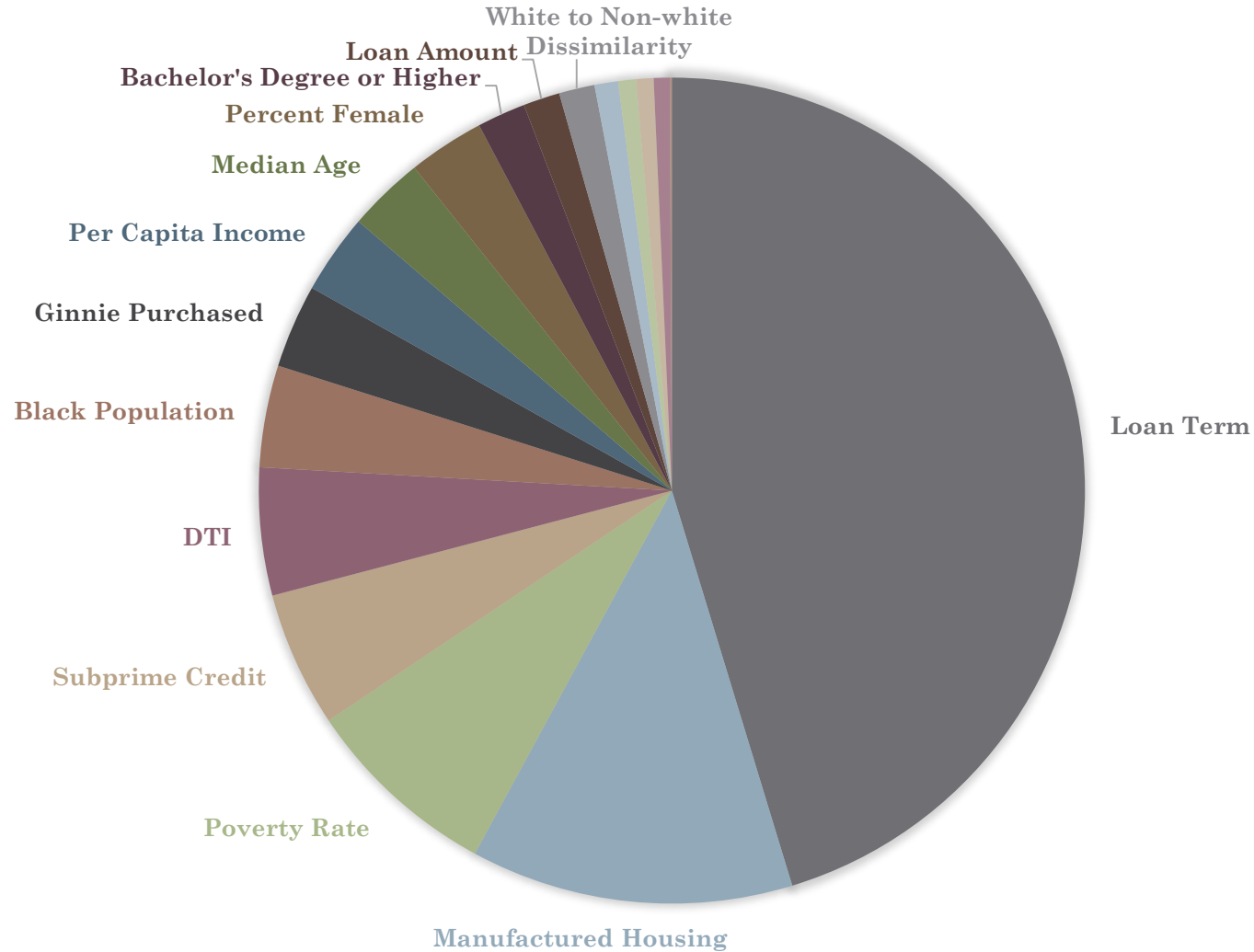|	Beta	Std. Error	t-stat	p-value	
(Intercept)	4.18E-01	3.01E-02	13.898	< 2E-16	***
Average Income	-8.86E-06	1.95E-06	-4.548	5.7E-06	***
Loan Amount	-9.32E-08	1.83E-08	-5.098	3.7E-07	***
Loan Term	-9.51E-04	8.61E-05	-11.04	< 2E-16	***
Percent Female	7.13E-02	1.62E-02	4.404	1.1E-05	***
Rate Spread	4.46E-03	2.93E-03	1.524	0.12753	
Origination Charges	-4.37E-06	2.31E-06	-1.893	0.05851	.
Fannie Purchased	-1.46E-01	1.55E-02	-9.445	< 2E-16	***
Freddie Purchased	-1.12E-01	1.99E-02	-5.65	1.8E-08	***
Ginnie Purchased	-1.73E-01	1.42E-02	-12.19	< 2e-16	***
DTI	1.01E-03	2.93E-04	3.429	0.00061	***
Per Capita Income	6.16E-07	1.06E-07	5.807	7.1E-09	***
Manufactured Housing	0.2128	0.01075	19.784	< 2E-16	***
Bachelor's Degree or Higher	-0.00109	0.000139	-7.814	7.8E-15	***
Poverty Rate	0.002002	0.000189	10.596	< 2e-16	***
White to Non-white Dissimilarity	-2.20E-04	5.44E-05	-4.039	5.5E-05	***
Black Population	6.87E-02	9.56E-03	7.186	8.5E-13	***
Median Age	4.92E-04	1.63E-04	3.024	0.00251	**
Subprime Credit	-7.63E-04	1.63E-04	-4.679	3E-06	***

Signifiance codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1

Multiple R-squared: 0.8487,   Adjusted R-squared: 0.8477

Contribution of Variables

- Since many of the variables have different magnitudes and interpretations, it be can be difficult to determine their relative significance looking solely at the regression output

- To account for this I graphed the magnitude of the effect of each variable for two sample counties: Durham and New York

- From this output, it is obvious that loan term is particularly impactful on loan denial rates
    - I also graph effects for regressions without loan term to better compare the other factors

- The significant factors differ greatly between the two counties
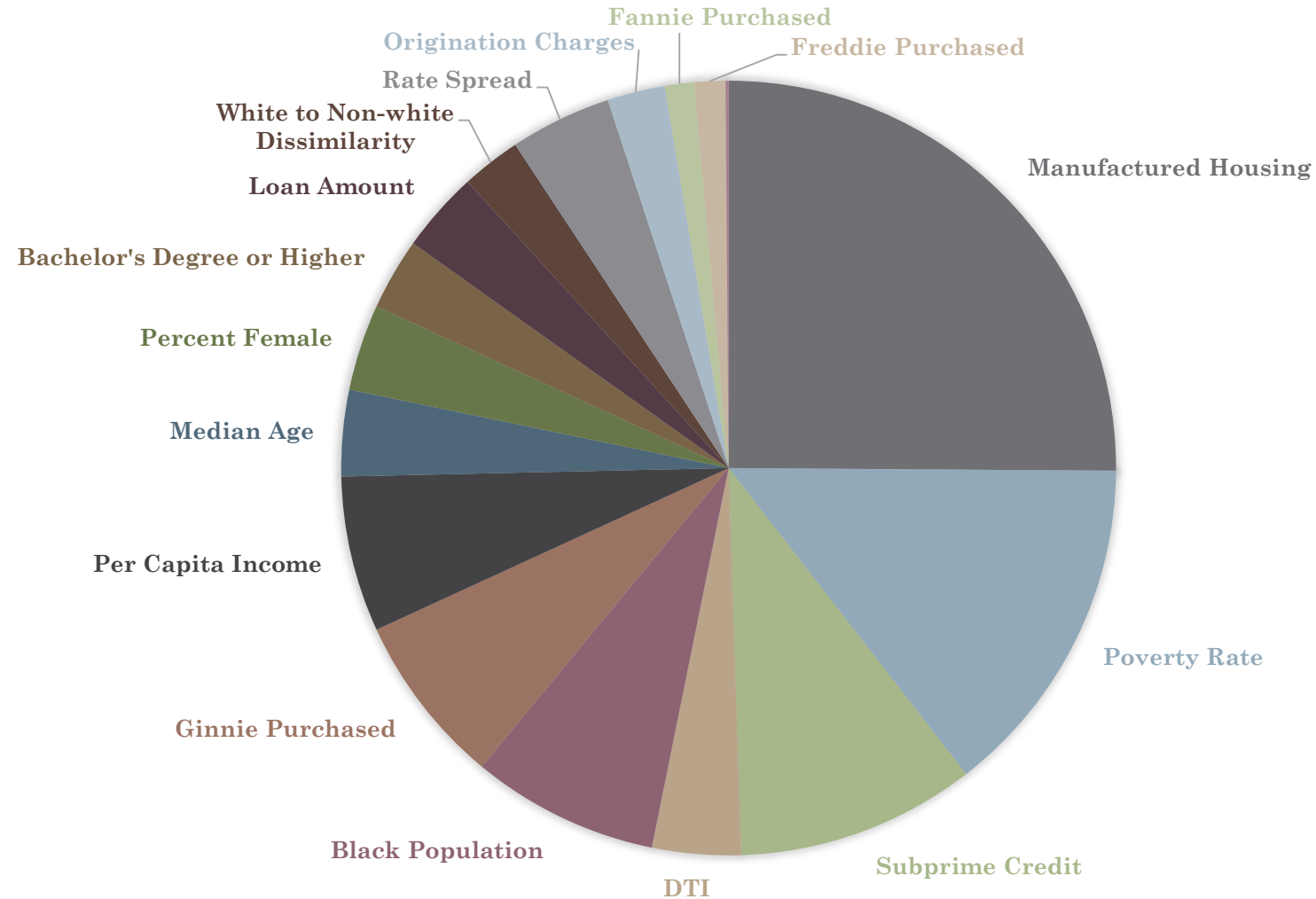   - For example, loan amount contributes much more to New York than Durham county
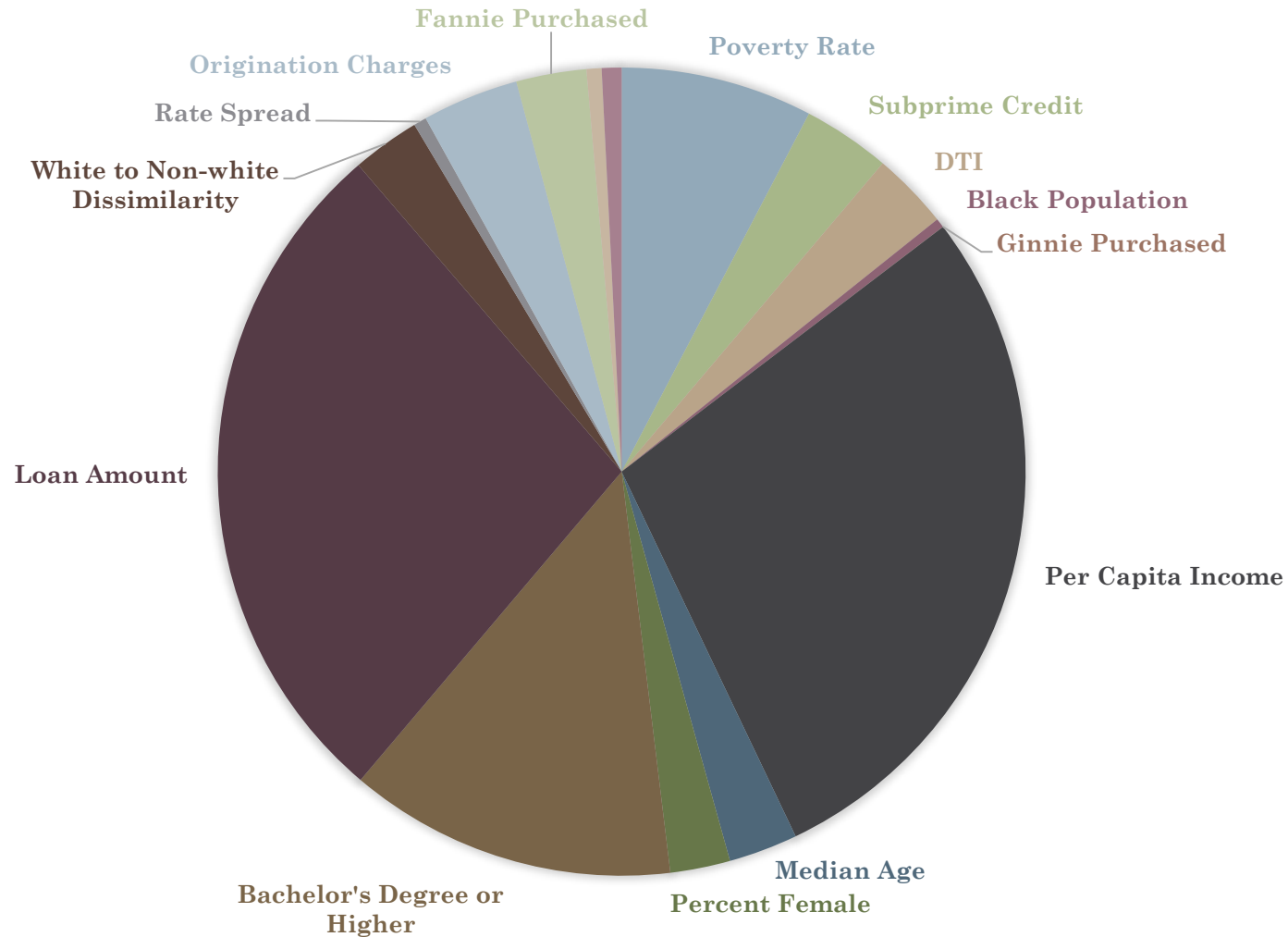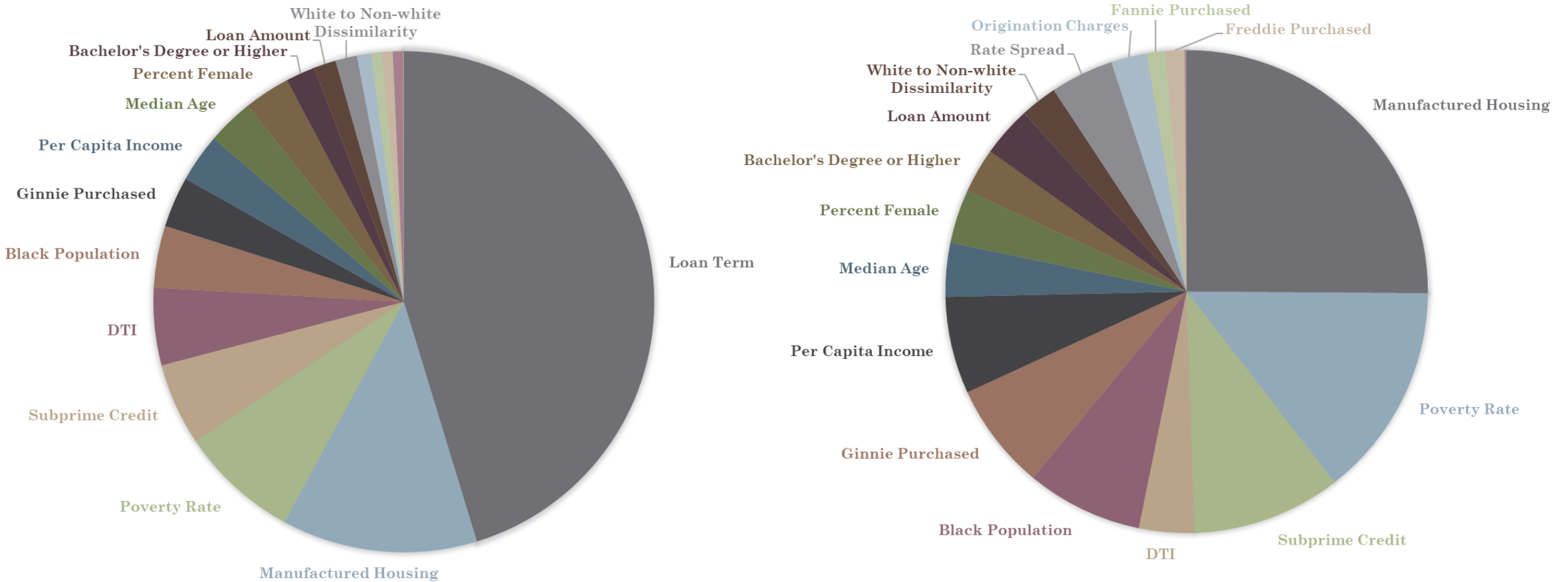
Durham

New York
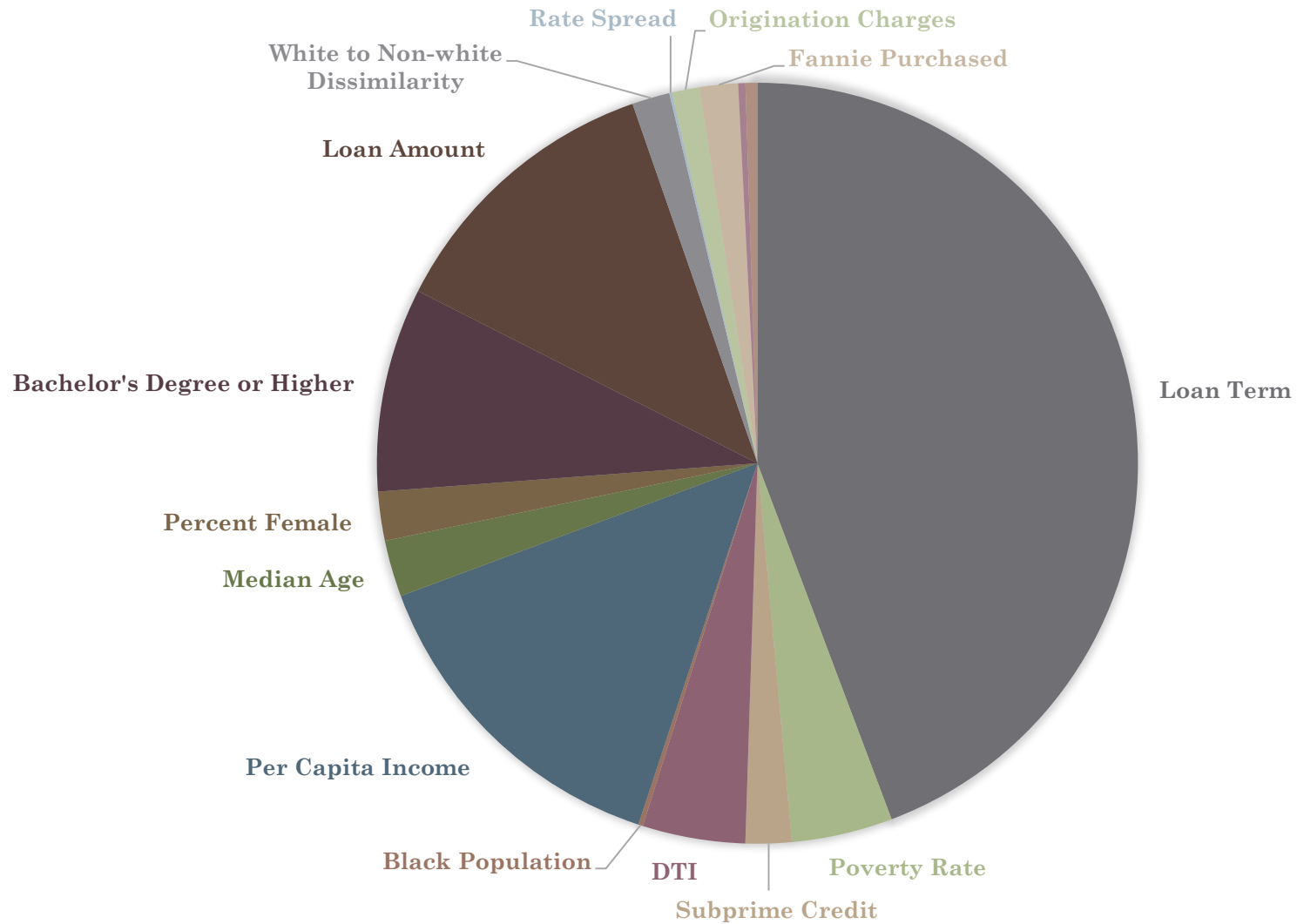
Durham Ex-Loan Term

New York Ex-Loan Term

Durham Comparison

County Ex-Loan Term

	Estimate	Std. Error	t-stat	p-value	
(Intercept)	1.33E-01	1.58E-02	8.436	< 2E-16	***
Average Income	-8.83E-06	1.99E-06	-4.442	9.24E-06	***
Loan Amount	-1.33E-07	1.83E-08	-7.284	4.19E-13	***
Percent Female	5.21E-02	1.64E-02	3.171	0.00154	**
Rate Spread	1.24E-02	2.90E-03	4.283	1.90E-05	***
Origination Charges	-9.16E-06	2.32E-06	-3.955	7.84E-05	***
Fannie Purchased	-1.60E-01	1.58E-02	-10.151	< 2E-16	***
Freddie Purchased	-1.40E-01	2.01E-02	-6.944	4.72E-12	***
Ginnie Purchased	-2.29E-01	1.35E-02	-16.933	< 2E-16	***
DTI	4.51E-04	2.95E-04	1.529	0.12632	
Per Capita Income	7.75E-07	1.07E-07	7.218	6.76E-13	***
Manufactured Housing	2.57E-01	1.02E-02	25.135	< 2E-16	***
Bachelor's Degree or Higher	-1.04E-03	1.42E-04	-7.346	2.67E-13	***
Poverty Rate	2.27E-03	1.91E-04	11.865	< 2E-16	***
White to Non-white Dissimilarity	-2.37E-04	5.55E-05	-4.270	2.02E-05	***
Black Population	8.16E-02	9.69E-03	8.421	< 2E-16	***
Median Age	3.60E-04	1.66E-04	2.168	3.03E-02	*
Subprime Credit	-8.70E-04	1.66E-04	-5.235	1.77E-07	***

Signifiance codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1
Multiple R-squared: 0.8421
Adjusted R-squared: 0.8412

Loan Term

- Without adjustment, loan term is negatively correlated with interest rate

  - This should not be the case considering longer-term mortgages have higher duration

- The factor driving this, and also the significance of loan term on denials, appears to be shorter-term mortgages are more likely to have higher rate spreads and to be from less financially secure counties

- Race is also significant indicating Black Americans are either choosing or are being pushed into shorter, higher-cost mortgages

	Estimate	Std. Error	t-stat	p-value	
(Intercept)	3.00E+02	3.41E+00	87.961	< 2E-16	***
Average Income	-2.30E-05	4.27E-04	-0.054	0.9571	
Loan Amount	4.21E-05	3.93E-06	10.711	< 2E-16	***
Percent Female	2.02E+01	3.52E+00	5.719	1.18E-08	***
Rate Spread	-8.42E+00	6.25E-01	-13.477	< 2E-16	***
Origination Charges	5.04E-03	4.97E-04	10.14	< 2E-16	***
Interest Rate	8.46E-02	9.25E-02	0.915	0.3605	
Fannie Purchased	1.44E+01	3.39E+00	4.245	2.25E-05	***
Freddie Purchased	2.89E+01	4.32E+00	6.696	2.58E-11	***
Ginnie Purchased	5.89E+01	2.90E+00	20.319	< 2E-16	***
DTI	5.83E-01	6.33E-02	9.21	< 2E-16	***
Per Capita Income	-1.66E-04	2.30E-05	-7.228	6.29E-13	***
Manufactured Housing	-4.61E+01	2.19E+00	-21.067	< 2E-16	***
Bachelor's Degree or Higher	-4.58E-02	3.04E-02	-1.505	0.1323	
Poverty Rate	-2.82E-01	4.11E-02	-6.864	8.18E-12	***
White to Non-white Dissimilarity	1.83E-02	1.19E-02	1.536	0.1246	
Black Population	-1.36E+01	2.08E+00	-6.527	7.91E-11	***
Median Age	1.39E-01	3.56E-02	3.919	9.11E-05	***
Subprime Credit	1.13E-01	3.56E-02	3.159	0.0016	**

Signifiance codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1
Multiple R-squared: 0.8302
Adjusted R-squared: 0.8291

Conclusions

- At the state level, the proportion of black applicants is not significant on overall denial rates after adjusting for applicant demographics and mortgage type
  - This indicates that there is not significant discrimination in mortgage availability on the basis of race at the state level
  - However, this does not provide evidence there is not discrimination at the local level

- At the county-level, the proportion of black applicants is significant
  - This holds after adjusting for applicant demographics including LTV, DTI and income as well as county level characteristics including poverty statistics, home values, subprime credit, etc.
  - Counties with higher denial rates are also associated with lower loan terms which are often loans with larger rate spreads

- Loan term drives a large share of the variation in denial rates between counties
  - This significance appears to be driven by the type of borrower that takes out these mortgages and the loan characteristics associated with non-30 year mortgages
  - This could indicate discrimination not only in credit availability but in mortgage types given to vulnerable and/or minority applicants
    - I intend to pursue this line of research more thoroughly for my thesis