DIFFERENTIAL ACCESS TO CAPITAL FROM FINANCIAL INSTIUTIONS BY MINORITY ENTREPRENEURS

by

Darius Palia[*]

January 2016

Abstract

This paper examines if Minority small business borrowers have the same access to loans from financial institutions as similar White borrowers. Using matching methods, we find that African-American borrowers are rejected at a higher rate (17-33% higher) than similar risk White-owned firms. We find that the impact of unobservable variables has to be greater than 50%-60% the impact of observable variables to show no discrimination. This bound seems to be a high number given that we have controlled for a comprehensive list of borrower credit risk variables that include their actual credit score and wealth. If the effect of the unobservable variables is less than this bound, then we have found a *causal impact* of discrimination for African-American borrowers. No such differential effect is found for Hispanic borrowers. Lumping both races together can give misleading results. Finding higher credit risks, higher rejection rates and equal expected losses for African-American-owned firms is consistent with the statistical or information based theory of discrimination.

_____

1. Introduction

The importance of small business to economic growth and employment has long been understood by both policy makers and academics. For example, Federal Reserve Chairwoman Janet Yellen recently stated "After the onset of the crisis, the Federal Reserve took extraordinary steps to stabilize the financial system and halt the plunge in economic activity. Since then the Fed has continued to use its monetary-policy tools to promote the recovery… Crucial to this process… is job creation. … Small businesses, of course, are responsible for a large share of these new jobs." (National Small Business Week event held at the U.S, Chamber of Commerce, May 15, 2014). But banks are a critical part of financing for small firms. According to the 2003 Survey of Small Business Finance, 57% of debt funding for U.S. small businesses is from banks.

But do entrepreneurs of different races (namely, African-American, Hispanic and White) have similar access to loans? The Equal Credit Opportunity Act of 1974 (ECOA), as amended, prohibits discrimination against any applicant with respect to any credit transaction. The Federal Institutions Examination Council (FIEC) manual issued by the regulators explains that ECOA 'applies to any extension of credit, including extensions of credit to small businesses.'"[1]

This paper makes the following contributions to the existing literature.[2] First, we believe we are the first paper to use causal inference methods to test if race-based discrimination exists in access to loans from financial institutions.[3] With regard to these causal inference methods Imbens

---

[1] See Federal Institutions Examination Council, Intraagency Fair Lending Examination Procedures, August 2009, page i.

[2] See Section 3 of this paper for details on the literature on discrimination.

[3] See Imbens and Ruben (2014) and Section 2 of this paper for a detailed explanation of these methods.

and Rubin (2014 p.7) state: "Another interesting comparison is to the "but for" concept in legal settings. Suppose someone committed an action that is harmful, and a second person suffered damages. From a legal perspective, the damage that the second person is entitled to collect is the difference between the economic position of the plaintiff had the harmful event not occurred (the economic position 'but for' the harmful action), and the actual economic position of the plaintiff. Clearly, this is a comparison of the potential outcome that was not realized and the realized potential outcome, this difference being the *causal effect* of the harmful action (emphasis added)." In our setting, we are able to mimic in actual data an experimental audit study wherein similar credit-risk firms are randomly assigned to lenders but differ only in one dimension, namely, the race of their principal owner (the Imbens and Rubin (2014) 'but for' argument). We test and find African-Americans to have significantly different characteristics/risks than White borrowers, making the probit/logistic regression coefficients biased (see Rubin (1973a,b, 1979), Imbens and Wooldridge (2008) and Section 2 of this paper for details). Additionally, in the criticism of experimental audit studies, Heckman and Seligman (1992) and Heckman (1998) argue that testers are either consciously or unconsciously trained to look for effects that are consistent with their beliefs, is generally not market data, and that it is extremely hard to erase all possible differences in the audit pair due to unobservable effects. Given that we are creating a pseudo-random experiment using actual market data and can measure the impact of unobservable variables on our results, we are not subject to these criticisms and have improved on the experimental audit study.

Second, and importantly, we check whether unobservable variables significantly bias our tests of discrimination. These unobservable variables could be quantifiable (such as the lenders screening methodology) but are not included as we do not have access to such data, or nonquantifiable (such as the borrower's bargaining power). Although we have included a large set

of covariates one cannot completely control for all unobservable variables. Therefore, we bound our point estimates. In doing so, we are the first paper to use the Rosenbaum bounds (2002)[4] to capture the impact of unobservable variables on the probability of being denied credit. Third, we separately analyze African-American-owned firms and Hispanic-owned firms rather than lumping them together as minorities. Finally, we use two databases, the Survey of Small Business Finance (SSBF) and the Kaufman Firm Survey (KFS), and find similar results for African-American borrowers. Therefore, our results, which are admittedly based on relatively small non-overlapping samples, are robust.

We find the following results. First, we find that African-American-owned firms have higher credit risks than White-owned firms. They also face a much higher denial rate than White-owned firms. Second, controlling for observable credit risk differences, we find that African American-owned firms have lower access to capital than White-owned firms. They are generally rejected at a higher rate (17-33% higher) than similar risk White-owned firms. Third, using the Rosenbaum (2002) bounds test, we find that the impact of unobservable variables has to be greater than 50%-60% the impact of observable variables to show no discrimination. This bound seems to be a high number given that we have controlled for a comprehensive list of borrower credit risk variables that include their actual credit score and wealth. If the effect of the unobservable variables is less than this bound, then we have found a causal impact of race on access to capital for African American-owned small business firms. Fourth, we find that Hispanic-owned firms do not generally have higher credit risks than White-owned firms. The only variable of statistical significance is that they have a lower wealth. They also do not face a higher denial rate than White-owned firms.

---

[4] See Section 5.3 of this paper for further details.

This suggests that controlling for credit risk, there is no causal impact of access to capital between Hispanic-owned and White-owned firms. Lumping Hispanic-owned with African-American owned firms as minority-owned firms is therefore not optimal. Fifth, our results are robust to three different causal inference methods, namely, propensity score matching, inverse probability weighting matching, and nearest neighbor matching. Sixth, we find a similar expected loss of default between African-American-owned and White-owned firms. Finding higher credit risks, higher rejection rates and equal expected losses for African-American-owned firms is consistent with the statistical or information-based theory of discrimination.

There are three limitations of this study. First, on the one hand, we have used Dun and Bradstreet's business credit score which is independently verified by KFS, and is not self-reported by the borrower or the lender and is therefore not subject to manipulation. On the other hand, the lending officer might rely on a credit bureau such as Equifax, Experian or Trans Union. There might not be a high correlation between these credit scores. Second, the number of African American and Hispanic principal owners who borrow from banks in our sample is small. Finding a statistical significant result in a small sample increases the chances that the actual impact might be higher if we had a larger sample. But having a small sample suffers from the limitation that these observations may not be representative of the full population. That said, what we need and have for good matching is a large sample of White borrowers for each Minority borrower to match to. Finally, our finding of no discriminatory effect for Hispanic borrowers relies on our classification that all Hispanics are the same. It is often argued that there are significant differences between the sub-categories of Hispanics such as Mexicans, Cubans, Salvadorans, etc. We are unable to make such distinctions due to data limitations.

The paper proceeds as follows. Section 2 explains the causal inference methods and Section 3 describes the related literature. Section 4 describes our data and variables used in the analysis, and Section 5 presents our results. Section 6 concludes.

2. Causal Inference Methods

Causal inference is inherently a comparison of potential outcomes (see Rubin (1974) and Holland (1986)). The causal effect for individual i is the comparison of individual i's outcome if the individual receives the treatment ($Y_i(1)$), and the same individual's outcome if she receives the control ($Y_i(0)$). Causality is hence defined as equal to $Y_i(1) - Y_i(0)$. In our setting, outcomes are if the borrower applicant is granted or denied the loan; and treatment is if the applicant is a minority borrower and control is if the applicant is a white borrower. The causal question is whether being a minority borrower increases your loan denial rate, ceteris paribus. One observes that minorities had higher loan denial rates on applying for a loan. The question of causality implies that had the borrowers not been minorities, ceteris paribus, they would not have been rejected at such high rates.

But estimating causality is challenging. The fundamental problem of causal inference (Holland 1986) is that for each individual borrower we can observe only one of these potential outcomes. In our setting, we can observe if a minority or white borrower was either denied or granted a loan, and not both. The estimation of causal effects can thus be thought of as a missing data problem (Rubin 1976). We would like to compare treated (R=1 for minority) and control (R=0) groups *who are very similar in observable risk variables X.*[5] We use three matching methods which we describe below.

---

[5] A number of papers including Cochran and Rubin (1973), Rubin (1973a,b, 1979), Heckman et al. (1998), Rubin and Thomas (2000) and Rubin (2001) have shown that linear regressions and their adjustments can increase the bias in the estimated treatment effect especially when there are significant differences between

Propensity score matching: The first method is the propensity score method. Let us define the propensity score $\rho(X)$ as the selection probability conditional on the observable risk variables X, namely, $\rho(X) = \Pr(R = 1 \mid X)$. Under two assumptions, one, the treatment assignment R is independent of potential outcomes $Y_i(1)$ and $Y_i(0)$, namely $Y_i(1), Y_i(0) \perp R \mid X$; and two, often called the "common support" assumption wherein there is a positive probability of receiving treatment for all values (i.e., $0 < \Pr(R=1 \mid X) < 1$). Rosenbaum and Rubin (1983) prove that conditioning on the propensity score is equal to having independence between the treatment indicator R and potential outcomes $Y_i(1), Y_i(0)$. In other words,

$$\Pr(R=1 \mid Y_i(1), Y_i(0), \rho(x)) = (Y_i(1), Y_i(0)) \perp R \mid \rho(x)$$

The above equation results in a substantial reduction in the dimensionality of the matching variables as one does not need to match on the covariates. In the actual estimation we use a logistic regression to estimate $\rho(X)$. Abadie and Imbens (2012) derive the standard errors for such an estimate.

Inverse probability weighting estimator: The inverse probability weighting estimator uses weighted averages of the observed outcome variable to estimate the means of the potential outcomes. Each weight is the inverse of the estimated probability that an individual is a minority. Outcomes of individuals who are more likely to be a minority borrower get a weight close to one, and outcomes of individuals who are more likely to be a white borrower receive a weight greater than one. As the propensity score $\rho(X)$ is the probability of being a minority borrower conditional on the observable risk variables X, the weights are therefore $\rho(X)$ and $1 - \rho(X)$.[6]

---

the means and variances between the treated and control groups. We show in section 5.1 of this paper that there are significant differences between such observable characteristics of our sample.

[6] See Busso, DiNardo and McCrary (forthcoming) for conditions wherein this method works well.

Nearest neighbor matching estimator: Rather than conditioning on the probability of being a minority, the nearest neighbor matching estimator conditions on the covariates X directly. This estimator is nonparametric in that it has no functional form (as opposed to the logistic regression we used in the above two methods), but comes with a price. It converges to the true value at a rate of square root of the sample size.[7] We use matching with replacement to reduce the bias.

3. Related Literature

The study of discrimination in law and economics began with the seminal paper by Becker (1957).[8] Under what is now often called the taste-based theory of discrimination, the motive that drives the discriminatory behavior is animus or prejudice towards a particular group. That is, lenders bear some psychic costs of prejudice when interacting with certain minority groups. Whereas the motive under Becker (1957) is explicit, Bertrand, Chugh, and Mullainathan (1995) offer a refinement of motive, namely that discrimination can take place through "implicit" attitudes which are unconscious mental associations towards agents of a certain group. These implicit associations are often fleeting and impulsive rather than deeply thought and rational. Under this theory, lenders inaccurately perceive Minority borrowers to have lower credit risks than White borrowers and therefore reject them more often. Therefore the loans that are approved to Minority borrowers have to have a higher profitability or a lower expected rate of default than loans that are approved to White borrowers. Accordingly, if taste-based discrimination is to be confirmed we

---

[7] See Abadi and Imbens (2006, 2011).

[8] The theoretical papers generally focus on the labor market between employers and employees. For ease of explanation, we rephrase their arguments in terms of lenders and borrowers.

would expect to find Minorities to have higher loan rejection rates and *lower* expected default rates than Whites.

Alternatively, under the information-based or statistical discrimination theory of Phelps (1972), Arrow (1972), and Bordalo, Gennaioli and Shleifer (2014), the motive that drives the agent's behavior is expected profit maximization. In an imperfect information world, economic agents discriminate against certain groups because they believe or speculate these groups have lower credit quality, which will reduce their profit. In Arrow (1972) and Phelps (1972), stereotypes are accurate as they fill up for missing information. In Bordalo, Gennaioli and Shleifer (2014), the decision maker only recalls the group's most representative or distinctive group characteristic. Stereotypes are inaccurate because the decision maker's beliefs overreact to information that confirms the stereotype (exhibiting 'confirmation bias') and their beliefs ignore any information that contradicts the stereotype (exhibiting 'base rate neglect'). Interestingly, in these models the lender's biased beliefs are confirmed in equilibrium, even though the ex ante probability of loss is identical. If information-based or statistical discrimination is to be confirmed, we would expect to find Minorities to have higher loan rejection rates and expected default rates that are equal to those of Whites.

Empirical studies have examined discrimination in different settings, from home mortgages, to the goods and services market, to labor market hiring. With respect to home mortgages, studies have found that the likelihood of denial by a lender is higher if the borrower is from a minority racial group (for example, Black et al. (1978), Holmes and Horvitz (1994), Berkovec et al. (1998), Munnell et al. (1996), Ross and Yinger (1999), Clarke, Roy and Courchane et al. (2009), and Hubbard, Palia and Yu (2012)).The above papers conduct regression analysis using actual market data. In contrast, Ayres and Siegelman (1995) present evidence from a paired

audit experiment that shows new car dealerships in Chicago quote significantly lower prices to White males than to African-American or female buyers. List (2004) finds that minorities received lower offers in the baseball sports card market than majorities. Zusmann (2013) finds evidence that fictitious advertisements of used cars by Arab sellers got lower responses than similar fictitious advertisements by Jewish sellers. Ayres, Banaji and Jolls (2011) find that baseball cards photographed held by an African-American hand sold for less than cards photographed held by a White hand in eBay auctions. Goldin and Rouse (2000) find that hiding the identity of a musician via a screen increased the probability of female musicians being hired by symphony orchestras. Fershtman and Gneezy (2001) find evidence of ethnic discrimination by Israeli Jewish males towards men of Eastern origin in a game of trust. In response to help-wanted ads, Bertrand and Mullainathan (2004) find that "black-sounding" names such as Lakisha and Jamal were less likely to be called back for job interviews than "white-sounding" names such as Emily and Greg. In different experiments, Gneezy, List and Price (2012) find discriminatory behavior against females, sexual orientation, disabled, non-whites and the elderly in different settings. They find support for information-based discrimination when the source of discrimination is uncontrollable (such as race, gender), and for taste-based discrimination when it is perceived to be controllable (such as sexual orientation).

More directly relevant to this study, a number of papers have examined whether AfricanAmerican and Hispanic small business borrowers have the same loan acceptance rates as White small business borrowers.[9] Bates (1997), Bostic and Lampani (1999), Cavalluzo and

---

[9] A couple of studies ask different questions using the KFS sample. Cheng (2014) finds that business closure rates for minority borrowers are not higher than white borrowers. Bates and Robb (2014) examine the impact of the 1977 Community Reinvestment Act, and find that firm location in a predominantly minority neighborhood is not a statistically significant predictor of loan application outcomes.

Cavalluzo (1998), Blanchflower, Levine and Zimmerman (2002), Cavullozo and Wolken (2005), and Blanchard, Zhao, and Yinger (2008) find evidence in support of discrimination as African American small business borrowers have a lower loan acceptance rate than White small business borrowers. All the above papers are limited in that they cannot claim a causal effect of discrimination, as the race variable is correlated with the credit-risk covariates, (a result we will show later). As they are not an experimental audit study, they are unable to create a random sample of observationally equivalent small business borrowers that only differ in their race. By using our causal inference methodology, we are able to create a pseudo-random experiment using actual market data that successfully addresses the criticism of experimental audit studies by Heckman and Seligman (1992) and Heckman (1998). Additionally, we ensure that our results are not generally impacted by unobservable variables.

4. Data and Variables

We use two databases, the 2003 Survey of Small Business Finance (SSBF03) and the Kaufman Firm Survey from 2008 to 2011 (KFS08). SSBF03 has a much larger set of variables than KFS08, but as we will see later, SSBF03 does not have enough observations for Hispanic borrowers. Unfortunately the Federal Reserve has discontinued the publication of the Survey of Small Business Finance. Accordingly, in order to examine the more recent period and Hispanic borrowers we use KFS08.

The Survey of Small Business Finance is the most extensive public data set available on small business finance. The survey asks small business respondents to report on their borrowing experiences within the preceding three years. Our SSBF03 sample consist of 1,743 observations, of which 42 are African-American, 6 are Hispanic, and 1,695 are White borrowers. As part of an

effort to understand the contribution of small business entrepreneurship and innovation to the U.S. economy and job creation, the Ewing Marion Kauffmann Foundation conducts a panel survey of new startup firms founded in 2004 and tracks them through 2011. The Kaufmann Firm Survey (KFS08) first surveyed about 5,000 firms and followed them for eight years. But data on whether a bank borrower applied for a loan and was accepted or rejected is available only from 2008 onwards, restricting our sample to the four years 2008-2011. Our KFS08 sample consist of 729 observations. Of this, 46 are African-American, 35 are Hispanic, and 648 are White borrowers.

For both the SSBF03 and KFS08 samples, we create a variable, *reject*, that is a dummy variable that is set to unity if loan application is denied, and set to zero if it is always approved or sometimes approved or denied. We also create a dummy variable *African-American*, which is set to unity if the borrower is African-American, and set to zero if the borrower is White. A similar dummy variable for Hispanic borrowers (*Hispanic*) is created. All variables are summarized in Table 1, with Panel A having the definitions for SSBF03, and Panel B the definitions for KFS08.[10]

*** Table 1***

To create our covariates, we begin by examining the more comprehensive database SSBF03. We first include the borrower's credit risk variable, *credit risk*, which we define as Dun & Bradstreet's business credit score. SSBF has the following categories: one equal to the most risky scores, namely 0-10; two equal to 11-25; three equal to 26-50; four equal to 51-75; five equal to 76-90; and six equal to the least risky scores, namely, 91-100. Note that in SSBF03, a higher value for *credit risk* denotes more risky firms. We control for the borrowing firm's size (*size*) which is defined as the natural logarithm of total assets, and age (*age*) defined as the number of

---

[10] We have a few borrowers who appear more than once in the KFS08 sample. When we only include the first observation for such borrowers, all of our results generally hold (results not reported but available from the author).

11

years the business has been established under the current owner. We set a dummy variable *company* that is set to unity if the organization form of the borrower is a limited liability company, and set to zero otherwise. To further control for the firm's financial health we include: the firm's profitability (*profit*), defined as the ratio of net income to total assets; *debt*, defined as the ratio of all liabilities excluding equity to total assets; *cash*, the ratio of cash holdings to total assets; and *f_bankrupt*, which is set to unity if in the past seven years the firm has declared bankruptcy, or if the firm in the past three years has had any business obligation due for 60 days or more, or has any business judgments rendered against them, and zero otherwise.

We also include the risk characteristics of the business owner. We begin by including a variable that captures the borrower's wealth (*wealth*), defined as the net worth of the borrower that does not include the value of her primary home or the current business. We also include a variable that captures the personal financial distress of the business owner. We define a variable *p_bankrupt*, which is set to unity if in the past seven years the business owner has declared bankruptcy, or if in the past three years the business owner has had any business obligation due for 60 days or more, or has any business judgments rendered against her, and zero otherwise.

To control for rents from future income, we include *education*, defined as the borrower's education level. SSBF03 has the following categories for the highest level of education the borrower has received: one equal to less than high school; two equal to high school graduate or GED; three equal to some college but no degree; four equal to associates degree in occupational or academic fields; five equal to associates degree in trade or vocational fields; six equal to bachelors; and seven equal to master's and other postgraduate degrees. We control for differences in the borrower's experience by including the number of years she has worked managing or owning a business (*experience*).We include a dummy variable *gender* that is set to unity if the borrower is

female, and zero if the borrower is a male.[11] It is possible that managerial agency issues of corporate ownership impacts the effort put forward by the borrower in making her firm successful. We therefore include as a control variable, *ownership*, defined as the percentage ownership of the borrower in the firm.

We also control for lender characteristics that might influence whether a small business owner is granted credit or not. Petersen and Rajan (1994) show that lending rates to small businesses are lower when the borrower and the bank have a longer relationship. They also show that this relationship is affected by the market power of the lending institution. Accordingly, we define a variable *relation* which is calculated as the number of months the borrowing firm has had a lending relationship with their lender; and a variable *hhi* equal to one if the Herfindahl Hirschman Index (HHI) is between 0 and 1000, equal to two if HHI between 1001 and 1799, and equal to three if HHI greater than or equal to 1800. Berger et al. (2005) shows that small banks are better in processing soft information of small business borrowers by being close to them geographically and communicating with them in person rather than the more distant large banks. We therefore include the variable *distance* defined as the distance in miles between the borrower's and lender's offices.

To control for industry effects we set a dummy variable equal to one if the two-digit SIC code is in mining (10-14); equal to two if the two-digit SIC code is in construction (15-19); equal to three if the two-digit SIC code is in transportation/public utilities (40-49); equal to four if the two-digit SIC code is in wholesale trade (50-51); equal to five if the two-digit SIC code is in retail trade (52-59); equal to six if the two-digit SIC code is in fire, insurance & real estate (60-69); equal

---

[11] All our results hold when we exclude the gender dummy variable.

to seven if the two-digit SIC code is in services (70-89); equal to eight if two-digit SIC code is in public administration (91-98); and equal to 10 if unclassified.

As we noted before, KFS08 has a smaller subset of covariates than SSBF08 and in many cases the two databases have different definitions. That said, the variables reject, African-American, Hispanic, company, gender, and ownership have similar definitions between the two databases. We now explain the different KFS08 definitions. KFS08 has the following six categories of credit score: one equal to 91-100 percentile; two equal to 71-90 percentile; three equal to 31-70 percentile; four equal to 11-30 percentile; five equal to 1-10 percentile; and six equal to open bankruptcy, out of business and highest risk. Note that in KFS08 a higher credit risk value suggests lower borrower risk, whereas in SSBF03 it suggests a higher credit risk. KFS08 defines a variable size which is equal to one if firm asset size is equal to $500 or less; equal to two if firm asset size is between $501 and $1,000; equal to three if firm asset size is between $1,001 and $3,000; equal to four if firm asset size is between $3,001 and $5,000; equal to five if firm asset size is between $5,001 and $10,000; equal to six if firm asset size is between $10,001 and $25,000; equal to seven if firm asset size is between $25,001 and $100K; equal to eight if firm asset size is between $100,001 and $1M; and equal to nine if firm asset size is greater than $1,000,000.

KFS08 has the following categories for both gross profit and loss of the business: zero equal to $0; one equal to $500 or less; two equal to $501-$1,000; three equal to $1,001-$3,000; four equal to $3,001-$5,000; five equal to $5,001-10,000; six equal to $10,001-$25,000; seven equal to $25,001-$100K; eight equal to $100,001-$1M; and nine equal to $1,000,001 and more. For each of them we use midpoint of each category and then take gross profits less losses (if any) to calculate *profit*. KFS08 has the following categories for *wealth* of the borrower: one equal to negative or zero net worth; two equal to $41-$50K; three equal to $51-$100K; four equal to

$100,001-\$250K; and five greater than \$250K. KFS08 has the following categories for *education*: one equal to less than 9th grade; two equal to some high school but no diploma; three equal to high school graduate or GED; four equal to technical or vocational degree; five equal to some college but no degree; six equal to associate's degree; seven equal to bachelor's degree; eight equal to some graduate school but no degree; equal to master's degree; and ten equal to professional school or doctorate. E*xperience* is defined as the number of years the business owner has worked in the same industry.

Unfortunately, KFS08 does not provide the actual SIC codes, but provides an industry variable that is a categorized industry variable. Accordingly, we create two dummy variables, *hightech* and *mid-tech* to control for industry differences. The dummy variable *high-tech* is set to unity if the business is in two-digit SIC codes consisting of chemical and allied products, industrial machinery and equipment, electrical and electronic equipment, and instruments and related products, and set to zero otherwise. The dummy variable *mid-tech* is set to unity if the business is in two-digit SIC codes consisting of petroleum and natural gas operations, cigarettes, miscellaneous textile goods, pulp, mills and miscellaneous converted paper products, petroleum refining and miscellaneous petroleum and coal products, nonferrous rolling and drawing, ordnance and accessories, motor vehicles and equipment, aircraft and parts, guided missiles, space vehicles and parts, miscellaneous transportation equipment, computer and data processing services, engineering and architectural services, research and testing services, management and public relation, and services (not elsewhere classified) and set to zero otherwise.

5. Tests and Results

5.1 Are there observable differences between Minority borrowers and White borrowers?

We begin by examining if there are differences between the observable credits risks and firm characteristics between minority African-American borrowers and White borrowers. The results of such an analysis are given in Table 2. In Panel A we present the results for differences between African-American borrowers and White borrowers using the SSBF03 sample, in Panel B differences between African-American borrowers and White borrowers using the KFS08 sample, and in Panel C the differences between Hispanic borrowers and White borrowers using the KFS08 sample. We use four types of tests. We explain them in the context of African-American borrowers and White borrowers using the SSBF03 sample, but they are similar for testing differences in the other samples.

*** Table 2***

The first is a simple t-test in order to check if the means between African-American borrowers and White borrowers are the same. It is given by $\frac{x_a - x_b}{(\frac{\sigma_a^2}{n_a} + \frac{\sigma_b^2}{n_b})^{1/2}}$ where x_a is the mean value for an African-American borrower, x_b is the mean value for a White borrower, σ_a is the standard deviation for an African-American borrower, σ_b is the standard deviation for a White borrower, and n_a and n_b is the sample size of African-American and White borrowers, respectively. In Panel A, we find that African-American borrowers have a statistically significant higher rejection rate (42.9%) than White borrowers (5.7%). African-American borrowers also have higher credit risks, higher personal bankruptcy problems, have a higher probability of being female, have less experience, have lower wealth levels, own a higher percentage of their firm, own firms that are smaller, younger but with higher cash holdings, and have less of a relationship with their lender, than White borrowers. African-American borrowers have equal: education levels, firm

profitability, firm bankruptcy issues, firm leverage and their lenders are in equally concentrated markets when compared to White borrowers. The above t-test might potentially suffer from finding statistical significance because the sample sizes n_a and n_b might be large. Therefore, Imbens and Wooldridge (2008) suggest using the normalized difference in means test, defined as $\frac{x_a - x_b}{(\frac{\sigma_a^2 + \sigma_b^2}{2})^{1/2}}$. They suggest that if the absolute value of the normalized difference is greater than 0.25, then it shows that the means are statistically significant. We find similar results using the normalized difference test as in the simple t-test, suggesting that sample size issues are not overstating our results for differences between African-American and White borrowers.

The third and fourth tests evaluate if the distributions are the same between African American and White borrowers. The difference in dispersion test checks if there are differences in the standard deviations between African-American and White borrowers. It is defined as $\ln(\sigma_a) - \ln(\sigma_b)$. The fourth test is the Kolmogorov-Smirnov test which is defined as $\sup|F_{a,}(x) - F_{b,n'}(x)|$, where $F_{a,n}$ and $F_{b,n'}$ are the empirical distribution functions of African-American and White borrowers, respectively. Both tests show that the distributions of the credit risk and firm variables are significantly different between African-American and White borrowers.

In Panel B, we repeat the analysis for African-American and White borrowers using the less comprehensive KFS08 sample. One again, we find that African-American borrowers have a statistically significant higher rejection rate (41.3%) than White borrowers (14.8%). On the reduced set of covariates we find that African-American borrowers still have higher credit risks, have lower wealth levels and own a higher percentage of their firm, than White borrowers. We now find an equal probability of African-American borrowers being female when compared to their White counterparts, and a similar result with regard to borrower experience.

We now repeat the analysis for differences between Hispanic and White borrowers in Panel C. We find no significant differences between the rejection rate for Hispanic and White borrowers. The only statistically significant difference in the t-test is that Hispanic borrowers have a lower wealth level than White borrowers.

The above results show that there are significant univariate differences between the credit risks and firm characteristics of African-American borrowers and White borrowers, but there are less significant differences between the credit risks and firm characteristics of Hispanic and White borrowers. We now test if these differences hold up in a multivariate setting using a logistic regression. Table 3 presents the results of such an analysis, wherein we present the marginal coefficients only.[12]

<div align="center">*** Table 3***</div>

For African-American borrowers, we present the results using the SSBF03 sample in column (1), and results using KFS08 sample in column (2). We still find that African-American borrowers have higher credit risks than White borrowers in both samples. In the KFS08 sample, we find that African-American borrowers have lower wealth levels, have higher education levels and own a higher percentage of their firm than White borrowers. No such significant effect is found for the SSBF03 sample. These results are robust to including industry dummy variables. Examining for differences between Hispanic and White borrowers in column (3) we find the former to have lower wealth levels. No other variables including credit risk is statistically significant between Hispanic and White borrowers.

---

[12] Although many studies present a pseudo R^2 measure, limited dependent regressions such as logistic are well known to have low pseudo R^2 measures econometrically (see Greene 2011). Additionally, it is challenging to find statistically significant results in a multivariate setting, given that we have 42 (46) African-American borrowers in the SSBF03 (KFS) sample, and 18 (11) independent variables, respectively.

In summary, the above results suggest that African-American borrowers have higher credit risks than White borrowers. This result is against the commonly held argument that those African American borrowers who apply for small business loans are better credit risks than Whites who apply for small business loans. Given that it is reasonable that those African-American borrowers who apply are likely to be better than those African-American borrowers who do not apply, it is likely that the latter group (whom we do not observe) have credit risks even worse than the White borrowers who apply. This common argument does not distinguish between the credit risks of White borrowers who do apply or not for a small business loan.

5.2 Causal inference tests to examine whether Minority borrowers are rejected more often than comparable White borrowers?

Before one can run the three causal inference methods, Imbens and Ruben (2014) suggest that we should check if the two distributions have common support. We use the min-max condition to find common support. For the SSBF03 sample, we find that the minimum predicted value for African-American borrowers is 0.001125 and the maximum predicted value is 0.484648. Similarly, we find the minimum predicted value for White borrowers is 0.00000 and the maximum predicted value is 0.505226. The min-max criteria for common support is therefore 0.001125, 0.484648. We find that 120 of the White borrowers were excluded due to the lack of common support which is confirmed in the kernel density plot of propensity scores in Figure 1. A similar analysis is undertaken using the KFS08 sample, and the results are given in Figures 2 and 3.

Figures 1, 2 and 3

We estimate three causal inference regressions to test whether Minority borrowers are rejected more often than comparable White borrowers. The results of such an analysis are given in Table 4. The first method, namely the propensity score method, matches the Minority borrower

with the closest White borrower by propensity score. We find that an African-American borrower is rejected at a 22-24% higher rate than a similar White borrower. No significant difference for loan rejection rates are found between similar Hispanic and White borrowers. We also estimate the inverse probability weighting estimator. Consistent with our previous result, we find that an African-American borrower is rejected at a 21-33% higher rate than a similar White borrower, whereas there is an insignificant difference for loan rejection rates between similar Hispanic and White borrowers. Using the nearest-neighbor matching procedure, we find that an African American borrower is rejected at a 17-33% higher rate than a similar White borrower, whereas there is an insignificant difference for loan rejection rates between Hispanic and White borrowers.

*** Table 4***

In summary, the above results suggest that an African-American borrower is rejected at a higher rate (between 17-33%) than a similar White borrower, suggesting differential access to capital for such borrowers. In contrast, there are no significant differences in loan rejection rates between similar Hispanic and White borrowers. This also suggests that one should not lump the two sub-classes of minorities together.

5.3 Impact of unobservable effects on the differential impact on loan rejection rates of African-American and White borrowers

It is possible that the discriminatory effect that we find for African-American borrowers is because there are unobservable variables that are not included in the two databases. For example, Hubbard, Kuttner and Palia (2002) find that low-capital banks charge higher loan rates to information-captured small borrowers when these banks are hit by a negative shock. Gan and Riddiough (2008) suggest that lenders possess proprietary credit quality information embedded in

their screening technologies that is not observable to empirical researchers. The two databases do not provide the name of the lender, and we therefore are not able to control for their characteristics. Such excluded or unobservable variables could impact our finding for differential access to capital. For the logistic distribution, Rosenbaum (2002) shows that the odds ratio of two matched individuals (i, j) of receiving treatment is bounded as follows: $\frac{1}{e^\gamma} \leq \frac{pi(1-pj)}{pj(1-pj)} \leq e^\gamma$ where p_j and p_i are the probability of being treated, and γ is the sensitivity of being treated to unobservable variables. If e^γ is equal to unity there is no bias, and increasing e^γ reflects the bias of overtreatment in our results. The overtreatment bias can be calculated using the Mantel and Haenszel test statistic. We provide the p-values of this statistic at different levels of e^γ in Table 5. At the ten-percent level of significance we find that the impact of unobservable variables has to be greater than 50% (SSBF03 sample), and greater than 60% (KFS08 sample) to have an impact on our results. This suggests that the impact of unobservable variables has to be greater than 50%60% to show no discrimination. The 50%-60% bounds seems to be a high number given that we have controlled for a comprehensive list of borrower credit risk variables that include their actual credit score and wealth. If the effect of the unobservable variables is less than these bounds, we have found a causal impact of race on access to capital for African-American-owned small business firms.

*** Table 5***

5.4 Tests of taste-based and information-based discrimination theories

The above results have shown that African-American borrowers are rejected at a higher rate than similar White borrowers. There are two possible theories of discrimination that can potentially explain this result. The first is Becker's (1957) taste-based discrimination, wherein lenders reject African-American borrowers due to animus and/or prejudice and are unwilling to

bear some psychic costs to interact with such borrowers. The second theory is statistical or information-based discrimination wherein lenders apply group stereotypes in evaluating loan applications (Arrow 1972, Phelps 1972, Bordalo, Gennaioli and Shleifer 2014). In order to distinguish between the taste-based and information-based theories, we create the expected loss of default for the more comprehensive SSBF03 sample.[13] We have shown that African-American borrowers have higher credit risks than White borrowers. . If taste-based discrimination exists, we should find higher rejections rates and *lower* expected default losses for African-American borrowers than similar White borrowers. Conversely, if statistical or information-based discrimination is to be confirmed we should find higher rejections rates and *equivalent* expected loss of default for African-American borrowers than similar White borrowers.

The expected loss of default is defined as the probability of default times the size of the loan. As in Dichev (1998) and Griffin and Lemmon (2002), we use Ohlson's (1980) O-score to proxy for the probability of default.[14] The variable O-score is defined as -1.32 – 0.407log(assets) + 6.03(liabilities/assets) - 1.43(working capital/assets) + 0.076(current liabilities/current assets) – 1.72(1 if liabilities > assets, 0 otherwise) – 2.37(net income/assets) – 1.83(funds from operations/liabilities) + 0.285(1 if a net loss for last two years, 0 otherwise) – 0.521(change in net income from current year to previous year/ absolute value of change in net income from current year to previous year). We find the average O-score for African-American borrowers to be 0.464, which is statistically significantly higher (at the five percent level) than the 0.283 average O-score for White borrowers. African-American borrowers are found to have higher probabilities of

---

[13] Due to data limitations we are unable to calculate the expected rate of return or the actual rate of default.

[14] We are unable to use other measures of default such as Altman's z-score because they require the market value of the firm. Our borrowers are small businesses who are not traded.

defaults at the fifth, tenth, twenty-fifth, fiftieth, and seventy-fifth percentile levels than White borrowers.[15] When we examine loan size, we find that African-American borrowers obtain smaller loans on average (approximately \$542K) than White borrowers (approximately \$1.74M).

We now conduct the three causal inference tests while replacing the dependent variable loan denial rates with the expected default loss rates. The results of such an analysis are given in Table 6. In all three tests we find that African-American borrowers have statistical insignificantly different expected default losses than White borrowers. Finding a higher rejection rate and equal expected loss for African-American-owned firms is consistent with the statistical or information-based theory of discrimination.

6. Conclusions

This paper uses two databases to examine if race-based discrimination exists in lending by financial institutions to small businesses. We are the first paper to use matching methods to test if there is a *causal impact* of race on access to entrepreneurial loan capital from financial institutions. These methods allow us to control for observable differences in credit risks between Minority- and White-owned startup firms. In other words, we are able to mimic in actual data, an experimental audit study wherein similar credit-risk firms are randomly assigned to lenders but differ only in one dimension, namely, the race of their principal owner. In the criticism of experimental audit studies, Heckman and Seligman (1992) and Heckman (1998) argue that testers are either consciously or unconsciously trained to look for effects that are consistent with their beliefs, and that it is extremely hard to erase all possible differences in the audit pair due to unobservable

---

[15] At the ninetieth percentile, both African-American and White borrowers have a probability of default close to unity.

effects. Given that we are creating a pseudo-random experiment using actual market data and can measure the impact of unobservable variables on our results, we are not subject to these criticisms and have improved on the experimental audit study.

We find that African-American-owned firms have higher credit risks than White-owned firms and have a higher loan rejection rate than White-owned firms. This suggests that the previous literature's probit/logistic regression coefficients are biased. Controlling for observable credit risk differences, we find that African-American-owned firms have lower access to capital than White-owned firms. They are generally rejected at a higher rate (17-33% higher) than similar risk White-owned firms. Third, using the Rosenbaum (2002) bounds test we find that the impact of unobservable variables has to be greater than 50%-60% the impact of observable variables to show no discrimination. This bound seems to be a high number given that we have controlled for a comprehensive list of borrower credit risk variables that include their actual credit score and wealth. If the impact of unobservable variables is less than this bound, we have found a causal impact of race for African-American small business borrowers. We also find no causal impact of access to loans from financial institutions between Hispanic-owned and White-owned firms. Lumping Hispanic-owned with African-American-owned firms as minority-owned firms is therefore not optimal. Our results are robust to three different causal inference methods, namely, propensity score matching, inverse probability weighting matching, and nearest neighbor matching. Finally, we find a similar expected loss of default between African-American-owned and White-owned firms. Finding higher credit risks, higher rejection rates and equal expected losses for African American-owned firms is consistent with the statistical or information-based theory of discrimination.

There are three possible legal remedies to deal with such disparate impact. The first is class action litigation. Finding a causal negative effect for access to loans from financial institutions for African-American borrowers is unlikely to be sufficient for plaintiffs to file a class action suit under the disparate impact theory of discrimination. The Supreme Court's decision in *Wal-Mart Stores, Inc. v. Dukes*[16] has made such legal theory based on statistical evidence harder to successfully litigate. In the last couple of years, class certification in three fair lending class actions were denied by district courts on the basis of the Dukes case. The first appeal concerned the denial of class certification in In re Countrywide Financial Mortgage Lending Practices Litigation.[17] The U.S. Court of Appeals for the Sixth Circuit held that the district court did not abuse its discretion by finding that *Dukes* foreclosed the plaintiffs' proposed class because "the mere presence of a range within which acts of discretion take place will not suffice to establish commonality."[18] The second appeal was from the decision to decertify the class following the *Dukes* ruling in *Barrett v. H&R Block, Inc,*[19] but instead of taking the appeal and ruling on its merits, the U.S. Court of Appeals for the First Circuit denied the petition.[20] The third appeal was from the decision to deny

---

[16] 131 S. Ct. 2541 (2011). See also Ropiequet, Navaja, and Noonan (2013) for extensive discussion of this case on fair lending litigation.

[17] No. 08-MD-1974, 2011 U.S. Dist. LEXIS 118695 (W.D. Ky. Oct. 11, 2011), *aff'd*, Miller v. Countrywide Bank, N.A. (In re Countrywide Financial Corp. Mortg. Lending Practices Litig.), 708 F. 3rd 704 (6th Cir. 2013); see also Ropiequet, Navaja, and Noonan (2013) p. 646-47.

[18] Miller v. Countrywide Bank, N.A., i*d.*

[19] No. 08-cv-10157-RWZ, 2011 U.S. Dist. LEXIS 30713 (S. Mass. Mar. 21, 2011), *class decertified sub nom.* Barrett v. Option One Mortg. Corp., No 08-cv-10157-RWZ, 2012 U.S. Dist. LEXIS 132775, at *13 (D. Mass. Sept. 18, 2012); see also Ropiequet, Navaja, and Noonan (2013) p. 645.

[20] Barrett v. Option One Mortg. Corp., No 12-8033 (1st Cir. Feb. 7, 2012) (judgment) (unpublished). That effectively ended the litigation. See Stipulation of Dismissal with Prejudice at 1, Barrett v. Option One Mortg. Corp. No. 08-cv-10157-RWZ (D. Mass. Aug. 21, 2013).

class certification in *Rodriguez v. National City Bank*,[21] where the parties had agreed to a class settlement and the court had scheduled a fairness hearing before the *Dukes* decision came down.[22] However, the U.S. Court of Appeals for the Third Circuit found no obstacle to affirming the denial of class certification,[23] noting that the case bore "a striking resemblance to *Dukes*."[24]

The second potential legal remedy is to file a case and use it to build a legal precedent to ultimately change the law at the Appellate or the Supreme Court level. This method is time consuming and requires a cohort of plaintiffs and lawyers who are resolute enough to change the law. The third potential legal remedy is to use regulatory institutions like the newly created Consumer Financial Protection Bureau or the Federal Reserve to empirically examine a much larger sample of loans to check for any disparate impact to African-American borrowers. Given the current legal case history and the costs of legal intervention, this legal remedy seems to be a logical first step.

---

[21] F.R.D. 148 (E. D. Pa. 2011), *aff'd*. 726 F. 3d 372 (3rd Cir. 2013); see also Ropiequet, Navaja, and Noonan (2013) p. 646-47.

[22] Rodriguez, v. National City bank, 726 F. 3d 372, 377-82 (3rd Cir. 2013) at 151-53.

[23] *Id.* evaluating this argument.

[24] *Id* at 384.

References

Abadie, A., and G.W. Imbens, 2006, "Large sample properties of matching estimators for treatment effects," *Econometrica*, 74, 235-67.

Abadie, A., and G.W.Imbens, 2011, "Bias-corrected matching estimators for average treatment effects," *Journal of Business and Economic Statistics*, 29, 1-11.

Arrow, K.J., 1972, "Models of job discrimination," in A. Pascal (ed.), *Racial discrimination in economic life*, Lexington MA: Lexington Books.

Ayres, I., Banaji, M. and C. Jolls, 2011, "Race effects on eBay," working paper, Yale University.

Ayres, I. and P. Siegelman, 1995, "Race and gender discrimination in bargaining for a new car," *American Economic Review*, 85(3), 304–21.

Bates, T., 1997, "Unequal access: Financial institution lending to black and white-owned small business start-ups," *Journal of Urban Affairs*, 19(4), 487-495.

Bates, T. and A. Robb, 2014, "Has the Community Reinvestment Act increased loan availability among small businesses operating in minority neighborhoods?" *Urban Studies,* 1-20.

Bordalo, P., Gennaioli, N., and A. Shleifer, "Stereotypes", working paper, Harvard University.

Berger, A., Miller, N., Petersen, M., Rajan, R., and J. Stein, 2005, "Does function follow organizational form? Evidence from the lending practices of large and small banks," *Journal of Financial Economics*, 76, 237-69.

Bostic, R., and P. Lampani, 1999, "Racial differences in patterns of small business finance," in J. Blanton, A. Williams and S. Rhine (Eds.), *Business Access to Capital and Credit*, Washington D.C.: Federal Reserve System, 149-79.

Becker, G.S., 1957, *The Economics of Discrimination*, Chicago, IL: The University of Chicago Press.

Bertrand, M. and S. Mullainathan, 2004, "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination," *American Economic Review*, 94(4), 991–1013.

Bertrand, M., Chugh, D., and S. Mullainathan, 2005, "Implicit discrimination," *American Economic Review*, 95(2), 94-98.

Berkovec, J.A., Canner, G.B., Gabriel, S.A., and T. Hannan, "Discrimination, competition, and loan performance in FHA mortgage lending," *Review of Economics and Statistics*, 870(2), 241250.

Black, H.A., Schweitzer, R.L., and L. Mundell, 1973, "Discrimination in mortgage lending," *American Economic Review*, 68(2), 186-91.

Blanchard, L., Zhao, B., and J. Yinger, 2008, "Do lenders discriminate against minority and women entrepreneurs?," *Journal of Urban Economics*, 63, 467–97.

Blanchflower, D., Levine, P., and D. Zimmerman, 2003, "Discrimination in the small business credit markets," *Review of Economics and Statistics*, 85, 930-943.

Busso, M., DiNardo, J., and J. McCrary, forthcoming, "New evidence on the finite sample properties of propensity score reweighting and matching estimates," *Review of Economics and Statistics.*

Cavalluzzo, K., and L. Cavalluzzo, 1998, "Market structure and discrimination: The case of small businesses," *Journal of Money, Credit, and Banking*, 30, 771–92.

Cavalluzzo, K., and J. Wolken, 2005, "Small business loan turn-downs, personal wealth and discrimination," *Journal of Business*, 78, 2153-78.

Cheng, S., 2014, "Potential lending discrimination? Insights from small business financing and new venture survival," *Journal of Small Business Management*. DOI: 10.1111/jsbm.12112.

Clarke, J.A., Roy, N., and M.S. Courchane, 2009, "On the robustness of racial discrimination findings in mortgage lending studies," *Applied Economics*, 41, 2279-97.

Cochran, W., and D. Rubin, 1973, "Controlling bias in observational studies; A review," *Sankhya: The Indian Journal of Statistics, Series A*, 35, 417-46.

Dichev, I., 1998, "Is the risk of bankruptcy a systemic risk?" *Journal of Finance*, 53(3), 1131-47.

Fershtman, C., and U. Gneezy, 2001, "Discrimination in a segmented society: An experimental approach," *Quarterly Journal of Economics*, 116, 351-377.

Gan, J., and T. Riddiough, 2008, "Monopoly and informational advantage in the residential mortgage market," *Review of Financial Studies,* 21(6), 2677-703.

Greene, W. H., 2011, *Econometric Analysis*, Prentice Hall, 7th edition.

Griffin, J., and M. Lemmon, 2002, "Book-to-market equity, distress risk, and stock returns," *Journal of Finance*, 57(4), 2317-36.

Gneezy, U., List, J., and M. Price, 2012, "Towards an understanding of why people discriminate: Evidence from a series of natural field experiments," NBER working paper 17855.

Goldin, C., and C. Rouse, 2000, "Orchestrating impartiality: The impact of "blind" auditions on female musicians," *American Economic Review*, 90(4), 715–41.

Heckman, J. and Siegelman, P., 1992. 'The Urban Institute audit studies: Their methods and findings," in M. Fix and R. Struyk, Eds., *Clear and convincing evidence: Measurement of discrimination in America.* Lanham, MD: Urban Institute Press, 187-258.

Heckman, J., 1998, "Detecting discrimination," *Journal of Economic Perspectives*, 12(2), 101-16.

Heckman, J., Ichimura, H., and P. Todd, 1998, "Matching as an econometric evaluation estimator," *Review of Economic Studies*, 65, 261-94.

Holland, P., 1986, "Statistics and causal inference," *Journal of the American Statistical Association*, 81, 945-60.

Holmes, A. and Horvitz, P., 1994, "Mortgage lending, race and model specification," *Journal of Finance*, 49(1), 81-99.

Hubbard, R.G., Palia, D., and W. Yu, 2012, "Analysis of discrimination in prime and subprime mortgage markets," working paper, Columbia University.

Hubbard, R.G., Kuttner, K. D. Palia, 2002, "Are there bank effects in borrowers' costs of funds? Evidence from a matched sample of borrowers and banks," Journal of Business, 75(4), 559-89.

Imbens, G., and J. Wooldridge, 2008, "Recent developments in the econometrics of program evaluation," NBER working paper 14251.

Imbens, G., and D. Rubin, 2014, *Causal Inference in Statistics, Biomedical and Social Sciences: An Introduction*. New York, Cambridge University Press.

List, J.A., 2004, "The nature and extent of discrimination in the marketplace: Evidence from the field," *Quarterly Journal of Economics*, 119(1), 49–89.

Munnell, A.H., Geoffrey, M.B., Tootell, Lynn E. Browne, and J. McEneaney, ''Mortgage lending in Boston: Interpreting HMDA data,'' *American Economic Review,* 86(1), 25–53.

Ohlson, J., 1980, "Financial ratios and the probabilistic prediction of bankruptcy," *Journal of Accounting Research*, 18(1), 109-31.

Phelps, E., 1972, "The statistical theory of racism and sexism," *American Economic Review*, 62(4), 659-61.

Petersen, M., and R. Rajan, 1994, "The benefits of lending relationships: Evidence from small business data," *Journal of Finance*, 49(1), 3-37.

Phelps, E.S., 1972, "The statistical theory of racism and sexism," *American Economic Review*, 62(4), 659–61.

Ropiequet, J.L., Naveja, C.S., and L.J. Noonan, "Fair lending developments: Enforcement Intensifies, Class Actions Diminish," *Business Lawyer*, 68(2), 637-48.

Rosenbaum, P., and D. Rubin, 1983, "The central role of the propensity score in observational studies for causal effects," *Biometrika,* 70, 41-55.

Rosenbaum, P., 2002, *Observational Studies*, 2nd Edition, New York, Springer.

Ross, S., and J. Yinger, 1999, "Does discrimination in mortgage lending exist? The Boston Fed study and its critics," in M. Turner and F. Skidmore, (Eds.), *Mortgage Lending Discrimination: A Review of Existing Evidence,* Urban Institute, Washington, DC, 43-99.

Rubin, D., 1973a, "Matching to remove bias in observational studies," *Biometrics*, 29, 159-84.

Rubin, D., 1973b, "The use of matched sampling and regression adjustment to remove bias in observational studies," *Biometrics*, 29, 185-203.

Rubin, D., 1974, "Estimating causal effects of treatments in randomized and nonrandomized studies," *Journal of Education Psychology*, 66, 688-701.

Rubin, D. and N. Thomas, 2000, "Combining propensity scores matching with additional adjustments for prognostic covariates," *Journal of the American Statistical Association*, 95, 57385.

Rubin, D., 2001, "Using propensity scores to help design observational studies: Application to the tobacco industry," *Health Services and outcomes research Methodology*, 2, 169-88.

Zussman, A., 2013, "Ethnic discrimination: Lessons from the Israeli online market for used cars," *Economic Journal*, 123, F433–68.

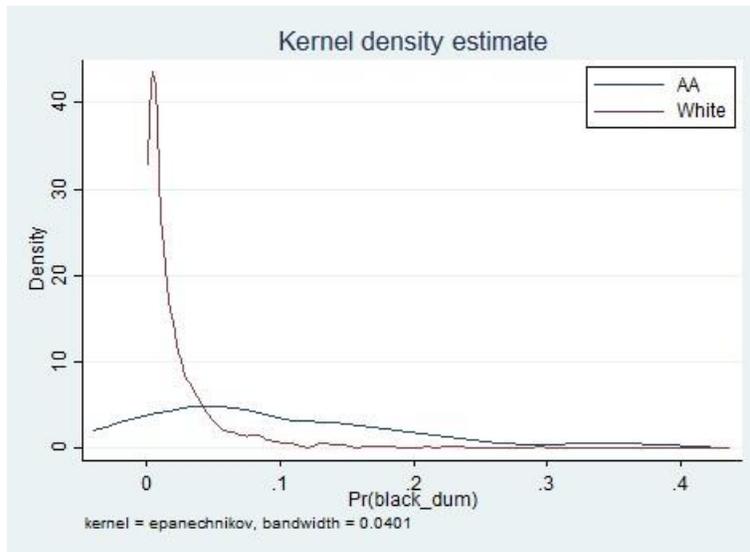**Figure 1. Differences between African-American and White borrowers (SSBF03 sample)**

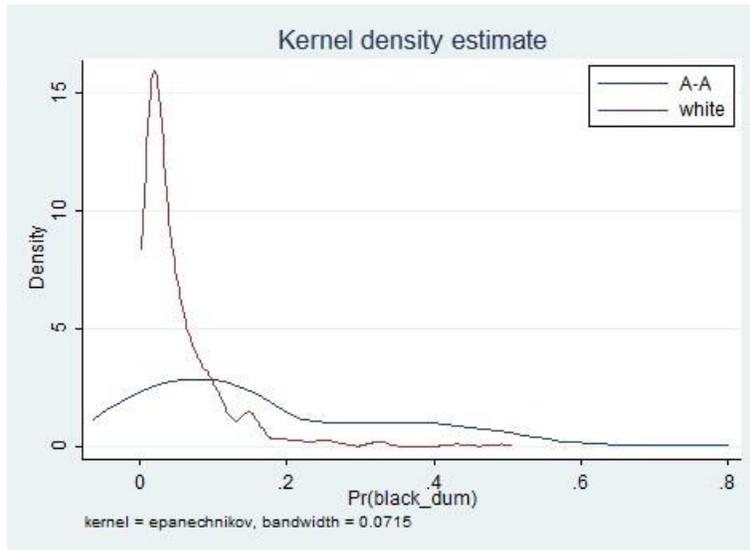**Figure 2. Differences between African-American and White borrowers (KFS08 sample)**

Figure 3. Differences between Hispanic and White borrowers (KFS08 sample)
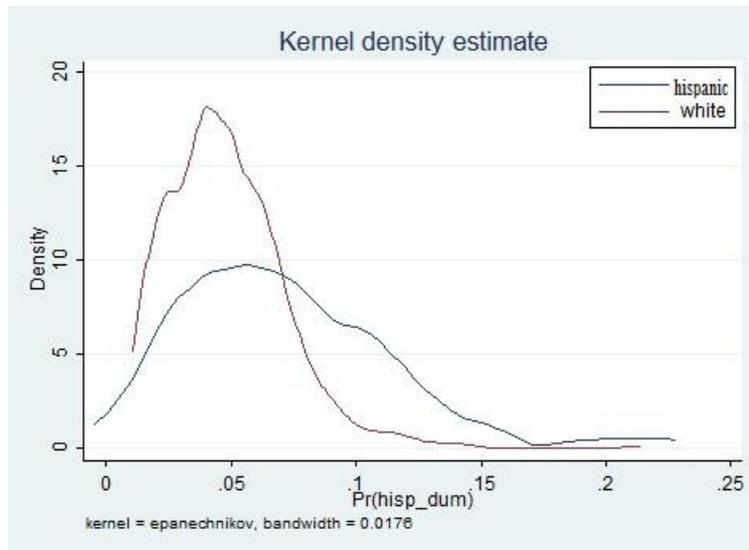
Table 1: Variable Names and Description

Name	Description
Panel A: Survey of Small Business Finances 2003(SSBF03)	
reject	Dummy variable set to unity if loan application is always denied, and set to zero if always approved or sometimes approved/denied
African-American	Dummy variable set to unity if borrower is African-American, and set to zero if borrower is White
credit risk	Dun & Bradstreet's business credit score. SSBF03 has the following categories: 1=most risky scores, namely 0-10; 2=11-25; 3=26-50; 4= 51-75; 5=76-90; 6= least risky scores, namely, 91100
size	Natural logarithm of total assets
age	Number of years the business has been established by current owner
company	Dummy variable set to unity if the business is set up as a corporation, and set to zero otherwise
profit	Ratio of net income to total assets
debt	Ratio of all liabilities excluding equity to total assets
cash	Ratio of cash holdings to total assets
f_bankrupt	Dummy set to unity if in the past seven years the firm has declared bankruptcy, or if the firm in the past three years has had any business obligation due for 60 days or more, or has any business judgments rendered against them
wealth	Natural logarithm of net worth. SSBF has the value of net worth of the owner that excludes the value of her primary home and current business
p_bankrupt	Dummy set to unity if in the past seven years the business owner has declared bankruptcy, or if in the past three years has had any business obligation due for 60 days or more, or has any business judgments rendered against her
education	Highest level of education the borrower has received. 1=less than high school degree; 2=high school or GED; 3= some college but no degree; 4=associate degree occupational/academic; 5= trade school/vocational; 6= bachelors; 7=graduate and post graduate degree
experience	Number of years the borrower has worked managing or owning a business, including this business
gender	Dummy variable set to unity if borrower is female, and set to zero if borrower is male
ownership	Percentage ownership of the borrower
relation	Number of months the borrowing firm has conducted business with lender

Table 1 (continued)

Name	Description
hhi	SSBF03 has the following categories: 1=if the Herfindahl-Hirschman Index (HHI) is between 0 and 999; 2= if HHI less than or equal to 1000 and less than 1800; 3= if HHI greater than or equal to 1800
distance	Distance in miles between borrower and lender offices
sic	=1 if two-digit SIC code is in mining (10-14); =2 if two-digit SIC code is in construction (15-19); =3 if two-digit SIC code is in transportation/public utilities (40-49); =4 if two-digit SIC code is in wholesale trade (50-51); =5 if two-digit SIC code is in retail trade (52-59); =6 if two-digit SIC code is in fire, insurance & real estate (60-69); =7 if two-digit SIC code is in services (70-89); =8 if two-digit SIC code is in public administration (91-98); =10 if unclassified

Panel B: Kaufmann Firm Survey 2008-2011 (KFS08)

Name	Description
reject	Dummy variable set to unity if loan application is always denied, and set to zero if always approved or sometimes approved/denied
African -American	Dummy variable set to unity if borrower is African-American, and set to zero if borrower is White
Hispanic	Dummy variable set to unity if borrower is Hispanic, and set to zero if borrower is White
credit risk	Dun & Bradstreet's business credit score. KFS08 has the following categories: 1=highest credit scores (91-100 percentile); 2= 71-90 percentile; 3-31-70 percentile; 4= 11-30 percentile; 5=1-10 percentile; 6=open bankruptcy, out of business and highest risk
size	KFS08 has the following categories for total assets: 1= $500 or less; 2= $501-$1,000; 3=$1,001$3,000; 4=$3,001-$5,000; 5=$5,001-10,000; 6=$10,001-$25,000; 7=$25,001-$100K; 8=$100,001-$1M; 9=$1,000,001 and more
company	Dummy variable set to unity if the business is set up as a corporation, and set to zero otherwise
profit	KFS08 has the following categories for both gross profit and loss of the business: 0=$0; 1= $500 or less; 2= $501-$1,000; 3=$1,001-$3,000; 4=$3,001-$5,000; 5=$5,001-10,000; 6=$10,001$25,000; 7=$25,001-$100K; 8=$100,001-$1M; 9=$1,000,001 and more. For each of them we use midpoint of each category and then take gross profits less losses (if any) to calculate net profits
wealth	KFS08 has the following categories for the net worth of the borrower: 1= negative or zero net worth; 2=41-$50K; 3=$50,001-$100K; 4=100,001-250K; 5=greater than $250K
experience	Number of years the borrower has worked in this industry
gender	Dummy variable set to unity if borrower is female, and set to zero if borrower is male
ownership	Percentage ownership of the borrower

Table 1 (continued)

Name	Description
high-tech	Dummy variable set to unity if the business is in 2-digit SIC codes consisting of chemical and allied products, industrial machinery and equipment, electrical and electronic equipment, and instruments and related products, and set to zero otherwise
mid-tech	Dummy variable set to unity if the business is in 2-digit SIC codes consisting of petroleum and natural gas operations, cigarettes, miscellaneous textile goods, pulp mills and miscellaneous converted paper products, petroleum refining and miscellaneous petroleum and coal products, nonferrous rolling and drawing, ordnance and accessories, motor vehicles and equipment, aircraft and parts, guided missiles, space vehicles and parts, miscellaneous transportation equipment, computer and data processing services, engineering and architectural services, research and testing services, management and public relation, and services (not elsewhere classified), and set to zero otherwise

Table 2: Univariate Differences between African-American and White Borrowers

Panel A: SSBF03 African-American v. White Borrowers

Variable	Means for African-American	Means for White	t-statistics for differences in means[1]	Norm-Differences[2]	Differences in dispersion[3]	Kolmogorov-Smirnov p-values[4]
reject	0.429	0.057	4.79[a]	0.556	0.768	0.000[a]
credit risk	3.000	3.917	3.92[a]	-0.400	0.022	0.002[a]
size	12.146	13.364	3.37[a]	-0.361	0.088	0.018[b]
age	2.308	2.603	2.48[b]	-0.252	-0.096	0.020[b]
company	0.310	0.309	0.08	0.009	0.017	1.000
profit	1.406	0.941	0.79	0.049	-0.897	0.290
debt	1.282	1.123	0.36	0.029	-0.531	0.000[a]
cash	0.220	0.120	2.02[b]	0.174	-0.416	0.127
f_bankrupt	0.262	0.219	0.62	0.071	0.073	1.000
wealth	11.253	12.678	2.54[b]	0.290	0.182	0.003[a]
p_bankrupt	0.333	0.088	3.31[a]	0.403	0.518	0.009[a]
education	4.690	4.696	0.01	-0.002	0.001	0.847
experience	2.660	2.998	3.60[a]	-0.359	-0.062	0.002[a]
gender	0.286	0.160	1.77	0.210	0.221	0.461
ownership	79.500	69.658	2.34[b]	0.246	-0.031	0.032[b]
relation	99.190	138.893	3.55[a]	-0.257	-0.642	0.237
hhi	1.810	2.018	1.66	-0.187	0.086	0.162
distance	17.000	14.212	0.30	0.301	-0.152	0.428
sic	7.214	5.443	6.87[a]	0.537	-0.317	0.000[a]

Panel B: KFS08 African-American v. White Borrowers

Variable	Means for African-American	Means for White	t-statistics for differences in means[1]	Norm-Differences[2]	Differences in dispersion[3]	Kolmogorov-Smirnov p-values[4]
reject	0.413	0.148	3.552[a]	0.397	0.337	0.003[a]
credit risk	3.239	2.647	3.412[a]	0.344	0.003	0.000[a]
size	6.652	7.668	3.00[a]	-0.366	0.632	0.004[a]
company	0.152	0.098	0.988	0.114	0.199	1.000
profit	50.42	156.54	1.554	-0.162	-0.042	0.098
wealth	3.391	4.208	3.195[a]	-0.365	0.365	0.003[a]
education	7.522	6.545	3.480[a]	0.341	-0.069	0.000[a]
experience	16.28	16.06	0.126	0.014	0.074	0.819
gender	0.174	0.193	0.331	-0.035	-0.031	1.000
ownership	83.72	75.94	2.711[a]	0.260	-0.170	0.012[b]
high-tech	0.261	0.271	0.148	-0.016	-0.002	1.000
mid-tech	0.391	0.196	2.625[a]	0.294	0.216	0.053

Panel C: KFS08 Hispanic v. White Borrowers

Variable	Means for Hispanic	Means for White	t-statistics for differences in means[1]	Norm-Differences[2]	Differences in dispersion[3]	Kolmogorov-Smirnov p-values[4]
reject	0.229	0.148	1.101	0.144	0.182	0.971
credit risk	2.857	2.647	1.040	0.128	0.027	0.284
size	7.400	7.668	1.408	0.162	-0.103	0.197
company	0.143	0.098	0.073	0.096	0.176	1.000
profit	18.25	156.54	1.599	0.198	0.072	0.182
wealth	3.514	4.208	2.836[a]	0.350	0.183	0.006[a]
education	6.429	6.545	0.342	0.042	-0.003	0.935
experience	14.11	16.06	1.107	0.130	-0.070	0.576
gender	0.171	0.193	0.329	0.040	-0.033	1.000
ownership	80.06	75.94	0.924	0.108	-0.079	0.086
high-tech	0.143	0.271	2.052[b]	0.219	-0.225	0.569
mid-tech	0.200	0.196	0.054	0.007	0.021	1.000

[a] statistically significant at the 1% level; [b] statistically significant at the 5% level; [c] statistically significant at the 10% level.

Table 3: Logistic Regression of Differences between Minority and White Borrowers

	African-American v. White		Hispanic v. White
	SSBF03	KFS08	KFS08
	(1)	(2)	(3)
credit risk	-0.004[b]	0.013[a]	0.004
	(-2.80)	(3.29)	(0.73)
size	0.001	-0.007	0.001
	(0.81)	(-1.64)	(0.18)
age	-0.000		
	(-0.08)		
company	0.006	0.021	0.024
	(1.42)	(1.03)	(1.00)
profit	-0.000	-0.000	-0.000
	(-0.73)	(-0.32)	(-0.92)
debt	-0.000		
	(-0.61)		
cash	0.016		
	(1.66)		
f_bankrupt	-0.005		
	(-0.95)		
wealth	-0.000	-0.012[a]	-0.013[b]
	(-0.92)	(-2.70)	(-2.26)
p_bankrupt	0.016[a]		
	(2.75)		
education	-0.001	0.010[a]	-0.000
	(-0.84)	(3.07)	(-0.11)
experience	-0.003	0.000	-0.000
	(-1.08)	(0.37)	(-0.52)
gender	0.002	-0.002	-0.014
	(0.47)	(-0.12)	(-0.76)
ownership	0.000	0.000[c]	0.000
	(1.34)	(1.82)	(0.68)
relation	-0.000		
	(-1.24)		
hhi	-0.004		
	(-1.50)		
distance	0.000		
	(0.81)		
sic	0.005[a]		
	(3.81)		
high-tech		0.031[b]	-0.030
		(2.03)	(-1.58)
mid-tech		0.038[b]	-0.006
		(2.47)	(-0.32)
observations	1,617		
pseudo R^2	0.166	0.167	0.057

[a] statistically significant at the 1% level; [b] statistically significant at the 5% level; [c] statistically significant at the 10% level.

Table 4: Causal Inference Tests for Differences in Loan Rejection Rates between Minority and White Borrowers

	African-American v. White		Hispanic v. White
	SSBF03	KFS08	KFS08
	(1)	(2)	(3)
Propensity score matching	0.238[b]	0.217[b]	0.057
			(0.57)
	(2.87)	(2.04)	
Inverse probability weighting matching	0.327[a]	0.211[a]	0.028
			(0.38)
	(4.05)	(2.80)	
	0.333[a]	0.174	
			0.014
Nearest neighbor matching	(3.58)	(1.59)	(0.15)

[a] statistically significant at the 1% level; [b] statistically significant at the 5% level; [c] statistically significant at the 10% level.

Table 5: Sensitivity of results to unobservable variables (Rosenbaum bounds, 2002)

| Value of bias (e^γ) | African-American v. White | | Hispanic v. White |
	SSBF03 (1)	KFS08 (2)	KFS08 (3)
1 (no bias)	2.100[b] (0.018)	2.283[b] (0.011)	0.554 (0.290)
1.05	2.013[b] (0.022)	2.193 (0.014)	0.479 (0.316)
1.10	1.918[b] (0.028)	2.096[b] (0.018)	0.404 (0.343)
1.15	1.828[b] (0.034)	2.002 (0.22)	0.332 (0.370)
1.20	1.742[b] (0.041)	1.914[b] (0.028)	0.263 (0.396)
1.25	1.660[b] (0.048)	1.829 (0.034)	0.197 (0.422)
1.30	1.581[c] (0.057)	1.748[b] (0.040)	0.134 (0.447)
1.35	1.501[c] (0.066)	1.670[b] (0.047)	0.074 (0.471)
1.40	1.433[c] (0.076)	1.595[c] (0.055)	0.015 (0.494)
1.45	1.363[c] (0.086)	1.523[c] (0.064)	-0.041 (0.516)
1.50	1.296[c] (0.096)	1.454[c] (0.072)	-0.096 (0.538)
1.55	1.231 (0.109)	1.387[c] (0.083)	-0.149 (0.559)
1.60	1.167 (0.122)	1.322[c] (0.093)	-0.200 (0.579)
1.65	1.106 (0.134)	1.260 (0.104)	-0.249 (0.598)
1.70	1.047 (0.147)	1.199 (0.115)	-0.297 (0.617)
1.75	0.990 (0.161)	1.141 (0.127)	-0.286 (0.613)
1.80	0.934 (0.175)	1.084 (0.139)	-0.242 (0.596)
1.85	0.881 (0.189)	1.028 (0.152)	-0.200 (0.579)
1.90	0.828 (0.204)	0.975 (0.164)	-0.158 (0.563)
1.95	0.778 (0.219)	0.922 (0.178)	-0.118 (0.546)
2.00 (large bias)	0.728 (0.233)	0.871 (0.192)	-0.078 (0.531)

[a] statistically significant at 1% level; [b] statistically significant at 5% level; [c] statistically significant at 10% level.

Table 6: Causal Inference Tests for Differences in Expected Default Losses between African American and White Borrowers

	SSBF03
Propensity score matching	-0.104
(coefficients x 10^6)	(-0.53)
Inverse probability weighting matching	-0.239
(coefficients x 10^6)	(-1.30)
Nearest neighbor matching	0.098
(coefficients x 10^6)	(0.80)

[a] statistically significant at the 1% level; [b] statistically significant at the 5% level; [c] statistically significant at the 10% level.