Color and Credit: Race, Regulation, and the Quality of Financial Services*

Taylor A. Begley†
Amiyatosh Purnanandam‡

First Draft: March 15, 2017
This Draft: May 19, 2017

Abstract

The incidence of mis-selling, fraud, and poor customer service by retail banks is significantly higher in markets with lower income and educational attainment. Further, areas with a higher share of minority population experience significantly worse outcomes even after controlling for factors such as income, education, and house price changes. Regulations aimed at improving access to credit to such areas are partly responsible for these findings. Specifically, low-to-moderate-income (LMI) areas targeted by the Community Reinvestment Act have significantly worse outcomes and this effect is magnified further for LMI areas with high-minority population. The results highlight an unintended adverse consequence of such *quantity*-focused regulations on the *quality* of credit to poor and minority customers.

Keywords: discrimination, product quality, financial sophistication, consumer protection, regulation

JEL Classification: G21, G28, L13, L14

---

*We are grateful to Scott Frame, Todd Gormley, Mark Leary, Uday Rajan, and seminar participants at Princeton and Washington University in St. Louis for helpful comments on the paper.

†Olin Business School, Washington University in St. Louis and London Business School; email: `tbegley@wustl.edu`.

‡Ross School of Business, University of Michigan; email: `amiyatos@umich.edu`.

1 Introduction

Policymakers have for long been concerned about unequal and unfair treatment of minority and poor customers by large financial institutions. The differential treatment manifests itself in several forms: excessive denial of credit, poor customer service, outright fraud, and mis-selling of financial products to name a few. Motivated by these concerns, policy-makers have enacted several consumer protection regulations to address discrimination and, in particular, promote equal access to credit in poor and minority neighborhoods.[1] The Department of Justice (DoJ) often enforces provisions of these legislations to protect minorities against discriminatory lending practices.[2] More generally, allegations of large-scale fraud in the mortgage market during the early parts of this century and anecdotal evidence of fraudulent banking practices by Wells Fargo Company during the early 2010s have made consumer protection concerns even more salient in recent years (see, e.g., Gurun, Matvos, and Seru, 2016; Griffin and Maturana, 2016).

A number of papers such as the famous Boston-Fed study on loan denials to minorities by Munnell, Tootell, Browne, and McEneaney (1996) have looked at the quantity or pricing of financial services provided to minority and poor customers. However, little is known about the other dimension of unequal and unfair treatment, namely, the *quality* of financial products and services received by these consumers. Our paper takes a first step in this direction by examining the incidence of fraud, mis-selling, and poor customer service – our measure of *quality* – in the consumer credit market. Piskorski, Seru, and Witkin (2015) and Griffin and Maturana (2016) document compelling evidence of fraud and mis-selling by banks to

[1] Some prominent examples include the Fair Housing Act, the Equal Credit Opportunity Act, and the Community Reinvestment Act.

[2] The DoJ website provides a number of examples of exploitation in the mortgage market that we focus on in this paper. Based on their recent case, DoJ states that "On January 20, 2017, the court entered a consent order in United States v. JPMorgan Chase Bank (S.D.N.Y.). The complaint, which was filed on January 18, 2017 by the United States Attorney's Office, alleged that the defendant violated the Fair Housing Act and the Equal Credit Opportunity Act when African American and Hispanic borrowers paid higher rates and fees for wholesale mortgage loans than similarly situated white borrowers. The consent order provides monetary relief of $53 million, including a civil penalty of $55,000." `https://www.justice.gov/crt/recent-accomplishments-housing-and-civil-enforcement-section`

their investors in the securitization markets. Our work, on the other hand, focuses on banks' customers in the retail markets who directly bear the costs of aggressive behavior.

We obtain a measure of the quality of financial services from a newly available dataset from the Consumer Financial Protection Bureau (CFPB). In 2010, the Dodd-Frank Act established the CFPB as a watchdog of the financial services industry. Dissatisfied customers can send their complaints against financial institutions to the CFPB using the bureau's online system, email, postal mail, fax, phone, or through referral from other agencies. By the very nature of this process, these are not typically minor complaints that are easily resolved between the customer and the financial institution. Rather, they range from a customer's allegation of serious failing in customer service to claims of egregious exploitative behavior by the financial institution. The incidence of consumer complaints against financial institutions for mortgage-related products in the CFPB dataset is our measure of the quality of financial services.[3] For many consumers, acquiring and choosing a home mortgage product involves difficult choices between various complex products. These transactions leave many potential borrowers at a substantial information disadvantage compared to sophisticated financial institutions (Campbell, 2006). Prominent examples of the nature of complaints include allegations of hidden or excessive fees, unilateral changes in contract terms after the purchase, aggressive debt collection tactics, and unsatisfactory resolution of mortgage servicing issues. Our data are from 2012-2016 and include over 175,000 mortgage-related complaints from 16,309 unique zip codes. All of our key empirical exercises are based on geographical variations across zip codes.

We find that there are substantially more complaints in zip codes with lower average income, lower educational attainment, and higher minority shares of population. While each of these characteristics are associated with more complaints in multivariate regressions, the

[3]The database also has complaints about other financial products (e.g., checking accounts and student loans). We focus on mortgage products because this is the category with most complaints in the dataset, and this is the product category that is economically large, involving many millions of homeowners and many trillions of dollars. Moreover, a home mortgage is often the single largest financial transaction for many households.

effect of high-minority status is two to three times larger than the effect of low income or low education. We include fixed effects at the three-digit zip code levels in these regressions, indicating that our results are unlikely to be driven by differences in local economic conditions or other region-specific variation. Thus, the minority effect that we document cannot be simply explained away by the lower income of these borrowers, their poor educational attainment, or differences in broad economic situations. A one-standard deviation increase in minority population leads to a 16% increase in complaints; the corresponding effect for area income is 2% and for area education is 6% . We also show that the effect of minority population on complaints increase non-linearly, with a sharp rise in the effect for zip codes with greater than 80% minority population.

We conduct a number of additional tests in the paper to establish the statistical and economic importance of our results. Since all of our main tests control for *three-digit zip code* fixed effects, our results cannot be explained by differences in economic conditions at this level. Yet, one may be concerned about differences in economic conditions at the five-digit zip code level within a three-digit zip code, particularly potential differences in house price appreciation. The concern is that consumers complain more when they experience a loss in their home value and are more likely to be underwater on their loan. If minority zip codes had disproportionately larger price declines and the complaints are driven only by these price drops, then our results could be spurious. We directly address for this concern by including two variables in the regression model: (a) zip-code-level house price changes in the past five years, and (b) the zip-code-level foreclosure rate during the sample period. Zip codes with larger declines in house prices and higher foreclosure rates do indeed have more complaints, but these factors appear to be orthogonal to our main effect as the relationship between minority population share and complaints do not change in any meaningful ways.

Another concern with our analysis may be geographical heterogeneity in the costs and benefits of complaining. If consumers in minority zip codes have lower cost of complaining or perceive a higher marginal benefit of complaining, higher complaint frequency would occur irrespective of banks' behavior. Note that these differences in net benefits would have to go beyond what can be explained by local income and educational attainment. To separate out the baseline "complainer" effect from our analysis, we include the number of complaints to a different government agency, the Federal Communications Commissions (FCC), in the given zip code in our regression model. Our results remain similar. Lastly, if a complaint is filed against a specialized mortgage servicer, the source of the discontent could be driven by the behavior of either the originating bank or the servicer itself. To make sure unscrupulous servicers are not driving the results, we drop companies whose primary business is loan servicing (e.g., Ocwen) from the tests and again find that the results are virtually unaffected.

What could be driving the robust relationship between minority neighborhoods and complaints? In a frictionless world characterized by fully informed consumers and no distortions in the supply of banking services, there should be no systematic differences in the incidence of fraud, mis-selling, and poor customer service across areas based on racial composition. However, retail financial markets are filled with information frictions: consumers are often at a significant information disadvantage compared to large banks, they often face large search costs, and there is limited scope of learning from past experiences since mortgage decisions are relatively infrequent. The association between lower income and educational attainment, and consumer complaints is consistent with these friction. Why, however, should minority consumers face even worse outcomes even after controlling for these influences?

There are two possible channels that can explain our results. The first one is purely a credit-demand side (i.e., consumer-driven) explanation: poor and minority customers are more likely to complain due to unobserved reasons beyond those directly examined in the analysis, and thus our findings have nothing to do with credit-supply-side (i.e., lender-driven) forces. Said differently, lenders are providing similar quality to all communities, but minority customers end up complaining more about fraud, mis-selling and poor service for reasons unobserved by the econometrician.

The other possible channel is related to the supply side of these services.[4] Specifically, it may be that lenders behave differently when they deal with poor and minority customers such as putting in less less effort in explaining the costs and benefits of various products or even actively engaging in predatory lending aimed these communities. For example, on May 15, 2017, the city of Philadelphia sued Wells Fargo for violations of the Fair Housing Act, alleging that since 2004 its employees were encouraged to push African-American and Hispanic borrowers toward riskier loans despite having credit scores that would warrant better loans for the borrowers.[5] Such behavior is consistent with findings in Gurun et al. (2016), who show that mortgage lenders were actively targeting minority consumers with misleading advertising to originate more high-priced mortgages. In a congressional testimony, Ginny Hamilton of the Fair Housing Center of Greater Boston also provides some compelling anecdotal evidence in support of unequal treatment of minority consumers by banks (U.S. House of Representatives, 2007) . Her organization performed "mystery shopper" tests in the mortgage market during 2005-2006 in Boston. They found that minority borrowers were treated systematically less favorably than white borrowers, even though the minority mortgage shoppers, by the design of the experiment, had better credit profiles. To further investigate the presence of a supply-side effect, we exploit an institutional feature of this market that allows us to examine differences in *quality* when there is regulatory pressure to increase the *quantity* of credit.

In light of concerns about discrimination in lending markets, a number of regulations have been enacted in the U.S. over the years to provide better access to credit to poor and minority consumers. These regulations make it illegal for lenders to discriminate against

[4]Some anecdotal evidence of the supply side driver of this behavior can be found in the recent allegations of aggressive sales tactics by the Wells Fargo Bank in which at least 2 to 3.5 million allegedly unauthorized accounts were opened from 2002-2017.

[5]City of Philadelphia v Wells Fargo & Co et al, U.S. District Court, Eastern District of Pennsylvania, No. 17-02203. Baltimore, Los Angeles, Memphis, Miami, and Oakland have filed similar lawsuits, though their ability to do so was legally in question. On May 1, 2017, the Supreme Court ruled that the city of Miami can sue Bank of America and Wells Fargo for "discriminatory conduct that led to a disproportionate number of foreclosures and vacancies in majority-minority neighborhoods, which diminished the city's property-tax revenue and increased the demand for police, fire, and other municipal services."

historically disadvantaged groups. Prominent examples of these regulations include the Fair Housing Act (FHA) of 1968, Equal Credit Opportunity Act (ECOA) of 1974, Home Mortgage Disclosure Act (HMDA) of 1975, and the Community Reinvestment Act (CRA) of 1977. The main focus of many of these laws is on ensuring access to credit products at fair lending rates. That is, these regulations focus on the *quantity* and *price* of credit. Subsequent empirical studies examine impact of these regulations on consumers focus exclusively on these more-easily measured metrics and whether they are equal for poor and minority customers and neighborhoods (see Ladd, 1998, for a literature review). The potential effects of such regulations on product *quality* are ambiguous. Regulations may improve the quality of financial services received by poor and minority borrowers if lenders are concerned about close monitoring of quality by regulators. Alternatively, regulation that focuses too much on ensuring the quantity of credit can provide incentives to dilute quality. For example, in a multi-tasking framework, Holmstrom and Milgrom (1991) underscore the importance of a dilution in the quality of output when agents are rewarded for the quantity of output. In our setting, quantity-based goals may unintentionally encourage banks to, for example, engage in aggressive sales tactics or make loans to uninformed borrowers without proper disclosure as they seek to satisfy their regulatory quantity requirements.

We focus on CRA regulations because of an attractive empirical feature that allows us to identify the effect of regulation-driven supply-side shock on quality. The CRA designates certain census tracts as “underserved” if the average income of the area is less that 80% of the median income of the Metropolitan Statistical Area (MSA) that the area belongs to. We refer these areas as *CRA target areas*. Banks that fail to lend enough to these target areas face repercussions including denial of M&A deals and branch opening applications, potential sanctions, and reputational harm. These consequence provide pressure for lenders to increase lending to poor and minority customers beyond what they normally would have. In terms of empirical design, the designation of an area into a target area has the attractive feature that it depends on the area’s *relative* income within an MSA, and not on its *absolute* level of income. Regional variation in MSA income provides several areas that are similar in terms many key characteristics including absolute income, but have different CRA designation. We compare the number of complaints about fraud, mis-selling, and poor service across the CRA target areas (treatment group) and their observationally similar non-target areas (control group) and find substantially more complaints in target areas. Specifically, the CRA target areas have 32% more complaints relative to the control areas. Thus, holding fixed many demographic characteristics of the zip codes, CRA-targeted areas have relatively poorer quality. We interpret this finding as an unintended consequence of quantity based regulation.

We further develop the matched sample test discussed above to tease out the supply side channel behind our main result linking minority zip codes to quality by examining the difference in complaints across CRA target and control groups separately for below- and above-median minority share areas. The difference in complaints between target and control areas come predominantly from zip codes with above-median minority population. Within neighborhoods with below-median minority share of population, complaints rate are higher by 5-10% (depending on the matching criteria) for the target areas as compared to observationally similar control areas. This difference increases to 40-50% within neighborhoods with above-median minority population. The difference in these two differences is economically large and statistically significant and suggests that the unintended consequences of the regulation are particularly severe for minority areas. These results are hard to explain by a demand-side (i.e., customer driven) difference across neighborhoods since the target and control areas are very similar on demographic characteristics such as income, education, population, and mortgage volume. Banks face pressure to increase quantity of lending in every area, but in high-minority areas, they effectively have two "boxes to check" for regulatory compliance – lending to poor and lending to minority customers. All of our results remain similar for a wide range of matching techniques and criteria.

Our paper provides important inputs to policy debates on equality in the lending market and strands of literature including the consequences of bank regulation, the economics of household finance, and the role of consumer sophistication on financial outcomes. Our work relates to the often controversial literature on equality of consumer treatment in mortgage markets. This literature has traditionally focused on racial disparities in access to credit (e.g., Munnell et al., 1996; Ross and Yinger, 2002) and the cost of borrowing (e.g., Haughwout, Mayer, and Tracy, 2009; Bayer, Ferreira, and Ross, 2014). Another stream of work has focused on "predatory" behavior in mortgage origination (e.g., Gurun et al., 2016; Di Maggio, Kermani, and Korgaonkar, 2016), but no work to our knowledge has examined the quality of financial products and services as reflected by consumers' experience. Our paper also contributes to the debate surrounding the effectiveness or unintended consequences of the Community Reinvestment Act (e.g., Bhutta, 2011; Agarwal, Benmelech, Bergman, and Seru, 2012) by providing evidence that CRA-targeted areas experienced higher incidence of fraud, mis-selling, and general dissatisfaction with their mortgage products and services. At a broader level, our work suggests that regulators should consider the effect of regulation on both the quantity and quality of financial services received by underserved customers, particularly those in areas with high minority populations.

2 Theoretical Motivation and Research Design

The underlying theoretical motivation behind our work is rooted in two streams of literature: (a) informational frictions between borrowers and lenders and (b) the economics of discrimination. Lack of complete information or knowledge about financial products is a key friction in consumer financial markets. In recent years there has been a lot of interest in developing theoretical models that focus on economic drivers of information obfuscation (see, e.g., Gabaix and Laibson, 2006; Carlin and Manso, 2011). These papers show that financial institutions can take advantage of uninformed customers in a variety of ways such as offering confusing products, selling bundled services, or by improper disclosure. Our first set of tests, relating the extent of complaints to income and education level, is designed to uncover the importance of these frictions in the credit markets. We expect relatively poor and less educated neighborhoods to be financially less sophisticated and thus more likely to experience higher incidence of fraud, mis-selling and poor service.

The second strand of literature that we connect to goes back to the seminal work on economics of discrimination by Becker (1957) and Arrow (1973). There can be at least two potential reasons behind providing poor-quality treatment to minorities, over and above the effects that can be attributed to differences in observed income and education. The first one is a result of "taste-based" bias against minorities, and the other a result of "statistical" discrimination. Taste-based discrimination arises from racial prejudice even at the expense of profits. For example, a loan officer may not provide adequate information to minority borrowers or may not assist them in navigating through costly search for correct mortgage products, regardless of profit motivations. Similarly, a loan officer may not expend enough time and resources in resolving genuine difficulties faced by minority borrowers after the loan has been made, even if it is not consistent with the bank's profit-maximizing behavior. As a result, minority consumers may end up with poor-quality treatment as compared to other consumers. Statistical discrimination, on the other hand, is motivated by profit concerns. For example, suppose a loan officer uses race as a proxy for unobservable (or simply costly to observe) borrower characteristics such as the profitability of future income from the client. If she thinks that minority customers are less likely to provide higher profits to the bank in the long run, she may be tempted to sell predatory or otherwise unsuitable lending products in the short run to maximize what can be extracted from the present transaction. Here the discrimination arises due to profit motivation. Regardless of the channel, both are harmful to minority consumers and, because they are driven by differences in race, illegal. We do not tease out the relative importance of these two particular channels in the paper; our goal is to establish clear empirical evidence of differential, worse results for minority communities.

We begin our examination by relating various demographic characteristics to complaints using standard linear regression techniques. After establishing strong relationship between poor and minority borrowers and quality of financial services, we focus our analysis on teasing out the effect of supply-side forces in shaping these outcomes. We do so by analyzing the effect of a regulation-induced shock to increase supply of credit to poor and minority communities.

2.1 Regulation-induced shock to lending incentives

The Community Reinvestment Act (CRA) was passed in 1977 with the goal of eliminating discrimination against lending activities in low- and moderate-income (LMI) areas to ensure that institutions meet the credit needs of the entire community. While there have been several adjustments to the details of the policy since, the primary focus on meeting the credit needs of "under-served" communities has remained unchanged. Regulators periodically evaluate lenders' performance in serving these areas, and use these evaluations in approval decisions regarding lenders' applications for branch opening, mergers and acquisition activities, or entering new lines of business. These factors, along with the potential reputational harm, provide strong incentives for the lender to perform well on CRA exams. Thus, the designation of an area as LMI serves as a shock to lenders' incentives to increase supply of lending to poor and minority borrowers in these areas.

LMI status is determined at the census-tract level by a simple rule. Tracts with a median family income less than 80% of the MSA-level median income are designated as LMI. Thus the LMI designation is based on *relative* income of an area, relative to the MSA it resides in. We exploit variation in median MSA income across the country to compare outcomes for areas that are similar on observable characteristics including income, but differ in regulatory designation and thus the pressure for lenders to supply credit to the area. As an example, consider two MSAs in Texas in 2010: Dallas-Plano-Irving ("Dallas") and San Antonio-New Braunfels ("San Antonio"). The Dallas MSA median income is $68,900, which means tracts with median family income below $55,120 are designated LMI. In San Antonio, the MSA median income is $57,800, so tracts with median family income below $46,240 are LMI. As shown in Figure 4, the shaded area represents an absolute income range where zip codes in Dallas are designated as LMI, while zip codes with identical income in San Antonio are not LMI designated. Most simply, our empirical design assumes that shaded areas in Dallas and San Antonio (those with income between $46,240 and $55,120) are similar on unobserved demand-side dimensions such as their marginal propensity to complain. Under this assumption, we attribute differences in complaints across the two areas to the CRA regulatory designation that motivates banks to increase credit supply.

LMI designation is at the census-tract level. Since our data and analyses are at the zip code level, we first define zip codes as LMI if the majority of its population resides in LMI census tracts. We call these these *treatment* zip codes, which are receiving the CRA-induced incentive shock. Using propensity score matching, we find a set of control zip codes that are very similar to the treatment groups in terms of not only income, but also population, outstanding mortgages, education, and house price change in recent past. We conduct a series of matched sample analyses that differ in terms of the precise matching criteria used, for example matching based on continuous values of these variables versus matching based on coarse bins. For expositional simplicity, we defer their detailed discussions to the later part of the paper when we present the results. After estimating the propensity score, we use a kernel-based weighting matching technique to construct a set of control zip codes for each treatment zip code in the sample. We estimate the average treatment effect on the treated (ATET) zip codes based on the difference in outcome variables across these two areas. The key identifying assumption behind this approach is that conditional on a host of demographic variables such as income, education, house price change, population, and volume of outstanding mortgages, CRA designation is close to randomly assigned. Thus, the supply-side incentives to lend is orthogonal to any omitted variables of concern that come from demand side (i.e., consumer-driven) such as unobserved borrower characteristics across different areas.

Note that our treatment and control groups, by construction, are relatively poor areas of the country. Thus, the difference in outcomes across treatment and control group provides us with the effect of the quantity-focused regulation on quality in poor areas. We extend this analysis further by conditioning our analysis on the racial composition of the LMI areas. Specifically, we break all zip codes into two groups: below- and above-median minority areas. Within each subset, we separately estimate the ATET of the regulation. This analysis allows us to examine whether there is heterogeneity in the effect of supply shock for areas with low- versus high-minority composition. Thus, the test allows us to tease out the effect of supply-side forces on the quality of services received by high-minority areas as compared to low-minority areas.

3 Data and Sample

The Consumer Financial Protection Bureau (CFPB) was created as an independent consumer watchdog agency under the umbrella of the Dodd-Frank Act in 2010. The Bureau officially began its operations in 2011 with a mandate to "protect consumers from unfair, deceptive, or abusive practices and take action against companies that break the law." The CFPB has instituted a system for consumer complaints where consumers can lodge their grievances against financial institutions using a simple online system on the CFPB website.[6] The CFPB then forwards these complaints to the respective institutions for explanation or resolution of the complaints. It is reasonable to expect that these complaints are not minor irritants that can get easily resolved at the branch level. By reaching out to a government agency for assistance, consumers often come for help for serious issues involving the quality of products and services they receive. Appendix A provides an example complaint for illustration.

Individuals first choose the financial product or service about which there is a problem

[6]There are a variety of ways that individuals can reach out to the CFPB for help including web, email, fax, and postal mail, but the primary means is through a website interface.

(e.g., mortgage, payday loan, bank account). The individual provides more details about the product and about the events that led them to filing a complaint and their desired resolution. They also list the company with which they have a problem along with account information or additional relevant documentation. Finally, the individual provides personal contact information (including their zip code) to ensure that they can be contacted regarding the company's response and resolution.

The database began in July 2011 with the coverage of credit card complaints first and mortgage-related complaints later in December 2011. It has since been expanded to cover other products such as payday loans, checking accounts and student loans. Our sample covers complaints made in 2012-2016. During this period, there are over 680,000 complaints in the database, with over 210,000 complaints about mortgages. We focus on mortgage-related complaints for a couple of reasons. First, it is economically less meaningful to compare quality across different products such as mortgages and credit card. Second, mortgage-related misconduct in the financial market has captured the center stage of many policy decisions and academic papers in recent years. Third, we need a reasonable "scaling" variable to compare the incidence of complaints across zip codes. Ideally, we want to evaluate the number of complaints regarding a product category while controlling for the number of transactions/interactions between banks and consumers in that category in the given area. It is difficult to find such a variable for transactions such as credit card complaints. We can, however, find such an appropriate variable for mortgages: number of tax filers with mortgage interest reported in their IRS tax filings in the zip code. Lastly, a home mortgage is one of the most significant financial products in the U.S. economy, involving trillions of dollars in outstanding loans and many millions of consumers across the country. The mortgage is often the single most significant and complex transaction that many households ever engage with.

The remainder of data comes from multiple sources. Demographics data are from the 2010 Census files. Data on average income at the zip code-level and the number of tax filers with mortgage interest in the zip code come from the 2012 IRS SOI database. Data on education is from the Census Bureau's American Community Survey 2012 5-year estimates. We measure education as the share of the adult population in the zip code with at least a bachelor's degree. Data on five-digit zip code median house price changes are from the Federal Housing Finance Agency. This data source covers the majority of zip codes in the sample. For those that have missing data, we impute a value based on the county where the majority of the housing units reside. Data on zip code level foreclosure rate are from Zillow.

Our final sample covers all mortgage-related complaints from 2012-2016 for which we have matched demographic data: over 175,000 complaints across 16,309 zip codes.[7] Table 1 provides summary statistics. Each observation represents a five-digit zip code in the sample. After winsorization at 1% tails, the mean (median) zip code has 10.33 (5.00) complaints from 1,973 (1,190) underlying mortgages. We have large cross-sectional variation in the complaints, ranging from a minimum of one to a maximum of seventy-one complaints in a zip code with an interquartile range from two to thirteen. To remove the effect of skewness from the dependent variable, in our regressions we use a log transformation of this variable. However, our results remain similar with the number of raw complaints as a dependent variable as well.

On the measures of consumer demographics, the average zip code has non-white population of 21%, again with a large cross-sectional variation. Median household income is about $51,000 and the median zip code has about 22% of its population with at least a graduate degree. These figures are representative of broad U.S. population. In terms of house price growth, we compute two measures: one based on five year house price change in the given zip code starting in 2007 and ending in 2012, i.e., starting before the great recession and ending just before our complaints sample. The median zip code experienced a -15.5% change in house prices during this window.[8] We use this measures to control for the effect of losses

[7]Since 2015, complaints in the database are allowed to have a narrative included along with the basic complaint information. Appendix A provides an example complaint for illustration. When a narrative is included (about 30% of the time for 2015-2016 when the option is available), the five-digit zip code is replaced with a three-digit zip code to provide an extra layer of anonymity for the consumer. From the total of 212,669 mortgage-related complaints, 182,175 complaints remain that have the five-digit zip code identified. Of this, we are able to find matching demographic data for 175,323 complaints.

[8]In untabulated tests, we also consider another measure based on a five year change between 2010 and

in home value on the propensity to complain. We use different estimation window to ensure that our results are not sensitive to the inclusion or exclusion of large drop in home value during the great recession.

4 Results

4.1 Variation in quality across demographics

We estimate the following regression model to tease out the effect of demographic characteristics on quality of financial services received by the consumers:

$$lnComplaints_{zip5} = \rho(IER_{zip5}) + \sum_{b=2}^{50} \left(Mort_{b,zip5} + Pop_{b,zip5}\right) + \zeta_{zip3} + \nu_{zip5} \tag{1}$$

The dependent variable is the log of complaints in the five-digit zip code (zip5). IER_{zip5} is our demographic variable that takes a value based on income (I), educational attainment (E), or racial composition (R) of the neighborhood. All continuous variables are winsorized at 1% to minimize the effects of outliers and are standardized by subtracting their respective means from the raw variable and then dividing them by their standard deviations. Thus, all reported estimates represent the effect of one standard deviation (s.d.) change in explanatory variables on (approximately) the percentage change in the number of complaints. Hence, we can directly compare the coefficients across regression specifications to assess the economic magnitude of various explanatory variables. We compute clustered standard errors at the level of three-digit zip codes.

Since zip codes vary considerably in terms of their population and mortgage activities, we need to account for these differences across zip codes in our analysis. We do so in an extremely flexible way as follows. We categorize all zip codes into one of fifty buckets based

2015, where the median zip code experienced a gain of 3.5%. We obtain similar results.

on their relative rank in terms of the number of outstanding mortgages and population. Based on these ranks we create two vectors of fifty indicator variables $Mort_b$ and Pop_b with an element equal to one for the respective mortgage quantity and population buckets where the zip code resides. We include these flexible controls in all of the regression estimates, and this allows us to separate out the baseline effects of mortgage volume and population on complaint frequency. The choice of fifty is admittedly arbitrary, and our results remain similar if we use other sensible techniques to separate out these effects such as using ten or 100 buckets or a flexible polynomial approach.

We also include fixed effects for three-digit zip codes (ζ_{zip3}) to remove the effects of local macroeconomic conditions and state regulations from affecting our results. Thus, our model captures variation in outcomes across five-digit zip codes within a given three-digit zip codes. Our 16,309 five-digit zip codes fall under 876 three-digit zip codes, providing us with enough variation within the three-digit zip codes to identify the effect of variation in demographic conditions after soaking away differences in economic and regulatory considerations.

Table 2 presents the estimates of the regression in equation (1). In column (1) we only include zip3 fixed effects in the model as explanatory variables, and find R^2 of 47% for the model. Column (2) that also includes fixed effects for the fifty mortgage buckets shows a dramatic increase in model fit, with R^2 increasing to 80%. Column (3) - (5) present our main results. One standard deviation decrease in income is associated with about 10% more complaints, whereas the corresponding effect for lower education is comparable at 8%. The effect of non-white population, on the other hand, is considerably higher at about 17%.[9] Needless to say, these demographic variables are correlated. Column (6) separates out the relative importance of each of these three variables by including them all in the model. While all three variables remain statistically significant, the non-white population of the zip code clearly dominates income and education in terms of economic magnitudes. The effect of the non-white variable is almost three times as large as that of education, and eight times as

[9]More precisely, the coefficient of 0.17 indicates an increase in complaints of $e^{0.17} - 1 = 19\%$.

large as that of income.

To further examine the difference in complaints along the minority dimension, we estimate the following regression:

$$lnComplaints_{zip5} = \sum_{r=2}^{5} \psi_r NWgroup_{r,zip5} + \\ \sum_{b=2}^{50} \left(Mort_{b,zip5} + Pop_{b,zip5} + Inc_{b,zip5} + CollEd_{b,zip5} \right) + \zeta_{zip3} + \nu_{zip5} \quad (2)$$

NWgroup represent indicator variables for minority population share in the zip code: 20-40%, 40-60%, 60-80%, and 80-100%. Zip codes with 0-20% form the omitted base category in the regression and $\psi_2 \ldots \psi_5$ represent the increase in frequency of complaints for their respective groups. $Mort_b$, Pop_b, Inc_b, and $CollEd_b$ respectively represent vectors of fifty indicator variables with an element equal to one for the respective number of mortgages outstanding, population, income, and education where the zip code resides. Figure 1 presents these results graphically, all of which are statistically different from zero and each other. These estimates reinforce the earlier results and demonstrate that the results are even stronger in areas with highest minority concentrations, with areas above 80% of minority population having nearly double the complaints compared to those with below 20% minority population.

4.2 Alternative channels

We now provide evidence that our key results relating minority consumers to the number of complaints is not driven by baseline differences in the propensity to complain, house price declines, or the behavior of mortgage servicers. To ease comparisons across these tests, we restrict the sample to observations with complete information on all of these additional variables.

4.2.1 Propensity to complain

An alternative explanation of our results linking minority neighborhoods to complaints could be that minority areas contain residents that are simply more likely to complain regardless of how they are treated or what products they are sold. If such variation is regional, then the baseline model already soaks away this heterogeneity with three-digit zip code fixed effects. To further address this potential confounder, we use the number of complaints made by consumers to a different government agency that is unrelated to mortgage business. We obtain the number of complaints filed with the Federal Communication Commission (FCC) by consumers about issues involving telecommunications billing and services. We use the log of FCC complaints (*lnFCC*) for each zip code in 2015-2016 (when the data became available) as a proxy for potential baseline differences in propensity to complaint. Because telecom-related complaints are unrelated to mortgage transactions, FCC complaints provide a reasonable control for a baseline "complainer effect" in the area.

Table 3 present the results, with column (1) reproducing our base case results for comparison. Column (2) presents the specification including *lnFCC*. We do find a positive and significant coefficient on FCC complaints variable, which suggests there are some common factors that explain complaints arising from different aspects of consumers' lives. However, this factor seems orthogonal to the main effect of the relationship between the minority share of the population and complaints.

4.2.2 House price decline

Another potential concern is whether the results could be driven by differences in house price changes across minority and non-minority areas within the three-digit zip code. If minority zip codes experienced disproportionately large price drop in the aftermath of the subprime mortgage crisis, and if the propensity to complain correlates with loss in home value, then our results could be explained away by this factor. Recall, our three-digit zip code fixed effects will capture a great deal of variation from such regional shocks, so this mechanism must be working within three-digit zip code. To test whether this is driving our results, we compute the price change over the past five years leading up to 2012 (*HPgrowth*) for every five-digit zip code in our sample from the Federal Housing Finance Agency for which we can gather such data. We include this house-price change as an additional variable in our regressions. For the few zip codes with no data at that level, we impute those observations with the house-price change of their respective county. Column (3) of Table 3 presents the results. While areas with higher price drop do have more complaints, the coefficient estimate on the minority variable is virtually unchanged.

Similarly, we investigate if our results could be explained away by differences in actual foreclosure rate. If borrowers are foreclosed upon with higher frequency in a zip code, it is likely there could be more complaints to the CFPB. Foreclosure decisions are often an outcome of bargaining process between the borrower and the lender. To the extent that a higher foreclosure rate is driven by discriminatory differences in forebearance across areas, this variable may contain variation driven by the channel we have been emphasizing in the paper. We compute the foreclosure rate at the five-digit zip code level using data from Zillow and control for this variable in column (4) of the Table 3. Areas with higher foreclosure rate do have more complaints, but again this effect does not explain away our key finding. Column (5) includes FCC complaints, house price drop and foreclosure rates together in the model and shows that our results remain intact.

4.2.3 Mortgage Servicing

Our complaints database also includes the identity of the company against which the complaint is filed. Because mortgage servicing rights may be sold to mortgage-servicing specialists, there may be a concern that our results are simply driven by unscrupulous servicers. Complaints to the CFPB about servicers could itself be a results of predatory lending by originating banks in the first place. Still, to make sure these firms are not driving all our results, we manually examine each institution with at least 200 complaints in the CFPB database and identify companies whose primary business is loan servicing (e.g., Ocwen). We exclude complaints made against specialized servicers and estimate the model with remaining data. As shown in column (6) of Table 3, our results are virtually unchanged.

4.3 The role of regulation

We now present our results using the CRA's low- or moderate-income (LMI) designation (or target area) as a shock to prioritize lending to poor and minority communities as discussed in section 2. We compare outcomes across LMI ("treatment") zip codes to similar non-LMI ("control") zip codes to tease out the effect of supply shock on quality. Figure 2 plots separately the income distribution of LMI and non-LMI zip codes across the country. By the very definition of this regulatory criteria, the treatment zip codes are concentrated in the left tail of income distribution. Thus, simply comparing outcomes across these two areas is not particularly meaningful. Since the LMI designation relies on an area's relative income within the MSA, there are many zip codes with similar absolute income as the areas in LMI group that do not carry the LMI regulatory designation. This overlap provides us with a meaningful set of treatment and control groups that are similar on many dimensions, but differ in the supply-side pressure on lenders to serve the area.

There are 3,049 LMI (treatment) zip codes in the sample spread all over the country including all 50 states and the District of Columbia. The vast geographical dispersion of treatment zip codes across the country makes our empirical design even more powerful: it ensures that our results are not driven by effects that are unique to a particular locale. To ensure that our treatment and control zip codes lie on common support of income distribution, we truncate the sample at the 1st percentile of the non-LMI income distribution and at the

99th percentile of the LMI distribution.[10] The remaining sample includes 1,987 treatment zip codes and 11,726 control zip codes which have average household incomes in the range of \$33,173 to \$112,484.

To find suitable control observations with which we can construct counterfactuals for each target area, we estimate the propensity score for LMI designation using a probit model, with the number of mortgages, population, income, education, and house price changes as key predictors. To provide greater flexibility to our matching, we include the continuous value of these variables, indicator variables for each decile of the respective dimension, and state of the zip code in our matching exercise.[11] In terms of matching methodology, our base estimator uses kernel-weighted propensity score matching to construct the counterfactual for target areas. For the kernel weighting, we use use a gaussian kernel with a bandwidth of 0.03.[12] Figure 3 shows the comparability of treatment and control observations before and after matching. The plot shows the standardized bias for each of the five matching variables, which is calculated as the difference in means across the treated and control group divided by the standard deviation of the respective variable. The matching procedure drastically reduces the bias for each covariate. While there is not a well-developed literature on formally assessing standardized bias, each covariate falls well below the 20% threshold that Rosenbaum and Rubin (1985) call "large" and comfortably within the bounds of the balancing tests (e.g., post-match covariate variance ratio) proposed in Rubin (2001).

Main matching results

Table 4 presents the base result in column (1). The treated zip codes have about 32% ($e^{0.28} - 1$) more complaints than control zip codes. Thus, the regulation-induced supply

[10] The results are not sensitive to using this particular criteria and are robust to examining the full, unrestricted sample. This restriction simply ensures the comparisons are most sensible and particularly avoids potential concerns about sparseness of control observations in the lower tail of the income distribution.

[11] In untabulated results, we find that requiring matches be in the same state, rather than using "state" as a predictor variable in propensity score, yields similar results.

[12] Robustness test shown later show that the results are not driven by these specific modeling choices.

shock results in substantially higher number of complaints about fraud, mis-selling, and poor customer service in the treated area.

Our matching criteria ensures comparability along a number of most important dimensions that can be correlated with unobserved factors related to complaints. One may still be concerned about differences in the cost of living across treatment and control zip codes since the matches, by construction, are located at different points in the MSA income distribution. For an extreme example consider San Francisco and Los Angeles MSAs in California. An area with median family income of $75,000 in San Francisco falls at the 25th percentile of income distribution in that MSA. Hence this area carries an LMI target area designation. In contrast, an area with the same level of income in Los Angeles falls at the 65th percentile of the MSA's distribution, and hence it does not fall under the LMI category. Suppose our matching exercise picks up the Los Angeles area as the control for the treated area in San Francisco. One may be concerned that cost of living differences are so high across the two areas that they do not constitute good matches.

To address this issue, we divide the sample into MSA-level income buckets that are $10,000 wide and require control zip codes to come from the same bucket as the treatment zip codes. We then repeat the analysis requiring matches to be in the same $5,000 bucket. Such requirements further ensure that we do not compare areas in high cost-of-living MSA (e.g., San Francisco in the example) to places that have relatively lower cost-of-living (Los Angeles). Columns (2) and (3) in Table 4 present the results, and the results are similar to the baseline test. We perform similar analysis by directly stratifying the sample by absolute income. This matching effectively places more weight in the matching on income (and thus less on other dimensions). Columns (4)-(5) present the results, which show slightly smaller point estimates, but still economically and statistical significant effects. These results paint a clear and consistent picture that perturbation on the supply side of credit leads to a substantial dilution in the quality of financial services provided by the lenders. Thus regulations aimed at increasing the quantity of lending to poor borrowers seem to have an adverse impact on the quality dimension.

Regulation and Race

We now turn to the effect of regulation on minority consumers, which allows us to separate the effect of supply shock on the quality of financial services received by these consumers. We break this analysis into two parts. In the first part, we investigate whether our results relating race to quality and then regulation to quality are coming from the same zip codes. Said differently, if the distinction between LMI and non-LMI groups is effectively divided along minority/non-minority lines, then the effects of race and regulation may be confounded. To isolate the effect of regulation beyond any effects of local minority population share, we extend our matching criteria to include the racial composition of the area as an additional covariate in the propensity score matching. Specifically, we include both the percentage of minority population and the rank decile of zip codes along this dimension as additional covariates in matching procedure. Thus, the new set of treatment and control groups are similar along several demographic dimensions such as income, education, and minority population. We present this result in column (2) of Table 5. Compared to the base case estimate of 0.28, the difference between treatment and control areas drops to 0.18 when we additionally match on racial composition. Thus, even after controlling for racial composition, areas with credit supply shock have about 20% ($e^{0.18} - 1$) more complaints than observationally similar control areas. Both these sources of variation – race and regulation – have independent explanatory power in explaining differences in quality across zip codes.

In the second part of the analysis, we explore the interactive effects: does the supply shock disproportionately hurt the minority consumers? We do so by estimating the matched sample test separately for areas with relatively large and small minority population. We break all zip codes into two groups based on whether they have below- or above-median share of minority population (median minority population is 12.1% for the matching sample).

Following the base matching technique, we now find a set of control firms for each treated firm within the same minority-population bucket. Said differently, target areas in above-median (below-median) minority population are matched with comparable control areas in above-median (below-median) minority population only. Table 5 presents the results in columns (3)-(4). A stark pattern emerges: in low-minority population areas, treatment areas have 9% more complaints than the control areas, whereas the corresponding difference is about four-times larger at 42% ($e^{0.35} - 1$) for the high minority areas. The difference-in-difference of 25% across low and high-minority population areas is significant at 1%. The results are even more stark when requiring matches within $10,000 MSA income strata as shown in columns (5)-(6). In this specification, above-median minority areas experience a 49% ($e^{0.42} - 1$) increase in complaints for LMI as compared to non-LMI areas, while the regulation effect is not statistically different form zero for below-median minority areas. In sum, the regulation-induced shock to the supply side of lending has disproportionately large detrimental impact on the quality of service received by minority population. Overall, these results provide strong evidence on the role of shocks to the supply-side of credit in affecting the quality of services received by minority consumers.

Placebo Tests and Alternative Matching Strategies

Our results so far exploit the institutional feature that areas below 80% of the median MSA-income are classified as target areas for CRA lending. We conduct two sets of placebo tests by varying the cut-off points in artificial ways to show that it is CRA's actual LMI cut-off point of 80% that drives our results. Specifically, we examine the complaints frequency around 70% and 90% thresholds.

For the 70% threshold analysis, we consider all zip codes that are below 80% cutoff, and assume that all zip codes below the 70% cutoff are in the CRA-target areas, whereas those above this threshold are not. Thus, we are artificially considering some of the zip codes that lie between 70% and 80% of the MSA median area as non-treated zip codes. In reality all these zip codes are below the actual threshold of 80%, and hence they are all treated by the CRA. For the placebo test at 90% threshold, we symmetrically only consider zip codes that are above 80% of the MSA-level median income. Thus, in this test we artificially consider zip codes between the 80%-90% of MSA income area as treated zip codes, whereas those above 90% are considered non-treated. Thus, each of the placebo tests allows us to examine nearby cutoffs (varying the income threshold) while making sure that we do not compare across the actual treatment-control threshold.

Using matching technique describe earlier, we present the results of placebo tests in Table 6. A stark pattern emerges from this analysis. There is some difference, though much smaller and nearly always statistically insignificant, in the quality of financial services across the placebo treatment and control group thresholds. This supports the notion that income is a driver of the quality of financial services. But there is a sharp discontinuity in this effect at the 80% threshold compared to the artificial thresholds at 70% and 90%. For example, when we match treatment and control zip codes in the most stringent specification, requiring MSA median income to be within the same $5,000 MSA income bucket, the difference in number of complaints across the treatment and control groups in placebo tests is 0.01 and statistically insignificant. In contrast, the corresponding difference at the actual threshold of 80% is 0.22 (p-value$<$0.01). Overall, these tests alleviate a number of concerns: (i) our results are unlikely to be driven by pure income differences across treatment and control areas; (ii) our results are unlikely to be driven by any unobserved difference in the characteristics of the two zip codes as long as they do not discontinuously jump at the threshold of 80%; and (iii) our results are unlikely to be driven by any correlation between cost-of-living differences (i.e., differences in relative income) across areas that differ slightly in relative income.

Our main tests use kernel-weighted propensity matching. The kernel-weighting allows us to efficiently use more data to construct our matched counterfactual in places where there are many possible matches. We also present several alternative estimates in Table 7 using different matching strategies to show that our results do not hinge on a particular matching scheme. Columns (1)-(2) vary the kernel bandwidth from 0.01 to 0.05 and show similar estimates to the base estimate using a bandwidth of 0.03. Columns (3)-(4) use nearest neighbor propensity score matching using one and three nearest neighbors, respectively. The results are similar. Finally, we dispense with propensity score matching and use Mahalanobis distance nearest neighbor matching.[13] Columns (5) and (6) present the results, with Column (6) using the additional constraint that the matched zip code be in the same $5,000 income strata. Again, our results indicate an economically significant increase in complaints for CRA-target zip codes as compared to their observationally similar counterparts in areas that lack the extra regulatory pressure to lend.

5 Discussion & Conclusions

Since the very beginnings of modern finance, there have been concerns about exploitation of low-income and minority consumers by large, sophisticated banks. Market failures such as banks' market power, high search costs, and asymmetric information problems make consumer finance an area that is particularly vulnerable to such behavior. Using mortgage-related consumer complaint data relating to fraud, mis-selling, and poor service from the Consumer Financial Protection Bureau, we show that areas with low income, low educational attainment, and high shares of minority consumers receive significantly worse quality of financial services. Most striking, the relationship between high-minority concentration and complaints is exceptionally strong even after controlling for income, education, and other potentially confounding factors such as house price changes.

As a result of concerns of unequal access to credit for low-income and minority areas, regulations such as the Community Reinvestment Act provide strong incentives to lenders

[13]For this matching technique, we match on mortgages, population, income, education and home price changes, but do not match on the deciles of these variables to keep the dimensions of matching manageable.

to serve areas that are defined as low- to moderate-income relative to the median MSA income. While such regulation may, at the margin, provide a higher *quantity* of credit to such target areas, we show that the average *quality* of products and services are substantially lower. Further, the dilution in quality is disproportionately larger for high-minority areas. Overall, our results show that consumers in low-income and minority areas experience worse outcomes along the quality dimension, and regulations that are mainly focused on increasing the quantity of lending to these borrowers only makes this effect worse.

References

Agarwal, Sumit, Efraim Benmelech, Nittai Bergman, and Amit Seru, 2012, Did the Community Reinvestment Act (CRA) lead to risky lending?, *NBER Working Paper* .

Arrow, Kenneth J, 1973, The theory of discrimination, in Orley Anshenfelter, and Albert Rees, eds., *Discrimination in Labor Markets*, 3–33 (Princeton University Press).

Bayer, Patrick, Fernando Ferreira, and Stephen L Ross, 2014, Race, ethnicity and high cost mortgage lending, *NBER Working Paper* .

Becker, Gary, 1957, *The Economics of Discrimination* (The University of Chicago Press).

Bhutta, Neil, 2011, The community reinvestment act and mortgage lending to lower income borrowers and neighborhoods, *The Journal of Law and Economics* 54, 953–983.

Campbell, John Y, 2006, Household finance, *Journal of Finance* 61, 1553–1604.

Carlin, Bruce Ian, and Gustavo Manso, 2011, Obfuscation, learning, and the evolution of investor sophistication, *Review of Financial Studies* 24, 754–785.

Di Maggio, Marco, Amir Kermani, and Sanket Korgaonkar, 2016, Partial deregulation and competition: Effects on risky mortgage origination, *Working Paper* .

Gabaix, Xavier, and David Laibson, 2006, Shrouded attributes, consumer myopia, and information suppression in competitive markets, *Quarterly Journal of Economics* 121, 505–540.

Griffin, John M, and Gonzalo Maturana, 2016, Did dubious mortgage origination practices distort house prices?, *Review of Financial Studies* 29, 1671–1708.

Gurun, Umit G, Gregor Matvos, and Amit Seru, 2016, Advertising expensive mortgages, *Journal of Finance* 71, 2371–2416.

Haughwout, Andrew, Christopher Mayer, and Joseph Tracy, 2009, Subprime mortgage pricing: the impact of race, ethnicity, and gender on the cost of borrowing, *Brookings-Wharton Papers on Urban Affairs* 2009, 33–63.

Holmstrom, Bengt, and Paul Milgrom, 1991, Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design, *Journal of Law, Economics, & Organization* 7, 24–52.

Ladd, Helen F, 1998, Evidence on discrimination in mortgage lending, *The Journal of Economic Perspectives* 12, 41–62.

Munnell, Alicia H, Geoffrey MB Tootell, Lynn E Browne, and James McEneaney, 1996, Mortgage lending in Boston: Interpreting HMDA data, *The American Economic Review* 25–53.

Piskorski, Tomasz, Amit Seru, and James Witkin, 2015, Asset quality misrepresentation by financial intermediaries: Evidence from the RMBS market, *Journal of Finance* 70, 2635–2678.

Rosenbaum, Paul R, and Donald B Rubin, 1985, Constructing a control group using multivariate matched sampling methods that incorporate the propensity score, *The American Statistician* 39, 33–38.

Ross, Stephen L, and John Yinger, 2002, The color of credit: Mortgage discrimination, research methodology, and fair-lending enforcement, *MIT Press Books* .

Rubin, Donald B, 2001, Using propensity scores to help design observational studies: application to the tobacco litigation, *Health Services and Outcomes Research Methodology* 2, 169–188.

U.S. House of Representatives, Committee on Financial Services Hearing, 2007, Rooting out Discrimination in Mortgage Lending: Using HMDA as a Tool for Fair Lending Enforcement. June, 25, 2007 Hearing before the Subcomittee on Oversight and Investigations.

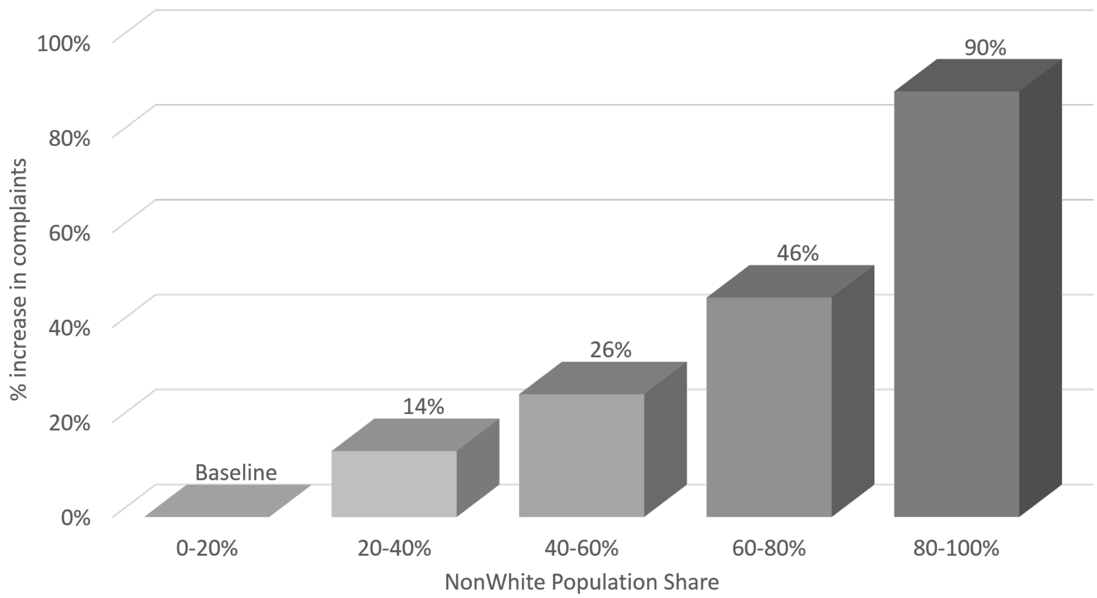

Figure 1: Minority Share and Quality
This figure presents a bar chart of the point estimates from the following regression:
$lnComplaints_i = \sum_{r=2}^{5} \psi_r NWgroup_{r,i} + \sum_{b=2}^{50} \left(Mort_{b,i} + Pop_{b,i} + Inc_{b,i} + CollEd_{b,i} \right) + \zeta_{zip3} + \nu_i$.
$NWgroup$ represents indicator variables for minority population share in the zip code: 0-20% (omitted base category in the regression), 20-40%, 40-60%, 60-80%, and 80-100%. *Mort*, *Pop*, *Inc*, *CollEd* respectively represent indicator variables for 50 equally populated buckets of the number of mortgages outstanding, population, income, and education for a zip code. Point estimates from the regression are translated into percent increase above the base category for this figure.

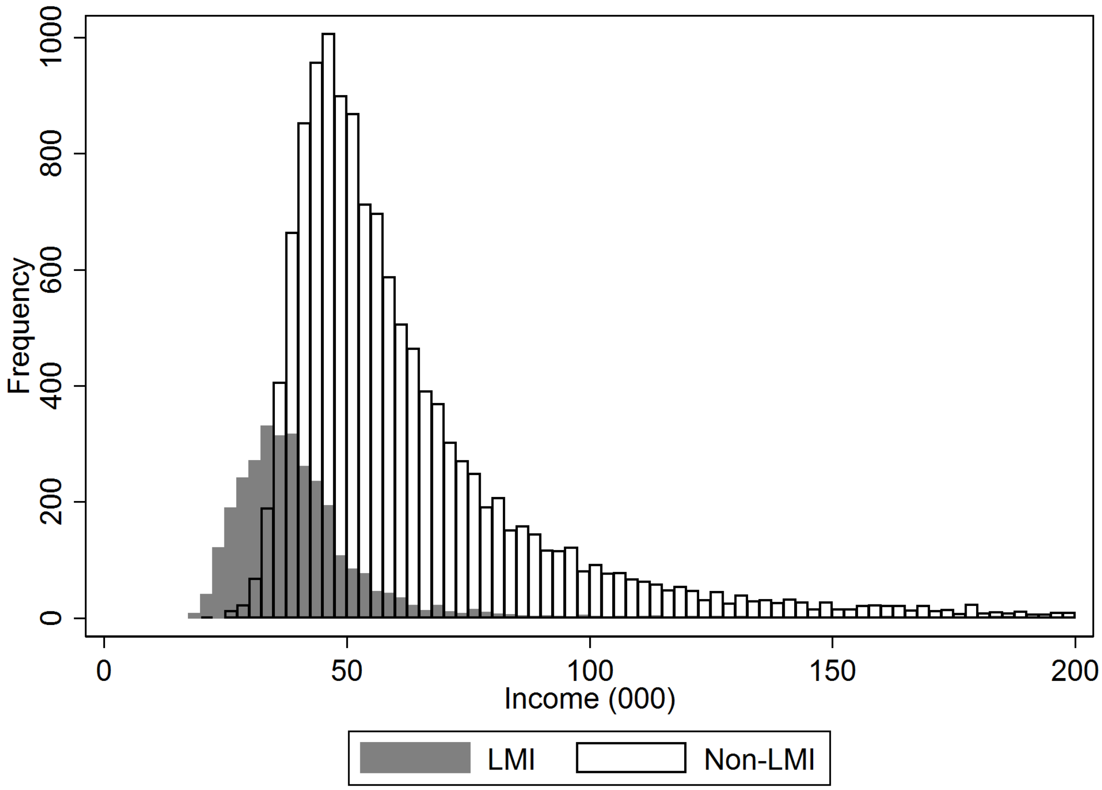

Figure 2: Area Income and Low- to Moderate-Income Status
This figure presents a histogram of income separately for low- to moderate-income (LMI) zip codes and non-LMI zip codes. Census tracts are designated LMI when the median family income is below 80% of the median MSA income. LMI zip codes are those with a majority of its population in low- to moderate-income tracts in 2010.

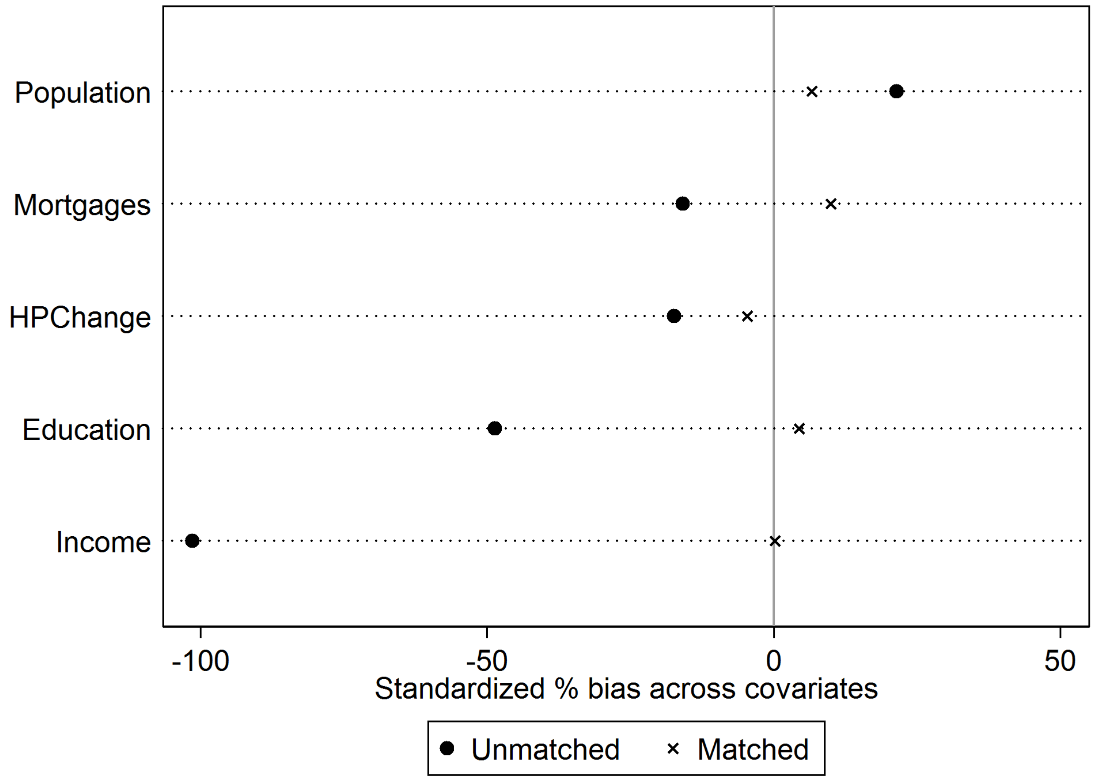

Figure 3: Covariate Balance for Raw and Matched Samples
This figure presents a the difference in means between low- to moderate-income (LMI) zip codes for the number of mortgages, population, income, education, and house price index changes. The circles (•) represent raw pre-match differences and the the x's (×) represent the post-match differences. The difference is measured in terms of standardized % bias, which is the percent difference in means divided by the sample standard deviation or the variable.

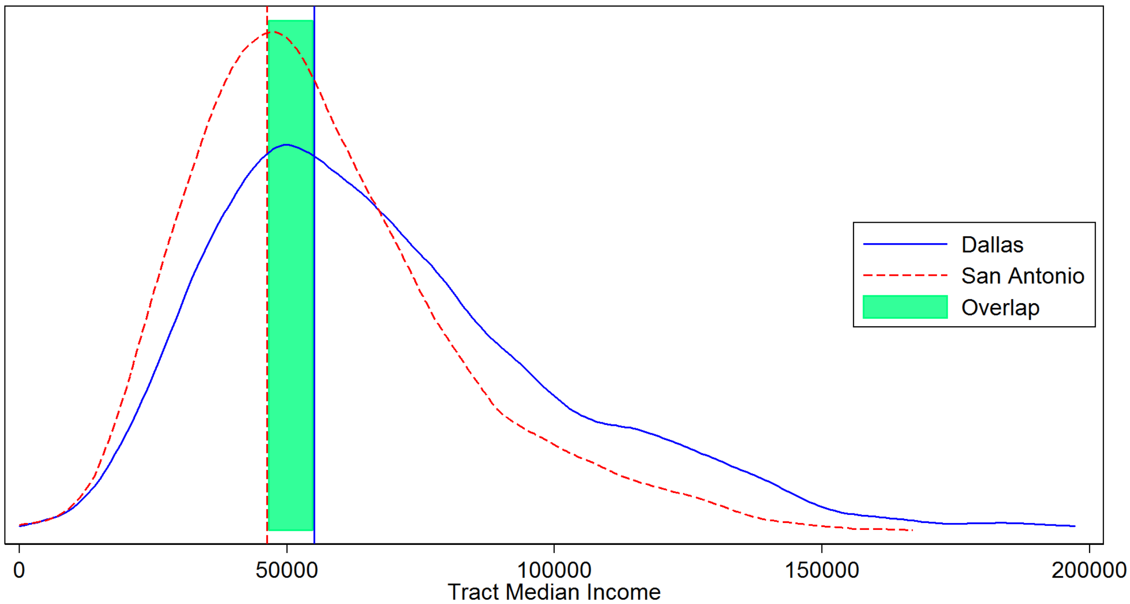

Figure 4: Example Geographical Variation in Low- to Moderate-Income Designation
This figure presents a kernel densities of 2010 census tract incomes for the Dallas-Plano-Irving ("Dallas") and San Antonio-New Braunfels ("San Antonio") metropolitan statistical areas (MSA) in Texas. The vertical lines at $55,120 and $46,240 represent 80% of the respective median MSA incomes for Dallas and San Antonio. The shaded "Overlap" region between those two lines indicate the income range where, despite identical median family income (e.g., $50,000), Dallas tracts are designated low- to moderate-income (LMI), but San Antonio tracts are not.

Table 1: Sample Summary Statistics

This table presents the sample summary statistics. *Complaints* is the number of mortgage-related complaints filed to the CFPB in a given five-digit zip code in a given year winsorized at the 1% tails, *lnComplaints* is the log of Complaints, *AGI* is the mean household adjusted gross income in the five-digit zip code for 2012, *lnAGI* is the log of adjusted gross income, *College Education* is the portion of the adult population in the five-digit zip code with at least a bachelor's degree in 2012. *NonWhite* is the share of the zip5 population that is a minority race for 2012, *LMI* is an indicator variable equal to one for zip codes with a majority of its population in low- to moderate-income tracts in 2010, *Mortgages* is the number of mortgages in the five-digit zip code in 2012 measured by IRS filings with reported mortgage interest, *Population* is the zip code population in 2010, $\%\Delta HP$ is the percentage point change in zip5 house price growth (county house price growth is used for observations with no zip code level house price data), *Foreclosures* is the total zip5 foreclosures for 2012-2016 as reported by Zillow, and *lnFCC* is the log of the number of complaints to the Federal Communications Commission from 2015-2016 (data begin in 2015). All variables are winsorized at the 1% level.

variable	mean	sd	min	p25	p50	p75	max	N
Complaints	10.33	13.25	1.00	2.00	5.00	13.00	71.00	16,309
lnComplaints	1.63	1.22	0.00	0.69	1.61	2.56	4.26	16,309
AGI Income (000)	64.06	52.97	18.65	42.05	51.23	67.61	1464.53	16,309
lnAGI	10.93	0.44	10.12	10.65	10.84	11.12	12.54	16,309
College Education	0.27	0.16	0.05	0.15	0.22	0.35	0.76	16,309
Nonwhite	0.21	0.21	0.01	0.05	0.13	0.30	0.90	16,309
LMI	0.19	0.39	0.00	0.00	0.00	0.00	1.00	16,309
Mortgages (000)	1.97	2.07	0.04	0.41	1.19	2.92	9.54	16,309
Population (000)	17.20	15.18	0.62	4.78	12.66	26.11	67.05	16,309
$\%\Delta\text{HP}_{2007-2012}$	-17.73	15.07	-58.3	-26.75	-15.50	-6.35	8.99	15,867
$\text{Foreclosures}_{2012-2016}$	254.69	253.92	0.00	59.08	182.04	368.37	1241.18	9,740
lnFCC	2.97	1.3	0	2.08	3.14	3.99	5.32	15,806

Table 2: Income, Education, and Race

This table presents OLS estimates from the regression of complaints (*lnComplaints*) for a given five-digit zip code (zip5) on measures of income, education, race, and various sets of fixed effects. *lnComplaints* is the log number of mortgage-related complaints filed to the CFPB in a given zip5 during the sample period (2012-2016) , *lnAGI* is the log of the average adjusted gross income of households in each zip5 for 2012, *CollEd* is the share of the zip5 adult population for 2012 with at least a bachelor's degree, and *NonWhite* is the share of the zip5 population that is a minority race for 2012. *MortBucket50* represents a set of dummy variables for 50 equally populated buckets of the number of mortgages outstanding in the zip code (e.g., 700-750 mortgages) measured by IRS filings with reported mortgage interest for 2012. Similarly, *PopBucket50* represents dummy variable for 50 zip code population buckets for 2012. All continuous independent variables are standardized to have a mean of zero and unit variance Standard errors are clustered by three-digit zip code (zip3).

	(1)	(2)	(3)	(4)	(5)	(6)
lnAGI			-0.10***			-0.02*
			(<0.01)			(0.10)
CollEd				-0.08***		-0.06***
				(<0.01)		(<0.01)
NonWhite					0.17***	0.16***
					(<0.01)	(<0.01)
MortBucket50 FE	No	Yes	Yes	Yes	Yes	Yes
PopBucket50 FE	No	No	Yes	Yes	Yes	Yes
zip3 FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16309	16309	16309	16309	16309	16309
R^2	0.47	0.80	0.81	0.81	0.82	0.82

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Focusing on Race: Alternative Channels

This table presents OLS estimates from the regression of complaints (*lnComplaints*) for a given five-digit zip code (zip5) on measures of income, education, race, and various sets of fixed effects. *lnComplaints* is the log number of mortgage-related complaints filed to the CFPB in a given zip5 during the sample period (2012-2016) , *NonWhite* is the share of the zip5 population that is a minority race for 2012, *lnFCC* is the log of the number of complaints to the Federal Communications Commission from 2015-2016 (data begin in 2015), $\%\Delta HP_{2007-2012}$ is the percentage point change in zip5 house price growth (county house price growth is used for observations with no zip code level house price data), $Foreclosures_{2012-2016}$ is the total zip5 foreclosures for 2012-2016 as reported by Zillow, *MortBucket50* represents a set of dummy variables for 50 equally populated buckets of the number of mortgages outstanding in the zip code (e.g., 700-750 mortgages) measured by IRS filings with reported mortgage interest for 2012. Similarly, *PopBucket50*, *IncomeBucket50*, *CollEdBucket50* respectively represents dummy variable for 50 zip code population, income, and education buckets. All tests in this table are limited to observations with full data on each variables for ease of comparability. Column (6) measures complaints excluding specialized mortgage servicers. All continuous independent variables are standardized to have a mean of zero and unit variance. Standard errors are clustered by three-digit zip code (zip3).

	All					NoServicers
	(1)	(2)	(3)	(4)	(5)	(6)
NonWhite	0.14***	0.15***	0.13***	0.13***	0.13***	0.12***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)
lnFCC		0.11***			0.10***	0.09***
		(<0.01)			(<0.01)	(<0.01)
$\%\Delta \mathrm{HP}_{2007-2012}$			-0.14***		-0.12***	-0.12***
			(<0.01)		(<0.01)	(<0.01)
$\mathrm{Foreclosures}_{2012-2016}$				0.06***	0.05***	0.05***
				(<0.01)	(<0.01)	(<0.01)
MortBucket50 FE	Yes	Yes	Yes	Yes	Yes	Yes
PopBucket50 FE	Yes	Yes	Yes	Yes	Yes	Yes
IncomeBucket50 FE	Yes	Yes	Yes	Yes	Yes	Yes
CollEdBucket50 FE	Yes	Yes	Yes	Yes	Yes	Yes
zip3 FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9234	9234	9234	9234	9234	9234
R^2	0.81	0.81	0.81	0.81	0.81	0.80

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Regulation and Quality

This table presents matching estimates for complaints (*lnComplaints*) for a given five-digit zip code (zip5) for LMI zip codes as compared to matched non-LMI zip codes, where *LMI* is an indicator variable equal to one for zip codes with a majority of its population in low- to moderate-income tracts in 2010. The matching method uses a kernel-weighted propensity score to construct counterfactuals to compute the average treatment effect on the treated (ATET). The kernel is gaussian with 0.03 bandwidth. The propensity score is estimating using probit regression and includes *lnMortgages*, *lnPopulation*, *lnAGI*, *CollEd*, $\%\Delta HP_{2007-2012}$, indicator variables for each decile of those five variables, and state indicator variables. Column (1) presents the base estimate. Columns (2)-(3) respectively divide the sample into \$10k and \$5k strata based on MSA-median income and require matches be within-strata. Columns (4)-(5) place similar restrictions using zip code income strata. The number of matched observations decreases with more stringent requirements.

		MSA Strata		Income Strata	
	Base	10k	5k	10k	5k
	(1)	(2)	(3)	(4)	(5)
LMI (atet)	0.28***	0.31***	0.22***	0.21***	0.21***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)
N	13713	13083	12521	11337	10151
Ntreat	1987	1864	1823	1891	1888
Ncontrol	11726	11219	10698	9446	8263

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Race and Regulation

This table presents matching estimates for complaints (*lnComplaints*) for a given five-digit zip code (zip5) for LMI zip codes as compared to matched non-LMI zip codes, where *LMI* is an indicator variable equal to one for zip codes with a majority of its population in low- to moderate-income tracts in 2010. The matching method uses a kernel-weighted propensity score to construct counterfactuals to compute the average treatment effect on the treated (ATET). The kernel is gaussian with 0.03 bandwidth. The propensity score is estimating using probit regression and includes *lnMortgages*, *lnPopulation*, *lnAGI*, *CollEd*, $\%\Delta \text{HP}_{2007-2012}$, indicator variables for each decile of those five variables, and state indicator variables. Column (1) presents the base estimate. Column (2) presents the estimate when *NonWhite* and indicator variables for each decile of *NonWhite* are used in the matching scheme. Columns (3)-(4) provide the base estimates when splitting the sample to below- and above-median *NonWhite* share (*Low NW* and *High NW*). Columns (5)-(6) perform the split-sample estimation while also requiring that matches be in the same $10k MSA-median income strata. The number of matched observations decreases with more stringent requirements.

		Match on	Base		MSA Strata	
	Base	NonWhite	Low NW	High NW	Low NW	High NW
	(1)	(2)	(3)	(4)	(5)	(6)
LMI (atet)	0.28***	0.18***	0.09*	0.35***	0.05	0.40***
	(<0.01)	(<0.01)	(0.07)	(<0.01)	(0.27)	(<0.01)
N	13713	13713	6705	6856	5131	6239
Ntreat	1987	1987	470	1517	437	1391
Ncontrol	11726	11726	6235	5339	4694	4848

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Placebo Tests: The Low- to Moderate-Income Threshold

This table presents matching estimates for complaints (*lnComplaints*) for a given five-digit zip code (zip5) for LMI zip codes as compared to matched non-LMI zip codes, where *LMI* is an indicator variable equal to one for zip codes with a majority of its population in low- to moderate-income tracts in 2010. The matching method uses a kernel-weighted propensity score to construct counterfactuals to compute the average treatment effect on the treated (ATET). The kernel is gaussian with 0.03 bandwidth. The propensity score is estimating using probit regression and includes *lnMortgages*, *lnPopulation*, *lnAGI*, *CollEd*, $\%\Delta\text{HP}_{2007-2012}$, indicator variables for each decile of those five variables, and state indicator variables. Row *80% (LMI)* presents the base estimates using the true LMI threshold of 80% of the median MSA income. Row *70%* presents estimates using the a placebo LMI threshold of 70% of the median MSA income. All observation in Row *70%* are below the true LMI threshold and are thus all in the treatment group in all other tests. Row *90%* presents estimates using the a placebo LMI threshold of 90% of the median MSA income. All observation in Row *90%* are above the true LMI threshold and are thus all in the control group in all other tests. *Base* tests estimate the results in the framework described above. *Within 10k MSA* and *Within 5k MSA* respectively divide the sample into $10k and $5k strata based on MSA-median income and require matches be within the same strata. *Nt* and *Nc* represent the number of treatment and control observations for each test.

	Base		Within 10k MSA		Within 5k MSA	
Threshold	ATET	Nt / Nc	ATET	Nt / Nc	ATET	Nt / Nc
70%	0.09 (0.19)	883 / 1102	0.09 (0.39)	784 / 993	0.01 (0.94)	784 / 993
80% (LMI)	0.28*** (<0.01)	1987 / 11726	0.31*** (<0.01)	1864 / 11219	0.22*** (<0.01)	1823 / 10698
90%	0.09** (0.01)	2085 / 9641	0.03 (0.48)	2048 / 9186	0.01 (0.78)	2018 / 8807

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Alternative Matching Strategies

This table presents matching estimates for complaints (*lnComplaints*) for a given five-digit zip code (zip5) for LMI zip codes as compared to matched non-LMI zip codes, where *LMI* is an indicator variable equal to one for zip codes with a majority of its population in low- to moderate-income tracts in 2010. The matching methods in Columns (1) and (2) use a kernel-weighted propensity score to construct counterfactuals to compute the average treatment effect on the treated (ATET) using gaussian kernels with 0.01 and 0.05 bandwidths, respectively. Columns (3) and (4) using nearest neighbor propensity score matching with one and three nearest neighbors, respectively. The propensity scores are estimating using probit regression and includes *lnMortgages*, *lnPopulation*, *lnAGI*, *CollEd*, $\%\Delta HP_{2007-2012}$, indicator variables for each decile of those five variables, and state indicator variables. Column (5) performs mahalanobis matching using the five continuous variables, and Column (6) further requires matches be in the same \$5,000 income strata. The number of matched observations vary according to the available matches using a particular strategy.

	Kernel Bandwidth		PS Nearest Neighbor		Mahalanobis	
Desc	bw=0.01 (1)	bw=0.05 (2)	PS-1NN (3)	PS-3NN (4)	NN (5)	NN, 5k strata (6)
LMI (atet)	0.28*** (<0.01)	0.28*** (<0.01)	0.33*** (<0.01)	0.27*** (<0.01)	0.16*** (<0.01)	0.17*** (<0.01)
N	13713	13713	13668	13663	13713	13650
Ntreat	1987	1987	1946	1941	1987	1986
Ncontrol	11726	11726	11722	11722	11726	11664

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A Appendix: Complaint Example and Further Tables

The following is an example complaint presented exactly as it appears in the Consumer Financial Protection Bureau (CFPB) Database. This particular example includes a consumer narrative, which consumers were given the option to display starting from 2015. For such data entries, the five-digit zip code is redacted to a three-digit zip code to protect the anonymity of the filer. The observations used in our analysis have no consumer narrative and so will include all the data items below with the consumer narrative left blank. The database can be viewed and downloaded from `http://www.consumerfinance.gov/data-research/consumer-complaints/`

Date received	8/9/2016
Product	Mortgage
Sub-product	Conventional fixed mortgage
Issue	Application, originator, mortgage broker
Sub-issue	[blank]
Consumer complaint narrative	I contacted Wells Fargo Home Mortgage to refinance my current mortgage. I informed them that I was behind and I filed bankruptcy about 6 years ago. They said no problem they could do it. My credit score was XXXX which they said was good. I applied and then they said that I had to pay {$11.00} for the credit report and {$530.00} for the appraisal in order to continue on with the application process. So I paid them the {$540.00}. I never received the appraisal on my home. They never contacted the appraisal company to schedule a date or time. They denied my application based on behind on mortgage and bankruptcy. I would like my {$530.00} back for the appraisal the I never received. I have called several times and left messages and no one has returned my calls. What does Wells Fargo do with all the money they get from people that don't qualify for refinance?
Company public response	Company has responded to the consumer and the CFPB and chooses not to provide a public response
Company	Wells Fargo & Company
State	ND
ZIP code	580XX
Tags	[blank]
Consumer consent provided?	Consent provided
Submitted via	Web
Date sent to company	8/9/2016
Company response to consumer	Closed with monetary relief
Timely response?	Yes
Consumer disputed?	No
Complaint ID	2050804