ESSAYS ON FINANCE, BUSINESS GROWTH, AND ENTREPRENEURSHIP

by

Mee Jung Kim
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Public Policy

Committee:

___________________________	John S. Earle, Chair
___________________________	Sita N. Slavov
___________________________	Anh Pham
___________________________	Constantine Yannelis, Ext. Reader
___________________________	Sita N. Slavov, Program Director
___________________________	Mark J. Rozell, Dean
Date: _____________________	Spring Semester 2020 George Mason University Fairfax, VA

Essays on Finance, Business Growth, and Entrepreneurship

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Mee Jung Kim
Master of Arts
Paul H. Nitze School of Advanced International Studies, Johns Hopkins University, 2011
Bachelor of Arts
University of Michigan - Ann Arbor, 2009

Director: John S. Earle, Professor
Scar School of Policy and Government

Spring Semester 2020
George Mason University
Fairfax, VA

DEDICATION

I dedicate my work to my loving grandfather Jung Chan Kim, my parents Tae Joon Kim and Nam Sook Han, whose love, support, and encouragement carried me through this journey.

ACKNOWLEDGEMENTS

There are so many people whom I want to thank for their support through this long journey. I am truly grateful for all people who provided indefinite love, support, and encouragement over the past seven years of my Ph.D. program. Nothing would have been possible without them.

First, I would like to express my deepest gratitude to my advisor, John S. Earle, for all of his support and guidance by providing opportunities to learn and grow intellectually as a scholar. It was him who introduced me to this world of empirical research. Under his supervision, I learned how to conduct rigorous empirical research with micro data, and most importantly how to think like a researcher and a scholar. He also inspired me with his passion toward research. I am indebted to him for his endless support, invaluable advice, and intellectual stimulus that helped to grow into the person and researcher I am today.

I am also grateful to have intellectual scholars and mentors who provided much-needed advices and guidance. I am especially thankful for my committee members Sita N. Slavov and Anh Pham, and my formal committee member, Connie L. McNeely for providing continued support, invaluable insights and encouragement. Without their support, I would not be able to finish this journey. I would also like to acknowledge and express gratitude to my external committee, Constantine Yannelis, for providing valuable advice for my dissertation.

I am incredibly grateful to have so much support and advice from economists and scholars at the Center for Economic Studies of the U.S. Census Bureau. My six years at the CES was full of amazing opportunities to work and interact with respected scholars. I would like to extend my gratitude to the CES team for their support that allowed me to conduct research using micro data at the Census. Special thanks to J. David Brown, Shawn Klimek, Emek Basker, Fariha Kamal and Emin Dinlersoz who provided thoughtful and kind guidance on my research and career.

I am grateful for my doctoral colleagues and staffs at George Mason University for inspiring me intellectually and providing support. I thank Kyungmin Lee, Solomiya Shpak, Monica Hyun-Ju Kim, Lisardo F. Bolaños, Lokesh Dani, not only for intellectual conversations but also emotional and moral support throughout the years. Your friendship encouraged me to take each next step throughout this long journey. I also thank Elizabeth

Eck Olchowski, Shannon Williams, and Michelle Ely for their support in facilitating the program's administrative work.

My research benefitted from financial support from the National Science Foundation (Grant 1262269 and 1719201), John E. Petersen Memorial Scholarship in Public Finance, Small Business Administration (SBA) Advocacy, and Schar School of Policy and Government at George Mason University. I am thankful for their financial support which enabled me to conduct my research more productively and effectively.

More than anything, I am truly thankful for my loving friends who have provided me with the strongest moral support over this long journey. I thank Catherine Y. Pyo, Sarah K.Yun, Eunice R. Kim, Hyun Kyong Lee, Jung Soo Kim, Han Jun Lee, Jin Kyung Min, Youlan No, and Shinyoung Lee for being there for me. Through the ups and downs, it was their love, support, and encouragement that kept me going.

Last but not least, I sincerely thank my loving family. I wish my grandfather was here to be with me, but I will always remember his love and support. I am truly blessed and thankful to have my parents, Tae Joon Kim and Nam Sook Han, and my brother Sungsoo Kim. I cannot imagine enduring through this process without their unconditional love and support. Thank you so much for always being there for me and being my family.

Finally, I would like to convey my heartfelt thanks to my loving fiancé Jun Sung Kim for being my best friend and supporter throughout this journey, especially in the hardest moments. I greatly appreciate his love, trust, support, patience, and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

American Community Survey	ACS
Annual Survey of Entrepreneurs	ASE
Business Register	BR
Business Research and Development and Innovation Survey	BRDIS
Characteristics of Business Owners	CBO
Community Innovation Survey	CIS
Community Reinvestment Act	CRA
Difference-in-Differences	DID
Dun and Bradstreet	D&B
Federal Deposit Insurance Corporation	FDIC
Federal Financial Institutions Examination Council	FFIEC
Federal Reserve Board	FRB
Home Mortgage Disclosure Act	HMDA
Kauffman Firm Survey	KFS
Linear Probability Model	LPM
Longitudinal Business Database	LBD
Lower- and Moderate Income	LMI
Median Family Income	MFI
Metropolitan Statistical Areas	MSAs
National Community Reinvestment Coalition	NCRC
Neighborhoods	NH
North American Industry Classification System	NAICS
Office of the Comptroller of the Currency	OCC
Office of the Thrift Supervision	OTC
Participant Statistical Areas Program	PSAP
Regression Discontinuity Design	RDD
Research and Development	R&D
Science, Technology, Engineering, and Mathematics	STEM
Survey of Business Owners	SBO
Survey of Small Business Finances	SSBF

ABSTRACT

ESSAYS ON FINANCE, BUSINESS GROWTH, AND ENTREPRENEURSHIP

Mee Jung Kim, Ph.D.

George Mason University, 2020

Dissertation Director: Dr. John S. Earle

This dissertation presents studies of the interactions of finance, business growth, and entrepreneurship. In the first essay, *Does the Community Reinvestment Act (CRA) Promote Small Business Growth in Lower-Income Neighborhoods*, I estimate the impact of the Community Reinvestment Act (CRA) on small business growth in "Low- and Moderate-Income" (LMI) neighborhoods in the United States. Using rich firm-level panel data on every U.S. employer, I use two principle identification strategies to estimate the effects on business employment. First, I exploit the sharp threshold cutoff for CRA eligibility based on median family income to use a regression discontinuity design (RDD) in an optimal bandwidth around the cutoff. Second, I exploit changes in CRA eligibility over time using difference-in-differences (DID) with firm fixed effects. Using both RDD and DID, I find that the firms located in the CRA eligible areas increase employment by about 0.8 percent compared to firms in non-CRA areas. The CRA effects are larger for young firms and firms in minority neighborhoods, which potentially face

higher barriers to access credit, with employment effects of about 1.5 and 1.9 percent, respectively. I also find that increases in lending in the CRA areas are related to the number of banks, implying that the channel is greater access to finance.

In the second essay, *The Financing of African American Entrepreneurship*, I study financing patterns by African American business owners more closely using rich data on U.S. businesses. I exploit unusually detailed owner characteristics and financing variables from the Annual Survey of Entrepreneurs (ASE) at the U.S. Census Bureau. The outcome variables include not only amount and source of startup capital but also source of new financing and other questions bearing on credit constraints. I estimate the differences in financing between African American and white business owners after controlling for a range of owner and firm characteristics, including age, gender, education, motivation, industry, and other choices owners make about their businesses. Although part of the gap is explained by different owner and firm characteristics, I still find that an unexplained gap in finance between African American and white entrepreneurs even with these controls, suggesting that African American entrepreneurs are more financially constrained.

In the third essay, *Immigrant Entrepreneurs and Innovation in the U.S. High-Tech Sector*, I and co-authors estimate differences in innovation behavior between foreign versus U.S.-born entrepreneurs in high-tech industries. Our data come from the Annual Survey of Entrepreneurs, a random sample of firms with detailed information on owner characteristics and innovation activities. We find uniformly higher rates of innovation in immigrant-owned firms for 13 of 14 different innovation measures; the only exception is for copyright/trademark. The immigrant advantage holds for older firms as well as for recent start-ups and for every level of the entrepreneur’s education. The size of the estimated immigrant-native differences in product and process innovation activities rises with detailed controls for demographic and human capital characteristics but falls for R&D and patenting. Controlling for finance, motivations, and industry reduces all coefficients, but for most measures and specifications immigrants are estimated to have a sizable advantage in innovation.

CHAPTER 1. DOES THE COMMUNITY REINVESTMENT ACT (CRA) PROMOTE SMALL BUSINESS GROWTH IN LOWER-INCOME NEIGHBORHOODS?[1]

1.1 Introduction

Financial access for individuals and businesses in lower-income communities has been a long-standing concern for policymakers and analysts. Small businesses in lower-income communities often lack sources of financing other than personal savings and wealth (Avery, Bostic, and Samolyk 1998). The Community Reinvestment Act (CRA) is a federal regulation to reduce credit discrimination and increase credit access in these areas. One of the major components of the CRA provides credits to small businesses. However, there is no study of the impact of the CRA on actual business performance at firm level in the CRA targeted areas.

In this paper, I estimate this impact using rich firm-level panel data on every U.S. employer between 2004 and 2015. My methods exploit the policy design of the CRA, where eligibility is designated to the "Lower- and Moderate Income" (LMI) areas based on the median family income at the Census tract level. About 30 percent of the tracts are

[1] I thank the National Science Foundation (NSF) for support (Grants 1262269 and 1719201 to George Mason University). Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the NSF or the U.S. Census Bureau or the NSF. All results have been reviewed to ensure that no confidential information is disclosed, and released under DRB Bypass Number CBDRB-FY19-570.

LMI tracts that are CRA eligible areas in the United States.[2] I use two principle identification strategies. First, I exploit the sharp threshold cutoff for CRA eligibility to use a regression discontinuity design (RDD). Second, I exploit the change in CRA eligibility over time using difference-in-differences (DID) with firm fixed effects. Combining RDD and DID, I estimate the impact of the CRA using firms located in switching eligibility tracts around the threshold cutoff.

I find that the CRA has a positive impact on business employment and employment growth. With RDD estimations, I find that the CRA increases annual employment growth by about 0.25 percent. Using RDD and DID along with firm fixed effects in a 5 percent bandwidth around the threshold, I find that firms in the CRA eligible tracts increase employment by about 0.8 percent compared to firms in non-CRA tracts. It is important to note that this is an intent-to-treat effect since I am unable to observe firms that actually received loans because of the CRA. The estimated CRA impact is on all firms in the tracts, regardless of whether they received loans due to the CRA, implying that the actual CRA impact could be larger than my estimated effect. I also provide evidence of the heterogeneous impact of the CRA across different groups of firms. I find larger CRA impacts on young firms and firms located in minority neighborhoods, which increase employment by about 1.5 and 1.9 percent, respectively, relative to firms in non-CRA neighborhoods. The results suggest that additional credit

[2] Using the U.S. Housing and Urban Development (HUD) data, LMI neighborhoods are defined by the Decennial Census 2010 and the author for this calculation.

access enabled by the CRA helps to grow the particular group of firms facing higher barriers to access to credit, such as young firms and firms in minority neighborhoods.

I also examine the mechanism underlying the estimated CRA impact. Using the same identification strategy, I find an increase in the number and amount of small loans and loans to small businesses, which suggests that financing is the channel of the CRA effect. This financing effect increases when the tract is part of a bank's assessment area examined by the CRA, which can be seen in the larger number of banks. As the banks play an essential role in financing, the CRA impacts can be larger in locations where more banks are available, such as large Metropolitan Statistical Areas (MSAs). Accordingly, I find that the firms located in CRA areas in large MSAs increase employment by about 1.4 percent relative to firms in non-CRA eligible areas. My analysis of the CRA thus demonstrates the impact of financing on business growth by providing evidence of causality running from finance to growth.

This paper contributes to the growing literature using firm-level data to examine the impacts of policies on business outcomes. For instance, Lelarge, Sraer, and Thesmar 2010, Bach 2013, and Brown and Earle 2017 study the impact of government programs on financial access, and Gamberoni et al. (2016) examine the impact of employment subsidies. Like these papers, my paper exploits policy design and the availability of firm-level data to estimate causal link from finance to growth.

No previous research on the CRA uses firm-level data, but there have been a number of studies using geographic aggregates. The impact of the CRA has been studied with different outcome measures, different data, and different methods. In particular, there are many studies examining the CRA impact on the supply of mortgage loans by looking at loan availability.[3] While a number of studies evaluate the CRA impact on mortgage loans in lower-income households and neighborhoods (Joint Center for Housing Studies 2002; Avery, Calem, and Canner 2003; Berry and Lee 2007; Bhutta 2011; Avery and Brevoort 2015), there are a relatively small number of studies on the CRA impact on small business lending. In addition to the findings of the positive association between the CRA and credit availability, this study demonstrates how increased credit accessibility affects business growth at the firm level.

Several papers have attempted to estimate the impacts of the CRA on small business outcomes, but these studies use aggregate data at the county level. Zinman (2002) looks at the causal effects of the CRA on small business lending and examines the real effect in terms of payroll and bankruptcy at the county level. He finds that the CRA increases small business lending as well as the county's overall payroll but decreases bankruptcies. More recently, Rupasingha and Wang (2017) examine the CRA impact on business growth outcome using county-level data, which is still at the aggregate level. While the CRA variation occurs at tract level, both of these studies examine the real outcome at the aggregated county level.

A few studies focus on the credit availability for small business lending in CRA eligible areas. Using the Kauffman Firm survey, Bates and Robb (2015) compare loan

[3] Studies are measured by the number and volume of loans in targeted areas and/or borrowers, and the findings on the impact of the CRA on mortgage loan accessibility in LMI neighborhoods and borrowers are mixed. While the Home Mortgage Disclosure Act (HMDA) data, which is the main data used for all of these studies, provides detailed information about borrowers, including race, gender, and income level, the CRA data only provides aggregate small business lending information at the tract level, making it difficult to investigate the impact of the CRA on businesses or particular groups of business owners.

availability between firms in minority neighborhoods and non-minority neighborhoods. However, they define the CRA neighborhoods based on minority share rather than the basis of CRA designation, which is income in the neighborhoods. Thus, it is hard to say that the difference is attributable to the CRA impact. Bostic and Lee (2017) estimate the impact of the CRA on small business lending with numbers and the volume of small business loans at the tract level in an RDD framework.

Several of these studies exploit the sharp income threshold, using RDD to estimate the CRA's impact (Avery, Calem, and Canner 2003; Berry and Lee 2007; Bhutta 2011; Avery and Brevoort 2015; Bostic and Lee 2017). While local RDD estimation creates credible "as if random" assignment around the threshold, the results are relevant only for neighborhoods close to the threshold, which makes it hard to draw conclusions about the broader population. Also, cross-sectional variation alone may be insufficient to tease out CRA impact. The status of CRA eligibility has been in place for ten years, implying that eligibility status may not necessarily represent the same lower-income neighborhoods the CRA targeted in the beginning. Thus, it is important to have an exogenous change in order to estimate the CRA impact before and after the change with time variation.

More recently, a number of studies (Ding and Nakamura 2017; Ringo 2017; Ding et al. 2019) use difference-in-differences to address the concern of local estimations only on the population around the threshold from the RDD framework. Using time variation allows them to employ whole neighborhoods, not just the neighborhoods around the threshold, to estimate the impact of the CRA, and the exogenous change helps to overcome potential endogeneity rising from the long time periods. However, all these previous studies exploit changes in CRA eligibility from a few MSA boundary changes, which provide relatively small variations over a short-term study period. When a major time variation in a CRA designation occurs during an update of income computation, tract boundaries are also changing, and these inconsistent tract boundaries make it nearly impossible to examine change at the tract level. The availability of firm-level microdata with current tract information enables me to use firm fixed effects to exploit a major time variation in CRA status from both tract boundaries change and recomputation of income in new designated areas. In this study, I use RDD and DID together to complement the potential concern of RDD and exploit exogenous major changes in eligibility with time variations to estimate the CRA impact.

I have organized this paper as follows: Section 2 describes the background and history of the CRA, and Section 3 describes the data. Section 4 explains the identification strategies, and Section 5 provides the main results and heterogeneity analysis. Lastly, Section 6 is a conclusion, followed by policy implications.

1.2 The Community Reinvestment Act (CRA)

The Community Reinvestment Act (CRA), enacted in 1977, is intended "to encourage depository institutions to help meet the credit needs of communities in which they operate, including Low- and Moderate-Income (LMI) neighborhoods."[4] With the

[4] The purpose of Community Reinvestment Act (CRA) is stated by the Federal Financial Institutions Examination Council at https://www.ffiec.gov/cra/history.htm

objective of meeting credit needs in localities, the CRA attempts to close the credit-need gap for residential mortgage loans and small business loans, which are the primary reasons that borrowers seek loans in LMI neighborhoods. An LMI neighborhood is defined at the census tract level as having a median family income of less than 80 percent of the median family income of the MSA to which it belongs. Following a major 1995 reform, the CRA requires each insured depository institution to meet the credit needs of its community by originating loans or purchasing loans originated by other lenders, and the institutions are assessed periodically by their federal supervisory agency to insure they are doing so.[5] All federally regulated depository institutions that are not small or specific-purpose institutions are required to report their lending activities[6] for data collection (Federal Financial Institutions Examination Council 2013). Also, public disclosure of a more quantitative analysis to evaluate each lender's performance is required. The CRA examination consists of three tests: a lending test, an investment test, and a service test. The lending test has the most weight, accounting for 50 percent of the overall rating.[7] This structure of examination induces bank institutions to take lending

5 There are four federal banking regulators: the Federal Reserve Board (FRB), the Office of the Comptroller of the Currency (OCC), the Federal Deposit Insurance Corporation (FDIC), and the Office of Thrift Supervision (OTS). An agency is assigned based on each lending institution's charter type.

6 CRA lending that is counted in the examination includes 1) mortgage, small business, farm, and consumer loans (here, small business lending is less than $1 million for business purpose); 2) loans to borrowers, businesses, or farms located in LMI neighborhoods as defined above; 3) loans to LMI borrowers, small businesses, or farms with revenues of $1 million or less; 4) loans with the purpose of serving community development; and 5) innovative and flexible lending practices. In essence, CRA lending is mainly evaluated in LMI neighborhoods and across LMI borrowers, including small businesses and farms. Further information is available at the official website of the Federal Financial Institutions Examination Council at https://www.ffiec.gov/cra/guide.htm

7 CRA exams consist of three tests: 1) a lending test that measures a bank's overall lending activities, including home mortgage, small business, and community development loans; 2) an investment test that examines a lender's qualified investments that benefit its assessment area and regional area; and 3) a service test that considers the scope of a lender's banking and community development service provision, such as service delivery through ATMs and branches and seminars and programs for financial literacy (Avery, Calem,

activities extra seriously in order to get better ratings from overall CRA assessments. More relevant to this study, banks are reviewed for the number of and amount of loans to small businesses in LMI neighborhoods, motivating them to expand lending activities in LMI neighborhoods.

The CRA examination also requires each bank to define its CRA assessment area. These geographic areas are what are actually accounted for by the examiners. Only a lender's activities in its assessment area will be counted toward the rating. The CRA assessment areas are the areas in which a bank has its main office, branches, or deposit-taking ATMs, as well as the surrounding areas in which the bank originates or purchases a substantial portion of its loans. Assessment areas are also periodically verified during the examinations to ascertain their validity.

A CRA rating may affect an institution's ability to expand, as "an institution's CRA record is taken into account in considering applications for deposit facilities, including mergers and acquisitions."[8] Delayed applications for these potential business activities can be costly for lenders, while an outstanding rating is considered to be "insurance" that limits challenges by regulatory agencies (Belsky, Lambert, and von Hoffman 2000; Immergluck 2004). Also, a rating determines the frequency of CRA examinations of an institution, with a better rating resulting in less frequent examination.[9]

and Canner 2003; Immergluck 2004). The assessment is rated either outstanding, high or low satisfactory, not satisfactory, need to improve, or not compliant. https://www.ffiec.gov/cra/default.htm

8 The purpose and background of the Community Reinvestment Act (CRA) can be found on its official website by the Federal Financial Institutions Examination Council at https://www.ffiec.gov/cra/history.htm.

9 Maintaining a greater than satisfactory rating limits examinations to every 24 to 36 months for institutions with total assets greater than $250 million, whereas a lower rating may lead to annual examinations. Institutions with total assets of $250 million are examined less frequently.

Moreover, public disclosure may result in pressure on financial institutions from neighborhood associations, civil-rights groups, and community development corporations. These community groups play an important role in lending institutions' participation in CRA agreements, which pledge future lending and activities to serve LMI and minority communities.[10] In sum, although there is no legal enforcement of CRA compliance, there are compelling incentives for lending institutions to comply and maintain good ratings and reputations.

1.3 Data

The main data for this analysis are from the Longitudinal Business Database (LBD) of the U.S. Census Bureau, which allows the tracking of firms over time. The LBD is an annual, longitudinally linked database covering all U.S. firms and establishments with payroll employment in the non-farm sector. The LBD employment measure is the number of employees in the period of March 12. While the LBD is based on the Business Register (BR) from the Census Bureau, it does not include tract information. Thus, I add the tract information of the LBD firms for the years 2004 to 2015. Moreover, the revenue-enhanced LBD permits me to use revenue information to define small businesses. According to the CRA, businesses with annual gross revenue of less than 1 million are small businesses. I focus on these small businesses by excluding

[10] The National Community Reinvestment Coalition (NCRC)'s CRA Commitments reports documented that banks made around $4.6 trillion of commitments in LMI and minority neighborhoods from 1992 to 2007.

big firms that always have more than 1 million in revenue across the 2004 to 2015 study period.

My sample includes both single-unit firms in single locations and multi-unit firms with multiple establishments in different locations. While more than 90 percent of the businesses are single-unit firms, it is still important to include multi-unit firms as a part of the analysis to reflect how firms can grow by adding establishments (Cao et al. 2019). However, the study is at the firm level, and it is essential to deal with these multi-unit firms, which have multiple locations of belonged establishments. According to the CRA Register, "the institution should record the loan location by either the location of the small business borrower's headquarters or the location where the greatest portion of the proceeds are applied, as indicated by the borrower" (Federal Register 2016). I assign the headquarter location of multi-unit firms across establishments by modal employment, where the largest employment establishment is located. After I assign the single headquarter location across establishments for multi-unit firms, I aggregate the employment at the firm level.

I link these data to publicly available Federal Financial Institutions Examination Council (FFIEC) Community Reinvestment Act (CRA) data. The CRA data provides tract level information. A tract's relative median family income, which is the most important information for this study, is used to define CRA eligible tracts in which a tract's median family income relative to its MSA is less than 80 percent. This also measures a tract's closeness to the income threshold of tracts below and above it. Next, the CRA provides the number and dollar amounts of small business loans at the tract level, allowing the measurement of differences in financing between CRA and non-CRA tracts. Finally, other tract level characteristics, such as population, share of population with Bachelor's degree or higher, share of minority population, and MSA median family income, are included in the regression as control variables. I exclude tracts in non-MSAs because most banks' assessment areas are in MSAs, where the CRA activities are actually examined by the regulators.

CRA tract eligibility is defined by the 2000 Census tract income levels with consistent tract boundaries for the years 2003 to 2011. Then, the same information is recomputed by the American Community Survey (ACS) 2006-2010 5-year estimates for the period between 2012 and 2015. The change in 2012 updated the median family income in the reference Census, which allows time variation in CRA eligibility, providing switchers and non-switchers in terms of CRA eligibility. Accordingly, there are "switching" firms, which are located in the switching tracts. This time-varying CRA status occurs in two sources:[11] 1) updates in tract-income or 2) changes in tract boundaries. I exploit these two sources of variation to estimate the impact of the CRA, which is described in the next section.

The main outcome variable is a firm's number of employees. In the RDD framework, annual employment growth[12] is also used as an outcome variable. As RDD is a cross-sectional analysis, it may solely capture the size effect when the average

[11] I dropped potential endogenous movers changing CRA eligibility due to change in business location before or after 2012.

[12] Here, this employment growth is computed where $g_{ijt} = \frac{emp_{ijt}-emp_{ijt-1}}{0.5(emp_{ijt}+emp_{ijt-1})}$, which is commonly referred to as a DHS growth rate (Davis et al 1996).

employment size for firms in the CRA eligible areas is slightly larger than the firms in the ineligible areas. Thus, employment growth, which captures flow, may be more appropriate to estimate the CRA effect in this setting. Conversely, in the time-varying framework of the DID setting, which is described in the next section, the number of employees is more appropriate to estimate the CRA impact from the differences in average employment before and after the change. An average difference in annual employment growth before and after the change would hardly capture the CRA impact, however, if growth occurs once after a change and then stays fairly steady in subsequent periods. Thus, I use employment as an outcome variable for the rest of the analysis in the DID setting.

1.4 Identification Strategy

1.4.1 Regression Discontinuity Design (RDD)

One of the main identification strategies for this study is a regression discontinuity design (RDD). A tract is eligible for the CRA if its percent of median family income is less than 80 percent of the median income of the MSA in which the tract is located.[13] The CRA eligible areas are relatively poorer areas compared to non-CRA areas. Accordingly, firms in these two areas can be systematically different, and the difference in outcomes between these two areas is more likely to be endogenous. To overcome this problem, I exploit the CRA regulatory framework with a sharp cutoff,

[13] If the tract is not located in the MSA, it is eligible if its MFI is less than 80 percent of the nonmetropolitan portion of the state in which the tract is located.

which allows an appropriate design for regression discontinuity where M_j is a tract's percent of median family income and D_j is defined by M_j as follows:

Equation 1

$$D_j = \begin{cases} 1 \ if \ M_j < 0.80 \\ 0 \ if \ M_j \geq 0.80 \end{cases}$$

The sharp discontinuity design allows the assumption that it is equivalent to random assignment (Angrist and Pischke 2009; Lee and Lemieux 2010) around the threshold, and the running variable is CRA eligibility of the tract where the firm is located. The key assumption for this RDD is the following:

Equation 2

$$\lim_{a \to 0} \left[E\left(Y_{ijt} \middle| 0.80 - a \ \leq M_j < 0.80\right) - E\left(Y_{ijt} \middle| 0.80 < M_j \leq 0.80 + a\right) \right] = 0$$

where the outcome should converge at the threshold point of 80 percent without the CRA, with areas slightly below and above the threshold expected to be similar in the absence of the CRA. Thus, the difference in outcome across the threshold between these two areas can be attributed to the CRA impact. For that reason, I assume smooth changes in tract characteristics between CRA and non-CRA tracts around the threshold except financial access and employment, which are my outcome variables. I conduct covariate balance test to verify this assumption, which will be discussed in a later section.

While the global polynomial provides a comprehensive approximation of the overall sample, the estimation can be driven by points far from the threshold points, which can lead to misleading estimation around the boundary (Cattaneo et al. 2018). Alternately, a local polynomial is employed a first-order function for the estimation around the cutoff threshold. This local polynomial estimation is not sensitive to outliers or any other extreme values, as it only uses the observations near the cutoff. I employ both the global polynomial and a local polynomial estimation to obtain overall and local estimations. Below is the RDD specification for the regression:

Equation 3

$$Y_{ijt} = \alpha + \beta D_{jt} + h(M_{jt}) + D_{jt} * h(M_{jt}) + X_{ijt}\theta + Z_{jt}\gamma + \rho_m + \delta_j + \tau_t + \varepsilon_{ijt}$$

where Y_{ijt} is the outcome variable measuring business growth, which is employment and annual employment growth, for firm *i* in tract *j* in year *t*. D_{jt} is a dummy equal to 1 if the firm *i* belongs to CRA eligible tract in year *t*, which is the variable of interest representing the CRA impact, and $h(M_{jt})$ is the function of the relative median family income ratio of the tract to its MSA in polynomial function. Then, the interaction term between D_{jt} and the income function of $h(M_{jt})$ addresses different functional forms for CRA and non-CRA tracts. Different orders of polynomials are used from first-order to fourth-order for each respective sample.[14] X_{ijt} includes firm age and age squared, and Z_{jt}

[14] Cattaneo et al. (2018) states that the global polynomial fit is an approximation using an unknown regression function based on the fourth- or fifth-order polynomial regression fit of the outcome, and linear function is more robust and less sensitive to boundary-related and overfitting problem, when it gets closer to the threshold.

is a set of tract characteristics, including population, MSA median family income, share of minority population, and share of population with Bachelor's or higher degree. Finally, ε_{ijt} is an idiosyncratic error term. The model also includes MSA (ρ_m), industry (δ_j), and time (τ_t) fixed effects. Here, the treated group is the firms in CRA eligible tracts and the control group is the firms in CRA non-eligible tracts. For local polynomial estimation, I use a 5 percent bandwidth, restricting to firms located in tract incomes below and above the threshold cutoff of 80 percent.[15]

1.4.2 Regression Discontinuity with Difference-in-Differences

Given the extensive history of the CRA, estimating to draw a clear causality from the CRA using cross-sectional variation in an RDD setting can be a concern. It is hard to tease out the CRA impact using only CRA eligibility when many other things have happened simultaneously over a long period. The defined CRA eligibility lasts for ten years, until its reference data is changed to recompute income and update the CRA designation. This means that the income of a tract may not precisely reflect its economic condition, and CRA eligibility does not necessarily represent lower-income neighborhoods in the beginning, particularly during later years of the cycle before an update. Also, local RDD estimation on the impact of the CRA that only focuses around

[15] More precisely, this bandwidth can be decided with an optimal bandwidth approach that considers trade-offs between bias and variance after looking at data for the exact sample size of each bin of the bandwidth (Cattaneo, Idrobo, and Titiunik 2018). Five percent bandwidth is often used as a local linear estimation, particularly as it also approaches the computed optimal bandwidth within my sample.

the threshold cannot address the potential heterogeneous impact on the neighborhoods further from the threshold.

Addressing potential limitations, I add a difference-in-differences analysis exploiting time variation to the RDD analysis. As I mentioned earlier, the CRA tract eligibility, defined by a tract's relative median family income relative to that of its MSA, is determined by the Decennial Census.[16] This reference data was changed in 2012, when each tract's relative income level was updated from Census 2000 to American Community Survey (ACS) 2006-2010. This update provides time variation for the CRA eligibility from two different sources: 1) an update in tract and MFI income and 2) change in tract boundaries. Not only the income update but also the tract boundaries change every 10 years, when the Participant Statistical Areas Program (PSAP) updates the boundaries of Census tracts by either splitting or merging them based on changes in population. Together, these changes create an exogenous variation in the CRA eligibility of the tracts where firms locate, providing firms in tracts that have changed in eligibility over time. For this longitudinal setting with time-varying treatment, I use firm fixed effects. Comparing outcomes between tracts with changes in eligibility, either non-eligible to eligible or eligible to non-eligible, and tracts with no eligibility change shows the impact of the CRA.

Using the time variation, only firms in switching tracts contribute to the DID estimation with the firm fixed effects setting. This estimation using switching firms

[16] Tract eligibility is determined every ten years by each Decennial Census, and eligibility is now determined every five years based on ACS 5-year estimates.

addresses the potential issue of RDD discussed earlier. However, it is still important to use RDD and DID together with the local bandwidth. There can be a concern that the change in CRA eligibility from income recomputation is not random. For instance, even among the switching firms, there are differences in the range of change in income. Estimation of the impact from firms in switching tracts with large changes in income can be affected by other confounding factors compared to the firms in switching tracts with relatively small changes in income. For instance, tracts that become CRA eligible when their income ratio changes from 90 to 30 percent are different from tracts that become CRA eligible when it changes from 85 to 79 percent. This is especially true when income recomputation that may reflect a gradual improvement of neighborhoods over time happens every ten years. Estimates can also reflect change in economic development of the neighborhoods rather than the impact of the CRA. In this regard, local estimation around the threshold is crucial to obtain estimates that address this concern. Bandwidth allows restriction of firms in tracts with similar incomes before and after a change. Thus, the changes in outcome are solely driven by changes in CRA designation.

Accordingly, I examine the impact of the CRA with tracts with changes in eligibility with the following specification:

Equation 4

$$Y_{ijt} = \alpha + \beta D_{jt} + h(M_{jt}) + D_{jt} * h(M_{jt}) + X_{ijt}\theta + Z_{jt}\gamma + \sigma_i + \tau_t + \varepsilon_{ijt}$$

where Y_{ijt} is the business outcome, which is employment for firm *i* in tract *j* in year *t*. When *t* is 2012, depending on the eligibility of the tract, D_{jt} can be changed to either 1 or

0 for firms in switching tracts, capturing the impact of change in CRA eligibility within the same tract if its eligibility changes. X_{ijt} includes firm age and age squares, and σ_i and τ_t are firm and time fixed effects, respectively. ε_{ijt} is an idiosyncratic error term. The 5 percent bandwidth for local polynomial estimation in this context is restricted to firms located in tracts with incomes between 75 to 85 percent before and after the changes. In this way, I can compare firms located in tracts with the same ranges of income across time periods, and the estimates do not reflect economic changes in neighborhoods over time.

1.4.3 Heterogeneous Effects

The CRA impact can vary depending on different firm or tract characteristics, such as firm age or the neighborhood where a firm locates. An increase in financing in the CRA eligible areas can be more impactful for firms facing high barriers to access to credit. The CRA impact may be more prominent for young firms for which financing is more binding compared to mature firms (Cabral and Mata 2003). In the same vein, given the CRA's concurrent purpose of reducing credit discrimination and removing redlining, its impact may be disproportionate across neighborhoods. Firms located in minority neighborhoods can be more affected by the CRA compared to non-minority neighborhoods. Considering another aspect, the impact of the CRA can be more pronounced in regions where the banks are more engaged in CRA lending, such as in big MSAs compared to smaller MSAs with smaller populations. I allow the coefficient of the

CRA to vary based on different firm or tract characteristics. Following is the specification for heterogeneous CRA contribution:

Equation 5

$$Y_{ijt} = \alpha + D_{jt}H_{jt}\beta + h(M_{jt}) + D_{jt} * h(M_{jt}) + X_{ijt}\theta + Z_{jt}\gamma + \sigma_i + \tau_t + \varepsilon_{ijt}$$

where $D_{jt}H_{jt}$ are the interaction terms between the CRA indicator and other firm or tract characteristics, such as the indicator for firm *i* being young or not, the indicator of firm location in minority neighborhoods or not, and the indicator of firm location in large MSAs or not. First, a young firm is defined as a firm less than five years old before the change in 2012, and I follow them over time. Second, following Bhutta (2011), a large MSA is defined as having more than 2 million population. Lastly, a minority tract is a tract where the share of minority populations is greater than or equal to 30 percent. This heterogeneity allows the estimation of varying CRA impacts for each group.

1.4.4 Financing Mechanism

While I estimate the CRA impact on business outcomes, understanding the channel of the impact is important. The purpose of the CRA is explicitly stated as the "increase in access to finance in needed areas,"[17] and this potential increase in financing is likely to improve firm outcomes in CRA eligible areas. Instead of a business outcome

[17] The purpose of the Community Reinvestment Act (CRA) is stated by the Federal Financial Institutions Examination Council at https://www.ffiec.gov/cra/history.htm

on the LHS, I estimate the impact of the CRA on financing outcome measured by a number and the amount of a loan with the following specification:

Equation 6

$$Loan_{jt} = \alpha + \beta D_{jt} + h(M_{jt}) + D_{jt} * h(M_{jt}) + X_{ijt}\theta + Z_{jt}\gamma + \rho_i + \tau_t + \varepsilon_{ijt}$$

where $Loan_{jt}$ is the size of a loan in tract *j* in time *t*, measured by either number or amount of small loans less than 1 million or number or amount of loans to small businesses with annual gross revenues of less than 1 million. X_{ijt} and Z_{jt} are the same as in previous specifications, ρ_i and τ_t are firm and time fixed effects, respectively, and ε_{ijt} is an idiosyncratic error term. The β estimates financing in CRA eligible areas compared to non-eligible areas. Tract fixed effects are not applicable when the tract boundary changes over time.

As mentioned earlier, each bank has its own assessment areas for the CRA examination, determined by the locations of its main office, branches, or deposit-taking ATMs. Banks are examined on their CRA activities in the assessment areas, implying that banks have more incentives to comply with CRA lending activities in these assessment areas. Accordingly, firms located in areas with more banks would have a higher probability of being financed due to their active engagement in CRA lending. Following this mechanism, I interact the number of assessment banks with the CRA dummy to estimate the additional impact of financing along with the number of assessment banks with the following equation:

Equation 7

$$Loan_{jt} = \alpha + \beta_1 D_{jt} + \beta_2 Num\_AA_{jt} + \beta_3 (D_{jt} * Num\_AA_{jt}) + h(M_{jt}) + D_{jt} * h(M_{jt}) + X_{ijt}\theta + Z_{jt}\gamma + \rho_i + \tau_t + \varepsilon_{ijt}$$

where D_{jt} is an indicator of whether tract j in time *t* is CRA eligible or not, and Num_AA_{jt} is the number of banks that are assessed by the CRA examination in tract *j* at time *t*. β_3 is the additional CRA impact with additional assessment bank on financing. This specification allows estimation of the relationship between the CRA and financing as well as the important role of banks as a financing channel that delivers the CRA impact to businesses.

1.5 Results

The following section shows descriptive statistics among firms in CRA and non-CRA tracts in the full sample as well as in the 5 percent bandwidth below and above the threshold of the 80 percent cutoff. Then, I estimate the CRA impact in the RDD framework, followed by combining it with the DID setting using firm fixed effects. The CRA impact is further analyzed across heterogeneous groups, and its financing mechanism is examined by looking at the relationship between loan availability and the CRA. Lastly, I present the results of the placebo test with different income thresholds.

1.5.1 Descriptive Statistics

Table 1 shows the summary statistics of firms in CRA and non-CRA tracts, where the unit of analysis is a firm-year and the core variable of interest is the CRA. The last two columns show the means of each variable for firm and tract characteristics in the 5 percent bandwidth below and above the income threshold, ranging from 75 to 85 percent. About 24 percent of the sample is in CRA eligible tracts, where about 9 percent is in CRA eligibility switching tracts. There are 4,222,000 firms in the full sample and 497,000 firms in the 5 percent bandwidth sample. More importantly, there are 278,000 and 127,000 switching firms in the full sample and 5 percent bandwidth, respectively, which provides the variation to estimate the CRA impact in the DID setting. At the tract level, there are about 69,000 tracts, 27.5 percent of which always stay in the CRA tracts throughout the study period. There are 5,100 switching tracts that were in the CRA areas before or after the change in 2012.

Table 1 Descriptive Statistics

	(1)	(2)	(3)	(4)	(5)
		All		Bandwidth 5%	
	All	CRA	Non-CRA	CRA	Non-CRA
Firm Characteristics					
CRA	0.244	1	0	1	0
CRA switching firms	0.092	0.174	0.066	0.323	0.278
CRA gain eligibility	0.057	0.083	0.049	0.150	0.218
Always stay in CRA	0.201	0.826	0	0.677	0
CRA lose eligibility	0.035	0.091	0.017	0.173	0.060
Always stay in non-CRA	0.707	0	0.934	0	0.722
Employment	5.653	6.188	5.48	5.996	5.883
Employment growth	0.122	0.115	0.124	0.110	0.110
Firm age	11.43	12.40	11.12	12.06	11.84
Large MSA	0.570	0.552	0.576	0.522	0.528
Multi-unit	0.008	0.010	0.008	0.009	0.009
Exit	0.0535	0.0552	0.0529	0.0518	0.0508
Primary sector	0.002	0.002	0.002	0.002	0.001
Construction	0.126	0.09	0.138	0.114	0.127
Manufacturing	0.044	0.062	0.038	0.055	0.05
Wholesale trade	0.049	0.048	0.049	0.044	0.042
Retail trade	0.102	0.139	0.090	0.124	0.115
Transportation	0.024	0.024	0.025	0.026	0.029
Information	0.013	0.010	0.013	0.010	0.011
Finance and insurance	0.046	0.040	0.048	0.044	0.045
Real estate	0.051	0.049	0.051	0.047	0.046
Business services	0.160	0.133	0.169	0.124	0.129
Administ. services	0.061	0.049	0.065	0.054	0.059
Educational services	0.015	0.013	0.016	0.012	0.013
Health care	0.128	0.129	0.128	0.136	0.134
Arts and entertainment	0.019	0.014	0.020	0.015	0.016
Food services	0.083	0.107	0.075	0.099	0.093
Other services	0.079	0.094	0.074	0.094	0.091
Observations	28,580,000	6,971,000	21,610,000	1,180,000	1,308,000

Note: The unit of analysis is firm-year. These are the mean of each variable in full sample and 5 percent bandwidth, and separately in the CRA and non-CRA tracts. The numbers are rounded for disclosure purposes.

While most firms are located in non-CRA tracts, firms in the CRA eligible tracts have an average of six employees and are slightly larger than firms in non-CRA tracts. Figure 1 shows the discontinuous jump in log employment around the income cutoff of 80 percent in the CRA tracts compared to non-CRA tracts.

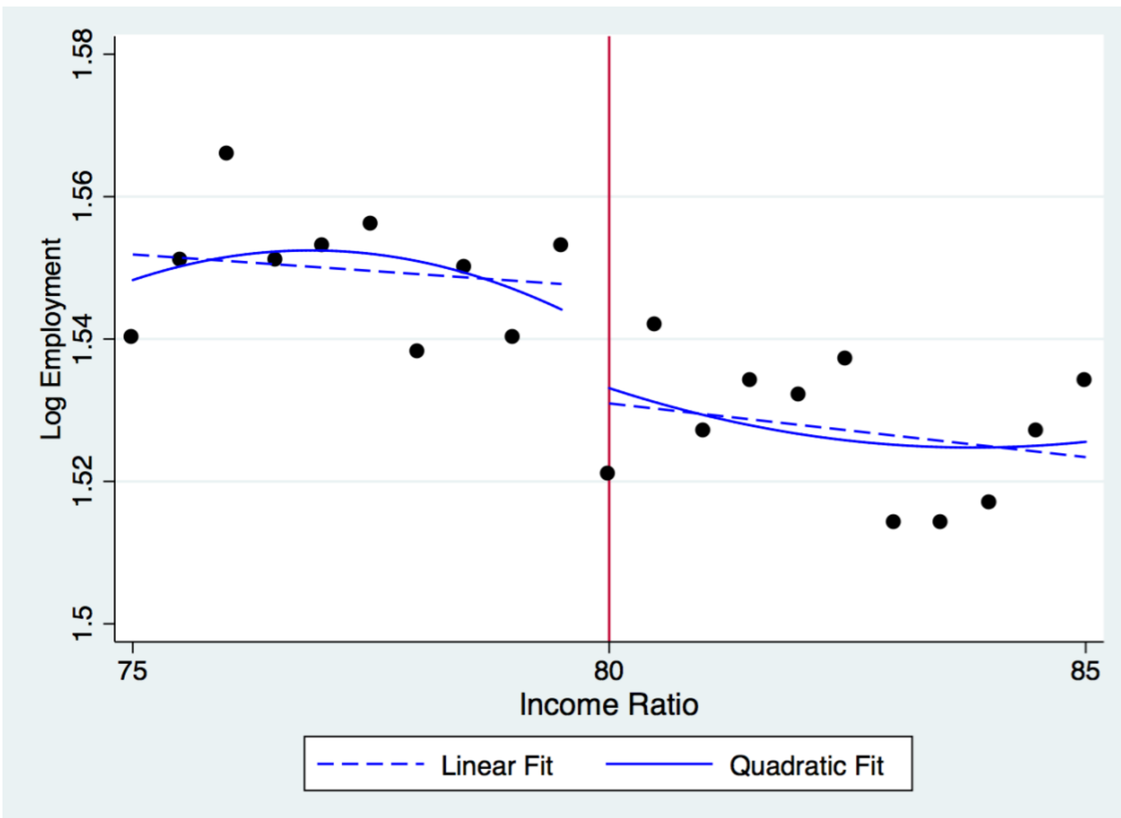

Figure 1 Average Employment on Income

Conversely, the average employment growth in the CRA tracts is slightly smaller than for firms in non-CRA tracts. Firms in the CRA tracts are an average of a year older, at 12 years old. In terms of industry, firms are more concentrated in manufacturing, the retail trade sector, and food services in the CRA tracts, whereas there are more firms in construction and administrative services in non-CRA tracts. In essence, firms have similar characteristics in a 5 percent bandwidth sample around the threshold.

Table 2 Descriptive Statistics with Tract-level Characteristics

	(1)	(2)	(3)	(4)	(5)
		All		Bandwidth 5%	
	All	CRA	Non-CRA	CRA	Non-CRA
Tract Characteristics					
Population	4424	3981	4624	4314	4392
Minority population	1564	2451	1165	1875	1641
Minority population share	0.358	0.607	0.246	0.412	0.356
Tract/MSA median family income (%)	102	59.81	121	77.5	82.55
MSA median family income	58420	58350	58460	57420	57800
BA and more population (%)	0.267	0.142	0.324	0.174	0.186
Number of firms	45	35	49	38	40
Number of employment	251	216	267	226	233
Number of CRA assessment banks	21	22	20	20	19
Loans					
Total number of small business loans	114	76	132	91	94
Total amount of small business loans	4199	3238	4632	3517	3552
Total number of loans to small business	46	29	53	36	38
Total amount of loans to small business	1625	1125	1850	1314	1356
Observations	648,000	201,000	447,000	32,000	34,000

Note: The unit of analysis is tract-year. These are the mean of each variable in full sample and 5 percent bandwidth, and separately in the CRA and non-CRA tracts. The numbers are rounded for disclosure purposes.

Looking at differences in tract characteristics between the CRA and non-CRA areas in Table 2, non-CRA areas have higher populations with higher education and smaller minority populations. Although tracts are designed to have similar populations of an average of 4,000 residents, populations are slightly higher in non-CRA areas. Certainly, non-CRA areas also have higher median family incomes. At the tract level, there are a higher number of larger firms in the non-CRA areas compared to CRA areas. There is an average about 21 CRA assessed banks between the two areas. In terms of financing, higher numbers and amounts of small loans are given in non-CRA areas relative to the CRA areas.

1.5.2 RDD Validation

As CRA eligibility is determined by the incomes of neighborhoods, the CRA areas are relatively poorer than non-CRA eligible areas. This means that firms located in these two areas can be systematically different across different characteristics. However, the RDD framework overcomes this problem, particularly close to the threshold. Within this local bandwidth, the firms are assigned to be "as if random," implying that underlying characteristics between CRA and non-CRA tracts in this bandwidth are not systematically different from each other if I control the income, which is a running variable. I assume smooth changes in tract characteristics between CRA and non-CRA tracts within the bandwidth except in my outcome variables, finance and employment. Since the CRA variation is at tract level, I conduct the covariate tests on difference in all other tract characteristics at the tract level.

The two additional constructed tract level variables are total number of firms and total number of employees. These may suggest how dynamic the business market is at the tract level, with the assumption that more “dynamic” tracts with more firms and employment may have a better chance of getting loans compared to tracts where businesses are slow and the market is not vitalized. Table 3 shows that the variables are not significantly different between the two areas, and I do not control these characteristics in my regressions because they are related to my outcome variables.

Table 3 Covariate Balance Test

	All	5%
MSA median family income	100.9	26.97
	(108.9)	(195.1)
Population	-24.94	97.20
	(18.08)	(30.57)
Minority population	129.1	12.33
	(13.26)	(27.15)
Share of population with BA or	0.0187	0.0015
higher degree	(0.0011)	(0.0018)
Number of firms	0.1589	-0.6958
	(0.4529)	(0.7165)
Number of employment	-2.528	-1.821
	(3.043)	(4.698)
Number of observations	648,000	66,000

Note: The unit of analysis is tract. These are OLS regressions. Coefficients are the CRA effect with row as a dependent variable. Polynomial for income fits quartic function for all sample, and a linear function within the 5 percent bandwidth.

The characteristics such as minority population or share of population with higher education are significantly different, but once I restrict to firms around the threshold, they are no longer significant. Although the population is still statistically different within the bandwidth, I control this variable in my regressions.

1.5.3 The CRA Impact on Business Outcomes

1.5.3.1 RDD on Employment

The main business outcome variable in this study is employment. Tables 4 and 5 show RDD estimates of the CRA impact on number of employees and annual employment growth. Each column shows additional impacts as I add controls and additional fixed effects. The CRA impact is consistently negative and insignificant on employment across specifications, even when restricted to the 5 percent bandwidth. The size of estimates between the full sample and the bandwidth is also similar. Conversely, the CRA estimates for annual employment growth are positive and significant, with an incremental increase in magnitude across the specifications for all global estimations. There is about a 0.2 to 0.3 percent increase in employment growth for firms in the CRA eligible areas compared to the ones in non-eligible areas. However, this increase loses significance as it gets close to the threshold in a 5 percent bandwidth. This may suggest that the CRA impacts on firms located at tracts further from the threshold are larger than for tracts close to the threshold.

However, the bandwidth in this cross-sectional RDD setting may not precisely reflect the income status of neighborhoods because the CRA designation based on neighborhood income is consistent for ten years until it is updated with recomputation by the new reference data. Overall, RDD estimates suggest that there is a significant CRA impact on the annual employment growth of firms in broader CRA neighborhoods, while the CRA impact is not significant around the local threshold.

Table 4 Regression Discontinuity Design (RDD): Employment

	(1) All	(2) All	(3) All	(4) All	(5) All	(6) All	(7) 5%
CRA	-0.0005	0.0005	-0.0016	-0.0010	-0.0007	-0.0016	-0.0017
	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0023)	(0.0023)	(0.0042)
Firm age	No	Yes	Yes	Yes	Yes	Yes	Yes
Tract chac.	No	No	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes	Yes
Industry FE	No	No	No	No	Yes	Yes	Yes
MSA FE	No	No	No	No	No	Yes	Yes

Note: The unit of analysis is firm-year. The number of observations for all sample is 28,580,000, and 2,488,000 firms for 5 percent sample. Income function for all is quartic function, and 5 percent bandwidth is a linear function. Firm age includes firm age and age squared, and tract characteristics include MSA median family income, population, share of minority population, and share of population with BA or higher degree. Standard errors clustered by firm are in parentheses.

Table 5 Regression Discontinuity Design (RDD): Employment Growth

	(1) All	(2) All	(3) All	(4) All	(5) All	(6) All	(7) 5%
CRA	0.0044	0.0021	0.0015	0.0018	0.0019	0.0025	0.0022
	(0.0009)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0014)
Firm age	No	Yes	Yes	Yes	Yes	Yes	Yes
Tract chac.	No	No	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes	Yes
Industry FE	No	No	No	No	Yes	Yes	Yes
MSA FE	No	No	No	No	No	Yes	Yes

Note: The unit of analysis is firm-year. The number of observations for all sample is 28,580,000, and 2,488,000 firms for 5 percent sample. Income function for all is quartic function, and 5 percent bandwidth is a linear function. Firm age includes firm age and age squared, and tract characteristics include MSA median family income, population, share of minority population, and share of population with BA or higher degree. Standard errors clustered by firm are in parentheses.

1.5.3.2 RDD-DID on Employment

To overcome the potential issue of RDD, I add time variation with firm fixed effects to RDD to estimate the impact of the CRA. Table 6 shows the RDD-DID with firm fixed effects estimates. Using the full sample for global estimation allows the use of all firms in the CRA eligibility switching tracts for the estimations, regardless of the range of income change. The CRA impact is insignificant and close to zero until I restrict the sample around the threshold. Then, as I add additional controls and year fixed effects, the magnitudes get smaller. Compared to local estimation, these global estimations include more switching firms but also more non-switching firms. The estimation can be driven by comparing the extreme cases between the firms in the wealthiest and the poorest tracts, which has a negative effect on the CRA.

Table 6 RDD with Difference-in-Differences (DID): Employment

	(1)	(2)	(3)	(4)	(5)
	All	All	All	All	5%
CRA	-0.0021	0.0016	0.0008	-0.0003	0.0081
	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0047)
Firm age	No	Yes	Yes	Yes	Yes
Tract characteristics	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes

Note: The unit of analysis is firm-year. The number of observations for all sample is 28,580,000, and 2,488,000 firms for 5 percent sample. Income function for all is quartic function, and 5 percent bandwidth is a linear function. Firm age includes firm age and age squared, and tract characteristics include MSA median family income, population, share of minority population, and share of population with BA or higher degree. Standard errors clustered by firm are in parentheses.

Thus, it is crucial to estimate the impact locally within the bandwidth, restricting firms located in tracts with similar incomes around the threshold before and after the change. In this regard, this estimation is from firms located in tracts with similar incomes over time, and the only changes are in the CRA eligibility. It also addresses potential concerns about reflecting changes or improvements in neighborhoods over time.

Looking at the local estimates around the bandwidth, the CRA designation increases about 0.8 percent of employment of the firms relative to firms located in non-CRA tracts at a 10 percent significance level. This firm fixed effects allows the estimation of the CRA impact from differences in employment of firms between treated and non-treated time using within variation. As I mentioned earlier, it is important to note that this impact is an intent-to-treat effect when I do not observe firms that actually received loans due to the CRA. This implies that the actual CRA impact can be larger than this estimated effect. The combination of RDD-DID with firm fixed effects within the bandwidth allows to estimate the CRA impact from firms in tracts with similar income before and after the change where the only change is its CRA eligibility.

1.5.4 Heterogeneous Effect of the CRA

1.5.4.1 Age

The CRA can be more impactful for young firms that face higher credit constraints, when young firms with a lack of credit history are often more likely to have credit constraints compared to older firms. Small and young firms face higher liquidity constraints compared to mature firms, affecting their growth paths (Cabral Mata 2003;

Oliveira and Fortunato 2006). Table 7 shows the heterogenous CRA effects on young firms and older firms. Here, young firms were less than or equal to five years old and older firms were more than five years old prior to the CRA designation "shock" in 2012. I flag firms as young if they were young before the change and then follow them to look at how the CRA affects their growth path throughout the life cycle.

Looking at Table 7, there are consistently positive and significant CRA impacts on young firms across specifications. While the addition of year fixed effects addresses the recession period over the study period, it decreases the size of impact by less than 30 percent, which is still significant. More importantly, local estimation shows that young firms increase their employment by about 1.9 percent compared to firms in non-CRA-eligible locations.

Table 7 RDD-DID: Age Heterogeneity

	(1)	(2)	(3)	(4)	(5)
	All	All	All	All	5%
Young firms*CRA	0.0229	0.0265	0.0269	0.0189	0.0185
	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0077)
Older firms*CRA	-0.0063	-0.0068	-0.0058	-0.0037	0.0059
	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0049)
Firm age	No	Yes	Yes	Yes	Yes
Tract charac.	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes

Note: The unit of analysis is firm-year. The number of observations for all sample is 28,580,000, and 2,488,000 firms for 5 percent sample. Income function for all is quartic function, and 5 percent bandwidth is a linear function. Firm age includes firm age and age squared, and tract characteristics include MSA median family income, population, share of minority population, and share of population with BA or higher degree. Standard errors clustered by firm are in parentheses.

Conversely, the CRA effects on employment for older firms are consistently negative and significant using the full sample, until they become positive and insignificant with local estimation using the 5 percent bandwidth. The difference in these two coefficients is statistically significant at the 5 percent significance level. This may suggest that young firms have more difficulty finding financing relative to older firms, so the additional financial access through the CRA has a larger impact on them.

1.5.4.2 Minority Neighborhoods

Access to finance is one of the pronounced obstacles for firm performance, and this is particularly true for minority groups. Most of the studies on minority discrimination in small business credit markets have examined the existence of statistical and taste-based discrimination (Asiedu, Freeman, and Nti-Addae 2012). Further, despite the desire to have more credit, a minority group of borrowers have been discouraged from applying for loans due to the fear of rejection (Blanchflower, Levine, and Zimmerman 2003; Blanchard, Zhao, and Yinger 2008). This implies that the CRA may lower the barrier to access to credit even more for this minority group. Thus, I expect a disproportionate impact of the CRA on firms in minority neighborhoods.

Although the share of minority population is not an explicit criterion in CRA examinations, minority support is an important aspect of the CRA. Particularly, when there is a high positive correlation between minority neighborhoods and LMI neighborhoods, minority neighborhoods are more subject to the CRA. Following the Federal Reserve Board report, a minority neighborhood is defined as a tract with more than 30 percent minority population. About 54 and 39 percent of CRA tracts are minority neighborhoods in the full sample and the bandwidth, respectively.

Table 8 shows negative insignificant CRA effects using the full sample, which become positive and significant for firms located in minority neighborhoods with the 5 percent bandwidth. The first specification suggests that there are a higher number of smaller firms in minority neighborhoods. While there are many older firms in the CRA areas, the effect on firms both in minority and non-minority neighborhoods becomes larger when I control firm ages. Additionally, when I control tract characteristics and year effects, the CRA impact disappears. However, since these are estimated using the overall sample, there could be other unobserved heterogeneity factors affecting the estimations.

Table 8 RDD-DID: Minority Heterogeneity

	(1) All	(2) All	(3) All	(4) All	(5) 5%
Firms in Minority NH	-0.0033	-0.0015	-0.0003	-0.0012	0.0146
*CRA	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0054)
Firms in non-Minority NH	-0.0009	0.0055	0.0030	0.0008	0.0013
*CRA	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0053)
Firm age	No	Yes	Yes	Yes	Yes
Tract characteristics	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes

Note: The unit of analysis is firm-year. The number of observations for all sample is 28,580,000, and 2,488,000 firms for 5 percent sample. Income function for all is quartic function, and 5 percent bandwidth is a linear function. Firm age includes firm age and age squared, and tract characteristics include MSA median family income, population, share of minority population, and share of population with BA or higher degree. Standard errors clustered by firm are in parentheses.

Conversely, the CRA increases by about 1.5 percent in employment for firms in CRA minority neighborhoods compared to firms in non-CRA areas at the 5 percent bandwidth. As discussed earlier, local estimation only uses firms located in similar income tracts around the threshold. The greater coefficient of firms in minority neighborhoods suggests that there are more credit-constrained firms in minority neighborhoods in the poorest tracts. When these are excluded, there are larger CRA impacts in minority neighborhoods around the threshold. The difference in the two coefficients is statistically significant at a 5 percent significance level. This suggests that the CRA plays an even more important role in firms in minority neighborhoods facing higher barriers to access to finance.

1.5.4.3 Large MSAs

Firm location in terms of size of MSAs may be an important factor that has a larger impact on the CRA for several reasons. There are few previous studies suggesting that banks in large MSAs are more likely to be responsive to the CRA compared to those in small MSAs (e.g., Berry and Lee 2007; Bhutta 2011). When banks are the key players that affect businesses through the CRA, it is important to understand each bank's behavior in relation to CRA responsiveness, which could accordingly affect businesses in LMI neighborhoods. Being located in large MSAs is influential when there tend to be more big banks, which have more incentives to comply and maintain high CRA scores. These banks may be more engaged in mergers and acquisitions and may have higher chances of conducting these expansion activities compared to small banks (Agarwal et al. 2012).

Also, large banks with more than $250 million in assets have a more frequent cycle of review and examination that occurs every two years compared to every five years for small institutions. Agarwal et al. (2012) found that banks increase lending to CRA-eligible tracts before CRA examinations, although lending to these tracts may result in more risky loans. More frequent examinations would offer banks persistent incentives to be more responsive to the CRA by providing sufficient CRA lending. Big banks need to have more extensive review in terms of both frequency and comprehensiveness. Regulators take MSA size and significance into account in terms of lending, investment, and service opportunities to decide whether the market should be subjected to a full-scope review of all lending, investment, and service tests. CRA performance in a large market means that greater weights affect the overall grade of an institution. Also, big banks are more likely to operate with a larger number of branches in large MSAs. Bhutta (2011) also found a disproportionately large CRA impact on mortgage loans in large MSAs with populations of more than 2 million.

As CRA eligibility is defined with a relative income ratio, there are more income variations in CRA eligible tracts in large MSAs compared to small MSAs. Also, banks may be more willing to lend money to LMIs in large MSAs because large MSAs tend to have higher median incomes. For instance, the LMI in New York is much bigger and has higher incomes in absolute terms compared to the LMI in Richmond, Virginia. Accordingly, banks may have more incentives to give out loans to businesses in the New

York LMI if they're given the same credit for the CRA examination. In this regard, banks may conduct more extensive CRA activities in large MSAs. For the same reason, Berry and Lee (2007) only looked at the CRA impacts in ten big MSAs.

Lastly, while CRA ratings for financial institutions are publicly available, pressures on banks from civil groups and community associations are more common in large MSAs, where these groups are more active. Active community groups play an influential role in lending institutions entering into CRA agreements, which typically only occur in big cities, that pledge future lending and activities to serve LMI communities.

Table 9 RDD-DID: Large MSA Heterogeneity

	(1) All	(2) All	(3) All	(4) All	(5) 5%
Large MSA*CRA	0.0005	0.0010	-0.0003	0.0001	0.0135
	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0053)
Small MSA*CRA	-0.0049	0.0022	0.0019	-0.0007	0.0020
	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0055)
Firm age	No	Yes	Yes	Yes	Yes
Tract characteristics	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes

Note: The unit of analysis is firm-year. The number of observations for all sample is 28,580,000, and 2,488,000 firms for 5 percent sample. Income function for all is quartic function, and 5 percent bandwidth is a linear function. Firm age includes firm age and age squared, and tract characteristics include MSA median family income, population, share of minority population, and share of population with BA or higher degree. Standard errors clustered by firm are in parentheses.

For all these reasons, I interact the CRA with the large MSA dummy, whether the firm is located in an MSA defined with more than 2 million population or not, following Bhutta (2011), to estimate the additional impact of the CRA on business growth located in large MSAs. Table 9 shows the estimates of the CRA impact on firms in large and small MSAs.

There are more smaller firms in small MSAs. When I control for tract characteristics, the coefficients decrease as they address heterogeneity in other tract characteristics related to the size of MSAs. Also, firms in large MSAs are affected more by heterogeneity across years. In sum, there are no significant effects across different specifications using the full sample, and effects are close to zero for firms in large MSAs.

However, restricted to the 5 percent sample, the local estimation shows a bigger CRA impact among firms in large MSAs. Firms located in the CRA eligible areas, and particularly in large MSAs, increase employment by about 1.4 percent more relative to the firms in the non-eligible areas. There are no significant increases in firms in the CRA eligible tracts in small MSAs. Among firms located in tracts with similar incomes around the threshold, the CRA effect is prominent in firms in large MSAs where credit is more accessible and more banks are engaged in CRA activities. The difference between these two coefficients is statistically significant at the 5 percent significance level. These results support the importance of bank presence in CRA areas. Thus, the examination of more banks in assessment areas is crucial to increase CRA lending activities. This mechanism is further analyzed in the next section.

1.5.5 Financing Mechanism

Several previous studies show that the CRA increases financing in CRA designated areas at the aggregated tract level. I estimate financing as an outcome variable to see whether the CRA increases financing measured in four different ways: 1) number of small loans of less than 1 million, 2) dollar amounts of small loans of less than 1 million, 3) number of loans to small businesses with annual gross revenues of less than 1 million, and 4) dollar amounts of loans to small businesses with annual gross revenues of less than 1 million. Panel A in Table 10 shows that the CRA designation increases the dollar amounts of small loans by about 1.5 percent and the number of loans to small businesses by about 3.1 percent. This positive and significant relationship between the CRA and financing supports the increase in access to credit as positive impact of the CRA on business outcomes.

More interestingly, as mentioned in an earlier section, financing is closely related to the presence of banks in a given area. Hence, more CRA banks examined in an area should increase CRA lending activities in that area. Panel B in Table 10 estimates the impact of the CRA and the number of banks assessed by the CRA examination. The interaction term between the CRA and number of banks represents an additional impact of banks in CRA areas. Note that while the outcome of loan variables and the number of assessment variables are all tract-level variables, this analysis is conducted at the firm level. This is because I want to use the time variation to estimate the change in financing between treated and non-treated time.

Table 10 RDD-DID: Credit Availability

	(1)	(2)	(3)	(4)
	5%	5%	5%	5%
	Number of small loans	Amount of small loans	Number of loans to small businesses	Amount of loans to small businesses
Panel A				
CRA	0.0008	0.0154	0.0308	-0.0032
	(0.0041)	(0.0071)	(0.0048)	(0.0075)
Panel B				
CRA	-0.0106	-0.0161	-0.0061	-0.0234
	(0.0050)	(0.0086)	(0.0058)	(0.0092)
Number of CRA	0.0005	0.0049	-0.0012	0.0068
AA banks	(0.0001)	(0.0002)	(0.0001)	(0.0002)
CRA*AA banks	0.0006	0.0016	0.0019	0.0010
	(0.0001)	(0.0002)	(0.0002)	(0.0002)
Firm age	Yes	Yes	Yes	Yes
Tract charac.	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes

Note: The unit of analysis is firm-year. All dependent variables are logged. The number of observations for all sample is 28,580,000, and 2,488,000 firms for 5 percent sample. Income function for 5 percent bandwidth is a linear function. Firm age includes firm age and age squared, and tract characteristics include MSA median family income, population, share of minority population, and share of population with BA or higher degree. Standard errors clustered by firm are in parentheses.

Given twenty assessed banks on average in a tract, the number of small loans increase about 0.1 percent and the number of loans to small business increases by about 3.2 percent relative to non-CRA eligible areas. More importantly, the dollar amounts of small loans increase by 1.6 percent. While the impact on dollar amounts of loans to small businesses is close to zero with twenty banks on average, it would increase in areas with more concentrated numbers of banks. Additionally, one more assessed bank in a CRA area would increase financing from about 0.1 to 0.2 percent. In sum, these results provide evidence of the relationship between the CRA and financing as a core mechanism for business growth. The banks play an important role when more financing is permitted in the CRA eligible tracts and there are more banks assessed by the CRA examinations.

1.5.6 Placebo Test

I conduct two placebo tests of the CRA impact with different income thresholds, one in CRA eligible tracts with a 60 percent income threshold and the other in CRA ineligible tracts with a 100 percent income threshold. If better business outcomes are due to the CRA, then the impact should be visible at the 80 percent income thresholds only. Estimates with these alternative thresholds should not show significant effects on employment.

These two alternative thresholds provide two samples, with 5 percent bandwidth around each threshold. One is a CRA placebo sample with firms with income between 55 to 65 percent, and the other is a non-CRA placebo sample with firms with income between 95 to 105 percent. Using these two placebo samples, Table 11 shows the estimates of my main specifications with DID with firm fixed effects and all other heterogeneous effects.

Table 11 RDD-DID: Placebo Tests

	(1) 60% threshold	(2) 100% threshold
Panel A		
CRA	-	-
Panel B		
Young firms*CRA	+	+
Older firms*CRA	-	+
Panel C		
Firms in minority NH*CRA	-	+
Firms in non-minority NH*CRA	-	-
Panel D		
Large MSA*CRA	+	-
Small MSA*CRA	-*	+
Firm age	Yes	Yes
Tract characteristics	Yes	Yes
Year FE	Yes	Yes
Firm FE	Yes	Yes

Note: The unit of analysis is firm-year. These are only qualitative results with signs and significance using different income thresholds 60 and 100 percent below and above 5 percent samples with linear income function. Firm age includes firm age and age squared, and tract characteristics include MSA median family income, population, share of minority population, and share of population with BA or higher degree. Standard errors clustered by firm are in parentheses.

It should be noted that due to further disclosure in the future, I am currently only able to release qualitative results for these estimates with their signs and significance. Each panel from Table 11 shows the estimates of the variable of interest, the CRA, and the heterogeneous impact of the CRA using DID with firm fixed effects regressions. In essence, it shows the insignificant impact of the CRA across all these estimates using two different placebo samples with alternative thresholds. One exception is the coefficient on the interaction term between small MSAs and the CRA in the sample using a 60 percent threshold, which is only marginally significant at a 10 percent significance level.

1.6 Conclusion

Despite its long history, the CRA is still undergoing policy discussion to "modernize" and reform for better application. The potential reform focuses on creating a more systematic examination in terms of clarity, consistency, and timeliness to verify compliance in providing sufficient financing and access to credit in needed areas. Using rich firm-level panel data on every U.S. employer, I estimate the impact of the CRA on the real outcomes of business growth in those areas. Exploiting the sharp income threshold from the policy and taking advantage of the change in CRA eligibility over time, I use regression discontinuity and difference-in-differences with firm fixed effects to estimate the CRA impact before and after the change.

The RDD estimation provides evidence of a positive CRA impact on annual employment growth of firms in CRA tracts compared to firms in non-CRA tracts, with an about 0.3 percent increase driving from the firms located in tracts further from the 80 percent threshold. Conversely, local estimations around the bandwidth suggest an insignificant CRA impact on both employment and employment growth of the firms in the CRA tracts. Exploiting time variation, RDD-DID local estimation around the threshold results suggest that the CRA designation of firm location increases employment by about 0.8 percent, while global estimations of the CRA are close to zero and insignificant.

The CRA impacts are larger for groups facing higher credit barriers, such as young firms or firms located in minority neighborhoods. The CRA designation increases employment by about 1.5 to 1.9 percent for these groups, respectively. Also, the CRA impact is larger for firms located in large MSAs having more than 2 million population, increasing about 1.4 percent in employment compared to firms in non-CRA areas. This may suggest an important role for easier access to credit from more banks in the area assessed by the CRA examinations, which are the main channel of the CRA's impacts on small businesses. In particular, CRA eligibility increases the amount of small loans and the number of loans to small businesses by about 1.5 to 3.1 percent, respectively, in the area.

While some previous studies looked at the impact of the CRA on loan availability in lower-income neighborhoods, this study takes a step further to estimate its impact on the real outcomes of businesses. The findings contribute to CRA impact evaluation and may suggest a direction for further CRA reform. Using the CRA as a channel of credit access in needed areas, the results provide new evidence of a causal link from finance to growth. These results shed additional light on the significance of financing, particularly for more credit-constrained groups.

CHAPTER 2: THE FINANCING OF AFRICAN AMERICAN ENTREPRENUERSHIP[18]

2.1 Introduction

Are African American entrepreneurs more financially constrained? While minority entrepreneurs, particularly African American entrepreneurs are frequently claimed to face higher credit constraints, we do not know much about detailed financing patterns such as their capital sources or amounts they use to start and run their businesses (Bates 1989; Cohn and Coleman 2001; Coleman 2002; Fairlie and Robb 2007). A number of studies claim that the potential financing gap between African American and white owners can be attributed to discrimination against African American business owners (Cavalluzzo and Cavalluzzo 1998, Blanchflower, Levine and Zimmerman 2003). However, these studies do not account for other related characteristics that may be related to finance and that might explain the gap between African American and white entrepreneurs, including differences in demographics, skills, motivations, industry choices, and other choices owners make about their businesses. Taking these factors into account would lead to a better understanding of the financing gap, and how much of it remains unexplained, which might reflect discrimination.

[18] I thank the National Science Foundation (NSF) for support (Grants 1262269 to George Mason University). Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the NSF or the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed, and released under DRB Bypass Number CBDRB-FY20-CES009-001 and CBDRB-FY20-CES009-002.

In this paper, I examine financing patterns by African American business owners more closely using rich data on businesses in the U.S. I exploit unusually detailed owner characteristics and financing variables from the Annual Survey of Entrepreneurs (ASE) at the U.S. Census Bureau. The outcome variables include not only amount and source of startup capital but also source of new financing and other question bearing on credit constraints. I estimate the differences in financing between African American and white business owners after controlling for a range of owner and firm characteristics, including age, gender, education, motivation, industry, and other choices owners make about their businesses. I examine not only whether African American entrepreneurs face higher credit constraints, but also whether the capital access gap reflects characteristics of business owners and their choices.

The preliminary results suggest that African American business owners do face higher financial constraints compared to white owners. They use lower amounts of capital to start their businesses relative to white business owners. In terms of capital source, African Americans are less likely to use external financing such as banks or other financial institutions for their startups, while they are more likely to use personal savings, credit cards, or grants as startup capital. They have about 2.5 percentage point lower propensity of sourcing startup capital from banks and 4.6 percentage point higher from personal savings, compared to white entrepreneurs. The gap between African American and white business owners still remains even after taking into account differences in related owner and firm characteristics, which explain between 17 and 30 percent of the gap. In addition to startup capital, I find that African Americans are less likely to receive additional new funding from banks, angel investors or venture capital relative to white business owners. Not only differences in demographics, human capital and motivations explain part of the financing gap between African American and white owners, but also the gap is explained by industry differences. However, even after controlling for differences in all of these characteristics, a financing gap still remains, and this gap may reflect potential discrimination faced by African Americans entrepreneurs.

I also examine subjective questions on the underlying reasons for not looking for additional financing and the importance of the access to finance for business profitability. Consistent with the previous literature (Blanchflower, Levine and Zimmerman 2003, Coleman 2004, Coleman 2005, and Blanchard, Yinger and Zhao 2008), African Americans are more likely to choose not to look for additional funding because they expect to be rejected by banks. Using the Federal Reserve Board's Survey of Small Business Finances (SSBF) data, all these studies have relatively small samples of firms and are not able to control relevant owner and firm characteristics other than credit history, whereas the current study uses a much larger sample of firms with detailed characteristics including human capital, motivations, and other controls. Also, I find that African Americans are more likely to consider access to finance as a critical factor in business profitability. This is consistent with Robb and Morelix (2016), who use the publicly available ASE and report that access to finance plays an even more critical role in generating profits for businesses owned by African Americans relative to white owners.

This research exploits the micro data from the ASE, providing detailed owner and firm characteristics for a large sample. The new data allow me to measure several sets of variables not available in the other sources such as the Survey of Business Owners (SBO), Kauffman Firm Survey (KFS), and SSBF data, to answer the same questions, such as owner's motivations to start business, and most relevantly detailed financing related variables. The ASE micro data allows me to compare the financing patterns between African American and white owners after taking into account those detailed firm and owner characteristics which may affect the demand for financing rather than sole differences in unconditional means that other research (Robb 2018) provides. It is important to address differences in those observables to estimate the remaining gap between two groups.

This study is related to the broad topic of finance and specifically to research on the differences in access to finance between minority and non-minority business owners, particularly African Americans. A general finding from the previous research is that African American entrepreneurs face worse credit conditions compared to white entrepreneurs (Bates 1989; Cavalluzzo and Cavalluzzo 1998; Cohn and Coleman 2001; Coleman 2002a and 2002b; Fairlie and Robb 2007). Using Characteristics of Business Owners (CBO) 1987, Bates compares the Black male, Asian male and non-minority male firms on different characteristics such as sales, industry, and financial capital, but only by providing unconditional mean differences among those racial groups. Fairlie and Robb also use the CBO 1992 data to find the gap in finance between African American- and white- owned businesses, and explain the gap and difference in business performance mainly due to disparities in personal wealth (Fairlie and Robb 2007). While personal wealth could be one of the important factors influencing the demand for finance, other owner or firm characteristics also may affect demand. In this paper, I use unusually detailed observable characteristics for both firm and owner from nationally representative data to examine the remained gap in finance between African American and white entrepreneurs.

Robb, Fairlie, and Robinson (2016) investigate a similar question using the KFS linked to Dun & Bradstreet (D&B) data on startups. They find that black-owned businesses have persistently smaller capital and face more constraints in raising external financing compared to white-owned businesses. They use 4,928 startups from 2004 and seven years of follow up data for the cohort from 2005 to 2011. While they control for credit history and some other firm and owner characteristics, they do not include variety of other related observables such as owner ages, citizenship, and motivations, industry, and choices which can affect the financing pattern of entrepreneurs.

The rest of the paper is organized as follows: Section 2 describes the data in more detail, Section 3 explains methods and Section 4 provides the results. Finally, Section 5 concludes.

2.2 Data

Using detailed microdata on business owners in the U.S., this chapter studies differences in finance between African American and white entrepreneurs. I use firm- and owner-level data from 2014 Annual Survey of Entrepreneurs (ASE), collected by the

U.S. Census Bureau. The ASE is an annual survey, containing a nationally representative random sample of non-farm businesses with at least one paid employee and receipts of $1,000 or more. The initial 2014 ASE sample was about 290,000 employer firms with a response rate of 74 percent (Foster and Norman 2016). The ASE provides detailed demographic characteristics on business owners (up to four), including gender, age, race, ethnicity, citizenship, and human capital such as type of education, prior business experience, and veteran status. Most importantly for the purpose of this study, this allows me to identify African American entrepreneurs and their corresponding characteristics. I generate a set of mutually exclusive racial dummy variables, consisting of one Hispanic race variable as well as non-Hispanic White, African American, Asian and other minorities. Gender and citizenship are dummy variables, and age is expressed as six categorical variables from less than 25, 25-34, 35-44, 45-54, 55-64 and 65 or over. Demographic controls also include a set of owner-team variables, comprised of several diversity variables and a family-owned business dummy. I construct diversity variables in terms of gender, race and ethnicity, and immigrant within the owner team. Human capital variables include educational attainment, prior business experience, and veteran status. Educational attainment is defined by the highest degree prior to owning the business as a categorical variable (less than high school graduate, high school graduate, vocational/some college/associate degree, Bachelor’s degree and graduate degree). Prior business experience and veteran status are dummy variables.

For the purposes of this study, the ASE contains unusually detailed measures of finance. The ASE provides information on the amount of startup capital as ten categorical variables from less than $5,000 to $3 million or more, as well as “none needed” and “don’t know”. Based on these categorical variables I construct a dummy variable for greater than $100k to define the use of large amount of startup capital. It also asks the source of startup capital used to start or initially acquire the business including personal savings, home equity loan, credit cards, bank loan, government loan, family loan, venture capital and grants. Each of these is a dummy variable indicating the source of capital. Also, there are detailed questions related to new funding relationships such as whether the business attempted to engage with banks, credit unions, other financial institutions, angel investors, venture capitalists, other investor businesses, or grants. Along with these, the ASE provides an information whether they received the total amount of the requested funding or not from each of those sources. I use this information to create a dummy variable whether they received the total amount of the funding requested from each of new funding relationship. All these variables allow me to closely examine the financing patterns of African American entrepreneurs, specifically whether they are less likely to seek external financing and less likely to receive the full amount of new funding for which they applied. Furthermore, the ASE allows me to identify discouraged borrowers as well as their corresponding reasons by providing information on whether they chose not to apply during 2014 despite a need for additional financing, and the reasons for not applying, such as expected non-approval. Thus, I define the discouraged borrower variable, which equals to one if the owner chose not to apply despite the need of additional funding and the reason for not applying is that they did not think the business would be approved by a lender. Lastly, I create a dummy variable which indicates whether the access to financial capital negatively impacts the profitability of the business.

Differences in the amount of finance and in the frequency or manner of the attempts to receive finance may result from differences in motivations for owning the businesses. In addition to detailed financing variables, the ASE asks how important is for owning the business for nine different motivations with the options of very important, somewhat important, or not important. I construct two dummy variables for each motivation, where very important equals to 1 otherwise equals to 0, and somewhat important equals to 1 otherwise equals to 0. These motivations include 1) “Best avenue for my ideas/goods/services” (*Ideas*); 2) “Opportunity for greater income/wanted to build wealth” (*Income*); 3) “Couldn’t find a job/unable to find employment” (*No Job*); 4) “Wanted to be my own boss” (*Own Boss*); 5) “Working for someone else didn’t appeal to me” (*Work for Self*); 6) “Always wanted to start my own business” (*Always Wanted*); 7) “An entrepreneurial friend or family member was a role model” (*Role Model*); 8) “Flexible hours” (*Flexible Hours*); 9) “Balance work and family” (*Balance Family*).

The ASE also includes variables representing owner choices. Job function is a categorical variable for the owner’s main role in the business including manager, good/service provider, financial controller, and none of these roles. Primary income is a dummy variable indicating whether this business is the owner’s primary income source. Hours worked is a categorical variable for ranges of average weekly hours the owner spends managing or working in the business. Home-based is a dummy variable indicating whether the business operates primarily from somebody’s home.

2.3. Methods

My analysis includes summary statistics on all finance variables for African American vs. White, and I estimate regressions of measures of finance on race dummies and alternative sets of control variables. The main variable of interest is dummy for an African American owner, where the reference category is white. The first specification I report is a regression with no controls, providing the unconditional mean difference for dependent variable between African American and white owners. The second specification is a base regression including only two categorical variables: the number of owners (2-4 owner nummy, 5 or more owner dummy and don't know number of owner dummy, where one owner dummy is the base group), and firm age (dummies for age 0-2, 3-5, 6-10 and 11-15, where 15 and more is the base category), plus race categories. White is always the reference group, so that the coefficient of African American captures the differences in financing outcomes between African American and white owners. The third specification adds demographic controls, including gender, age and citizenship, and team variables. Team variables include indicators of having racial and ethnic diversity, gender diversity, or immigrant diversity within the team, and whether the business is family-owned. This specification would show how much of the financing difference is explained through demographic characteristics. The fourth specification adds human capital controls, including educational categories, prior business experience, and veteran status. The fifth adds motivations, based on nine survey questions about reasons for owning a business, which was described in detail above. The sixth adds 4-digit NAICS industry to look at within industry variation between African American and white owners. The last one adds other choice variables which are potentially related to business type or size and ultimately financing. This includes job function of the owner, indicators of whether the business is the source of primary income, and whether the business is a home-based.

I estimate the series of regressions using linear probability models where the financing outcomes are binary variables. The regressions are estimated with weighted least squares, and standard errors are clustered at firm. The firm-owner observations are weighted by ownership shares first, and firms are weighted by the ASE sample weight to be representative of the U.S. population. Equation (1) is specified below, describing the first regression in which I compare the unconditional mean of each outcome variable between African American and white business owners.

Equation 8

$$F_i = \alpha + \beta A_i + \sum_k \delta^k G_i^k + u_i$$

where F_i represents the various finance measures in the ASE which are described above in the data section. A_i is an indicator for an African American owner, where β captures the differences in financing between African American versus white owners. G_i^k is a dummy variable for a race/ethnicity group k (e.g., Asian, other minority, and Hispanics). Here, β reflects the mean difference in financing variable between African American and white entrepreneurs.

Then, the next baseline regression is specified in the following equation.

Equation 9

$$F_i = \alpha + \beta A_i + \sum_k \delta^k G_i^k + \delta' T_j + \gamma' N_j + u_i,$$

T_j is a vector of dummies for the number of owners for firm *j*, and N_j is a vector of firm age categories for firm *j*. I control firm age since firms are at different ages, and financing including demand and constraints is related with firm age.

Equation (2) estimates the raw financing gap without controlling for observable characteristics between the two racial groups. In order to understand how much of the gap can be explained by demographic characteristics, I estimate regressions with a set of owner characteristics, specified in the following equation:

Equation 10

$$F_i = \alpha + \beta A_i + \sum_k \delta^k G_i^k + \delta' T_j + \gamma' N_j + \boldsymbol{X}_i \boldsymbol{\gamma} + u_i,$$

where $\boldsymbol{X}_i$ is a vector of characteristics of owner *i*. This vector includes demographic variables, including owner age, gender, citizenship, and ownership team variables, indicators of having racial and ethnic diversity, gender diversity, or immigrant diversity within the owner-team and whether the business is family-owned. Here, the β estimated from equation (3) provides an unexplained gap in financing outcomes after adjusting for differences in demographic characteristics between African-American and white owners.

In addition to demographic characteristics, I include other human capital variables including education, prior business, and veteran status, to explain the gap.

Equation 11

$$F_i = \alpha + \beta A_i + \sum_k \delta^k G_i^k + \delta' T_j + \gamma' N_j + \boldsymbol{X}_i \boldsymbol{\gamma} + HK_i + u_i,$$

where HK_i is a vector of human capital characteristics of owner *i*. This includes educational attainment, prior business experience, and veteran status. Previous research finds that owners with higher educational attainment are less likely to be rejected from loan application (Coleman 2004), suggesting that educational differences may partially explain the differences in financing outcomes between African American and white owners. In this regard, controlling for a set of human capital variables provides information on how much gap is explained by differences in these characteristics between African American and white entrepreneurs.

African American owners may differ from white owners in terms of their motivations for owning a business. Most small business owners start their businesses for non-pecuniary reasons without intention to grow or innovate (Hurst and Pugsley 2011). African American owners may have different motivations to own their business, which may influence their financing patterns. I estimate this difference with the following specification:

Equation 12

$$F_i = \alpha + \beta A_i + \sum_k \delta^k G_i^k + \delta' T_j + \gamma' N_j + \boldsymbol{X}_i \boldsymbol{\gamma} + HK_i + M_i + u_i$$

where M_i is the set of motivation variables. As described in the data section, the survey questionnaire asks about the reasons for owning a business. The nine motivation variables available include; new idea, income, no job, own boss, work for self, always wanted, role model, flexible hours, and balance work and life. The survey asks respondents to indicate whether each motivation is "not important", "somewhat important", or, "very important". In this specification, I control dummies for "somewhat important" and "very important" for each motivation.

Different choices of industry may affect financing patterns between African American and white owners, and controlling for industry allows me to compare them within the same industry. If African American owners are more likely to own businesses in industries requiring more capital, the financing gap in previous specifications may be influenced by different choices of industries. I estimate the regression with the following specification:

Equation 13

$$F_i = \alpha + \beta A_i + \sum_k \delta^k G_i^k + \delta' T_j + \gamma' N_j + \boldsymbol{X}_i \boldsymbol{\gamma} + HK_i + M_i + S_i + u_i,$$

where S_i is the set of vectors of 4-digit NAICS industry dummies. Industry decision may be endogenously determined with finance, thus it is important to control for industry to find the remaining gap in finance between African American and white business owners.

Finally, different choices that owners make for their businesses may explain the gap in finance between African American and white entrepreneurs. Below is estimated with the following equation:

Equation 14

$$F_i = \alpha + \beta A_i + \sum_k \delta^k G_i^k + \delta' T_j + \gamma' N_j + \boldsymbol{X}_i \boldsymbol{\gamma} + HK_i + M_i + S_i + C_i + u_i,$$

where C_{ij} include the set of owner choice variables including job function in business, whether the businesses is the source of primary income or not, the number of hours worked, and whether the business is home-based or not. Together, I am able to explore the potential difference in financing, which can be explained by observed characteristics including demographic characteristics, human capital, motivation, industry choice (4-digit NAICS), or other choices described above. Estimations from equation (1) through (7) helps to understand how much of the gap in finance between African Americans and white business owners is explained by those observable characteristics. Any remaining gap after controlling for such variables, estimation from equation (7), may reflect other unobserved characteristics or discrimination.

2.4 Results

2.4.1 Descriptive Statistics

Tables 12 through 19 provide descriptive statistics for all variables used in the regressions. The first column provides the mean for the full sample, and the next two columns are sub-samples for African Americans and white business owners. Table 12 provides the share of owners across each racial group in the sample. The main variable of interest is non-Hispanic Black, called African American in this study, representing about 2 percent, whereas non-Hispanic white owners make up roughly 85 percent of the full sample. Followed by white, Non-Hispanic Asians and Hispanics are next largest share of the sample, and non-Hispanic other race group is the smallest at less than one percent.

Table 12 Summary Statistics: Owner Race and Ethnicity

	All	Black	White
Owner Race/Ethnicity			
Non-Hispanic Black	0.0172	1.0000	0.0000
Non-Hispanic White	0.8354	0.0000	1.0000
Non-Hispanic Asian	0.0879	0.0000	0.0000
Non-Hispanic other race	0.0080	0.0000	0.0000
Hispanic	0.0515	0.0000	0.0000

Note: N = 288,000 individual owners of 184,000 employer-firms. Owners are weighted by their ownership share in the firm and by the ASE weights, so the sample is representative of all employer-firms in the U.S. non-farm private sector.

Table 13 shows the summary descriptive for financial variables. African American owners tend to use less capital to start their business, and they are more likely to use less than $100k in startup capital. African Americans are also less likely to report that no capital is needed for their startups compared to white owners. In terms of startup capital source, African American business owners tend to use personal savings, home equity loans, credit cards, or government loans, whereas white owners are more likely to use bank loans or family loans. At less than one percent, venture capital and grants are less likely to be used for starting a business by either group, but African Americans tend to use much more grants relative to white owners.

Table 13 Summary Statistics: Finance

	All	Black	White
Startup Capital Amount			
No capital needed	0.0832	0.0696	0.0891
Capital under 5k	0.1545	0.1912	0.1595
Capital 5k to 10k	0.0852	0.1034	0.0848
Capital 10k to 25k	0.1216	0.1544	0.1196
Capital 25k to 50k	0.0959	0.1082	0.0929
Capital 50k to 100k	0.1033	0.1189	0.0986
Capital 100k to 250k	0.1027	0.0788	0.0974
Capital 250k to 1m	0.0675	0.0517	0.0661
Capital 1m to 3m	0.0148	0.0087	0.0147
Capital 3m and more	0.0055	0.0044	0.0055
Don't know amount	0.1657	0.1109	0.1717
More than 100k	0.1905	0.1436	0.1838
Startup Capital Source			
Personal savings or other assets	0.6908	0.7445	0.6782
Home equity loans	0.0745	0.0793	0.0727
Personal/business credit cards	0.1266	0.1993	0.1221
Bank loans	0.1837	0.1539	0.1902
Government loans	0.0232	0.0346	0.0230
Family loans	0.0520	0.0328	0.0518
Venture capital	0.0050	0.0060	0.0050
Grants	0.0019	0.0061	0.0017
New Funding received in 2014			
Banks or other financial institutions	0.0958	0.0786	0.0987
Angel investors or VC	0.0028	0.0025	0.0023
Other investor businesses	0.0027	0.0028	0.0022
Grants	0.0024	0.0043	0.0018
Financial Constraints			
Expected lender would not approve	0.0464	0.1489	0.0425
Lack of capital reduced profits	0.1068	0.2726	0.0964

Note: N = 288,000 individual owners of 184,000 employer-firms. Owners are weighted by their ownership share in the firm and by the ASE weights, so the sample is representative of all employer-firms in the U.S. non-farm private sector.

Not only for the startup capital but also for new funding, African American owners are less likely to use banks or other financial institutions. On the other hand, they tend to use more angel investors or venture capital, other investor businesses or grants compared to white owners, but these all make up a small share of the activity, at less than one percent. Together, looking at the source of financial capital, African American entrepreneurs are less likely to use external financing relative to white entrepreneurs.

Finally, they have a much higher propensity to be discouraged from seeking additional funding because they did not think that their request would be approved, although they have much higher propensity to think access to finance is a critical role in generating profits for their businesses. This may suggest the existence of financial constraints against African Americans despite the greater importance of access to finance on their business profitability.

Table 14 Summary Statistics: Number of Owners and Firm Age

	All	Black	White
Number of Owners			
Single owner	0.5847	0.6851	0.5809
2 to 4 owners	0.3775	0.2915	0.3810
5 or more owners	0.0332	0.0187	0.0339
Don't know number of owners	0.0046	0.0047	0.0042
Firm Age (years)			
0 to 2	0.1423	0.2242	0.1287
3 to 5	0.1460	0.1986	0.1343
6 to 10	0.2138	0.2404	0.2074
11 to 15	0.4707	0.3180	0.4998
15 or older	0.0273	0.0189	0.0298

Note: N = 288,000 individual owners of 184,000 employer-firms. Owners are weighted by their ownership share in the firm and by the ASE weights, so the sample is representative of all employer-firms in the U.S. non-farm private sector. Firm age is defined as number of years since the first employee was hired.

In Table 14, descriptive statistics for number of owners and firm age are presented. African American-owned firms tend to be smaller as they have a higher propensity to be owned by a single person. White-owned firms are more likely to have two or more owners. Regarding firm age, African American-owned firms tend to be younger, with more than half of them being less than 11 years old, while more than half of white-owned firms are older than 10 years.

Table 15 Summary Statistics: Demographics

	All	Black	White
Gender			
Female	0.2804	0.3793	0.2688
Male	0.7196	0.6207	0.7312
Owner Age			
Less than 35	0.0532	0.0528	0.0488
35 to 44	0.1662	0.2088	0.1503
45 to 54	0.2898	0.3187	0.2822
55 to 64	0.3102	0.2711	0.3229
65 or more	0.1807	0.1486	0.1957
Immigrant	0.1549	0.2005	0.0655
Team			
Race/ethnicity diversity	0.0311	0.0857	0.0184
Gender diversity	0.0392	0.0426	0.0371
Immigrant diversity	0.0335	0.0389	0.0257
Family	0.2128	0.1539	0.2154

Note: N = 288,000 individual owners of 184,000 employer-firms. Owners are weighted by their ownership share in the firm and by the ASE weights, so the sample is representative of all employer-firms in the U.S. non-farm private sector.

Table 15 provides demographic characteristics of business owners. Female business owners are more present in African American-owned firms, while white-owned firms tend to be owned by men. The age categories show an inverse U shape relationship between owner age and propensity of owning business, as it increases up to 55 to 64 groups and decreases with older than 65 group. African American owners tend to be younger than white owners. Surprisingly, they are much more likely to be immigrants compared to white owners. African American-owned businesses tend to be more diverse in terms of race and ethnicity, gender, and immigrant status within the team. There is a lower share of family-owned business for African Americans compared to white owners.

Table 16 Summary Statistics: Human Capital

	All	Black	White
Education			
Less than high school	0.0334	0.0265	0.0248
High school	0.1862	0.1322	0.1875
Vocational/Some college	0.2636	0.2636	0.2717
Undergraduate	0.2774	0.2391	0.2832
Graduate	0.2394	0.3386	0.2329
Prior business experience	0.3220	0.2725	0.3224
Veteran status	0.0999	0.1258	0.1105

Note: N = 288,000 individual owners of 184,000 employer-firms. Owners are weighted by their ownership share in the firm and by the ASE weights, so the sample is representative of all employer-firms in the U.S. non-farm private sector.

Table 16 provides summary statistics for human capital of business owners. In terms of education, African Americans tend to have a lower level of education compared to white owners up to the undergraduate level. However, a much higher share of African American business owners hold graduate degrees, as more than 30 percent of African Americans hold graduate degree, compared to about 20 percent of white owners. While the general education level of African American is lower than whites[19], this may suggest that highly educated (higher than college) African Americans appear to have a higher probability of becoming business owners compared to whites. African American owners have a lower propensity to have prior business experience, and a higher propensity of being a veteran compared to white owners.

Table 17 Summary Statistics: Motivations for Business Ownership

	All	Black	White
Wanted to be own boss	0.5663	0.6087	0.5679
Wanted flexible hours	0.4376	0.5274	0.4297
Balance work and family	0.4762	0.5550	0.4657
Opportunity for greater income	0.5421	0.6259	0.5363
Best avenue for ideas/goods/service	0.4992	0.5776	0.4935
Unable to find employment	0.0670	0.0907	0.0593
Want to work for self	0.2743	0.2768	0.2754
Wanted to start business	0.4144	0.5799	0.3944
Entrepreneurial role model	0.2402	0.2788	0.2340

Note: N = 288,000 individual owners of 184,000 employer-firms. Owners are weighted by their ownership share in the firm and by the ASE weights, so the sample is representative of all employer-firms in the U.S. non-farm private sector.

In Table 17, African American owners have a higher propensity to own businesses for all nine motivations. The difference in propensity between two groups is particularly larger for the motivations of "No job" and "Wanted to own a business."

[19] Looking at educational attainment between African Americans and whites from American FactFinder using 2013-2017 American Community Survey (ACS) 5-year estimates, share of high school graduate or higher population between African American and white is 84 percent vs. 89 percent, respectively. (https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_17_5YR_S1501&src=pt)

Income is the largest share of motivation for owning a business for African Americans, while a desire to be one's own boss is the largest share of motivation for whites.

In Table 18, I show the differences in industry (2-digit NAICS sectors) between African American and white business owners. Primary sector includes the agriculture and mining sectors, as these industries have a small number of firms in the data. African American-owned businesses are more prevalent in the transportation, health and education sectors.

Table 18 Summary Statistics: Industry

	All	**Black**	**White**
Primary sector	0.0098	0.0037	0.0110
Construction	0.1253	0.0706	0.1369
Manufacturing	0.0471	0.0128	0.0509
Wholesale trade	0.0552	0.0194	0.0549
Retail trade	0.1151	0.0592	0.1109
Transportation	0.0293	0.0501	0.0294
Information	0.0118	0.0134	0.0125
Finance	0.0453	0.0511	0.0483
Real estate	0.0485	0.0301	0.0521
Professional and management	0.1628	0.1707	0.1682
Administrative and support	0.0612	0.0884	0.0632
Education	0.0106	0.0214	0.0103
Health	0.1118	0.2749	0.1024
Art and entertainment	0.0165	0.0171	0.0181
Accommodation and food	0.0779	0.0475	0.0615
Other services	0.0670	0.0665	0.0636
Missing sector	0.0051	0.0030	0.0058

Note: N = 288,000 individual owners of 184,000 employer-firms. Owners are weighted by their ownership share in the firm and by the ASE weights, so the sample is representative of all employer-firms in the U.S. non-farm private sector. Primary sector includes agriculture and mining sector. Manufacturing comprises NAICS 31-33. Retail trade comprises NAICS 44-45. Transportation comprises NAICS 48-49. Professional and management comprises NAICS 54-55.

Lastly, Table 19 includes descriptive statistics for choices that owners make for their businesses. While African American owners tend to be managers or producers, white owners are more likely to serve as financial controller. In terms of working hours, African American owners tend to work fewer hours in general, but a higher share of African Americans work more than 60 hours relative to white owners. African Americans and white owners have a similar propensity to choose the business as their primary source of income and home-based.

Table 19 Summary Statistics: Owner Choice

	All	Black	White
Owner Role in Business			
Manager	0.7979	0.8253	0.7985
Producer	0.6238	0.6714	0.6330
Financial controller	0.7288	0.7100	0.7480
None listed	0.0627	0.0489	0.0623
Average Hours Per Week Owner Works in Business			
None	0.0571	0.0404	0.0576
Less than 20 hours	0.1354	0.1154	0.1365
20 to 39 hours	0.1478	0.1436	0.1486
40 hours	0.1519	0.1329	0.1462
41 to 59 hours	0.3019	0.2776	0.3091
60 hours or more	0.2059	0.2901	0.2019
Business is primary source of income	0.7282	0.7085	0.7260
Home-based	0.2381	0.2515	0.2501

Note: N = 288,000 individual owners of 184,000 employer-firms. Owners are weighted by their ownership share in the firm and by the ASE weights, so the sample is representative of all employer-firms in the U.S. non-farm private sector.

2.4.2 Regression results

Table 20 shows the regression results for the amount of startup capital using the seven specifications described earlier. (1) displays the mean differences between African American and white entrepreneurs, (2) adding number of owner and firm age categories, (3) adding demographic controls, (4) adding measures of human capital, (5) adding reported motivations for business ownership, (6) adding industries, and (7) adding owner's choices. Additional controls used in each regression are labeled at each column of the table.

Table 20 Regression Results: Startup Capital Amount (Greater than 100K)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	No Controls	+N owner Firm age	+Demog.	+Human K	+Motivation	+Industry	+Choice
	-0.0402	-0.0361	-0.0394	-0.04	-0.043	-0.0424	-0.0457
	(0.0065)	(0.0066)	(0.0066)	(0.0065)	(0.0065)	(0.0063)	(0.0062)

Note: N = 288,000 individual owners of 184,000 employer-firms. Owners are weighted by their ownership share in the firm and by the ASE weights, so the sample is representative of all employer-firms in the U.S. non-farm private sector. These are results from LPM regressions. Number of owners and firm age are included as category variables described in table 14. Demographics includes all variables from table 15, and human capital includes all variables from table 16. Motivations includes all variables from table 17. Industry includes four-digit NAICS industry dummies, and choice variables are all variables from table 19.

The results show that African Americans have about a 4.6 percentage point lower propensity to use startup capital amounts greater than $100,000 compared to white owners. Less negative effect after controlling for the number of owners and firm age suggests that the size and the age of firm affects financing. African American-owned firms tend to be smaller and younger, where larger firms are more likely to use larger amounts of startup capital (Brown et al 2019). Differences in demographics, human capital and motivation do not help to close the gap in finance, while industry differences slightly explain the gap. However, the gap in finance between African American and white owners still remains after controlling for all sets of observable characteristics.

Table 21 provides the results on different sources of startup capital between African American and white-owned businesses. Consistent with previous findings (Robb 2018), African Americans are more likely to use personal savings or credit cards to finance their startups. Not only differences in demographics explain the differences in sources partially, but also the differences in motivations (African American firms are more likely to start the business to make greater income) explain the differences as the gap gets smaller with motivation controls.

Table 21 Regression Results: Startup Capital Source

	(1) No Controls	(2) +N owner Firm age	(3) +Demog.	(4) +Human K	(5) +Motivation	(6) +Industry	(7) +Choice
Personal savings	0.0663	0.0534	0.0461	0.0465	0.0415	0.0454	0.0461
or other assets	(0.0083)	(0.0083)	(0.0083)	(0.0083)	(0.0083)	(0.0082)	(0.0082)
Home Equity	0.0066	0.0082	0.0072	0.0083	0.0067	0.005	0.0034
	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0049)
Personal/business credit cards	0.0772	0.0651	0.0619	0.0619	0.0590	0.0572	0.0545
	(0.0075)	(0.0075)	(0.0075)	(0.0075)	(0.0075)	(0.0075)	(0.0075)
Bank or financial institutions	-0.0363	-0.0205	-0.0116	-0.0159	-0.0190	-0.0234	-0.0245
	(0.0067)	(0.0067)	(0.0067)	(0.0067)	(0.0067)	(0.0064)	(0.0063)
Government loan	0.0115	0.0128	0.0127	0.0119	0.0113	0.01	0.0096
	(0.0032)	(0.0032)	(0.0033)	(0.0033)	(0.0032)	(0.0032)	(0.0032)
Family loan	-0.0191	-0.0176	-0.0171	-0.0165	-0.0163	-0.0155	-0.0159
	(0.0032)	(0.0032)	(0.0033)	(0.0033)	(0.0033)	(0.0033)	(0.0033)
Venture capital	0.001	0.0008	0.0006	0.0006	0.0003	0.0004	0.0002
	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)
Grants	0.0044	0.0042	0.0039	0.0037	0.0037	0.0024	0.0023
	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0014)	(0.0014)

Note: N = 288,000 individual owners of 184,000 employer-firms. Owners are weighted by their ownership share in the firm and by the ASE weights, so the sample is representative of all employer-firms in the U.S. non-farm private sector. These are results from LPM regressions. Number of owners and firm age are included as category variables described in table 14. Demographics includes all variables from table 15, and human capital includes all variables from table 16. Motivations includes all variables from table 17. Industry includes four-digit NAICS industry dummies, and choice variables are all variables from table 19.

On the other hand, African American entrepreneurs are less likely to use external financing such as loans from banks or other financial institutions for their startups. As expected, firm size and age affect getting finance from banks, which explains the gap against African American-owned firms for being smaller and younger compared to white-owned firms. Notably, differences in demographics and human capital between the two groups explain much of the gap. With demographic controls, the coefficient rises (becomes less negative) from -0.0205 to -0.0116. This reflects a higher propensity of African Americans for characteristics which are negatively associated with financing, such as being female, younger, having lower educational attainment and less prior business experience. However, controlling for motivation and industry does not help to reduce the gap in getting finance from banks between African Americans and white business owners. In fact, the gap falls back to -0.0234 with industry controls. Nonetheless, all these set of characteristics together explain about 30 percent of the gap. Besides bank financing, African Americans are less likely to use capital from family loans. This may happen from a difference in family wealth between the two groups if African American family initially does not have enough financing capacities. Differences in demographics, human capital, motivation, and industry all partially explain the gap, and this holds across all specifications.

The differences in using venture capital to start or initially acquire a business are nearly zero, and are insignificant between African American and white business owners. African American owners are more likely to use government loans or grants for their startup capital, and differences in human capital and demographics partially explain the gap in each source, respectively. However, once industry is controlled, the significance disappears for grants, suggesting African American-owned firms more often choose industries with more grants.

Table 22 displays results for financial variables measuring successful attempts to establish new funding in 2014 from different sources (such that the total amount of the funding requested was received). The coefficients and standard errors of this table are multiplied by 100 for ease of reading. Differences in demographics, human capital, motivation, and industry explain the financing gap between African American and white business owners in receiving total amount of new funding from banks. The story is similar to the finding for the source of startup capital, where differences in demographics and human capital explain much of the financing gap from the banks, as these characteristics of African Americans negatively affect external financing. For instance, female entrepreneurs are more likely to face credit constraints (Blanchard, Zhao and Ying 2008, Coleman 2002); younger entrepreneurs face higher financial constraints (Cabral and Mata 2003); and education level is highly correlated to the level of debt capital (Bates, 1990, Coleman 2004). All these factors are more heavily represented in African American business owners, explaining the gap between African American and white owners. African Americans are also less likely to receive new funding from angel investors or venture capitals, and industry differences explain the half of the gap, although the estimation is almost zero and not statistically significant.

Table 22 Regression Results: New Funding Source

	(1) No Controls	(2) +N owner Firm age	(3) +Demog.	(4) +Human K	(5) +Motivation	(6) +Industry	(7) +Choice
Banks, credit unions, or	-2.007	-1.781	-1.315	-1.161	-1.482	-1.473	-1.601
Other financial institutions	(0.494)	(0.495)	(0.497)	(0.497)	(0.496)	(0.502)	(0.500)
Angel investors or	0.0244	-0.0276	-0.0595	-0.0523	-0.0576	-0.0241	-0.0468
Venture capitals	(0.0824)	(0.0828)	(0.0839)	(0.0839)	(0.0850)	(0.0869)	(0.0870)
Other investor businesses	0.0658	0.0220	0.0001	0.0048	0.0052	-0.0023	-0.0133
	(0.0991)	(0.0992)	(0.0993)	(0.0993)	(0.0994)	(0.1019)	(0.1020)
Grants	0.245	0.2279	0.2051	0.2029	0.2044	0.1623	0.1519
	(0.1078)	(0.1082)	(0.1090)	(0.1089)	(0.1093)	(0.1122)	(0.1121)

Note: N = 288,000 individual owners of 184,000 employer-firms. Owners are weighted by their ownership share in the firm and by the ASE weights, so the sample is representative of all employer-firms in the U.S. non-farm private sector. These are results from LPM regressions. Number of owners and firm age are included as category variables described in table 14. Demographics includes all variables from table 15, and human capital includes all variables from table 16. Motivations includes all variables from table 17. Industry includes four-digit NAICS industry dummies, and choice variables are all variables from table 19.

Differences in new funding from other investor businesses between African American and white owners are nearly zero, whereas African American owners are more likely to receive new funding from grants. For both sources, industry differences explain the gap between African American and white entrepreneurs' financing. In fact, except for external financing from banks, credit unions, and other financial institutions, the gap in new funding between African American and white entrepreneurs from other sources is fairly small, close to zero, and insignificant.

Table 23 Regression Results: Avoid Additional Funding

(1)	(2)	(3)	(4)	(5)	(6)	(7)
No Controls	+N owner Firm age	+Demog.	+Human K	+Motivation	+Industry	+Choice
0.1064	0.1015	0.0995	0.1007	0.0987	0.0995	0.0969
(0.0067)	(0.0067)	(0.0067)	(0.0067)	(0.0067)	(0.0067)	(0.0067)

Note: N = 288,000 individual owners of 184,000 employer-firms. Owners are weighted by their ownership share in the firm and by the ASE weights, so the sample is representative of all employer-firms in the U.S. non-farm private sector. These are results from LPM regressions. Number of owners and firm age are included as category variables described in table 14 Demographics includes all variables from table 15, and human capital includes all variables from table 16. Motivations includes all variables from table 17. Industry includes four-digit NAICS industry dummies, and choice variables are all variables from table 19.

Table 23 and 24 shows the regression results for subjective questions on the underlying reasons for avoidance of new funding and credit constraints. Across all specifications, African Americans have about 10 percentage point higher probability of being discouraged borrowers compared to white business owners, with a mean of 4.6 percent. The gap is partially explained by differences in industry of businesses. In Table 23, I find that African American owners are less likely to receive external financing from banks or other financial institutions, and this finding suggests that the underlying reason for not applying for further loans is that they did not expect to be approved by lenders.

Table 24 Regression Results: Access to Finance Negatively Impacts Negatively

(1)	(2)	(3)	(4)	(5)	(6)	(7)
No Controls	+N owner Firm age	+Demog.	+Human K	+Motivation	+Industry	+Choice
0.1762	0.1678	0.164	0.167	0.1615	0.1648	0.1596
(0.0083)	(0.0083)	(0.0083)	(0.0083)	(0.0083)	(0.0083)	(0.0083)

Note: N = 288,000 individual owners of 184,000 employer-firms. Owners are weighted by their ownership share in the firm and by the ASE weights, so the sample is representative of all employer-firms in the U.S. non-farm private sector. These are results from LPM regressions. Number of owners and firm age are included as category variables described in table 14. Demographics includes all variables from table 15, and human capital includes all variables from table 16. Motivations includes all variables from table 17. Industry includes four-digit NAICS industry dummies, and choice variables are all variables from table 19.

At the same time, African American entrepreneurs have about a 16 percentage point higher propensity of considering access to finance as an important factor of the profitability of their businesses. This finding holds across all specifications and is statistically significant, implying African American entrepreneurs bear credit constraints relative to white entrepreneurs. Compared to earlier results, the gap size in these measures is much greater, indicating a larger gap in access to finance between African American and white entrepreneurs.

2.5 Conclusion

This paper attempts to answer whether African Americans are more financially constrained than their white counterparts. Previous literature suggests the existence of a financing gap between African American and white entrepreneurs using small sample sizes and limited finance variables with little information on other business owners and firm characteristics. In this study, I am able to study African Americans' financing patterns and the gap relative to white owners using a large nationally representative sample with much more detailed owner and firm variables. The outcome variables include capital amount and the source of startup capital. Not only startups but also the source of additional funding is examined to study difference in source of new funding relationship between African Americans and white entrepreneurs. Moreover, I examine discouraged borrowers to explore underlying reason for not seeking further loans, and how important access to finance plays in generating profits for businesses.

To estimate the gap in financing between African American and white business owners, I use an extensive set of firm and owner characteristics, including gender, age, citizenship, diversity within team, education attainment, prior business experience and veteran status. Also, I control for differences in motivations of owners to own the businesses, differences in industry and the choices that owners make on their businesses. Controlling for all of these variables is important in understanding how much of the financing gap is explained by those variables, and it allows me to estimate the remaining gap that is not explained by those observable characteristics, which may reflect discrimination in capital markets.

The results suggest that the gap in finance between African Americans and white entrepreneurs still remains after I control a wide variety of owner and firm characteristics which may affect the financing patterns of two groups. African American entrepreneurs are more likely to use a lower amount of startup capital, and they are less likely to use external financing from banks or other financial institutions relative to white owners. They are more likely to use personal savings, credit cards, and government loans to start their businesses. This finding is robust across all specifications, and demographics and human capital explain much of the gap between African Americans and white entrepreneurs, while industry and choice differences do partially. Also, African Americans have a higher propensity for not seeking further loans from banks or other financial institutions, because they expect to be rejected. At the same time, they are more likely to consider access to finance as a crucial factor in determining the profitability of their businesses, implying that they face higher credit constraints compared to white owners. These results are also robust across all specifications. Findings from this study provide additional evidence that the unexplained gap in finance between African American and white entrepreneurs still exists, suggesting that African American entrepreneurs are more financially constrained.

CHAPTER 3 IMMIGRANT ENTREPRENEURS AND INNOVATION IN THE U.S. HIGH-TECH SECTOR (WITH J. DAVID BROWN, JOHN S. EARLE, AND KYUNG MIN LEE)[20]

3.1 Introduction

How much do immigrants contribute to innovation? Popular accounts of U.S. science, engineering, and high-tech business creation tend to cast immigrants in a starring role, and anecdotes on exceptional immigrants are easy to find, but systematic evidence is rare. A number of studies have examined immigrants as individual inventors, as employees of high-tech firms, and as scientists, engineers, and self-employed (e.g., Wadhwa et al. 2007a and 2007b, Kerr and Lincoln 2010, Hunt 2011).[21]

However, there have been few studies of immigrant entrepreneurs, and most of those focus on firm size and growth.[22] Only Hart and Acs (2010) examine innovation measures – research and development and patenting – at the firm-level, using a survey of

[20] We thank the National Science Foundation (NSF) for support (Grants 1262269 and 1719201 to George Mason University). Any opinions and conclusions expressed herein are ours only and do not necessarily reflect the views of the NSF or the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. The Disclosure Review Board bypass numbers are CBDRB-2018-CDAR-087, DRB-B0017-CED-20181126, and DRB-B0025-CED-20181219.

[21] Other contributions to these topics include Stephan and Levin (2001), Peri (2007), Hunt and Gauthier-Loiselle (2010), and Kerr (2013).

[22] As far as we can determine, the only studies of job creation by immigrant entrepreneurs using broad, representative samples are Fairlie and Lofstrom (2014) and Kerr and Kerr (2017, 2018). Brown et al. (2018) analyze immigrant status among other founder characteristics in a study of high-growth entrepreneurship. Our focus on innovation outcomes is different, but we build on this work and provide some comparisons with our approach below. A few other studies focus on particular industries, regions, or immigrant ethnicities.

1,300 "high impact" high-tech companies.[23] They report little difference between firms with and without immigrant founders, but they consider a sample of firms already at the right tail of the firm performance distribution. Such data do not permit research to draw any inferences on the relative innovativeness of typical high-tech businesses owned by immigrants and natives, which is the question we address in this paper.

This paper aims to contribute to understanding the innovation impact of immigrant entrepreneurship on the U.S. high-tech sector using a much larger and richer data set than those heretofore available. We analyze the Annual Survey of Entrepreneurs (ASE), a new database from the U.S. Census Bureau covering about 11,000 owners of 7,400 high-tech employer businesses based on a random sample of all nonfarm businesses. Like the well-known Surveys of Business Owners (SBO), the ASE questionnaire contains detailed information on the four largest owners and some characteristics of the business, which provide us with control variables for measuring immigrant-native differences conditional on other characteristics including demographics, human capital, and ownership team. Unlike the SBO, however, and crucially for this paper, the ASE includes many innovation measures that form the outcome variables in our study, including reported innovation activities in both products and processes, research and development, trademarks, and patents.[24]

[23] Saxenian (2002) and Wadhwa et al. (2007a) examine immigrants as owners but do not measure innovation at their firms.
[24] The random sampling for the ASE contrasts with the usual approach in "innovation surveys," including the Business Research and Development and Innovation Survey (BRDIS) in the U.S., where the sample is principally based on firms known or expected to be carrying out R&D.

The ASE also contains a number of variables that permit more disaggregated analysis. Data on race/ethnicity permit some examination of immigrant country of origin. Data on educational attainment allow us to estimate separately by education group. We are also able to examine immigrant-native differences in the roles played by a number of factors that may be jointly determined with innovation outcomes, including ownership motivations, start-up capital, and choice of industry. For all of these variables, we are interested both in characterizing immigrant relative to native entrepreneurs and in measuring how they influence or mediate the immigrant-native entrepreneur differences in innovation performance.

The subject of our study lies at the intersection of several large areas of research. To start with, there is a voluminous literature on the economic effects of immigrants. Most of this research focuses on the consequences of immigration for native worker wages and treats immigrants as a qualitatively similar factor of production, so that immigration represents a labor supply shock to a particular region or education-experience group (e.g., Card 1990, 2001; Borjas and Doran 2015; Borjas and Monras 2017; Ottaviano and Peri 2012; Peri 2012, 2015). Other immigration research focuses on the disadvantage faced by immigrants in U.S. labor markets and the extent and pace of immigrant-native convergence in wages, or "assimilation" (Borjas 1985, 2015; Chiswick, Lee, and Miller 2005; Chiswick 2009). Some studies of immigrants consider the possibility that immigrants have certain advantages and document higher rates of science, technology, engineering, and mathematics (STEM) workforce participation, patents, publication citations, and Nobel Prize winners among immigrants (Kerr and Lincoln 2010, Stephan and Levin 2001, Hunt and Gauthier-Loiselle 2010).

Yet much innovation takes place within firms, and our study relates to research on firm-level Research and Development (R&D), patenting, and other aspects of innovation. As is widely recognized, however, R&D and patents both have limitations as measures of innovation, much of which takes place without formal R&D or patenting. Some surveys, including the Community Innovation Surveys (CIS) in Europe and the Business Research and Development and Innovation Survey (BRDIS) in the U.S., attempt to fill this gap with qualitative questions on product and process innovations (Mairesse and Mohnen 2010). These surveys have documented the incidence of such activities and demonstrated their correlation with productivity (e.g., Griffith et al. 2006, Parisi et al. 2006, Hall 2011). But the data in these studies are usually based on small samples (for example, only 5,000 receive the full questionnaire for the BRDIS) that are non-randomly selected to focus on firms with known R&D activity. Still more importantly for our purposes, they contain no information on the firm's founders or owners.

Such characteristics have been extensively analyzed in the literature on self-employment determinants, including immigration status (e.g., Fairlie and Lofstrom 2014). But they are seldom measured for owners of firms, as distinguished from own-account (employee-less self-employed) workers. And a rich set of owner-founder characteristics has never before been linked to the kind of innovation measures that have become common in firm-level studies.

We find uniformly higher rates of innovation in immigrant-owned firms for 13 of 14 different measures. In most but not all cases the differences are statistically significant, and in most cases they survive detailed controls for other demographic and human capital characteristics of the entrepreneurs, as well as the size and family composition of teams. In many cases, they also remain significant in specifications controlling for start-up finance, motivations, and industry. The immigrant-native difference holds for both recent start-ups and older firms and at all levels of the entrepreneur's education. The main exception is owning a copyright or trademark, the most marketing-related activities measured here. Otherwise, the data imply a robust immigrant advantage in innovation.

The rest of the paper is organized as follows. Section 2 describes the data, and Section 3 the methods. Section 4 contains results, and Section 5 concludes.

3.2 Data

We exploit new confidential microdata from the Census Bureau's 2014 Annual Survey of Entrepreneurs (ASE). The ASE is an annual survey that supplements the Survey of Business Owners (SBO), conducted every five years, providing detailed demographic characteristics on business owners and their motivations to start a business, as well as economic characteristics of their firms. Of particular importance for this paper, it includes a rich set of innovation measures, which are the main outcome variables in our study.

The ASE sample contains non-farm businesses with at least one paid employee and receipts of $1,000 or more. Using the Census Business Register (BR) as the sampling frame, the ASE sample is stratified by the 50 most populous Metropolitan Statistical Areas (MSAs), state, and the firm's number of years in business.[25] The ASE sample is randomly selected, except for large companies in each stratum, which are selected with certainty based on volume of sales, payroll, or number of paid employees. The initial 2014 ASE sample was about 290,000 employer firms, and the response rate was 74 percent.

For this paper, we restrict the full ASE sample to firms in the high-tech sector as defined by the share of Science, Technology, Engineering, and Mathematics (STEM) employment in the industry.[26] This represents about 5.31 percent of firm-owner observations in the ASE. We also exclude businesses where no individual owns at least 10 percent of the equity, because detailed owner information is not provided for such businesses. We drop owners who choose the same answers for every motivation question (all very important, all somewhat important, or all not important), because those answering patterns may not reflect the true intensity for each question, as well as firm-owner observations that have missing values for any of the variables used in the regressions. Our final sample consists of about 11,000 owners of 7,400 firms. We weight each owner by their ownership equity share, adjusting them to sum up to one within each

[25] See Foster and Norman (2016) for further details about the ASE.

[26] We define high-tech sector based on the share of STEM employment in the industry, using Bureau of Labor Statistics data; for the exact definition, see Goldschlag and Miranda (2016, p. 58).

firm, and we weight each firm by ASE survey weights to make the sample representative for the U.S. economy.

Our main variable of interest is an indicator for whether the owner is an immigrant, defined in the ASE as a noncitizen at birth.[27] As we examine the differences in the propensity to innovate between immigrant and native owners, we control for various other owner and firm characteristics. The owner demographic characteristics consist of gender, age, race and ethnicity, type of education, prior business experience, and veteran status. We also include the relationships among business owners in firms with multiple owners, whether they are couple-owned, non-couple family-owned, or multi-generation. Variable construction is similar to the procedures in Brown et al. (2018).

The ASE asks about nine different motivations for owning the business, including 1) "Best avenue for my ideas/goods/services" (*Ideas*); 2) "Opportunity for greater income/wanted to build wealth" (*Income*); 3) "Couldn't find a job/unable to find employment" (*No Job*); 4) "Wanted to be my own boss" (*Own Boss*); 5) "Working for someone else didn't appeal to me" (*Work for Self*); 6) "Always wanted to start my own business" (*Always Wanted*); 7) "An entrepreneurial friend or family member was a role model" (*Role Model*); 8) "Flexible hours" (*Flexible Hours*); and 9) "Balance work and family" (*Balance Family*). These questions ask how important the reason is: not

[27] This definition reflects a change in practice relative to previous surveys such as the SBO which asked about birthplace (whether in the U.S.). The difference is in people who were born outside the U.S. but as citizens (i.e., because at least one parent was a citizen at the time). We nonetheless retain the conventional labels "immigrant" and "native" in our analysis.

important, somewhat important, or very important. In the descriptive statistics, we collapse the variables for a particular motivation into a single variable equaling 0 if not important, 1 if somewhat important, and 2 if very important, while in the regressions we include separate dummies for somewhat important and very important for each motivation.

In some specifications we also use amount of start-up capital and 4-digit NAICS industries as controls. Amount of finance used to start or initially acquire the business includes all sources: savings, other assets, and borrowed funds. Finance is expressed as ten categorical variables from less than $5,000 to $3 million or more, as well as “none needed” and “don’t know”.

Descriptive statistics for owner and firm characteristics are provided in Table 25. Almost 20 percent of owners of high-tech firms are immigrants, which is higher than the shares of immigrants (defined as born non-citizen) in the general population, at about 13.0 percent, in the adult population, about 15.7 percent, and in self-employment, about 17.9 percent, based on our calculations from the 2014 Current Population Survey March Supplement. The 20 percent of owners within high-tech is also higher than the 16 percent of immigrant owners in the full ASE sample that includes all industries, and higher than Hart and Acs’ (2010) estimate for their “high-impact” sample of high-tech firms, again 16 percent. But it is lower than reported by Saxenian (2002) for immigrant ownership of high-tech firms in Silicon Valley, at 24 percent, Wadhwa et al.’s (2007a and 2007b) estimate of 25 percent, and Kerr and Kerr’s (2017) of 24 percent. Each of these sources draws on different types of sample and definitions.

Table 25 shows the fraction of the owners in the sample having each characteristic and the fraction for immigrants and the native-born separately. We distinguish Hispanics, and among non-Hispanics, whites, Asian Indians, Chinese, Other Asians, and others. Among high-tech entrepreneurs, immigrants have a higher share than natives in the Hispanic, Asian Indian, Chinese, and Other Asian populations. The largest difference is for Asian Indians, who account for 36 percent of all immigrant owners, and only 1 percent of native owners.

Table 25 Descriptive Statistics: Demographic Characteristics

VARIABLES	All	Immigrant	Native
Immigrant	19.79	100.00	0.00
Race/ethnicity			
Hispanic	3.59	6.81	2.79
White (non-Hispanic)	80.55	33.58	92.14
Asian Indian (non-Hispanic)	7.93	36.46	0.89
Chinese (non-Hispanic)	2.72	10.38	0.83
Other Asian (non-Hispanic)	2.80	9.52	1.14
Other Minority (non-Hispanic)*	2.41	3.25	2.20
Education			
Less than Bachelor's Degree	23.71	9.55	27.21
Bachelor's Degree	43.55	37.20	45.11
Graduate Degree	32.74	53.24	27.68
Observations	11,000	2,000	9,000

Note: These are percentages of owners by characteristics from the ASE high-tech sample. Non-Hispanic African Americans are included with Other Minority (non-Hispanic) because the number of immigrants in this category is too small to disclose.

Table 25 also shows differences in educational attainment. Immigrants are less likely to have only a Bachelor’s degree, and they are much less likely to have less than a Bachelor’s degree: only about one-third as likely as natives. But more than half of immigrant owners hold an advanced degree, and they are much more likely than natives – nearly twice as likely – to have graduate education.

Do immigrants differ from natives in their reported motivations for entrepreneurship? Table 26 contains the means of the motivation variables on a 0-1-2 scale, as discussed above, for the full sample and for immigrants and natives separately. Immigrant owners report a higher propensity to cite inability to find a job as their motivation (although this motivation is uncommon for both groups in this high-tech sample), and a higher share of them say they have always wanted to own the business as a lifelong dream compared to natives. More relevant to innovation, immigrants have a slightly higher propensity to own the business because it is "the best avenue for their ideas, goods, or services." Overall, however, the differences in patterns of motivation appear slight.

Table 26 Descriptive Statistics: Motivations for Owning the Business

VARIABLES	All	Immigrant	Native
Idea	1.49	1.51	1.48
Income	1.49	1.46	1.50
No Job	0.10	0.14	0.09
Own Boss	1.47	1.35	1.50
Work for Self	0.90	0.79	0.92
Always Wanted to Own Business	1.18	1.32	1.14
Role Model	0.62	0.63	0.62
Flexible Hours	1.26	1.21	1.27
Balance Work/Family	1.28	1.28	1.28
Observations	11,000	2,000	9,000

Note: These are means of motivation variables measured on a scale where 0 is not important, 1 is somewhat important, and 2 is very important.

Concerning the amount of start-up capital, Table 27 shows that the immigrant-native differences exhibit a J-shaped relationship such that immigrants are slightly more likely to be in the lowest category of start-up capital and substantially more likely to be in the highest categories. Immigrants are 43 percent more likely than natives to have finance in the range $1-3mln, and for more than $3mln they are 60 percent more likely.

Table 27 Descriptive Statistics: Start-up Capital and Firm Age

VARIABLES	All	Immigrant	Native
Finance			
No capital needed	10.73	9.27	11.09
Capital under 5k	26.35	31.05	25.19
5k to 10k	11.54	12.80	11.22
10k to 25k	14.06	14.98	13.83
25k to 50k	7.77	7.70	7.79
50k to 100k	6.75	5.73	7.00
100k to 250k	5.80	5.14	5.96
250k to 1m	3.50	3.85	3.42
1m to 3m	1.17	1.54	1.08
3m and more	0.50	0.72	0.45
Don't know start-up capital	11.84	7.21	12.98
Firm age			
Young (age<=5)	39.66	50.50	36.99
Old (age>5)	60.34	49.50	63.01
Observations	11,000	2,000	9,000

Note: These are percentages of owners by characteristics from the ASE high-tech sample.

We also consider firm age as a possible correlate of innovation behavior. Table 27 shows that immigrants typically own younger firms (here defined as five years or less since first hiring) than do natives. Just over half of the immigrant-owned high-tech firms started up within the previous five years, while 63 percent of the native-owned firms are older than 5 years.

Table 28 High-Tech Industries: Definition and Composition

High-Tech Industry	Share of Sample	Share of Immigrants	Share of Natives
Oil & Gas Extraction (2111)	2.29	D	D
Pharmaceutical & Medicine Manufacturing (3254)	0.54	17.63	82.37
Computer & Peripheral Equipment Manufacturing (3341)	0.39	D	D
Communications Equipment Manufacturing (3342)	0.44	D	D
Semiconductor and Other Electronic Component Manufacturing (3344)	1.01	18.41	81.59
Navigational, Measuring, Electromedical, & Control Instruments Manufacturing (3345)	1.38	16.94	83.06
Aerospace Product & Parts Manufacturing (3364)	0.32	D	D
Software Publishers (5112)	1.44	23.25	76.75
Wired Telecommunications Carriers (5171)	0.71	21.00	79.00
Other Telecommunications (5179)	0.94	D	D
Data Processing, Hosting, & Related Services (5182)	2.46	17.67	82.33
Other Information Services (5191)	2.17	17.27	82.73
Architectural, Engineering, & Related Services (5413)	39.07	12.19	87.81
Computer Systems Design & Related Services (5415)	43.67	28.55	71.45
Scientific Research & Development Services (5417)	3.18	23.18	76.82

Notes: "D" means suppressed to ensure that no confidential information is disclosed.

Nearly three-quarters of the firms in this high-tech sample are in two four-digit NAICS industries: Architectural, Engineering, and Related Services (5413), and

Computer Systems Design and Related Services (5415). As shown in Table 28, immigrant-owned firms are disproportionately located in the latter and under-represented in the former. No other industry accounts for as much as 3 percent of the sample, and the immigrant-native differences in all these other industries are small and statistically insignificant.[28]

Our outcome variables include detailed innovation, research and development (R&D), and intellectual property measures. The ASE asks whether the business conducted twelve different product or process innovation activities in the last three years (2012-2014). We create a binary variable for innovation to indicate whether a firm conducted any product or process innovation in the last three years. We also calculate an innovation count by summing the number of product and process innovation activities. We make binary indicator variables for each type of product and process innovation activities. Product innovations include 1) sold a new good or service that no other business has ever offered before; 2) sold a new good or service that this business has never offered before; 3) improved a good or service's performance by making changes in materials, equipment, software, or other components; 4) developed a new use for a good or service; 5) added a new feature to a good or service; and 6) made it easier for customers to use a good or service. Process innovations include 1) applied a new way of purchasing, accounting, computing, maintenance, inventory control, or other support activity; 2) reduced costs by changing the way a good or service was distributed; 3)

[28] While there are 15 4-digit high-tech industries, some sectors have too few observations for the results to be disclosed.

upgraded a technique, equipment, or software to significantly improve a good or service; 4) made a significant improvement in a technique or process by increasing automation, decreasing energy consumption, or using better software; 5) decreased production costs by improving the materials, software, or other components; and 6) changed a delivery method to be faster or more reliable.

Table 29 Descriptive Statistics: Innovation Measures

VARIABLES	All	Immigrant	Native
Innovation Activities			
Innovation dummy	69.39	72.01	68.74
Innovation count	3.58	3.89	3.50
Production Innovation dummy	56.90	60.55	56.00
Process Innovation dummy	60.30	61.61	59.98
R&D Activities			
R&D activity (any type)	23.11	28.02	21.90
Work toward patent	13.40	16.98	12.52
Developed prototypes	13.29	17.18	12.34
Applied scientific/technic. knowled.	11.16	15.26	10.14
Produced publishable findings	9.68	12.55	8.97
Created generalizable research	11.34	15.73	10.26
Work to discover scientific facts	6.02	9.27	5.22
Work to extend understanding of scientific facts	10.51	14.37	9.56
Intellectual Property			
Copyright or Trademark	20.03	16.79	20.83
Patent granted or pending	6.60	8.50	6.13
Observations	11,000	2,000	9,000

Note: These are percentages of owners by innovation measures (except for innovation count) from the ASE high-tech sample.

Table 29 shows means of these innovation activities. About 69 percent of firms report they conducted at least one innovation, and the average number of innovation types is 3.6 in our high-tech sample. Although not shown in the table, the most common product innovation is improving a good or service's performance (44.3 percent) and making it easier for customers to use good or service (41.7 percent), and upgrading a technique, equipment, or software to significantly improve a good or service (50.9 percent) is the most frequent process improvement.

The ASE asks business owners whether their business carried out seven different R&D activities in 2014. We create an indicator for whether the business conducted any of these types of R&D. We also construct binary variables for each of the activities separately. We classify the following activities as "Applied R&D:" 1) conducted work that might lead to a patent; 2) developed and tested prototypes that were derived from scientific research or technical findings; and 3) applied scientific or technical knowledge in a way that has never been done before. We classify "Basic R&D" as activities that 1) produced findings that could be published in academic journals or presented at scientific conferences; 2) created new scientific research or technical solutions that can be generalized to other situations; 3) conducted work to discover previously unknown scientific facts, structures, or relationships; and 4) conducted work to extend the understanding of scientific facts, relationships or principles in a way that could be useful to others. In Table 29, 23.1 percent of firms conducted at least one of these R&D activities in 2014, and the most frequent R&D activity is work that might lead to a patent.

In general, the average rate of conducting R&D activities is lower than the innovation activities above.[29]

The last set of outcome variables concerns intellectual property. The ASE asks whether the business owns one or more of each of the following in 2014: copyright, trademark, patent (granted), and patent (pending). We use a dummy variable for owning either a copyright or trademark, and another for ownership of a patent granted or pending. Looking at Table 29, about 20 percent of firms within the high-tech sector own a copyright or trademark, while less than 5 percent of firms own patents either pending or granted.

A striking result from Table 29 is the consistently stronger innovation performance of immigrant- compared to native-owned firms. Immigrants are more likely to carry out 13 of the 14 measures of innovation. The exceptions are copyrights and trademarks, where native-owned firms have the advantage. Examining the statistical significance of these differences and how they change when other variables are taken into account are the subjects of the next sections.

3.3 Methods

We use the sample of owners and firms to estimate a series of regression models for each firm-level innovation outcome conditional on the owner's immigrant status. To take into account firms with multiple owners, we weight firm-owner observations by

[29] The lower R&D propensity could be partly due to the fact that the R&D questions are about activity in just one year, while the innovation activities are over three years.

ownership shares. Given that the ASE is a random sample of employer businesses drawn from the BR, this implies our results are representative of the firm population. We use a linear probability model for binary innovation outcomes and a Poisson regression model for innovation count. Standard errors are clustered at the firm level. Our base specification is:

Equation 15

$$Y_{ij} = \beta M_{ij} + f(Age_j) + u_{ij},$$

where M_{ij} is an immigrant owner indicator for owner i of firm j. The dependent variables are each type of product innovation, process innovation, R&D activity, and intellectual property. Since businesses are of different ages, and innovation may be correlated with firm age, in every specification (including the base) we control for a quadratic function of firm age, $f(Age_j)$. The coefficient on the immigrant owner indicator (β) captures the differences in innovation outcomes, essentially the raw gaps controlling only for firm age, between immigrant and native owners.

The purpose here is simply to describe differences in innovation behavior between immigrant and native owners. Just as in an analysis of gender differences in wages, for example, there is no issue of causality: we do not interpret the results as the impact of turning a random native into an immigrant (just as the interpretation placed on a female coefficient is not the impact of changing a male into a female). But it is also of interest to know whether there are observable differences that might account for the raw gap estimated by equation 15. For this purpose, we estimate another specification with owner characteristic controls as:

Equation 16

$$Y_{ij} = \beta M_{ij} + f(Age_j) + X_{ij}\gamma + u_{ij},$$

where $\boldsymbol{X}_{ij}$ is a vector of characteristics of owner i of firm j. The vector includes demographic variables (gender, age, and race/ethnicity), proxies for human capital (education, veteran, and prior business), and ownership team variables (size and family relationships). Arguably, these variables are pre-determined with respect to innovation behavior. The β estimated from equation 16 is a measure of the innovation gap between native and immigrant owners adjusted for personal characteristics.

In addition, immigrants may differ from natives in ways that are less clearly exogenous and indeed may be jointly determined with innovation: motivations, start-up capital, and industries as shown in the following specification:

Equation 17

$$Y_{ij} = \beta M_{ij} + f(Age_j) + X_{ij}\gamma + Q_{ij}\alpha_Q + K_j\alpha_K + S_j\alpha_S + u_{ij},$$

where Q_{ij} is the set of motivation variables, K_j is the set of vectors of the amount of start-up finance categories, and S_j is the set of vectors of 4-digit NAICS industry dummies. Most small business owners start their businesses due to non-pecuniary motives with no intention to grow or innovate (Hurst and Pugsley 2011). Given the selection process to come to the U.S., immigrant owners may have different motivations to own their businesses, which may influence their innovation outcomes. The importance of access to finance for business start-ups is well documented in the literature (e.g. Evans and Jovanovic 1989; Evans and Leighton 1989), and immigrant-owned businesses also tend to have higher start-up capital amounts than those owned by natives (Fairlie 2012). Higher start-up finance among immigrant owners may account for the differences in innovation outcomes between immigrant and native owners. Finally, immigrants may select into specific industries. Immigrants may be more or less likely to own businesses in industries with more innovation activities (e.g., certain parts of the high-tech sector), and this specification controls for this choice, comparing immigrants and natives within industries.

We also examine the heterogeneity of relative innovation performance of immigrant owners along three dimensions: education categories, race/ethnicity, and firm age. The literature on high-skilled immigrants (those with bachelor's degree or more) provides evidence that they are more likely to hold patents (e.g. Hunt and Gauthier-Loiselle 2010; Kerr and Lincoln 2010). However, the role of education in immigrant entrepreneurship has been less studied. We therefore examine heterogeneous innovation outcomes by owner education, distinguishing three groups: those with less than a bachelor's degree, those with a bachelor's degree, and those with advanced degrees.

Previous research has also examined immigrants by country of origin. Saxenian (2002) and Wadhwa et al. (2007) report higher shares of Indian and Chinese immigrants (Asian) in high-tech sectors, for example, showing an especially high share for Indians.

Although the ASE does not ask for country of origin, we use race/ethnicity to reflect the region of origin. We distinguish Hispanics, and among non-Hispanics, whites, Asian Indians, Chinese, Other Asians, and others.

Finally, we investigate whether the relative innovation performance of immigrant owners varies with the age of the firm. Although all specifications control for firm age, it is interesting to ask whether any immigrant advantage in innovation holds only during the early, entrepreneurial phase of a firm's development or also during more mature phases. For this purpose, we permit the immigrant owner coefficient to vary based on whether the firm is five or fewer years old or not.

The specification for heterogeneous immigrant contributions is:

Equation 18

$$Y_{ij} = Z_{ij}M_{ij}\delta + f(Age_j) + X_{ij}\gamma + \varepsilon_{ij},$$

where $Z_{ij}M_{ij}$ are the interaction terms between owner characteristics Z_{ij} (education categories, race/ethnicity, or firm age) and the immigrant indicator M_{ij} for owner i of firm j.

3.4 Results

Tables 30 and 31 display regression results for each measure of innovation using the three specifications described above: (1) base (no controls other than firm age), (2) adding demographic controls, and (3) adding motivations, finance, and industry controls.

The different types of product and process innovation activities, including the dummy for any activity and the count of the number of activities are in Table 30. Table 31 contains the different types of R&D as well as the intellectual property measures (copyright or trademark, and patent granted or pending).

Table 30 Product and Process Innovation by Immigrants

VARIABLES	Base	+Demographics	+ Motivations, Finance & Industry
Innovation Activities			
Innovation dummy	2.883	4.669	2.539
	(1.469)	(1.788)	(1.748)
Innovation count	0.090	0.146	0.081
	(0.031)	(0.036)	(0.036)
Product Innovation	3.488	6.438	3.055
	(1.588)	(1.921)	(1.870)
Process Innovation	1.632	4.606	2.887
	(1.582)	(1.964)	(1.950)
Observations	11,000	11,000	11,000

Note: Results from LPM estimation of Equation 15 at firm age 1. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions include firm age and age squared. The second column ("+Demographics") includes demographic variables (gender, age, and race/ethnicity), proxies for human capital (education, veteran, and prior business), and ownership team variables (size and family relationships). The last column includes motivations from Table 26, start-up finance from Table 27, and 4-digit NAICS industry dummies from Table 28. Standard errors clustered by firm are in parentheses.

The results show that immigrant-owned firms have higher propensities to conduct product and process innovation as well as R&D activity. The inclusion of demographic controls generally raises the immigrant association with innovation activities, suggesting that immigrant owners tend on average to have other characteristics that are negatively associated with product and process innovation. Demographic controls attenuate the immigrant associations with R&D activities, however.

Table 31 R&D, Copyright, Trademark, and Patents by Immigrants

VARIABLES	Base	+Demographics	+Motivation, Finance & Industry
R&D activity			
R&D activity (any type)	5.580	4.653	3.720
	(1.426)	(1.828)	(1.767)
Work toward patent	3.714	2.886	2.297
	(1.175)	(1.514)	(1.450)
Developed prototypes	4.729	3.885	3.169
	(1.180)	(1.565)	(1.492)
Applied scientific/technic.	4.528	3.698	3.358
knowledge	(1.114)	(1.453)	(1.407)
Produced publishable	3.342	1.667	1.877
finding	(1.019)	(1.334)	(1.267)
Created generalizable	4.772	4.102	3.654
research	(1.122)	(1.451)	(1.399)
Work to discover	3.749	2.754	3.009
scientific facts	(0.895)	(1.150)	(1.103)
Work to extend	4.574	3.062	3.346
understanding scientific facts	(1.084)	(1.405)	(1.341)
Intellectual Property			
Copyright or Trademark	-3.343	-0.150	-2.201
	(1.199)	(1.592)	(1.555)
Patent granted or pending	2.362	0.035	-0.330
	(0.858)	(1.051)	(1.009)
Observations	11,000	11,000	11,000

Note: Results from LPM estimation of Equation 15 at firm age 1. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions include firm age and age squared. The second column ("+Demographics") includes demographic variables (gender, age, and race/ethnicity), proxies for human capital (education, veteran, and prior business), and ownership team variables (size and family relationships). The last column includes motivations from Table 26, start-up finance from Table 27, and 4-digit NAICS industry dummies from Table 28. Standard errors clustered by firm are in parentheses.

Differing motivations, levels of start-up capital, and/or choices of industry explain much of the immigrant association with innovation activities, but not R&D activities, as evidenced by the significant attenuation of the immigrant coefficients when including those controls in the innovation activity regressions and more modest attenuation or even intensification when adding them to the R&D regressions.[30]

The immigrant effect is positive across all R&D activities, though after adding controls it becomes insignificant for producing publishable findings. Immigrant ownership is generally not associated with owning intellectual property, and the association is actually negative and significant in two of the three trademark specifications. The only positive and significant association is with patent pending in the specification without controls.

To investigate whether the immigrant advantage varies with firm age, we permit the immigrant indicator to vary with firm age in two categories: up to 5 years old and more than 5 years old. Regression estimates are shown in Table 32. The propensity to engage in innovation activities is similar for both young and older firms owned by immigrants. The point estimates are higher for immigrant-owned older firms for R&D activity and ownership of intellectual property. Among native-owned firms, the propensity to conduct R&D activities is higher for young firms, but for innovation activities a positive young firm effect disappears once adding controls, and differences are insignificant for intellectual property ownership. Both immigrant-owned firm age categories exhibit higher propensities to engage in innovation and R&D than either

[30] In results not shown here, the effect varies considerably across innovation measures. It is especially strong for developing a new use for a good or service. Immigrants have a higher propensity to develop goods or services that no other firm offers, but not goods or services that are new only to this firm. The former is a more radical form of innovation. Among process innovations, the immigrant association is insignificant for new way to support activity and upgrading a technique/equipment/software, while it is quite strong for increased automation/used better software.

native-owned firm age category across most specifications, while differences are generally insignificant for intellectual property ownership. These results suggest the immigrant advantage is maintained or even increases with firm age.

To investigate whether the immigrant advantage varies with firm age, we permit the immigrant indicator to vary with firm age in two categories: up to 5 years old and more than 5 years old. Regression estimates are shown in Table 32. The propensity to engage in innovation activities is similar for both young and older firms owned by immigrants. The point estimates are higher for immigrant-owned older firms for R&D activity and ownership of intellectual property. Among native-owned firms, the propensity to conduct R&D activities is higher for young firms, but for innovation activities a positive young firm effect disappears once adding controls, and differences are insignificant for intellectual property ownership. Both immigrant-owned firm age categories exhibit higher propensities to engage in innovation and R&D than either native-owned firm age category across most specifications, while differences are generally insignificant for intellectual property ownership. These results suggest the immigrant advantage is maintained or even increases with firm age.

Table 32 Innovation by Immigrants – Firm Age Heterogeneity

VARIABLES	Base	+ Demographics	+ Motivations, Finance & Industry
Innovation dummy			
Old*Immigrant	4.086	4.630	2.700
	(2.022)	(2.230)	(2.147)
Young*Native	3.982	1.099	1.298
	(1.308)	(1.379)	(1.342)
Young*Immigrant	5.385	4.518	3.530
	(2.032)	(2.393)	(2.341)
Innovation count			
Old*Immigrant	0.127	0.157	0.084
	(0.042)	(0.046)	(0.044)
Young*Native	0.105	0.004	-0.004
	(0.028)	(0.029)	(0.027)
Young*Immigrant	0.164	0.122	0.069
	(0.043)	(0.050)	(0.048)
R&D activity (any type)			
Old*Immigrant	8.592	7.051	5.337
	(1.968)	(2.211)	(2.117)
Young*Native	3.665	3.021	1.924
	(1.195)	(1.250)	(1.207)
Young*Immigrant	6.383	4.899	3.862
	(1.947)	(2.362)	(2.274)
Copyright or Trademark			
Old*Immigrant	-2.572	-0.127	-2.165
	(1.719)	(1.965)	(1.938)
Young*Native	-2.186	-3.296	-3.780
	(1.142)	(1.187)	(1.147)
Young*Immigrant	-7.064	-4.577	-6.434
	(1.581)	(1.985)	(1.900)
Patents (granted or pending)			
Old*Immigrant	3.756	1.582	0.490
	(1.252)	(1.302)	(1.246)
Young*Native	0.464	0.848	0.284
	(0.684)	(0.709)	(0.675)
Young*Immigrant	1.336	-0.069	-0.951
	(1.099)	(1.331)	(1.235)
Observations	11,000	11,000	11,000

Note: Results from LPM estimation of Equation 15 at firm age 1. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions include firm age and age squared. The second column ("+ Demographics") includes demographic variables (gender, age, and race/ethnicity), proxies for human capital (education, veteran, and prior business), and ownership team variables (size and family relationships).

Table 33 Innovation by Immigrants - Education Heterogeneity

VARIABLES	Base	+ Demographics	+ Motivations, Finance & Industry
Innovation dummy			
Below BA*Immigrant	6.834	6.655	3.921
	(4.056)	(4.003)	(3.809)
BA*Native	3.273	1.887	0.878
	(1.487)	(1.493)	(1.439)
BA*Immigrant	4.472	6.105	3.225
	(2.456)	(2.664)	(2.556)
Graduate*Native	4.681	4.089	2.181
	(1.664)	(1.684)	(1.639)
Graduate*Immigrant	6.288	7.503	4.405
	(2.155)	(2.468)	(2.428)
Innovation count			
Below BA*Immigrant	0.180	0.186	0.119
	(0.081)	(0.081)	(0.077)
BA*Native	0.119	0.087	0.068
	(0.033)	(0.033)	(0.030)
BA*Immigrant	0.081	0.150	0.070
	(0.053)	(0.058)	(0.056)
Graduate*Native	0.117	0.114	0.087
	(0.036)	(0.036)	(0.034)
Graduate*Immigrant	0.241	0.302	0.218
	(0.045)	(0.051)	(0.049)
Observations	11,000	11,000	11,000

Note: Results from LPM estimation of Equation 15 at firm age 1. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions include firm age and age squared. The second column ("+ Demographics") includes demographic variables (gender, age, and race/ethnicity), proxies for human capital (education, veteran, and prior business), and ownership team variables (size and family relationships). The last column includes motivations from Table 26, start-up finance from Table 27, and 4-digit NAICS industry dummies from Table 28. Standard errors clustered by firm are in parentheses.

Table 34 R&D, Copyright, Trademark, and Patents by Immigrants - Education Heterogeneity

VARIABLES	Base	+ Demographics	+ Motivations, Finance & Industry
R&D activity (any type)			
Below BA*Immigrant	6.141	5.916	5.498
	(3.662)	(3.762)	(3.747)
BA*Native	4.738	4.037	2.988
	(1.164)	(1.178)	(1.145)
BA*Immigrant	3.350	6.395	4.245
	(2.011)	(2.376)	(2.261)
Graduate*Native	16.89	15.62	11.94
	(1.473)	(1.485)	(1.429)
Graduate*Immigrant	19.86	21.76	17.27
	(2.099)	(2.432)	(2.328)
Copyright or Trademark			
Below BA*Immigrant	2.493	3.522	1.435
	(3.337)	(3.398)	(3.241)
BA*Native	5.151	4.248	3.049
	(1.190)	(1.192)	(1.142)
BA*Immigrant	-2.528	2.411	-0.421
	(1.725)	(2.043)	(1.998)
Graduate*Native	10.46	9.201	7.876
	(1.423)	(1.425)	(1.375)
Graduate*Immigrant	4.867	8.811	5.640
	(1.803)	(2.151)	(2.091)
Observations	11,000	11,000	11,000

Note: Results from LPM estimation of Equation 15 at firm age 1. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions include firm age and age squared. The second column ("+ Demographics") includes demographic variables (gender, age, and race/ethnicity), proxies for human capital (education, veteran, and prior business), and ownership team variables (size and family relationships). The last column includes motivations from Table 26, start-up finance from Table 27, and 4-digit NAICS industry dummies from Table 28. Standard errors clustered by firm are in parentheses.

Table 35 Innovation by Immigrants - Race Heterogeneity

VARIABLES	Base	+ Demographics	+ Motivations, Finance & Industry
Innovation dummy			
Hispanic*Immigrant	-1.417	-1.955	-1.735
	(5.235)	(5.119)	(4.932)
White*Immigrant	6.816	6.111	4.101
	(2.135)	(2.162)	(2.088)
Asian Indian*Immigrant	0.474	-0.320	-4.872
	(2.394)	(2.478)	(2.488)
Chinese*Immigrant	5.087	4.170	0.628
	(4.126)	(4.053)	(4.047)
Other Asian*Immigrant	-0.819	-1.831	-3.927
	(4.428)	(4.437)	(4.222)
Other Minority	1.046	0.249	-4.786
*Immigrant	(7.565)	(7.353)	(7.141)
Innovation count			
Hispanic*Immigrant	0.031	0.017	0.014
	(0.113)	(0.106)	(0.100)
White*Immigrant	0.217	0.204	0.132
	(0.042)	(0.042)	(0.041)
Asian Indian*Immigrant	0.008	-0.025	-0.172
	(0.052)	(0.053)	(0.051)
Chinese*Immigrant	0.143	0.133	0.024
	(0.088)	(0.084)	(0.081)
Other Asian*Immigrant	-0.079	-0.108	-0.156
	(0.092)	(0.090)	(0.089)
Other Minority	-0.056	-0.072	-0.238
*Immigrant	(0.174)	(0.165)	(0.170)
Observations	11,000	11,000	11,000

Note: Results from LPM estimation of Equation 15 at firm age 1. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions include firm age and age squared. The second column ("+ Demographics") includes demographic variables (gender, age, and race/ethnicity), proxies for human capital (education, veteran, and prior business), and ownership team variables (size and family relationships). The last column includes motivations from Table 26, start-up finance from Table 27, and 4-digit NAICS industry dummies from Table 28. Standard errors clustered by firm are in parentheses.

Table 36 R&D, Copyright, Trademark, and Patents by Immigrants - Race Heterogeneity

VARIABLES	Base	+ Demographics	+ Motivations, Finance & Industry
R&D activity (any type)			
Hispanic*Immigrant	1.493	-0.645	1.376
	(4.612)	(4.575)	(4.290)
White*Immigrant	12.99	9.360	8.009
	(2.329)	(2.291)	(2.184)
Asian Indian*Immigrant	-0.457	-4.143	-5.077
	(2.175)	(2.185)	(2.160)
Chinese*Immigrant	13.54	7.047	3.559
	(4.441)	(4.287)	(3.892)
Other Asian*Immigrant	-2.030	-3.537	-5.770
	(3.636)	(3.521)	(3.112)
Other Minority	-0.107	-3.195	-6.906
*Immigrant	(6.737)	(6.292)	(6.233)
Copyright or Trademark			
Hispanic*Immigrant	-4.912	-6.061	-5.087
	(3.841)	(3.797)	(3.843)
White*Immigrant	4.280	2.592	-0.095
	(2.097)	(2.064)	(1.975)
Asian Indian*Immigrant	-9.049	-10.02	-11.94
	(1.625)	(1.685)	(1.784)
Chinese*Immigrant	0.486	-2.116	-4.284
	(3.749)	(3.672)	(3.578)
Other Asian*Immigrant	-8.466	-9.400	-9.797
	(2.863)	(2.766)	(2.636)
Other Minority *Immigrant	D	D	D
Observations	11,000	11,000	11,000

Note: Results from LPM estimation of Equation 15 at firm age 1. Coefficients and standard errors are multiplied by 100 for ease of reading. All regressions include firm age and age squared. The second column ("+ Demographics") includes demographic variables (gender, age, and race/ethnicity), proxies for human capital (education, veteran, and prior business), and ownership team variables (size and family relationships). The last column includes motivations from Table 26, start-up finance from Table 27, and 4-digit NAICS industry dummies from Table 28. Standard errors clustered by firm are in parentheses. "D" means suppressed to ensure that no confidential information is disclosed.

Regarding variation in the immigrant effect with educational attainment, we specify the equation so that the reference category is natives with less than a bachelor's degree. As shown in Table 33, the propensity to carry out any product or process innovation activity is increasing in education for native-owned firms, but not immigrant-owned firm. For innovation count, there is a higher association with innovation for native-owned firms where the owner has at least a bachelor's degree, but there is little difference between bachelor's and advanced degrees. The coefficients exhibit a U-shape with educational attainment for immigrant-owned businesses. Firms with advanced degree-immigrants have the highest innovation count propensities and those with less than bachelor's degree natives have the lowest.

Having a graduate degree is strongly associated with R&D activity for both native- and immigrant owned-firms, and the immigrant effects within the graduate degree category are larger. For copyrights and patents, it is firms with native owners with graduate degrees that distinguish themselves. Across all innovation measures, the immigrant advantage is generally largest for owners with less than a bachelor's degree.

Finally, we use race and ethnicity to examine differences in the immigrant innovation advantage across region of origin. Results with white natives as the reference group are shown in Table 35 and 36. Sample sizes get thin, so results are less precisely estimated. One striking result is that firms owned by Asian Indians, despite their high prevalence in the sample, tend to produce less of all types of innovation when full controls are included.

3.5 Conclusion

Much of the research on immigration assumes that natives and immigrants are similar factors of production, in various cases conditional on geographical region, education, and experience. An influx of immigrants is analyzed as a labor supply shock to the region or the skill group. Another large and long-standing body of research focuses on the difficulties immigrants face in adjusting to their new environments, measuring rates of "assimilation," usually defined as degree of convergence to otherwise similar native workers.

A much smaller literature takes a different approach, treating immigrants as potentially advantaged rather than either similar or disadvantaged relative to natives. Much of this research has focused on individual immigrants in science, the STEM workforce, and entrepreneurship. With some variation, the results suggest disproportionate contributions to some measures of innovation, with immigrants more likely to hold patents, work in STEM, achieve high citation indices, and receive Nobel Prizes. (Hunt 2011; Kerr 2013; Kahn and MacGarvie 2016). One interpretation of these results is that immigrants self-select from the right tail of the ability distribution and perhaps that the distribution has a fatter right tail than that of natives (see for instance the discussion in Kahn et al. 2017).

Our premise is similar to this literature, asking whether immigrants tend to be more innovative than natives. But our focus is on firms founded and operated by immigrants in comparison to those owned by natives. There has been a lot of "hype" about immigrant entrepreneurs in the U.S. high-tech sector, but relatively little evidence on the extent to which they contribute disproportionately to innovation. This paper provides such evidence drawing upon a large representative sample of high-tech businesses and using detailed information on owner characteristics, motivations, and start-up capital, as well as an extensive set of innovation measures. We focus on the high-tech sector because of its prominence in U.S. growth.

The results suggest higher innovation activities by immigrants for nearly all the innovation measures we are able to analyze. The measures range from detailed product and process innovation, to several forms of R&D, to intellectual property rights associated with innovation, including patents. The only measures where immigrants have notably lower performance compared with natives is for copyrights and trademarks.

Immigrant entrepreneurs tend to be much better educated than their native counterparts in the high-tech sector, on average, but the immigrant advantage persists when we control for education and other owner characteristics, and we find an immigrant advantage at all levels of education, again with the exception of copyright or trademark. Immigrant entrepreneurs also tend to operate younger firms, and while we find firm age is negatively correlated with innovation, again the immigrant advantage exists when we control for firm age (as we do in all specifications). Moreover, we find an immigrant advantage in innovation for both younger and older firms.

Future research could expand on these findings by broadening both the population under consideration and the set of outcome variables to be analyzed. A sample including other industries could shed light on the relative innovativeness of immigrant entrepreneurs outside of the high-tech sector. Rather than confining attention to the nativity of individual owners, the analysis could be extended to the possible effects of combining immigrant and native human capital within entrepreneurial teams. Finally, the roles of immigrant entrepreneurs in job creation and productivity growth could be examined in a broader assessment of the contributions of immigrants to innovative entrepreneurship in the U.S. We hope to report our findings on these issues in the near future.

REFERENCES

Arestis, P., & Sawyer, M. C. (2010). *21st Century Keynesian Economics*. Houndsmills, Basingstoke, Hampshire: Palgrave Macmillan.

Agarwal, Sumit, Efraim Benmelech, Nittai Bergman, and Amit Seru. 2012. “Did the Community Reinvestment Act (CRA) Lead to Risky Lending?” National Bureau of Economic Research. No. 18609. http://www.nber.org/papers/w18609.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. Mostly Harmless Econometrics: An Empiricist’s Companion. Princeton: Princeton University Press.

Apgar, William, and Mark Duda. 2002. “The 25th Anniversary of the Community Reinvestment Act: Access to Capital in an Evolving Financial Services System, Prepared for the Ford Foundation.” Joint Center for Housing Studies, Harvard University.

Asiedu, Elizabeth, James A. Freeman, and Akwasi Nti-Addae. 2012. “Access to Credit by Small Businesses: How Relevant Are Race, Ethnicity, and Gender?” *American Economic Review.* 102 (3): 532–37.

Avery, Robert B., Raphael Bostic, and Katherine A. Samolyk. 1998. “The Role of Personal Wealth in Small Business Finance.” *Journal of Banking & Finance.* 22 (6–8): 1019–61.

Avery, Robert B., and Kenneth P. Brevoort. 2015. “The Subprime Crisis: Is Government Housing Policy to Blame?” *The Review of Economics and Statistics.* 97 (2): 352–63.

Avery, Robert B., Paul S. Calem, and Glenn B. Canner. 2003. “The Effects of the Community Reinvestment Act on Local Communities.” *Proceedings of " Seeds of Growth: Sustainable Community Development: What Works, What Doesn't, and Why,"* Conference sponsored by the Federal Reserve System.

Bach, Laurent. 2013. “Are Small Businesses Worthy of Financial Aid? Evidence from a French Targeted Credit Program.” *Review of Finance.* 18(3):877-919.

Bates, Timothy. 1989. "The Changing Nature of Minority Business: A Comparative Analysis of Asian, Nonminority, and Black-Owned Businesses." *The Review of Black Political Economy*, 18(2): 25-42.

Bates, Timothy, 1990, "Entrepreneur Human Capital Inputs and Small Business Longevity," *Review of Economics and Statistics*, 72(4), 551-559.

Bates, Timothy and Alicia M. Robb. 2013. "Greater Access to Capital Is Needed to Unleash the Local Economic Development Potential of Minority-Owned Businesses." *Economic Development Quarterly*, 27(3):250-259.

Bates, Timothy, and Alicia Robb. 2015. "Has the Community Reinvestment Act Increased Loan Availability among Small Businesses Operating in Minority Neighborhoods?" *Urban Studies.* 52(9):1702-1721.

Becker, Gary. 1957. The Economics of Discrimination. Studies in Economics of the Economics Research Center of the University of Chicago. Chicago: University of Chicago Press.

Belsky, Eric S., Matthew Lambert, and Alexander von Hoffman. 2000. "Insights into the Practice of Community Reinvestment Lending." Joint Center for Housing Studies of Harvard University.

Berry, Christopher R., and Sarah L. Lee. 2007. "The Community Reinvestment Act: A Regression Discontinuity Analysis." Harris School of Public Policy, University of Chicago.

Bhutta, Neil. 2011. "The Community Reinvestment Act and Mortgage Lending to Lower Income Borrowers and Neighborhoods." *Journal of Law and Economics.* 54(4): 953–83.

Blanchard, Lloyd, Bo Zhao, and John Yinger. 2008. "Do Lenders Discriminate against Minority and Woman Entrepreneurs?" *Journal of Urban Economics.* 63(2): 467–97.

Blanchflower, David G., Phillip B. Levine, and David J. Zimmerman. 2003. "Discrimination in the Small-Business Credit Market." *The Review of Economics and Statistics.* 85(4): 930–43.

Borjas, George J. 1985. "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants." *Journal of Labor Economics* 3 (4): 463-489.

Borjas, George J. 2015. "The Slowdown in the Economic Assimilation of Immigrants: Aging and Cohort Effects Revisited Again." *Journal of Human Capital* 9 (4): 483-517.

Borjas, George J., and Kirk B. Doran. 2015. "Cognitive Mobility: Labor Market Responses to Supply Shocks in the Space of Ideas." *Journal of Labor Economics* 33 (1): 109-145.

Borjas, George J., and Joan Monras. 2017. "The Labour Market Consequences of Refugee Supply Shocks." *Economic Policy* 32 (91): 361-413.

Bostic, Raphael W., and Hyojung Lee. 2017. “Small Business Lending Under the Community Reinvestment Act.” *Journal of Policy Development and Research.* 19(2): 63–84.

Brown, J. David, and John S. Earle. 2017. “Finance and Growth at the Firm Level: Evidence from SBA Loans.” The Journal of Finance 72 (3): 1039–80.

Brown, J. David, John S. Earle, Mee Jung Kim, Kyung Min Lee. 2018. “High-Growth Entrepreneurship.” IZA Discussion Paper No. 11662.

Brown, J. David, John S. Earle, Mee Jung Kim, and Kyung Min Lee. 2019. "Start-ups, Job creation, and Founder Characteristics." *Industrial and Corporate Change* 28 (6): 1637-1672.

Cabral, Luís M. B., and José Mata, 2003, “On the Evolution of the Firm Size Distribution: Facts and Theory,” *American Economic Review* 93(4), 1075–90.

Cao, Dan and Hyatt, Henry R. and Mukoyama, Toshihiko and Sager, Erick. 2019. “Firm Growth Through New Establishments.” Available at SSRN: https://ssrn.com/abstract=3361451

Card, David. 1990. "The Impact of the Mariel Boatlift on the Miami Labor Market." *ILR Review* 43 (2): 245-257.

Card, David. 2001. "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration." *Journal of Labor Economics* 19 (1): 22-64.

Cattaneo, Matias D., Nicolás Idrobo, and Rocio Titiunik. 2018. “A Practical Introduction to Regression Discontinuity Designs: Volume I.” Cambridge Press: Quantitative and Computational Methods for Social Science.

Cavalluzzo, Ken S. and Linda C. Cavalluzzo. 1998. "Market Structure and Discrimination The Case of Small Business." *Journal of Money, Credit, and Banking*, 30(4): 771-792.

Chiswick, Barry R., Yew Liang Lee, and Paul W. Miller. 2005. "A Longitudinal Analysis of Immigrant Occupational Mobility: A Test of the Immigrant Assimilation Hypothesis." *International Migration Review* 39 (2): 332-353.

Chiswick, Carmel U. 2009. "The Economic Determinants of Ethnic Assimilation." *Journal of Population Economics* 22 (4): 859.

Coleman, Susan and Richard, Cohn. 2001. "Borrowing Behavior of Small Black-Owned Firms." *The Journal of Applied Management and Entrepreneurship*, 6(2): 68-81.

Coleman, Susan. 2002a. "The Borrowing Experience of Black and Hispanic-Owned Small Firms: Evidence from the 1998 Survey of Small Business Finances." *The Academy of Entrepreneurship Journal*, 8(1): 1-20.

Coleman, Susan. 2002b. "Borrowing Patterns for Small Firms: A Comparison by Race and Ethnicity." *Journal of Entrepreneurial Finance*, 7(3): 77-97.

Coleman, Susan. 2004. "Access to Debt Capital for Women- and Minority-Owned Small Firms: Does Educational Attainment Have an Impact?" *Journal of Developmental Entrepreneurship*, 9(2):127-143.

Coleman, Susan. 2005. "Is There a Liquidity Crisis for Small, Black-Owned Firms?" *Journal of Developmental Entrepreneurship*, 10(1):29-47.

Davis, Steven J, John C Haltiwanger, and Scott Schuh. 1996. *Job Creation and Destruction*. The MIT Press.

Ding, Lei, and Leonard I. Nakamura. 2017. "Don't Know What You Got Till It's Gone — the Effects of the Community Reinvestment Act (CRA) on Mortgage Lending in the Philadelphia Market," Federal Reserve Bank of Philadelphia, Working Paper. No. 17-15.

Ding, Lei and Lee, Hyojung and Bostic, Raphael. 2018. Effects of the Community Reinvestment Act (CRA) on Small Business Lending (2018-12-03). FRB of Philadelphia Working Paper. No. 18-27.

Evans, David S., and Boyan Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *Journal of Political Economy* 97(4): 808-827.

Evans, David S., and Linda S. Leighton. 1989. "Some Empirical Aspects of Entrepreneurship." *American Economic Review* 79(3): 519-535.

Fairlie, Robert W. and Alicia M. Robb. 2007. “Why are Black-Owned Businesses Less Successful then White-Owned Businesses: The Role of Families, Inheritances, and Business Human Capital.” *Journal of Labor Economics*, 289-323.

Fairlie, Robert W. 2012. "Immigrant Entrepreneurs and Small Business Owners, and Their Access to Financial Capital." *Small Business Administration*: 1-46.

Fairlie, Robert W., and Magnus Lofstrom. 2014. “Immigration and Entrepreneurship.” In *Handbook of the Economics of International Migration.* Vol. 1B. (edited by Barry Chiswick and Paul Miller). North Holland: 877-911.

Federal Financial Institutions Examination Council. 2017. “Census File for Use with HMDA/CRA 2015 Demographics.” Federal Financial Institutions Examination Council.
———. 2013. “A Guide to CRA Data Collection and Reporting.” FFIEC.

Federal Register. 2016. “Community Reinvestment Act: Interagency Questions and Answers Regarding Community Reinvestment.” Federal Register Vol. 81, No.142. Rules and Regulations. Federal Financial Institutions Examination Council.

Foster, Lucia, and Patrice Norman. 2016. “The Annual Survey of Entrepreneurs: An Introduction.” CES Discussion Paper 15-40R.

Gamberoni, Elisa, Katerina Gradeva, and Sebastian Weber. 2016. "Firm Responses to Employment Subsidies: a Regression Discontinuity Approach to the 2012 Spanish Labour Market Reform." European Central Bank, Working Paper. No.1970.

Ganguli, Ina, Shulamit Kahn, and Megan MacGarvie. 2018. "Introduction to The Role of Immigrants and Foreign Students in Science, Innovation, and Entrepreneurship." In *The Role of Immigrants and Foreign Students in Science, Innovation, and Entrepreneurship*. (edited by Ina Ganguli, Shulamit Kahn, Megan MacGarvie). University of Chicago Press. (forthcoming).

Goldschlag, Nathan, and Javier Miranda, 2016, “Business Dynamics Statistics of High Tech Industries.” CES Discussion Paper 16-55.

Griffith, Rachel, Elena Huergo, Jacques Mairesse, and Bettina Peters. 2006. "Innovation and Productivity across Four European Countries." *Oxford Review of Economic Policy* 22 (4): 483-498.

Hall, Bronwyn H. 2011. *Innovation and Productivity*. National Bureau of Economic Research. No. w17178.

Hart, David M., and Zoltan J. Acs. 2011. "High-tech Immigrant Entrepreneurship in the United States." *Economic Development Quarterly* 25 (2): 116-129.

Hunt, Jennifer. 2011. "Which Immigrants are most Innovative and Entrepreneurial? Distinctions by Entry Visa." *Journal of Labor Economics* 29 (3): 417-457.

Hunt, Jennifer, and Marjolaine Gauthier-Loiselle. 2010. "How Much Does Immigration Boost Innovation?" *American Economic Journal: Macroeconomics* 2(2): 31-56.

Hurst, Erik, and Benjamin W. Pugsley. 2011. "What Do Small Businesses Do?" *Brookings Papers on Economic Activity*: 73.

Immergluck, Daniel. 2004. *Credit to the Community: Community Reinvestment and Fair Lending Policy in the United States.* Cities and Contemporary Society.

Kahn, Shulamit, and Megan J. MacGarvie. 2016. "How Important is US Location for Research in Science?" *Review of Economics and Statistics* 98(2): 397-414.

Kahn, Shulamit, Giulia La Mattina, and Megan J. MacGarvie. 2017. "Misfits, Stars, and Immigrant Entrepreneurship." *Small Business Economics* 49(3): 533-57.

Kerr, William R. 2013. *High-skilled Immigration, Innovation, and Entrepreneurship: Empirical Approaches and Evidence.* National Bureau of Economic Research. No. w19377.

Kerr, William R., and Sari Pekkala Kerr. 2017. "Immigrant Entrepreneurs." In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*. (edited by John Haltiwanger, Erik Hurst, Javier Miranda, and Antoinette Schoar). Chicago: University of Chicago Press.

Kerr, Sari Pekkala, and William R. Kerr. 2018. "Immigrant Entrepreneurship in America: Evidence from the Survey of Business Owners 2007 & 2012." National Bureau of Economic Research. No. 24494.

Kerr, William R., and William F. Lincoln. 2010. "The Supply Side of Innovation: H-1B Visa Reforms and US Ethnic Invention." *Journal of Labor Economics* 28(3): 473-508.

Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*. 48 (2): 281–355.

Lelarge, Claire, David Sraer, and David Thesmar. 2010. “Entrepreneurship and Credit Constraints: Evidence from a French Loan Guarantee Program.” *International differences in entrepreneurship*, University of Chicago Press.

Mairesse, Jacques, and Pierre Mohnen. 2010. "Using Innovation Surveys for Econometric Analysis." In *Handbook of the Economics of Innovation*, Vol. 2, p. 1129-1155. North-Holland.

Oliveira, Blandina, and Adelino Fortunato. 2006. "Firm growth and liquidity constraints: A dynamic analysis." *Small Business Economics*. 27(2-3): 139-156.

Ottaviano, Gianmarco IP, and Giovanni Peri. 2012. "Rethinking the Effect of Immigration on Wages." *Journal of the European Economic Association* 10 (1): 152-197.

Parisi, Maria Laura, Fabio Schiantarelli, and Alessandro Sembenelli. 2006. "Productivity, Innovation and R&D: Micro Evidence for Italy." *European Economic Review* 50 (8): 2037-2061.

Peri, Giovanni. 2007. “Higher Education, Innovation and Growth.” In *Education and Training in Europe*. (edited by Giorgio Brunello, Pietro Garibaldi, and Etienne Wasmer). Oxford: Oxford University Press.

Peri, Giovanni. 2012. "The Effect of Immigration on Productivity: Evidence from US States." *Review of Economics and Statistics* (1): 348-358.

Peri, Giovanni, Kevin Shih, and Chad Sparber. 2015. "STEM Workers, H-1B Visas, and Productivity in US Cities." *Journal of Labor Economics* 33 (1): 225-255.

Ringo, Daniel R. 2017. "Mortgage Lending, Default and the Community Reinvestment Act." *Default and the Community Reinvestment Act.*

Robb, Alicia M and Arnobio Morelix. 2016. “Startup Financing Trends by Race: How Access to Capital Impacts Profitability.” Ewing Marion Kauffman Foundation

Robb, Alicia M, Robert W. Fairlie, and David T. Robinson. 2016. “Black and White: Access to Capital among Minority-Owned Startups.” Stanford Institute for Economic Policy Research Discussion Paper 17-03.

Robb, Alicia. 2018. “Financing Patterns and Credit Market Experiences: A Comparison by Race and Ethnicity for U.S. Employer Firms.” U.S. Small Business Administration, Office of Advocacy.

Rupasingha, Anil, and Kyungsoon Wang. 2017. "Access to capital and small business growth: evidence from CRA loans data." *The Annals of Regional Science.* 59(1): 15-41.

Saxenian, AnnaLee. 2002. "Silicon Valley's New Immigrant High-Growth Entrepreneurs." *Economic Development Quarterly* 16(1): 20-31.

Squires, Gregory D., and Sally O'Connor. 2001. *Color and Money: Politics and Prospects for Community Reinvestment in Urban America*. SUNY Press.

Stephan, Paula E., and Sharon G. Levin. 2001. "Exceptional Contributions to US Science by the Foreign-Born and Foreign-Educated." *Population Research and Policy Review* 20 (1-2): 59-79.

Wadhwa, Vivek, AnnaLee Saxenian, Ben A. Rissing, and Gary Gereffi. 2007a. "America's New Immigrant Entrepreneurs: Part I." Duke Science, Technology & Innovation Paper No. 23. Available at SSRN: https://ssrn.com/abstract=990152

Wadhwa, Vivek, Ben Rissing, AnnaLee Saxenian, and Gary Gereffi. 2007b. "Education, entrepreneurship and immigration: America's new immigrant entrepreneurs, Part II." Available at SSRN: https://ssrn.com/abstract=991327 or http://dx.doi.org/10.2139/ssrn.991327

Yinger, John. 1995. *Closed Doors, Opportunities Lost: The Continuing Costs of Housing Discrimination*. Russell Sage Foundation.

Zinman, Jonathan. 2002. "The Efficacy and Efficiency of Credit Market Interventions: Evidence from the Community Reinvestment Act." Joint Center for Housing Studies of Harvard University.

BIOGRAPHY

Mee Jung Kim graduated from Paul H. Nitze School of Advanced International Studies (SAIS), Johns Hopkins University with an Master of Arts in International Economics and Southeast Asia Studies in 2011. She received her Bachelor of Arts in Economics from University of Michigan in Ann Arbor in 2009. She worked at East Asia and Pacific Region, Financial and Private Sector Development at the World bank. She is currently employed as a Pathway Intern at Center for Economic Studies at the U.S. Census Bureau. Her research centers on applied microeconomics, focusing on the topic of finance, business growth, and entrepreneurship.