Essays on Banking and Housing

by

Jiayi Xu

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
May 1, 2021

Keywords: Banking; minority depository institutions; credit
markets; loan rates; housing price

Approved by

James R. Barth, Chair, Lowder Eminent Scholar, Department of Finance
Jitka Hilliard, Co-chair, Professor, Department of Finance
Justin Benefield, Associate Professor, Department of Finance
Zhonglin Qin, Assistant Professor, Department of Finance
Xuan Shen, Senior Quantitative Risk Analyst, Regions Bank

Abstract

In 1865, the first minority bank in the United States was established. Over time, banks owned or controlled by minorities have grown in number. Yet, one hundred and fifty years later, they still account for only 2.8 percent of all banks. The contribution of this paper is fourfold. First, it provides a comprehensive assessment of the role and importance of different categories of MDIs in the banking industry. Second, it contributes to an understanding of the extent to which the diversity of ownership/control affects the performance and riskiness of firms, but is among the few that does so in terms of minority ownership/control of banks. Third, the paper examines the recent performance of MDIs from the perspective of whether their disproportionately small role in the banking industry is due to their relatively poorer and riskier performance as compared to non-MDIs. Fourth, a check on the robustness of the results is provided, including for the first time using two different databases on MDIs. The results indicate that MDIs, in contrast to most previous studies, are not significantly less profitable or riskier than non-MDIs.

The second essay uses nationwide census tract-level data to examine the relationship between credit market competition change and loan rates. Results show that an increase in competition of the banking industry, measured by the presence of a new bank branch, in a census tract, causes a decrease in the average loan rates in the following six months. On the other hand, a decrease in competition, measured by the occurrence of a branch closure in a census tract, causes an increase in the average loan rates after three quarters. The effects are community bank openings and closures. In addition, evidence supports the herd behavior of new bank branches documented by previous literature.

Large plant openings have been shown to impact the region in which they are located in a variety of ways. Traditional agglomeration economic arguments posit a host of potential benefits that might accrue to the new plant and incumbent plants in the area. When firms decide to open plants in areas without the benefit of agglomeration economies, the reason is often assigned as the firm responding to municipal or state incentives. The question naturally arises as to whether those incentives, financed via local or state taxes, are being offered in a considered way. Housing impacts are often overlooked as a possible if the partial, justification of those industrial incentives. This study analyzes housing market impacts in a rural area from the opening of a large manufacturing plant. Results of the third essay indicate that a large plant opening in a rural county is associated with an increase in housing prices relative to prices observed in a comparable control county.

Acknowledgments

I would like to thank all of those individuals who selflessly gave your time and attention to support me in the successful completion of this dissertation project.

I am deeply indebted to my advisor and committee chair, Dr. James R. Barth – I could not have imagined a better person to guide me through graduate school and continuously present me with so many wonderful opportunities. I am honored to have such a distinguished advisor to mentor me and I sincerely appreciate your time and guidance on my works.

I am very grateful to my committee co-chair Dr. Jitka Hilliard, and committee Dr. Justin Benefield, Dr. Zhonglin Qin, and Dr. Xuan Shen. Also, It has been a great pleasure to collaborate with Dr. Ash Abebe. Thanks to all of them for the immense amount of support they have given me. Their guidance and suggestions have helped strengthen all of the work presented here. Thanks for investing their time in my education – both as a student and a researcher.

I greatly appreciate Dr. Jimmy Hilliard for establishing and chairing the Ph.D. program. I would also like to express my sincere gratitude to the whole Finance Department at Auburn University, for their continued support of my academic and professional development.

This long journey might not be possible if my fellow friends were not next to me. I appreciate all my friend relations here at Auburn and cherish all my memories in my life. The friendship made here also became part of my treasure in the United States, notably with Madeline Arnall, Rui Chen, Yawei Chen, David Ding, Song Gao, Charles Liang, Yaru Lin, Ting Shen, Kang Sun, Shuting Tian, Joe Wang, Xiaozhu Wang, Hao Wu, Jolly Zhang, Jie Zhang, Ling Zhang, and Xinyu Zhao. Also, I must say thank you to Jing Wu and Yachi Kuo for supporting and encouraging me throughout the whole Ph.D. process.

I owe immense gratitude to my parents and parents-in-law for their unconditional belief in me. Finally, thanks go to my lifelong partner (and husband), Liang Tang, for his love, support, and scarifies along this meaningful journey.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACF	Autocorrelation function
ACS	American Community Survey
CB	Community Bank
CD	Certificates of deposit
CRA	Community Reinvestment Act
EPA	Environmental Protection Agency
FDIC	Federal Deposit Insurance Corporation
FFIEC	Federal Financial Institutions Examination Council
FIRREA	Financial Institutions Reform, Recovery, and Enforcement Act
FRB	Federal Reserve Board
GLS	Generalized least squares
LMI	Low- or moderate-income
MBDP	Minority Bank Deposit Program
MDI	Minority depository institutions
MDP	Million Dollar Plants
MHP	Median housing price
MSA	Metropolitan statistical area
NCUA	National Credit Union Administration
OCC	Comptroller of the Currency
OECD	Organisation for Economic Co-operation and Development
PSM	Propensity Score Matching

REML	Random-Effects Meta-Analysis
REMI	Regional Economic Models, Inc.
ROA	Return on assets
USPS	United States Postal Service
ZCTA	ZIP Code Tabulation Area

Chapter 1

U.S. Minority Banks: Why So Few — After 150 Years?[1]

1.1. Introduction

On March 3, 1865, the first minority bank, the Freedmen's Savings Bank and Trust Company (an African American bank), was approved to operate by the US Congress and signed into law by President Lincoln.[2] Over time, banks owned or controlled by minorities, known as minority depository institutions (MDIs), have grown in number. MDIs are owned/controlled by Blacks or African Americans, Asians or Pacific Islander Americans, Hispanic Americans, and Native Americans or Alaskan Native Americans.[3] These four categories of MDIs account for 100 percent of all MDI offices and total assets, as of December 2018. Although their numbers have grown over time, more than one hundred and fifty years after the first one was established, they still account for only 2.8 percent of all banks.[4] This is a disproportionately small percentage when compared to the minority shares of the population (38 percent) and business firms (19 percent). In the case of African American banks, moreover, they account for only 23 of 5,797 banks or less than one-half of one percent of the total number.

Importantly, the Financial Institutions Reform, Recovery, and Enforcement Act (FIRREA) of 1989, as amended by the Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank Act) of 2010, requires the Secretary of the Treasury to consult with the Chairman of the Board of Governors of the Federal Reserve System (FRB), the Comptroller of the Currency

---

[1] The authors are grateful to Kenneth Kelly, Joseph M. Santos and Harold Black for helpful comments.

[2] See Baradaran (2017), who points out Freedmen's Savings Bank and Trust Company was the only bank to be chartered by the US Congress, except for the First Bank and Second Bank of the United States. Also, see Ammons (1996) for an informative discussion of the evolution of black-owned banks between the 1880s and 1990s.

[3] See the following website for more information on MDIs: https://www.fdic.gov/regulations/resources/minority/mdi.html, accessed June 18, 2019.

[4] The corresponding share of all bank assets is 1.3 percent, as of the fourth quarter of 2018.

(OCC), the Chairman of the National Credit Union Administration (NCUA), and the Chairperson of the Board of Directors of the Federal Deposit Insurance Corporation (FDIC) on methods for best achieving the following: (1) preserving the present number of MDIs, and (2) promoting and encouraging the creation of new MDIs.[5] Also, the federal banking agencies are required to file an annual report to the US Congress containing a description of actions taken to carry out the requirements of FIRREA.[6] Beyond these governmental activities, there have been academic studies examining the performance of MDIs as compared to Non-MDIs in the banking industry. Such studies provide information on the progress made with respect to the preservation and expansion of MDIs. However, most of the studies are now somewhat dated and therefore may not reflect more recent developments regarding the role and performance of MDIs.

The contribution of this paper is to provide a more current assessment of the comparative performance of MDIs to Non-MDIs within the community banking industry. Specifically, we make four contributions to the existing literature. First, we provide a comprehensive and current assessment of the role of different categories of MDIs over time in the banking industry. Second, we contribute to understanding the extent to which the diversity in the ownership/control of firms affects performance and riskiness[7], but are among the few that do so in terms of actual minority ownership/control of banks, and for different minority groups. Third, we examine the performance

---

[5] See https://www.federalreserve.gov/publications/files/preserving-minority-depository-institutions-2018.pdf, accessed July 26, 2019. In this regard, Ricks and Consumer Financial Protection Bureau (2018, p. 14) state that "The Office of Minority Banks and Section 367 [of the Dodd Frank Act] have been in place for over ten years, the effectiveness of these measures appears at best to be minimal."

[6] In March 1969, the Minority Bank Deposit Program was created in response to an Executive Order by President Nixon, which established a national program supporting minority business enterprise. Section 308 of FIRREA includes provisions supporting the intent of the MDBP. See https://fiscal.treasury.gov/mbdp/about.html, accessed July 26, 2019.

[7] For example, in one of the earliest empirical studies, Carter, Simkins and Simpson (2003) examine the relationship between board diversity and firm value for Fortune 1,000 firms. They find a significant positive relationship between the fraction of women or minorities on the board and firm value. In a related paper, Erhardt, Werbel and Shrader (2003) find that there is a positive relationship between the percentage of women and minorities on boards of directors for 127 large US companies and financial indicators of firm performance.

of the MDIs from the perspective of whether their disproportionately small role in the banking industry can be explained by their relatively poorer profitability and riskier performance as compared to non-MDIs, controlling for various bank variables and taking into account characteristics of the local communities in which the different categories of banks operate. Fourth, we check the robustness of our results in various ways, including for the first time using two different databases on MDIs, one from the FDIC and the other from the FRB, and by comparing the performance and riskiness of MDIs and non-MDIs operating in the same communities to control for common demographic and economic factors that affect both categories of institutions.

Our results indicate that MDIs are not less profitable or display greater riskiness than non-MDIs.[8] This finding is contrary to many previous studies and suggests they may be no less desirable as investment opportunities. Indeed, in one of the earliest and important articles on MDIs, Brimmer (1971), focuses exclusively on the performance of black-owned banks and finds that they are less profitable than non-black-owned banks. A few years later, Boorman and Kwast (1974) examine the performance of eight minority-owned banks and find that their loan losses are about twice as high as non-minority banks. As a result, they also find that the large loan losses impair the ability of minority-owned banks to generate operating profits.[9] Still, later, Kwast and Black (1983) compare black-owned banks with a matched set of non-minority-owned banks and find that higher loan loss rates appeared to be the single most important factor in explaining the less profitable performance of black banks.[10] Also, Clair (1988) finds that compared with other banks

[8] Fairchild, et al. (2020) also find that MDIs are not systematically less efficient than their non-MDI counterparts.
[9] In a related paper, Henderson (1999) argues that the poor performance of Black MDIs may be attributable to inadequate assessment of risk as measured by adjustments to the provision for loan loss.
[10] An interesting paper, Black, Collins, and Cyree (1997) find that only black-owned banks utilize applicant race in the mortgage credit decision. In addition, they find that black-owned banks are more likely than white-owned banks to reject similarly situated black applicants, but do not attribute this result to discrimination against black borrowers. They suggest that their results indicate that the cultural affinity hypothesis may run from white-owned banks to white applicants and not from black-owned banks to black applicants.

serving the same areas, black-owned banks have a slightly lower return on assets. In the same year, Meinster and Elyasiani (1988) also examine the performance of minority-owned banks and find that their aggregate poor profitability as compared to non-minority-owned banks is due to black-owned banks. Price (1990) then finds that smaller minority-banks, on average, are less profitable than non-minority-banks of similar size. Lawrence (1997) also finds that black banks generally are less profitable than their non-minority peers competing in the same marketplace. At the same time, he finds that there is no significant difference in profitability in the case of Hispanic and Asian banks.[11]

Several more recent articles are by Breitenstein, et al. (2014), Hoque, et al. (2015), Toussaint-Comeau and Newberger (2017) and Breitenstein, et al. (2019). The article by Breitenstein, et al. (2014) compares the performance of MDIs and non-MDI community banks and finds that MDIs are significantly less profitable than non-MDI community banks based on t-tests. Toussaint-Comeau and Newberger (2017) also document that MDIs significantly underperform their peers in similar markets in terms of profitability. Hoque, et al. (2015) in a study of all banks and approximately 100 minority-owned banks, moreover, find that many minority-owned banks tended to be less profitable than other banks.[12] Furthermore, according to Elyasiani and Mehdian (1992, p. 946), "… the profitability shortfall of the MOBs [MDIs] is most probably due to … factors such as clientele profile and neighborhood characteristics, and thus, changes in management will not necessarily solve their profitability and viability problems." In a quite recent

---

[11] According to the Census Bureau website, "ZIP Code Tabulation Areas (ZCTAs) are generalized areal representations of United States Postal Service (USPS) ZIP Code service areas. The USPS ZIP Codes identify the individual post office or metropolitan area delivery station associated with mailing addresses. USPS ZIP Codes are not areal features but a collection of mail delivery routes." Based on this definition, we decided that a geographic unit with areal features is more appropriate for this study. See https://www.census.gov/geo/reference/zctas.html, accessed July 19, 2019.

[12] In a study of bank efficiency, Iqbal, et al. (1999) find that MDIs are less efficient than comparable non-MDIs in maximizing outputs from a given set of inputs. They further argue that this finding is yet another explanation for poor earnings performance of the MDIs.

article, Breitenstein, et al. (2019) find that MDIs tend to outperform non-MDI community banks in revenue generation, including net interest income and noninterest income. However, despite comparatively better revenue generation, MDIs have much higher noninterest expenses, especially among smaller MDIs, which tend to be predominantly African American and Native American MDIs. As a result, they conclude that MDIs have long underperformed other community and non-community banks when measured by efficiency ratios.[13]

Several reasons one might expect MDIs to underperform non-MDIs is that MDIs tend to locate and operate in minority neighborhoods as well as tend to serve minority customers. Such customers, moreover, tend to have lower incomes and less wealth accumulation. Also, households in minority neighborhoods tend to experience higher levels of poverty and unemployment. Furthermore, members of minority groups may be less financially literate and therefore experience more difficulties in managing financial products. Over time, these factors may have changed and become less significant in explaining the underperformance of MDIs. Of course, these categories of factors may affect different categories of MDIs differently, which makes separate analyses of different MDIs important.

Overall, we contribute to the existing banking literature by assessing whether the general finding of these previous studies that MDIs perform worse than non-MDIs still holds and find it does not. To the extent that there has been a general distrust and lack of confidence in the management of banks owned and/or controlled by minorities due to the belief that they underperform compared to banks owned and/or controlled by non-minorities, this could help

[13] Breitenstein, et al. (2014, p. 2) state, "… MDIs appear to underperform non-MDI institutions in terms of standard industry measures of financial performance …." In addition, according to Elyasiani and Mehdian (1992, p. 946), "… the profitability shortfall of the MOBs [MDIs] is most probably due to … factors such as clientele profile and neighborhood characteristics, and thus, changes in management will not necessarily solve their profitability and viability problems."

explain the extremely low participation rate of MDIs in the banking industry. If so, our findings may help contribute to more diversity in banking and thereby even lead to more and bigger MDIs that could generate greater benefits for lower-income communities with higher concentrations of minorities, which could help reduce the wealth gap in the country. Of course, there may be other factors beyond performance that may explain the relatively few MDIs, but identifying any other factors that explain the fewness in numbers is beyond the scope of our paper.

An important issue that arises is whether it matters that MDIs play a relatively minor role in the banking industry. In this regard, the FDIC (2017, p. 30) states MDIs "… play a vital role in the U.S. economy by providing responsive banking services to those who might not otherwise have access to a financial institution." They also "… tend to maintain offices in underserved communities that often have a higher concentration of low- or moderate-income (LMI) census tracts… and to minority borrowers compared to non-MDI institutions." It is therefore clear that more and bigger MDIs can contribute to providing more financial services to more individuals, and more individuals in underserved and lower-income communities. In this case, MDIs could also contribute to providing more credit to minority borrowers. Of course, as already stated, an important reason there are so few MDIs may be that they are perceived to be underperformers and riskier than non-MDIs, thereby serving as a disincentive at attempts to increase their numbers and size. This is the focus of our paper.

The remainder of the paper proceeds as follows. Section 1.2 provides an overview of the MDI industry. Section 1.3 discusses the research questions addressed and describes the data used in the empirical analysis. Section 1.4 introduces the specific models and associated variables used in assessing the determinants of the location of MDI offices as well as the performance and riskiness of MDIs as compared to non-MDIs. Section 1.5 reports and discusses the empirical results. Section 1.6 presents the results of robustness tests. Finally, Section 1.7 concludes and provides suggestions for future research.

1.2. Overview of the MDI Industry

An issue that immediately arises when focusing on MDIs is exactly how to distinguish these institutions from other depository institutions. In this regard, Section 308 of the FIRREA defines the term "minority depository institution" as any depository institution where "Black Americans, Asian Americans, Hispanic Americans, or Native Americans" own 51 percent or more of the stock. US citizens or permanent legal US residents must own the voting stock, moreover, in determining minority ownership. In addition to institutions that meet the ownership test, institutions are minority depository institutions if a majority of the Board of Directors is the minority and the community that the institution serves is predominantly minority. Institutions not already identified as minority depository institutions can request such a designation by certifying that they meet the above definition. The FDIC issued a Policy Statement that more simply defines a "minority depository institution" as any federally insured depository institution where minority individuals own 51 percent or more of the voting stock. [14]

The FRB, in contrast, identifies MDIs based on the list of institutions participating in the Treasury's Minority Bank Deposit Program (MBDP). According to Price (1990), the list is distributed to public and private organizations to encourage the use of minority-bank services. Under this program, moreover, participating MDIs are entitled to preference as depositories for federal government funds. [15] The potential deposits available to MDIs include federal agency

[14] See FDIC Definition of Minority Depository Institution, https://www.fdic.gov/regulations/resources/minority/mdi-definition.html, accessed July 19, 2019.

[15] In October 1977, President Carter signed a memorandum for all Heads of Departments and agencies promoting the use of minority-owned enterprises by placing deposits in minority banks. See https://fiscal.treasury.gov/mbdp/about.html , accessed July 29, 2019.

deposits of public money, cash advances to federal contractors and grantees, and certain independent demand deposits such as Postal Service deposits, and certificates of deposit (CDs).[16] This presumably enables MDIs to obtain deposits at a lower cost than having to compete for them in their local banking markets.[17] However, only 49 of the 155 MDIs had federal government deposits at yearend 2017.[18] Nevertheless, 125 of the MDIs did have state and local government deposits.[19] Also, MDIs may benefit from technical assistance, training, and educational programs provided by the banking regulatory agencies that are unavailable to other insured depository institutions. Furthermore, under the Community Reinvestment Act (CRA), non-MDI financial institutions may be encouraged to provide support to MDIs to meet the requirements of the Act concerning the lending, investment, and service tests.[20]

The eligible participants in MBDP include banks that are minority-owned or minority-controlled; banks that are owned, controlled, and operated by women; and low-income credit unions designated by the NCUA. The term "minority ownership" refers to banks where members of minority groups own more than 50 percent of the institution's outstanding stock.[21] In 2015, the FRB started including the MDIs identified by the FDIC in its list of MDIs, which resulted in an expansion in the number of MDIs by including those institutions not participating in MBDP.

---

[16] It might be noted that the National Bankers Association, which was incorporated in 1972, is a national trade association for MDIs. According to Price (1990), "[I]t provides a forum for sharing information and resources and also actively solicits deposits from government agencies and major corporations for its members."

[17] However, Price (1994) finds that government deposits are expensive, and that deposits received through the MBDP may have the effect of increasing risk in the asset portfolio. Price (1995, p. 300) in a subsequent paper argues that "… MBDP is a cost ineffective way to simulate BCB [Black Commercial Bank] entry, and should therefore be abandoned." More recently, Kashian, et al. (2017) find that government deposits did not adversely effect efficiency among MDIs, but may have improved survival rates for them after the financial crisis.

[18] In a study of Black MDIs covering the early 1970s, Bates and Bradford (1980) find that such banks hold federal government deposits that are a relatively large proportion of their total deposits.

[19] Federal government deposits as a percentage of total deposits for MDIs (non-MDIs) was 0.11 (0.03), while the corresponding percentage of state and local government deposits for MDIs (non-MDIs) was 8.49 (4.49).

[20] See Breitenstein, et al. (2014).

[21] Minority Bank Deposit Program (MBDP), https://www.fiscal.treasury.gov/fsservices/gov/rvnColl/mnrtyBankDep/rvnColl_mbdp_started.htm, accessed July 19, 2019.

In this study, although we primarily rely on the annual list of MDIs provided by the FDIC, we do compare the number of MDIs on both the FDIC and FRB lists before 2015 in the latter part of the paper.[22] Figure 1.1 Panel A shows that the number of MDIs increased sharply from 2001 to 2008, from 164 to 215, an increase of 25 percent, and then trended downwards over the remaining years. Specifically, the number declined from 215 in 2008 to 149 in 2018, a decline of 33 percent. At the end of the period, there were 15 fewer MDIs than at the beginning of the period. Figure 1.1 Panel A also shows the distribution of the different categories of MDIs over the same period. Over the period, all the different categories of MDIs increased in number except for Black MDIs. Indeed, the number of Black MDIs continuously decreases from 2001 to 2018, reaching a low of 23 at the end of the period.

[Insert Figure 1.1 Panel A About Here]

Figure 1.1 Panel B shows the distribution of the different categories of MDIs offices, including headquarters, over the same period as Figure 1.1 Panel A. The number of Hispanic MDI offices was the largest throughout the entire period and accounted for the largest percentage of all four categories of MDI offices. However, their percentage of the total number of MDI offices decreased from 53 percent in 2001 to 46 percent in 2018. The reason was due to the number of Asian MDI offices increasing to 660 in 2018 from 334 in 2001. As a result, their share of all MDI offices increased over the same period from 29 percent to 43 percent. The black share of offices over the period was quite low, always less than 15 percent of the total. Although there was a decline in the number of headquarters from 2001 to 2018, the number of MDI offices increased to 1,524 from 1,191, an increase of 333 offices.

[Insert Figure 1.1 Panel B About Here]

[22] We also perform a robust check on our empirical results using the MDIs on the FRB list.

Figure 1.1 Panel C shows the total assets of the different categories of MDIs and their shares in the MDI industry over the period from 2001 to 2018. Total assets of Asian MDIs increased substantially, from $19 billion in 2001 to $122 billion in 2018. These MDIs took the lead in total assets in 2016 when they held $102 billion in assets. Hispanic MDIs were the largest in terms of total assets before 2016. In yearend 2018, Asian MDIs accounted for 52 percent of total MDI assets, while Hispanic MDIs ranked second with 45 percent. Total assets of Black MDIs remained relatively small over the entire period with a high of $7 billion in 2008, before declining to $5 billion in 2018. Their share of the assets of all MDIs was 2 percent at the end of the period.

[Insert Figure 1.1 Panel C About Here]

Even though MDIs account for a small percentage of the total number and assets of banks nationwide, their importance is somewhat greater when one takes into account that they are concentrated in relatively limited geographical areas (see Figure 1.2). In particular, most Hispanic MDI offices tend to be located in Texas, New Mexico, Southern California, and Miami. Native American MDI offices tend to be mainly located in Oklahoma, Minnesota, and North Carolina. Black MDI offices are located almost exclusively east of the Mississippi River, ranging from the southern to the northern coasts of the country. Asian MDI offices, in contrast, are the most widely dispersed, being located mainly in the western coastal states, central southern states, and the eastern coastal states, though 51 percent are in California. Interestingly, the different categories of MDIs generally tend to have offices not overlapping with one another. This enables one to examine better the presence, performance, and riskiness of such institutions as compared to non-MDIs in the same and more localized geographical areas.

[Insert Figure 1.2 About Here]

Table 1.1 provides information on the number of all banks and bank offices as well as the subset of MDIs and MDI offices over the period 2001 to 2017. It shows there is a significant downward trend in the number of banks in the US, from 9,757 in 2001 to 5,797 in 2017. The number of bank offices, which includes bank headquarters and bank branches, generally increased over the period from 86,069 to 89,857 but declined every year after reaching a peak of 99,550 in 2009. [23] It also shows the matching of each of the offices with a specific census tract.[24] The matching procedure was based on the latitude and longitude of each office. In most cases, this locational data was available from the FDIC, while geocoding was used when it was not available to convert the addresses of bank offices into latitude and longitude. One could then combine this locational information with TIGER/Line Shapefiles from the US Census Bureau to identify the census tract for each bank office. Based on the procedures followed, the rate of failure to match offices with a specific census tract is always less than 0.1 percent, as shown by the rate of failure match in the table.

[Insert Table 1.1 About Here]

The table also shows that the number of census tracts in which there is a bank office. The lowest percentage of census tracts with a bank office is 47 percent in both 2016 and 2017, while the highest percentage is 53 percent in both 2007 and 2008. Of course, the percentage of census tracts with an MDI office is substantially lower than these figures, at no greater than 2 percent. This means there are roughly half of all census tracts without any banking offices and nearly all without any MDI offices. This situation may impose difficulties for individuals, especially poorer

[23] Branch data are only available in June of each year. Therefore, the data on the number of banks and bank branches covers the period June 2001 to June 2017.

[24] Census tracts are small, relatively permanent statistical subdivisions of counties designed to be relatively homogeneous units with respect to population characteristics, economic status, and living conditions. They average about 4,000 inhabitants. See https://www.iowadatacenter.org/aboutdata/tracts, accessed June 18, 2019.

and older individuals, who depend on local access to banking offices despite the growing importance of online banking. The situation may be even worse for minorities given the scarcity of census tracts with MDI offices.

Table 1.1 also provides additional information indicating that the number of census tracts with MDI headquarters reached a low of 140 in 2017 after reaching a high of 180 in both 2007 and 2008. At the same time, the number of census tracts with at least one MDI office increased from a low of 885 in 2001 to a high of 1,345 in 2010, before declining to 1,056 in 2017. The table also shows the number of census tracts with both MDI and non-MDI headquarters. The number of census tracts, in this case, trended downwards over the period to a low of 19 in 2017 from a high of 40 in 2001. Interestingly, every year the number of non-MDI headquarters exceeds the number of MDI headquarters in the same census tract. The number of MDI headquarters in this situation, moreover, is always one-third or less than the total number of MDI headquarters. Stated another way, two-thirds of the MDI headquarters are in census tracts with no other bank headquarters.[25]

Table 1.2 shows the basis for change of the number of MDIs over time in terms of whether institutions were entering, exiting, or acquired when exiting. As may be seen, the decline in the number of MDIs are from 2001 to 2017 was not due simply to institutions exiting the MDI industry. Indeed, over the period, 132 institutions became MDIs at one time or another. At the same time, 141 institutions exited the MDI industry, 39 of which were acquired when exiting. There are 46 MDIs that were established between 2001 and 2017. Fifteen MDIs changed from an MDI to a non-MDI. Eighty-seven non-MDIs changed to MDIs during the period, and among these 7 after switching returned to non-MDIs. There were 123 banks that were MDIS at one time or another during 2001- 2017 but disappeared by the end of the period. One-hundred-seventeen were

[25] Table 1.1 also shows that the average number of bank offices for non-MDIs exceeds the average number for MDIs in every year from 2001 to 2017.

merged into or acquired by other banks. Among the mergers or acquisitions, 61 were merged/acquired by an MDI. Over the period, 41 MDIs entered the industry and were still in existence at yearend 2017. Interestingly, the oldest MDI, CBW Bank, that still exists was established on January 1, 1892, while the newest, Urban Partnership Bank, was established in August 2010. In short, the table shows there was not simply a steady decline of MDIs by 9 institutions from 2001 to 2017.[26] In the next section, we discuss various studies examining the performance of MDIs as compared to non-MDIs.

[Insert Table 1.2 About Here]

1.3. Research Focus and Dataset

Regarding the identification of MDIs, the FDIC provides an annual MDI list starting in 2001. All the bank balance sheet and income statement information are from publicly available FDIC quarterly call reports.[27] FDIC also provides the branch office deposits of banks, on an annual basis every June 30. Since our analysis focuses on community banks, a list of such banks is also provided by the FDIC. Demographic data are available from the 2000 census and the 2009- to 2017- 5-year estimates from the American Community Survey (ACS) provided by US Census Bureau.[28] Linear interpolation allows us to obtain data for the intervening years from 2001 to 2008. Our analysis, therefore, covers the period 2001 to 2017. Also, we can determine whether a census tract is a low- or moderate-income tract from the Federal Financial Institutions Examination Council (FFIEC) Census Report.[29]

[26] See Kashian and Drago (2017) for a study of bank failures for 2009 to 2014. They find that failures among MDIs, compared to a similar sample of community banks, were disproportionately located among Asian MDIs and particularly the Black MDIs. This is consistence with Figure 1.1 Panel A, which shows a decline in these two categories of MDIs over the period 2009 to 2014.
[27] See https://www7.fdic.gov/sdi/download_large_list_outside.asp, accessed February 3, 2021.
[28] See https://data.census.gov/cedsci/, accessed February 3, 2021.
[29] See https://www.ffiec.gov/censusapp.htm, accessed July 11, 2019.

To match the demographic data and bank branches to the same census tracts, as already noted, we rely on the TIGER/Line Shapefiles for each year. To elaborate on the earlier discussion, the Shapefiles are available for 1990, 2000, 2010, and each year from 2013 to 2018. While census tracts are relatively permanent statistical geographic units, their geographic area does change occasionally. To ensure the accuracy of matching the location of bank offices with specific census tracts, the office addresses in a year are matched with the latest available file for the same year. That is, the 2000 Shapefile is used to match addresses from 2001 to 2009, the 2010 Shapefile is used to match addresses from 2010 to 2012, and then 2013 to 2017 addresses are matched with each corresponding year's Shapefile. The high success rate of the matching process (up to 99.9 percent, see Table 1.1) better ensures the completeness of obtaining matches for the 1,593,077 bank offices and census tracts over the seventeen years examined.

Our empirical work relies on a subset of MDIs and non-MDIs in census tracts as shown in Table 1.3. The primary reason for differences between Tables 1.1 and 1.3 is that the latitude and longitude information provided by the FDIC in some cases we believe identified an incorrect county for some bank offices. When this happened, the bank offices were removed from our sample. As a result, the number of MDI headquarters and MDI offices declined, as may be seen when comparing the figures in Table 1.3 with the comparable figures in Table 1.1. In a relatively few other cases, necessary demographic information was not available that led to the exclusion of some census tracts. The bottom line is that our sample of MDI headquarters and MDI offices was reduced as a result of these exclusions by an average of 19 MDI headquarters and 168 MDI offices over the period 2001 to 2017. However, over the period 2001 to 2009, the averages were 32 for MDI headquarters and 234 for MDI offices, while over the period 2010 to 2017 the comparable figures were substantially lower at 4 and 93, respectively. As we indicate below, we, therefore, run our regressions for the entire period as well as the latter period with fewer exclusions as a check and find no essential differences in results.[30]

[Insert Table 1.3 About Here]

In our analysis, we also include information as to whether MDI headquarters and MDI offices are located in low- and moderate-income (LMI) census tracts.[31] Table 1.3 shows that an average of 22 percent of all MDI headquarters are located in low-income census tracts, while an average of 27 percent of all MDI headquarters is located in moderate-income census tracts. In terms of MDI offices, an average of 11 percent of them are in low-income census tracts, while an average of 25 percent of them are located in moderate-income census tracts. More generally, roughly half of all MDI headquarters and slightly more than one-third of all MDI offices are in LMI census tracts. These figures are consistent with the earlier statement by the FDIC that MDIs "… maintain offices in underserved communities which often have a higher concentration of low- or moderate-income (LMI) census tracts".

1.4. Models and Variables

Our model involves a comparison of the performance and risk level of MDIs to non-MDIs. The specific model estimated is as follows:

(1) $\text{Bank performance}_{it} = \beta_0 + \beta_1 \text{MDI}_{it} + \sum_{j=2}^{6} \beta_j \text{Bank Control Variables}_{j,it} + \sum_{j=7}^{12} \beta_j \text{Demographic Variables}_{j,it} + \text{Time Fixed Effects} + \varepsilon_{it}$,

(2) $\text{Risk level}_{it} = \beta_0 + \beta_1 \text{MDI}_{it} + \sum_{j=2}^{6} \beta_j \text{Bank Control Variables}_{j,it} + \sum_{j=7}^{12} \beta_j \text{Demographic Variables}_{j,it} + \text{Time Fixed Effects} + \varepsilon_{it}$,

[30] It is also important to note that the panel regression results in Tables 1.8 and 1.9 that we discuss below are based on census tracts with both MDI headquarters and non-MDI headquarters in which case there is no difference in the numbers in Tables 1.1 and 1.3.

[31] A low-income census tract is one for which median family income is > 0% and < 50% of the MSA/MD median family income, while a moderate-income census tract is one for which median family income is >= 50% and < 80% of the MSA/MD median family income. See https://www.ffiec.gov/censusapp.htm, accessed July 10, 2019.

where bank performance is measured by return on assets (ROA) and risk level is measured by ln z-score.[32] We construct the z-score, Z_{it}, as follows:

(3) $$Z_{it} = \frac{(ROA_{it} + \frac{E_{it}}{A_{it}})}{\sigma_{ROA_{it}}},$$

where E_{it}/A_{it} is bank i's equity to total asset ratio in period t, while $\sigma_{ROA_{it}}$ is the variance of the distribution of ROA_{it}, respectively. We use the sample variance of quarterly ROA over the previous 12 quarters as estimates of $\sigma_{ROA_{it}}$. The z-score measures the distance to default and indicates the number of standard deviations ROA would need to decline to wipe out equity. The higher the z-score the lower risk of a bank.

MDI_{it} is a dummy variable equal to 1 if institution *i* is an MDI at time *t*, and 0 otherwise. A significantly positive coefficient on this dummy variable indicates that the ROA (ln z-score) of MDIs is higher (less risky) than for non-MDIs. We estimate equations (1) and (2) for all MDIs as well as each category of MDI.

The bank-specific control variables include: (1) log of total assets (Log_TA), (2) loan to total assets (Loan/TA), (3) liquidity to total assets (Liquidity/TA), defined as the sum of cash and due from, Federal Funds purchased and securities purchased under reverse repurchase agreements, and securities held to maturity plus securities available for sale, divided by total assets, (4) nonperforming loans (NPLs/TL), defined as the sum of loans past due 30 - 89 days and still accruing interest, loans past due 90 days or more and still accruing interest, nonaccrual loans, and OREO (other real estate owned), divided by total loans, and (5) equity-to-total assets (Equity/TA).

The specific demographic variables for each of the census tracts include the following: the fraction of minority population (Minority), the fraction of each of the four different minorities

[32] Because actual MDI failures are infrequent, we rely on these two performance measures. For our sample period, 2001 to 2017, only 2 MDIs failed.

(Black, Asian, Hispanic, and Native), the fraction of population below the poverty level (Poverty), vacancy rate (Vacancy), the fraction of population over 25 years old whose educational attainment is high school (Highsch), log median household income (Logincomeh) and fraction of population aged 65 years or over (Over 65). Based upon a logit analysis, not reported, we find that census tracts with a higher fraction of minority population, greater poverty rate, lower median household income, and an LMI tract are more likely to have MDI offices located in them.

Table 1.4 contains summary statistics for all the variables used in the empirical estimation. The table shows that MDIs have significantly higher ROAs than non-MDIs, but are riskier on average, as measured by the z-score. Also, MDIs are significantly bigger in terms of assets than non-MDIs and have significantly higher loan-to-asset ratios and equity-to-asset ratios. However, MDIs on average have significantly higher nonperforming loans-to-loan ratios. MDIs, moreover, are located in communities in which there are significantly higher poverty rates, lower median household incomes, higher percentages of individuals for which the highest level of education is a high school diploma, and a smaller percentage of individuals aged 65 and over.

[Insert Table 1.4 About Here]

In Table 1.5, the correlations among all the variables are presented, with information about their statistical significance. The table indicates that there is a high degree of statistical significance for the majority of the correlations. This is not that surprising given the choice of variables used in our analysis. In the next section, we move beyond bivariate correlations to examine the profitability and riskiness of MDIs and compared them to non-MDIs when controlling for a variety of factors.

[Insert Table 1.5 About Here]

1.5. Empirical Results

Table 1.6 presents the panel regression results for the performance (measured by ROA) of the MDIs (headquarters) as well as the four different categories of MDIs (headquarters). The regression results for MDIs are in the first two columns, while those for the four categories of MDIs are in the last eight columns. Also, some columns present results excluding demographic control variables, while other columns include them, but all columns include time fixed effects. The sample includes all MDIs, each of the four categories of MDIs, and all others. In the first two columns, the MDI_dummy variable equals 1 if the institution is an MDI, and 0 otherwise. In the last eight columns, the MDI dummy variables equal 1 if the institution is either a Black MDI, Asian MDI, Hispanic MDI, or Native American MDI, and 0 otherwise.

[Insert Table 1.6 About Here]

Bank control variables show various significance levels for different regressions. For all MDIs, the coefficients of Log_TA, Loan/TA, and Liquidity/TA are positive, while the coefficients of NPLs/TL and Equity/TA are negative. Also, all the variables are statistically significant. Importantly, and given the focus of this study, the coefficient of the MDI dummy variable is positive and significant for Black MDIs, while insignificant for all MDIs and Asian MDIs. In terms of economic significance, a Black MDI bank as compared to a non-black bank has a ROA that is 56 basis points higher. In the case of Hispanic MDIs and Native American MDIs, the coefficient is negative while insignificant. As compared to a non-Hispanic bank and non-Native American bank. This finding and especially the finding that Black MDIs perform better than non-black banks are contrary to the results of earlier studies, as noted above.

The regression results with the ln z-score as the measure of the riskiness of a bank are reported in Table 1.7. All the explanatory variables are the same as in Table 1.6. The regression for all MDIs indicates that all bank variables are significant. Four variables, Log_TA, Loan/TA, Liquidity/TA, and Equity/TA, have positive coefficients, indicating they are associated with lowering bank risk. However, the coefficient of NPLs/TL is negative and significant, indicating higher NPLs/TL is associated with greater risk. The most notable feature in Table 1.7 is that all MDI dummy variables are not significant, including all MDIs, Black MDIs, Asian MDIs, Hispanic MDIs, and Native American MDIs.[33] This indicates the various categories of MDIs are no riskier than non-MDIs.

[Insert Table 1.7 About Here]

Table 1.8 presents the panel regression results for the performance (measured by ROA) of the MDIs (headquarters) as well as the four different categories of MDIs (headquarters). [34]The sample for all MDIs includes those MDIs and non-MDIs located in the same census tract, while the sample for each of the four categories of MDIs includes only those census tracts with one or more of a specific category of MDI and at least one other bank. For example, the sample for Black MDIs consists of census tracts with at least one Black MDI and at least one non-black bank. In the first two columns, the MDI_dummy variable equals 1 if the institution is an MDI, and 0 otherwise. In the last eight columns, the MDI dummy variables equal 1 if the institution is either a Black MDI, Asian MDI, Hispanic MDI, or Native American MDI, and 0 otherwise.

[Insert Table 1.8 About Here]

Bank control variables show various significance levels for different regressions. For all MDIs, the coefficients of Log_TA and Loan/TA are positive, while the coefficients of Liquidity/TA, NPLs/TL, and Equity/TA are negative. Only Log_TA and NPLs/TL are statistically

---

[33] However, the dummy variable for Asian MDIs without demographic control variables is significantly negative at the 10% level.
[34] The dummy variable for Hispanic MDIs without demographic control variables is significantly positive at the 10% level.

significant for all MDIs, however. Importantly, and given the focus of this study, the coefficient of the MDI dummy variable is not significant for all MDIs. This indicates that there is no difference in the ROAs for MDIs and non-MDIs located in the same census tracts. This finding is contrary to the results of earlier studies, as noted earlier.

Turning to the different categories of MDIs, and focusing on the MDI dummy variable, we find that the coefficients on the Black MDIs, Asian MDIs, Hispanic MDIs, and Native American MDIs are not statistically significant. These results indicate that there is no difference in ROAs between each of the different categories of MDIs and other banks located in the same census tracts. More generally, these findings are contrary to the results of earlier studies, as noted above.

The regression results with the ln z-score as the measure of the riskiness of a bank are reported in Table 1.9. All the explanatory variables are the same as in Table 1.8. The regression for all MDIs indicates that all bank variables are significant, except Liquidity/TA. The three variables, Log_TA, Loan/TA, and Equity/TA, are significant and have positive coefficients, indicating they are associated with lowering bank risk. However, the coefficient of NPLs/TL is negative and significant, indicating higher NPLs/TL is associated with greater risk. The most notable feature in Table 1.9 is that the MDI dummy variables are insignificant for all MDIs, Black MDIs, Asian MDIs, and Native American MDIs, but significant for Hispanic MDIs. This indicates that three of the four categories of MDIs are no riskier than other banks. In the case of Hispanic MDIs, the results indicate they are less risky than non-Hispanic banks located in the same census tracts.

[Insert Table 1.9 About Here]

1.6. Robustness Tests

As a further check on our empirical results, we re-estimate our basic equation for all MDIs using only those MDIs that were in continuous existence throughout the sample period. Seventy-four MDIs satisfied this criterion, so we matched these banks with non-MDIs in the same census tracts. In the case of ROA, Table 1.10 shows that there is a positive and significant coefficient on the MDI dummy variable when including demographic variables, but insignificant when excluding these variables. This indicates that the MDIs perform significantly better than the non-MDIs, with a ROA that is, on average, 53 basis points higher in terms of this performance measure. As regards the ln z-score, the results indicate that the MDIs are less risky than non-MDIs, with a natural logarithm z-score that is higher by 0.41.

[Insert Table 1.10 About Here]

We also have access to a list of MDIs provided by the FRB. Interestingly, this list covers the period from the first quarter of 1960 to the fourth quarter of 2018. As shown in Figure 1.3, the number of MDIs increase substantially, albeit from a low base, from the early 1960s to the late 1980s. According to Price (1990), "[m]inority institutions experienced significant growth during the 1970s—growth that was assisted, in part, by social legislation."[35] Starting in 1989, the number tended to gradually decline, but still ended the period higher than in any year before the 1990s. The list of MDIs from the FDIC is included in the figure for purpose of comparison. As may be seen, the two lists of MDIs from the two bank regulatory authorities do differ until 2015, when the FRB started including the FDIC list in its list. Given the difference in the number of MDIs over our sample period, we re-estimated our basic results reported for ROA and ln z-score using the

[35] The social legislation he refers to is the establishment of MBDP mentioned in footnote 4.

FRB list for all MDIs. In this case, we find that the MDIs perform better than the non-MDIs in the same census tracts. We also find that they are less risky.

[Insert Figure 1.3 About Here]

Also, we compare all MDIs to non-MDIs located in the same zip code as well as the same city. The results indicate that MDIs perform better in terms of ROA when both categories of banks are operating in the same zip code. MDIs and non-MDIs perform no differently in terms of ROA when both categories of banks are operating in the same city. As regards the riskiness of banks, we find that the MDIs are less risky than non-MDIs operating in either the same zip or city, with a natural logarithm z-score that is higher by 0.53 and 0.23, respectively.

As the last robustness test, we construct pairs of MDIs and non-MDIs generated by propensity score matching (PSM), a commonly used matching method (Rosenbaum and Rubin, 1983). Table 1.11 provides information on the imbalance measures with and without matching MDIs and non-MDIs based on demographic variables. As may be seen, the imbalance measures are substantially lower when matching MDIs with non-MDIs.

[Insert Table 1.11 About Here]

Tables 1.12 and 1.13 present the panel regression results using this technique and including the same variables as in Tables 1.6 and 1.7. We employ a one-to-one without replacement matching method using logit regressions with demographic variables to generate control groups of non-MDIs whose performance and riskiness can then be compared to all MDIs and each category of MDI. As regards the MDI dummy variables, which is the focus of our study, they are not significant in nine of the ten regressions in the table, indicating once again no difference in ROAs between MDIs, and four categories of MDIs, and other banks. However, in the case of Hispanic MDIs, the results indicate they perform worse than non-Hispanic banks, with a ROA that is, on average, 26 basis points lower when including demographic variables. As regards the ln z-score, the findings indicate that all MDIs, Black MDIs, Asian MDIs, Hispanic MDIs, and Native American MDIs are no riskier than other banks.

[Insert Tables 1.12 and 1.13 About Here]

1.7. Concluding Remarks

The purpose of our paper is to motivate and contribute to a better understanding of the role and comparative performance of MDIs, and different categories of MDIs, to non-MDIs in the US community banking industry. The findings in some cases indicate that Black MDIs perform significantly better in terms of profitability but are no riskier than non-black banks. There is no difference, moreover, between MDIs, Asian, Hispanic, and Native American MDIs and non-MDIs, non-Asian, non-Hispanic, and non-Native American banks, respectively.

As regards differences between MDIs and non-MDIs whose headquarters are located in the same census tracts, our findings indicate that MDIs, as well as each category of MDI, perform no differently than non-MDIs in terms of ROA and overall riskiness as measured by the z-score, except in the case of Hispanic MDIs, which generally perform better in terms of both profitability and riskiness when demographic variables are omitted. As a check on these results, several different robustness tests are conducted. We generally confirm the results for all MDIs when comparisons are made with non-MDIs using the FRB data, at a zip-code level of analysis, and a city-level of analysis. Lastly, we match the MDIs with non-MDIs using PSM and find no difference in terms of profitability and riskiness, except in some cases of Hispanic MDIs.

We hope our study will stimulate further research on the role and performance of MDIs in the financial marketplace. More work could focus on uncovering reasons for the relatively small role that MDIs play in the banking industry and thereby limited diversity in the ownership/control of banks by minorities within the banking industry. Our focus is on reassessing the many earlier studies that find MDIs tend to underperform and/or are riskier than non-MDIs. Contrary to these studies, our findings generally indicate no difference in either profitability or riskiness.[36] Indeed, the results indicate that MDIs provide investors with essentially the same opportunities as non-MDIs. If greater investment in MDIs were to take place, this could lead to their increased growth and thereby the ability to provide more credit to individuals and firms located in low- and moderate-income communities and thereby helping reduce the income and wealth gaps in the country.

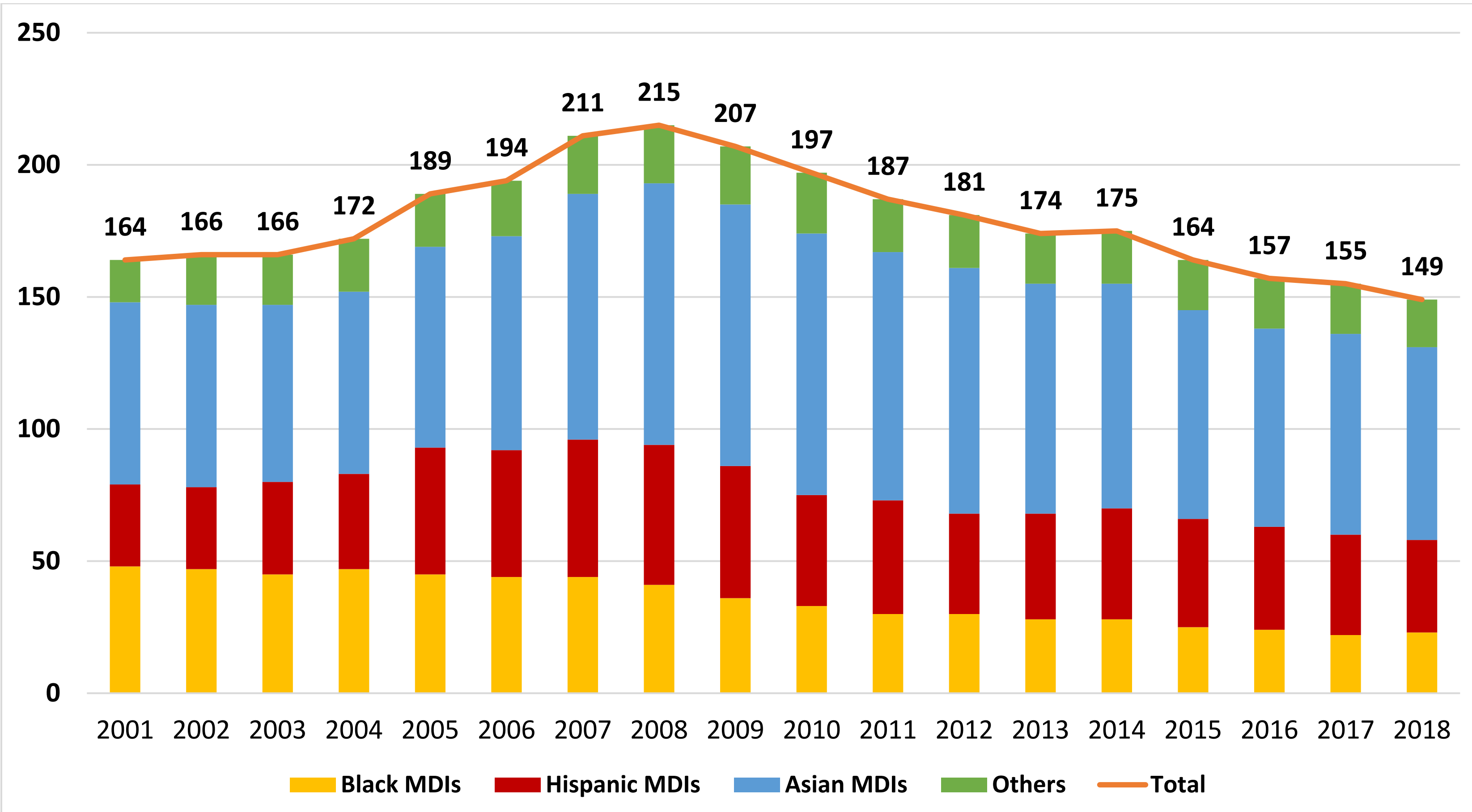
Finally, future research could examine the reasons for the small percentage of federal government deposits at MDIs that we document, given the purpose of the MBDP. More generally, given the importance for the local communities in which MDIs are located, it is also important to consider public policies that would be appropriate to further expand access to the services of banks as well as diversity in banking through the growth and expansion of MDIs.[37]

[36] In a private conversation, Harold Black, Professor Emeritus of Finance at the University of Tennessee, suggests that better personal networks among non-minorities interested in entering banking may have played a role in the establishment of so many non-MDIs.
[37] In this regard, Barth and Betru (2019) suggest allowing opportunity zone investments in MDIs.

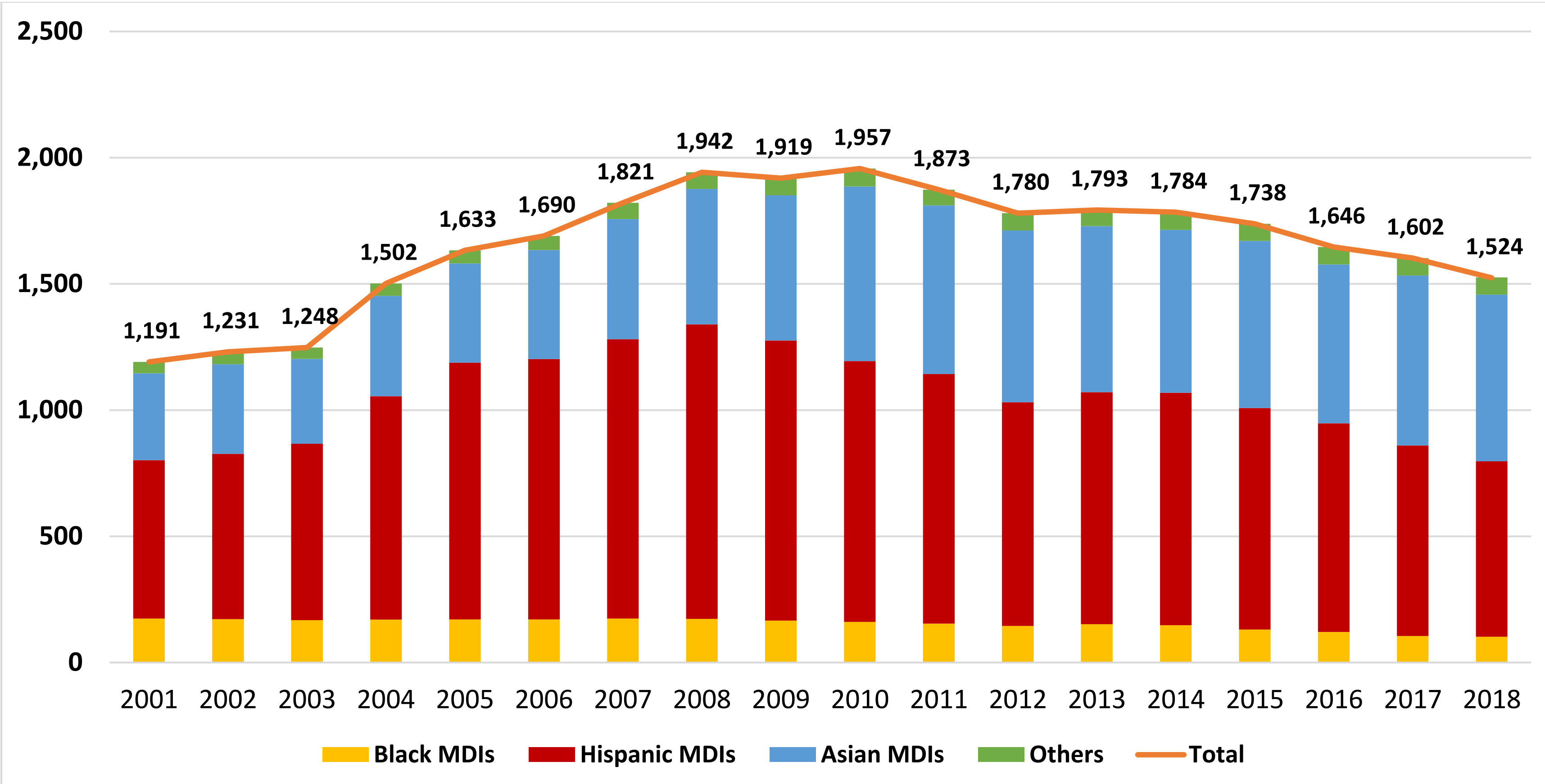
Figures and Tables

Figure 1.1 Minority Depository Institutions by Category

Panel A. Number of Minority Depository Institutions by Category



Panel B. Bank Offices of Minority Depository Institutions by Category



Panel C. Total Assets of Minority Depository Institutions by Category ($ Billion)

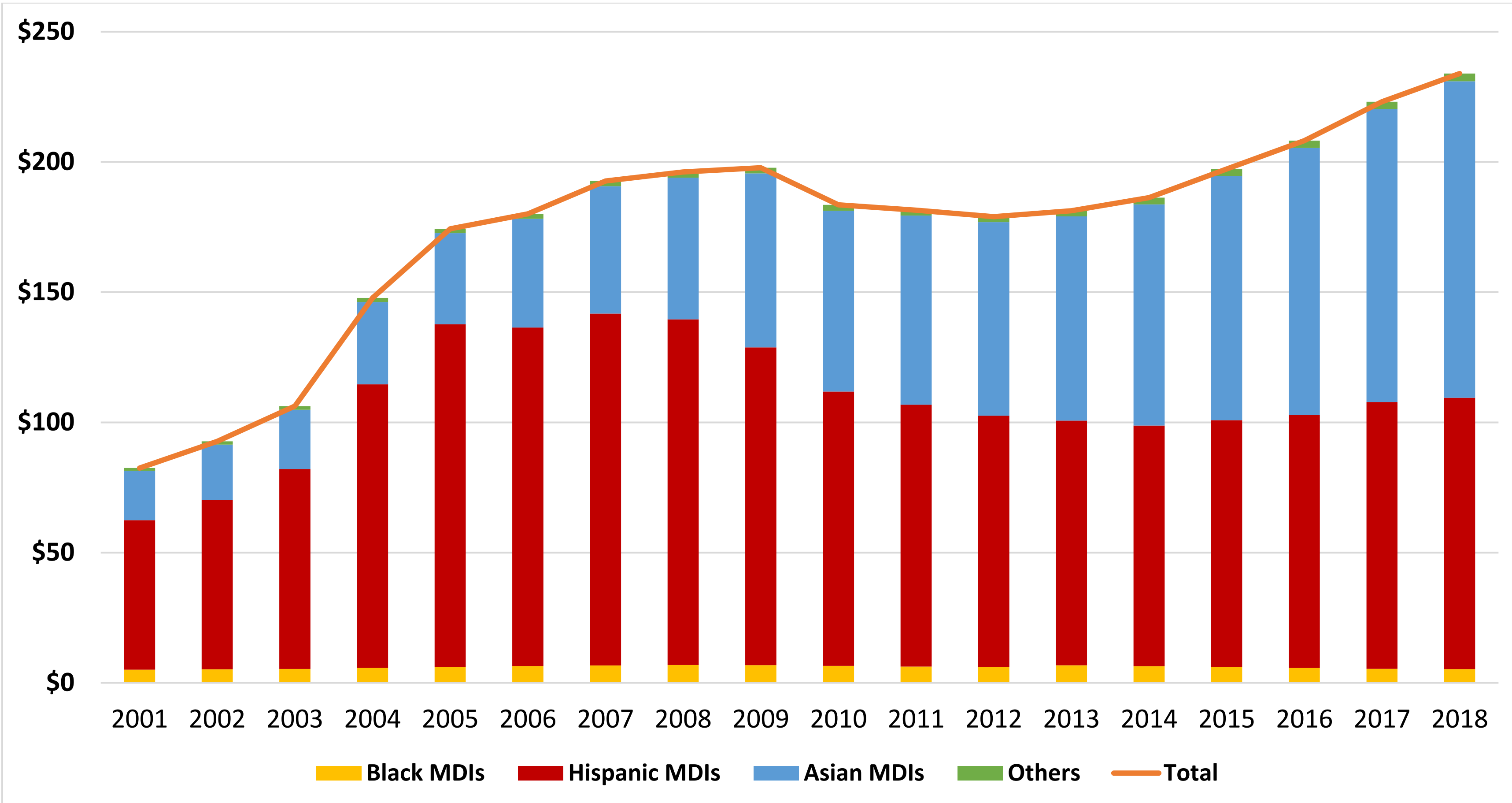


Figure 1.2 Location of Minority Depository Institution Offices by Category, as of June 30, 2018

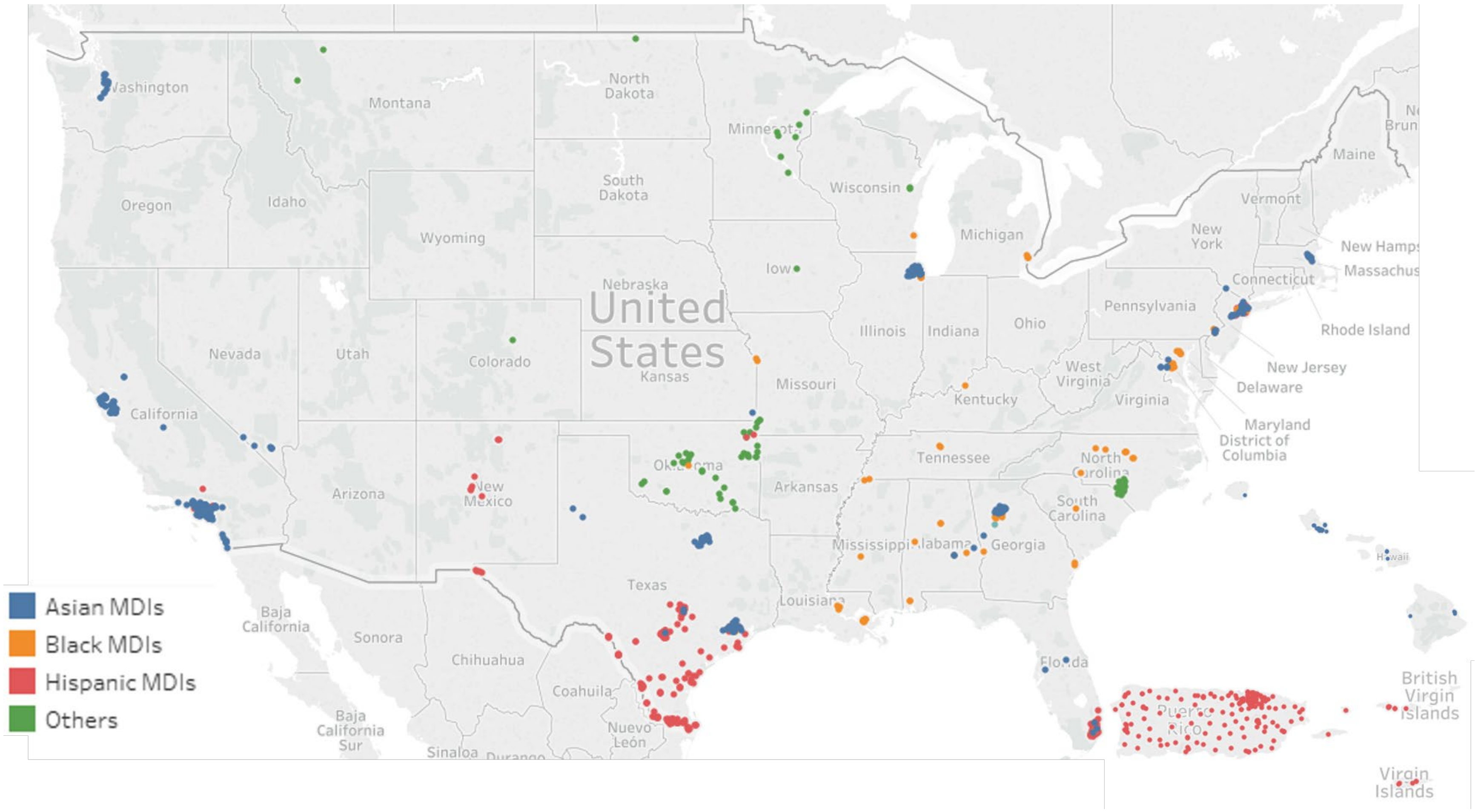


Table 1.1 Selected Information on Banks and Bank Offices, and MDIs and MDI Offices in Census Tracts Nationwide

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009
Total Number of Census Tracts	66,688	66,688	66,688	66,688	66,688	66,688	66,688	66,688	66,688
Number of Census Tracts with Bank Offices	33,551	33,625	33,947	34,174	34,544	34,869	35,132	35,136	34,119
Percentage of Census Tracts with Bank Offices	50%	50%	51%	51%	52%	52%	53%	53%	51%
Total Number of Bank Offices	86,069	86,577	87,785	89,785	92,043	94,752	97,274	99,164	99,550
Number of Bank Offices, Fail to Find a Census Tract	47	61	68	210	40	31	32	38	51
Rate of Failure Match (%)	0.055	0.070	0.077	0.234	0.043	0.033	0.033	0.038	0.051
Total Number of Banks	9,757	9,474	9,256	9,066	8,856	8,767	8,605	8,440	8,185
Number of Census Tracts with MDI Offices	885	898	894	1,018	1,126	1,152	1,241	1,236	1,272
Percentage of Census Tracts with MDI Offices	1%	1%	1%	2%	2%	2%	2%	2%	2%
Number of Census Tracts with MDI HQs	151	153	150	154	163	167	180	180	177
Number of MDI HQs	164	166	166	172	189	194	211	215	207
Number of MDI Offices	1,191	1,231	1,248	1,502	1,633	1,690	1,821	1,942	1,919
Number of Census Tracts with Both MDIs & Non-MDIs HQs	40	38	37	36	35	34	38	28	29
--Number of Non-MDIs HQs within These Census Tracts	105	94	86	85	80	84	89	85	83
--Number of MDIs HQs within These Census Tracts	49	46	44	41	40	38	45	32	37
Average Number of Bank Offices for Non-MDIs	8.85	9.17	9.52	9.93	10.43	10.86	11.37	11.82	12.24
Average Number of Bank Offices for MDIs	7.26	7.42	7.52	8.73	8.64	8.71	8.63	9.03	9.27

Year	2010	2011	2012	2013	2014	2015	2016	2017
Total Number of Census Tracts	73,803	73,803	73,803	73,749	73,880	73,881	73,873	73,873
Number of Census Tracts with Bank Offices	36,851	36,724	36,395	35,936	35,862	35,404	35,062	34,752
Percentage of Census Tracts with Bank Offices	50%	50%	49%	49%	49%	48%	47%	47%
Total Number of Bank Offices	98,519	98,193	97,340	96,339	94,725	93,272	91,834	89,857
Number of Bank Offices, Fail to Find a Census Tract	98	12	92	7	18	23	42	42
Rate of Failure Match (%)	0.099	0.012	0.095	0.007	0.019	0.025	0.046	0.047
Total Number of Banks	7,821	7,523	7,255	6,950	6,669	6,358	6,068	5,797
Number of Census Tracts with MDI Offices	1,345	1,180	1,139	1,136	1,175	1,131	1,122	1,056

Percentage of Census Tracts with MDI Offices	2%	2%	2%	2%	2%	2%	2%	1%
Number of Census Tracts with MDI HQs	169	163	159	153	155	141	141	140
Number of MDI HQs	197	187	181	174	175	164	157	155
Number of MDI Offices	1,957	1,873	1,780	1,793	1,784	1,738	1,646	1,602
Number of Census Tracts with Both MDIs & Non-MDIs HQs	27	26	25	25	22	22	19	19
--Number of Non-MDIs HQs within These Census Tracts	76	64	63	46	44	51	49	44
--Number of MDIs HQs within These Census Tracts	38	35	33	32	31	34	25	24
Average Number of Bank Offices for Non-MDIs	12.67	13.13	13.51	13.95	14.31	14.78	15.26	15.64
Average Number of Bank Offices for MDIs	9.93	10.02	9.83	10.30	10.19	10.60	10.48	10.34

Table 1.2 Changing Number of MDIs Over Time Based on FDIC Information

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009
MDIs	164	166	166	172	189	194	211	215	207
Entering		6	9	11	24	17	20	12	5
Exiting		4	9	5	7	12	3	8	13
Acquired when Exiting		1	0	0	0	0	1	2	7

Year	2010	2011	2012	2013	2014	2015	2016	2017
MDIs	197	187	181	174	175	164	157	155
Entering	8	6	1	6	6	0	0	1
Exiting	18	16	7	13	5	11	7	3
Acquired when Exiting	11	7	2	1	2	3	1	1

Table 1.3 Sample of Banks and Bank Offices and MDIs and MDI Offices

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009
Total Number of Census Tracts	29,816	29,937	30,315	30,531	30,828	31,102	31,339	31,329	30,344
Number of Census Tracts with MDI Offices	757	765	769	884	964	981	1,050	1,044	1,060
Number of Census Tracts with MDI HQs	128	128	126	129	135	137	147	155	149
Low Income MDI HQ Census Tracts	32	31	32	32	32	32	32	28	28
Moderate Income MDI HQ Census Tracts	33	35	43	45	45	45	48	49	43
Low Income MDI Office Census Tracts	78	79	90	91	94	96	97	106	92
Moderate Income MDI Office Census Tracts	170	172	240	271	286	285	308	325	291
Percentage of Low Income MDI HQ Census Tracts	25%	24%	25%	25%	24%	23%	22%	18%	19%
Percentage of Moderate Income MDI HQ Census Tracts	26%	27%	34%	35%	33%	33%	33%	32%	29%
Percentage of Low Income MDI Office Census Tracts	10%	10%	12%	10%	10%	10%	9%	10%	9%
Percentage of Moderate Income MDI Office Census Tracts	22%	22%	31%	31%	30%	29%	29%	31%	27%
Number of MDI HQs	139	140	139	144	152	155	169	183	173
Number of MDI Offices	1,008	1,040	1,062	1,291	1,397	1,445	1,549	1,666	1,572
Number of MDI HQs in Low Income Census Tracts	38	37	41	40	38	37	39	38	39
Number of MDI HQs in Moderate Income Census Tracts	34	36	45	48	48	48	51	52	46
Number of MDI Offices in Low Income Census Tracts	116	120	126	132	130	132	137	155	164
Number of MDI Offices in Moderate Income Census Tracts	207	220	314	375	392	396	434	512	408
Percentage of Low Income MDI HQs	27%	26%	29%	28%	25%	24%	23%	21%	23%
Percentage of Moderate Income MDI HQs	24%	26%	32%	33%	32%	31%	30%	28%	27%
Percentage of Low Income MDI Offices	12%	12%	12%	10%	9%	9%	9%	9%	10%
Percentage of Moderate Income MDI Offices	21%	21%	30%	29%	28%	27%	28%	31%	26%
Number of Census Tracts with Both MDIs & Non-MDIs HQ	40	38	37	36	35	34	38	28	29
--Number of Non-MDIs within These Census Tracts	105	94	86	85	80	84	89	85	83
--Number of MDIs within These Census Tracts	49	46	44	41	40	38	45	32	37

Year	2010	2011	2012	2013	2014	2015	2016	2017
Total Number of Census Tracts	36,374	36,352	36,051	35,597	35,506	35,065	34,787	34,493

Number of Census Tracts with MDI Offices	1,302	1,156	1,117	1,111	1,130	1,088	1,084	1,017
Number of Census Tracts with MDI HQs	165	162	158	151	153	139	137	137
Low Income MDI HQ Census Tracts	21	18	31	29	24	26	26	22
Moderate Income MDI HQ Census Tracts	30	28	37	40	46	38	35	36
Low Income MDI Office Census Tracts	95	93	144	138	140	127	118	113
Moderate Income MDI Office Census Tracts	267	246	290	307	315	279	278	265
Percentage of Low Income MDI HQ Census Tracts	13%	11%	20%	19%	16%	19%	19%	16%
Percentage of Moderate Income MDI HQ Census Tracts	18%	17%	23%	26%	30%	27%	26%	26%
Percentage of Low Income MDI Office Census Tracts	7%	8%	13%	12%	12%	12%	11%	11%
Percentage of Moderate Income MDI Office Census Tracts	21%	21%	26%	28%	28%	26%	26%	26%
Number of MDI HQs	190	184	178	170	171	160	153	152
Number of MDI Offices	1,833	1,778	1,679	1,676	1,701	1,660	1,573	1,527
Number of MDI HQs in Low Income Census Tracts	26	24	38	37	26	29	32	26
Number of MDI HQs in Moderate Income Census Tracts	33	32	41	43	50	44	38	41
Number of MDI Offices in Low Income Census Tracts	137	136	232	230	220	214	196	185
Number of MDI Offices in Moderate Income Census Tracts	353	357	406	421	447	401	391	386
Percentage of Low Income MDI HQs	14%	13%	21%	22%	15%	18%	21%	17%
Percentage of Moderate Income MDI HQs	17%	17%	23%	25%	29%	28%	25%	27%
Percentage of Low Income MDI Offices	7%	8%	14%	14%	13%	13%	12%	12%
Percentage of Moderate Income MDI Offices	19%	20%	24%	25%	26%	24%	25%	25%
Number of Census Tracts with Both MDIs & Non-MDIs HQ	27	26	25	25	22	22	19	19
--Number of Non-MDIs within These Census Tracts	76	64	63	46	44	51	49	44
--Number of MDIs within These Census Tracts	38	35	33	32	31	34	25	24

Table 1.4 Summary Statistics

The sample includes only community banks. All bank variables are winsorized at the 1- and 99-percentile levels.

Variables	All Community Banks			Non-MDIs			MDIs			t-test of the mean (non-MDI) – mean (MDI)
	Mean	Median	Std.dev.	Mean	Median	Std.dev.	Mean	Median	Std.dev.	
ROA (%)	0.80	0.88	1.02	0.81	0.89	1.00	0.29	0.58	1.50	48.93
Z-score	74.31	54.72	71.28	74.95	55.32	71.51	44.54	30.18	51.27	40.98
Asset (000's $)	247,664	125,217	371,681	247,056	125,026	370,473	276,216	133,679	423,674	-7.54
Loan/TA	0.64	0.66	0.16	0.64	0.66	0.16	0.67	0.69	0.15	-19.68
Liquidity/TA	0.30	0.28	0.16	0.30	0.28	0.16	0.26	0.23	0.14	28.62
NPLs/TL	0.03	0.02	0.04	0.03	0.02	0.04	0.05	0.03	0.06	-49.92
Equity/TA	11.08	10.18	3.82	11.08	10.17	3.80	11.33	10.46	4.56	-6.35
Minority (%)	0.21	0.12	0.23	0.20	0.12	0.21	0.69	0.77	0.27	-220
Black (%)	0.08	0.02	0.16	0.08	0.01	0.15	0.20	0.05	0.30	-71.24
Asian (%)	0.02	0.00	0.06	0.02	0.00	0.04	0.19	0.06	0.25	-290
Hispanic (%)	0.11	0.03	0.18	0.11	0.03	0.17	0.25	0.14	0.27	-77.22
Native American (%)	0.01	0.00	0.05	0.01	0.00	0.04	0.04	0.00	0.12	-57.16
Poverty (%)	0.15	0.13	0.10	0.15	0.12	0.10	0.22	0.20	0.13	-72.67
Vacancy (%)	0.13	0.11	0.09	0.13	0.11	0.09	0.13	0.11	0.09	2.42
Highsch (%)	0.33	0.34	0.10	0.33	0.35	0.10	0.24	0.23	0.09	89.11
Median Household Income ($)	45,037	40,715	20,290	45,073	40,752	20,199	43,310	37,576	24,108	8.35
Over 65 (%)	0.16	0.16	0.06	0.16	0.16	0.06	0.14	0.13	0.07	35.49
MDI	0.02	0	0.14							
MDI_B	0.00	0	0.07							
MDI_A	0.01	0	0.10							
MDI_H	0.00	0	0.06							
MDI_N	0.00	0	0.05							

Table 1.5 Correlation Matrix

	ROA	Ln Z-score	Log_TA	Loan/TA	Liquidity/TA	NPLs/TL	Equity/TA	Minority	Black	Asian	Hispanic	Native	Poverty	Vacancy	HighSch	Logincomeh	Over65	MDI	MDI_B	MDI_A	MDI_H
ROA	1																				
Ln Z-score	0.34*	1																			
Log_TA	0.07*	0.12*	1																		
Loan/TA	0.03*	-0.10*	0.22*	1																	
Liquidity/TA	0.05*	0.16*	-0.13*	-0.92*	1																
NPLs/TL	-0.35*	-0.35*	-0.03*	-0.03*	-0.02*	1															
Equity/TA	-0.05*	0.21*	-0.21*	-0.23*	0.15*	-0.08*	1														
Minority	-0.07*	-0.09*	0.20*	-0.04*	0.02*	0.07*	0.01*	1													
Black	-0.03*	-0.03*	0.10*	-0.02*	0.00	0.11*	0.00	0.66*	1												
Asian	-0.11*	-0.11*	0.16*	0.08*	-0.08*	0.00	0.03*	0.34*	-0.02*	1											
Hispanic	-0.01*	0.02*	0.14*	-0.06*	0.06*	-0.03*	0.02*	0.51*	-0.02*	0.05*	1										
Native	0.05*	0.03*	0.01*	-0.04*	0.04*	-0.03*	0.01*	0.20*	-0.05*	-0.02*	0.21*	1									
Poverty	0.03*	0.03*	0.10*	-0.05*	0.05*	0.06*	0.02*	0.51*	0.47*	0.02*	0.21*	0.13*	1								
Vacancy	0.03*	-0.01*	-0.10*	-0.10*	0.08*	0.06*	0.03*	0.09*	0.12*	-0.10*	0.06*	0.07*	0.25*	1							
HighSch	0.15*	0.17*	-0.29*	-0.13*	0.14*	0.00	0.03*	-0.25*	-0.03*	-0.36*	-0.11*	-0.01*	0.10*	0.11*	1						
Logincomeh	-0.11*	-0.10*	0.08*	0.07*	-0.0*	-0.00	-0.01*	-0.29*	-0.35*	0.12*	-0.05*	-0.06*	-0.77*	-0.27*	-0.37*	1					

Over 65	0.05*	0.05*	-0.09*	-0.04*	0.06*	-0.03*	0.01*	-0.29*	-0.16*	-0.10*	-0.13*	-0.04*	-0.11*	0.22*	0.15*	-0.08*	1				
MDI	-0.07*	-0.09*	0.02*	0.03*	-0.04*	0.07*	0.01*	0.31*	0.11*	0.40*	0.11*	0.08*	0.11*	-0.00	-0.13*	-0.03*	-0.05*	1			
MDI_B	-0.06*	-0.05*	-0.02*	-0.01*	0.00	0.11*	-0.03*	0.17*	0.25*	0.02*	-0.02*	-0.01*	0.10*	0.04*	-0.05*	-0.06*	-0.04*	0.47*	1		
MDI_A	-0.04*	-0.07*	0.03*	0.05*	-0.06*	0.01*	0.03*	0.22*	-0.00	0.55*	0.07*	0.00*	0.05*	-0.04*	-0.12*	0.01*	-0.03*	0.69*	-0.07*	1	
MDI_H	-0.04*	-0.03*	0.03*	0.01*	-0.01*	0.04*	0.01*	0.14*	-0.02*	0.01*	0.18*	0.07*	0.05*	0.01*	-0.06*	-0.01*	-0.02*	0.39*	-0.00*	-0.01*	1
MDI_N	0.00	-0.01*	-0.02*	-0.00	-0.00	0.01*	-0.01*	0.03*	-0.02*	0.00	-0.00*	0.17*	0.02*	0.01*	-0.00	-0.01*	-0.01*	0.35*	-0.00	-0.01*	-0.00

Note: * indicates statistically significant at the 1% level.

Table 1.6 Panel Regression Results for ROA: MDI vs. Non-MDI Performance Based on Bank Financial Variables, with Demographic Variables

This table presents results for panel regressions with a dummy variable included indicating whether a bank is an MDI (coded 1) or not (coded 0). The regressions use quarterly data from 2001 to 2017. The dependent variable is the return on total assets (ROA). All regressions include size (Log_TA), loans (Loan/TA), liquidity (Liquidity/TA), non-performing loans (NPLs/TL), and equity (Equity/TA) as control variables. All bank variables are winsorized at the 1- and 99-percentile levels. Time fixed effects are included, and robust standard errors are clustered at the bank level. *, **, and *** indicates statistically significant at the 10%, 5%, and 1% levels, respectively.

Dependent Variable is ROA	MDI	MDI	Black MDI	Black MDI	Asian MDI	Asian MDI	Hispanic MDI	Hispanic MDI	Native American MDI	Native American MDI
MDI_dummy	0.09	0.11	0.53**	0.56**	0.15	0.17	-0.20	-0.19	-0.16	-0.16
Log_TA	0.42***	0.43***	0.42***	0.43***	0.42***	0.43***	0.42***	0.42***	0.42***	0.42***
Loan/TA	2.57***	2.58***	2.58***	2.58***	2.58***	2.58***	2.58***	2.58***	2.58***	2.58***
Liquidity/TA	1.63***	1.62***	1.63***	1.63***	1.63***	1.63***	1.63***	1.63***	1.63***	1.63***
NPLs/TL	-9.14***	-9.12***	-9.14***	-9.13***	-9.13***	-9.12***	-9.13***	-9.12***	-9.14***	-9.13***
Equity/TA	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***
Demographic Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered S.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.10	0.12	0.10	0.11	0.10	0.12	0.10	0.12	0.10	0.11
Observations	452,885	452,885	452,885	452,885	452,885	452,885	452,885	452,885	452,885	452,885

Table 1.7 Panel Regression Results for Risk: MDI vs. Non-MDI Performance Based on Bank Financial Variables, with Demographic Variables

This table presents results for panel regressions with a dummy variable included indicating whether a bank is an MDI (coded 1) or not (coded 0). The regressions use quarterly data from 2001 to 2017. The dependent variable is Ln Z-score. All regressions include size (Log_TA), loans (Loan/TA), liquidity (Liquidity/TA), non-performing loans (NPLs/TL), and equity (Equity/TA) as control variables. All bank variables are winsorized at the 1- and 99-percentile level. Time fixed effects are included, and robust standard errors are clustered at the bank level. *, **, and *** indicates statistically significant at the 10%, 5%, and 1% levels, respectively.

Dependent Variable is Ln Z-score	MDI	MDI	Black MDI	Black MDI	Asian MDI	Asian MDI	Hispanic MDI	Hispanic MDI	Native American MDI	Native American MDI
MDI_dummy	-0.06	-0.02	0.16	0.19	-0.23*	-0.20	0.08	0.09	0.25	0.24
Log_TA	0.38***	0.39***	0.38***	0.38***	0.38***	0.38***	0.38***	0.38***	0.38***	0.38***
Loan/TA	1.17***	1.16***	1.17***	1.17***	1.17***	1.17***	1.17***	1.16***	1.17***	1.17***
Liquidity/TA	1.10***	1.10***	1.10***	1.10***	1.10***	1.10***	1.10***	1.10***	1.10***	1.10***
NPLs/TL	-5.26***	-5.25***	-5.26***	-5.26***	-5.26***	-5.26***	-5.26***	-5.26***	-5.26***	-5.26***
Equity/TA	0.06***	0.06***	0.06***	0.06***	0.06***	0.06***	0.06***	0.06***	0.06***	0.06***
Demographic Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered S.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.18	0.21	0.18	0.20	0.18	0.21	0.18	0.21	0.18	0.20
Observations	451,576	451,576	451,576	451,576	451,576	451,576	451,576	451,576	451,576	451,576

Table 1.8 Results for ROA: MDI vs. Non-MDI Performance Based on Bank Financial Variables, with Demographic Variables (same census tract)

This table presents results for panel regressions with a dummy variable included indicating whether a bank is an MDI (coded 1) or not (coded 0). The regressions use quarterly data from 2001 to 2017. The dependent variable is the return on total assets (ROA). All regressions include size (Log_TA), loans (Loan/TA), liquidity (Liquidity/TA), non-performing loans (NPLs/TL), and equity (Equity/TA) as control variables. All bank variables are winsorized at the 1- and 99-percentile level. Time fixed effects are included, and robust standard errors are clustered at the bank level. *, **, and *** indicates statistically significant at the 10%, 5%, and 1% levels, respectively.

Dependent Variable is ROA	MDI	MDI	Black MDI	Black MDI	Asian MDI	Asian MDI	Hispanic MDI	Hispanic MDI	Native American MDI	Native American MDI
MDI_dummy	0.37	0.32	0.75	0.77	0.14	0.17	0.83*	0.68	0.27	0.28
Log_TA	0.45***	0.52***	0.87***	0.84***	0.38**	0.42**	0.25	0.35**	0.24	0.54
Loan/TA	1.29	1.22	-1.80	-1.80	1.06	1.38	-0.37	-0.27	6.84***	7.97***
Liquidity/TA	-0.97	-1.14	-2.44	-2.13	-0.65	-0.59	-2.20	-2.74	5.21***	5.16***
NPLs/TL	-14.37***	-14.34***	-9.02***	-9.28***	-21.37***	-21.12***	-15.44***	-14.43***	-4.90	-3.36
Equity/TA	-0.06*	-0.06*	0.04	0.04	-0.09***	-0.09***	-0.14***	-0.13***	-0.02	-0.01
Demographic Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered S.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.16	0.19	0.23	0.20	0.34	0.34	0.32	0.36	0.34	0.44
Observations	3,978	3,906	1,094	1,073	1,635	1,635	1,133	1,100	490	472

Table 1.9 Results for Risk: MDI vs. Non-MDI Performance Based on Bank Financial Variables, with Demographic Variables (same census tract)

This table presents results for panel regressions with a dummy variable included indicating whether a bank is an MDI (coded 1) or not (coded 0). The regressions use quarterly data from 2001 to 2017. The dependent variable is Ln Z-score. All regressions include size (Log_TA), loans (Loan/TA), liquidity (Liquidity/TA), non-performing loans (NPLs/TL), and equity (Equity/TA) as control variables. All bank variables are winsorized at the 1- and 99-percentile level. Time fixed effects are included, and robust standard errors are clustered at the bank level. *, **, and *** indicates statistically significant at the 10%, 5%, and 1% levels, respectively.

Dependent Variable is Ln Z-score	MDI	MDI	Black MDI	Black MDI	Asian MDI	Asian MDI	Hispanic MDI	Hispanic MDI	Native American MDI	Native American MDI
MDI_dummy	0.14	0.15	0.26	0.26	0.03	0.08	0.43***	0.34**	0.14	0.19
Log_TA	0.45***	0.43***	0.54**	0.61**	0.39***	0.42***	0.49***	0.52***	0.17	0.26
Loan/TA	0.94*	0.88**	1.42*	1.21	0.22	0.40	0.99	1.17	2.16***	3.43***
Liquidity/TA	-0.23	-0.40	0.69	0.26	-1.68***	-1.66***	-0.05	-0.14	3.02***	2.98***
NPLs/TL	-3.31***	-3.35***	-3.11***	-3.09**	-3.38**	-3.23**	-4.62***	-4.16***	-0.69	0.05
Equity/TA	0.05***	0.05***	0.07***	0.07***	0.05***	0.05***	0.04***	0.04***	0.08***	0.08***
Demographic Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered S.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.20	0.20	0.37	0.39	0.21	0.22	0.23	0.28	0.44	0.60
Observations	3,953	3,882	1,084	1,063	1,627	1,627	1,124	1,092	490	472

Table 1.10 Results for Robustness Checks

This table presents results for panel regressions with a dummy variable included indicating whether a bank is an MDI (coded 1) or not (coded 0). The regressions are based on quarterly data from 2001 to 2017. The dependent variables are return on total assets (ROA) and Ln Z-score. All regressions include size (Log_TA), loans (Loan/TA), liquidity (Liquidity/TA), non-performing loans (NPLs/TL), and equity (Equity/TA) as control variables. All bank variables are winsorized at the 1- and 99-percentile level. Time fixed effects are included, and robust standard errors are clustered at the bank level. *, **, and *** indicates statistically significant at the 10%, 5%, and 1% levels, respectively.

	ROA	ROA	ROA	ROA	ROA	ROA	Ln Z-score	Ln Z-score	Ln Z-score	Ln Z-score	Ln Z-score	Ln Z-score
	Entire	Entire	FRB	FRB	Same Zip	Same City	Entire	Entire	FRB	FRB	Same Zip	Same City
MDI_dummy	0.48	0.53*	0.44**	0.54**	0.34*	0.29	0.41**	0.41*	0.54**	0.61***	0.53**	0.23*
Log_TA	0.84***	0.80***	0.62***	0.68***	0.63***	0.78***	0.56***	0.58***	0.63***	0.65***	0.55***	0.55***
Loan/TA	1.16	1.00	3.35*	3.16	5.42**	3.13***	0.33	0.52	0.89*	0.85*	1.78***	1.37***
Liquidity/TA	0.10	0.07	1.48	1.49	2.76	1.81***	-0.89	-0.83	-0.03	-0.17	1.07**	0.99***
NPLs/TL	-11.23***	-10.86***	-13.55***	-13.13***	-12.50***	-11.22***	-4.70***	-4.36**	-2.15*	-1.83	-3.59***	-3.84***
Equity/TA	-0.02	-0.01	-0.02	-0.02	-0.08***	-0.09***	0.05***	0.05***	0.06**	0.06***	0.05***	0.05***
Demographic Controls	No	Yes	No	Yes	No	No	No	Yes	No	Yes	No	No
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered S.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.07	0.07	0.12	0.13	0.25	0.14	0.17	0.16	0.10	0.10	0.18	0.20
Observations	2,216	2,174	3,416	3,345	6,407	53,396	2,206	2,164	3,380	3,310	6,344	52,935

Figure 1.3 Number of Minority Depository Institutions using both FDIC and FRB Lists

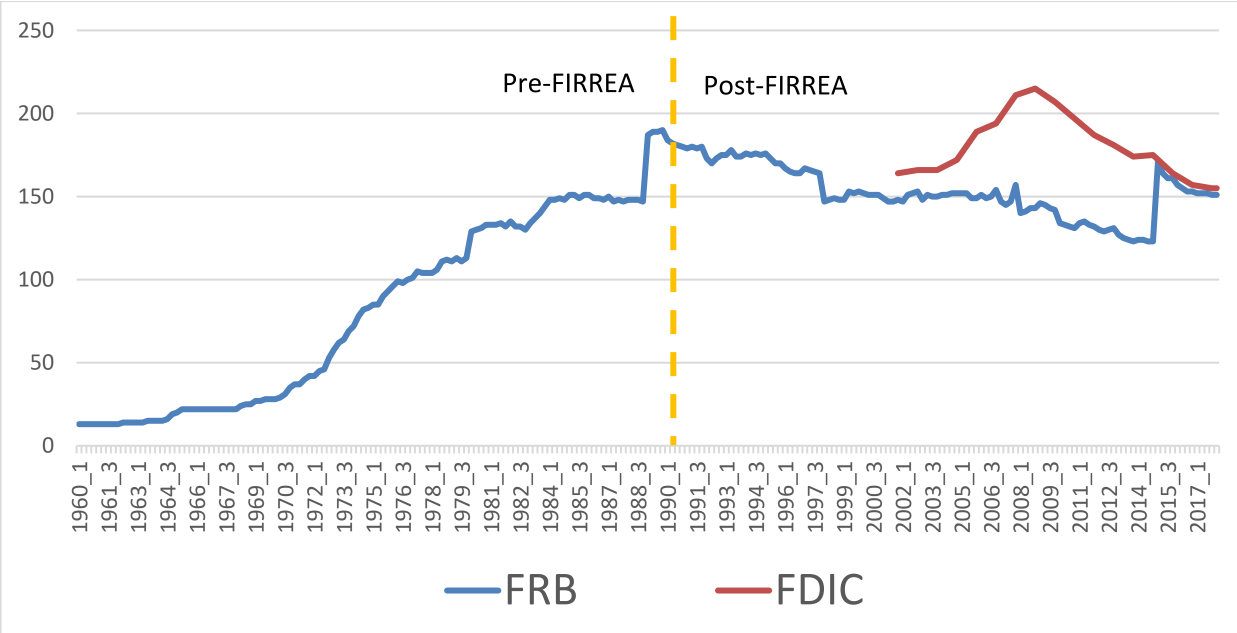


Table 1.11 Mean in Imbalance Measure Based on Demographic Variables with and without Matching (PSM)

Period: 2001Q1 – 2017Q4

	MDI	Black MDI	Asian MDI	Hispanic MDI	Native American MDI
Without Matching	0.99	1.00	1.00	1.00	1.00
With Matching - PSM	0.83	0.75	0.85	0.70	0.67

Table 1.12 Results for ROA using Propensity Score One-to-One No-Replacement Matching Sample

This table presents results for panel regressions with a dummy variable included indicating whether a bank is an MDI (coded 1) or not (coded 0) based on the sample generated by propensity score matching methodology. The analyses use annual data from 2001 to 2017. The dependent variable is the return on total assets (ROA). One to one without replacement propensity score matching method generates control groups for MDIs and different categories of MDIs. All regressions include controls as indicated in Table 9. Time fixed effects are included, and robust standard errors are clustered at the bank level. *, **, and *** denote an estimate that is statistically different from zero at the 10%, 5%, and 1% levels, respectively.

Dependent Variable is ROA	MDI	MDI	Black MDI	Black MDI	Asian MDI	Asian MDI	Hispanic MDI	Hispanic MDI	Native American MDI	Native American MDI
MDI_dummy	-0.06	-0.01	0.14	0.16	0.04	0.03	-0.23	-0.26*	0.15	0.16
Log_TA	0.36***	0.39***	0.37***	0.39***	0.40***	0.44***	0.27***	0.28***	0.11***	0.15***
Loan/TA	2.74***	2.70***	0.96	1.01	3.34***	3.31**	0.75	0.92	3.10***	3.10***
Liquidity/TA	2.20***	2.09***	1.35*	1.34*	1.89***	1.82***	0.84	0.92	3.38***	3.30***
NPLs/TL	-9.94***	-9.86***	-7.83***	-7.72***	-10.47***	-10.42***	-10.16***	-10.17***	-8.15***	-8.08***
Equity/TA	-0.05***	-0.05***	0.01	0.01	-0.06***	-0.05***	-0.06***	-0.06***	-0.01	-0.01
Demographic Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered S.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.27	0.29	0.26	0.27	0.34	0.36	0.32	0.34	0.22	0.25
Observations	18,833	18,833	4,142	4,142	9,179	9,179	2,956	2,956	2,358	2,358

Table 1.13 Results for Risk using Propensity Score One-to-One No-Replacement Matching Sample

This table presents results for panel regressions with a dummy variable included indicating whether a bank is an MDI (coded 1) or not (coded 0) based on the sample generated by propensity score matching methodology. The analyses use annual data from 2001 to 2017. The dependent variable is Ln Z-score. One to one without replacement propensity score matching method generates control groups for MDIs and different categories of MDIs. All regressions include controls as indicated in Table 9. Time fixed effects are included, and robust standard errors are clustered at the bank level. *, **, and *** denote an estimate that is statistically different from zero at the 10%, 5%, and 1% levels, respectively.

Dependent Variable is Ln Z-score	MDI	MDI	Black MDI	Black MDI	Asian MDI	Asian MDI	Hispanic MDI	Hispanic MDI	Native American MDI	Native American MDI
MDI_dummy	-0.06	-0.04	0.12	0.11	0.01	0.01	-0.20	-0.22	-0.11	-0.06
Log_TA	0.37***	0.39***	0.45***	0.46***	0.36***	0.38***	0.31***	0.31***	0.19***	0.22***
Loan/TA	1.02***	0.99***	1.46**	1.52***	1.45***	1.43***	0.59	0.69	0.66	0.58
Liquidity/TA	1.27***	1.23***	2.06***	2.15***	0.99***	0.94**	1.14	1.22*	1.69***	1.56***
NPLs/TL	-4.92***	-4.86***	-4.61***	-4.40***	-3.58***	-3.60***	-6.36***	-6.20***	-5.02***	-5.06***
Equity/TA	0.05***	0.05***	0.09***	0.08***	0.05***	0.05***	0.06***	0.06***	0.05***	0.06***
Demographic Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered S.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.22	0.24	0.32	0.33	0.25	0.25	0.24	0.26	0.27	0.27
Observations	18,750	18,750	4,130	4,130	9,119	9,119	2,946	2,946	2,356	2,356

Chapter 2

Effect of Local Bank Industry Competition on Loan Rates

2.1. Introduction

Supply and demand theory predicts an increase in supply will cause a decrease in price and vice versa, all else equal. Does this hold in the local credit market, and is the effect symmetric? This paper examines whether a change in supply in the local credit market results in a change in the price of credit. More specifically, it tests the effect of bank branch openings and closures within a census tract on the tract-level loan rates s (e.g., two-year used car loans with a maturity of 36 months). The Census Bureau defines census tracts as “small, relatively permanent statistical subdivisions of a county,” whose populations generally range between 1,200 and 8,000 people, with an optimum size of 4,000 people.

Though bank branches provide services that online banks cannot provide, banks are closing physical branches as more financial services move online. The decrease in bank branches is mainly attributable to fewer newly opened branches and more branch closures since 2009 (Figures 2.1 and 2.2). This drop imposes difficulties for older, poorer, and less-educated individuals who depend on local access to banking branches and may be unable to do their banking online. The situation is worse for communities of color, given the scarcity of census tracts with minority-owned banks (and Barth and Xu (2021)).

From 2001 to 2019, the number of bank branches in the United States increased by just 0.3 percent, from 86,069 to 86,367. The number peaked at 99,550 in 2009 but then fell every successive year. That is a loss of 13 percent of bank branches, or 13,183, across the country over the past decade. How do banks choose which branch to close? And how do they choose the location to open a new branch? Will they consider the competition in the local credit market? Do they prefer to have a new branch located in the same community with other bank branches, or without any bank branch? What are the effects of the bank branch closures and openings? Will these events decrease competition and result in a higher loan rate in the local credit market?

Despite the industry-wide shift, community banks (CBs), which focus on providing local services, have been able to buck the trend. CB branch closures occur much less often than do branch closures of non-CBs, with a monthly average of 51 nationwide for the former versus a monthly average of 90 for the period 2010–2019 (Figures 2.3 and 2.4). What are the effects of community bank branch closures and openings? Are the effects the same for non-community banks?

This paper will explore three research questions. First, what are the identifying factors that affect branch presence? Second, what's the role of community banks in their neighborhoods? Third, what is the relationship between changes in competition due to branch openings/closures and consumer loan rates? Does the relationship follow the supply and demand theory? In particular, will bank branch openings decrease the average loan rates in the local credit market, and will bank branch closures increase the average loan rates? In the next section, we present some relevant background literature and the motivation for the study.

2.2. Related literature

The banking literature explores many factors that may affect the presence of bank branches. Caskey (1994) uses data on bank branch locations in five cities from 1970 through 1989 and finds mixed results to the question of whether they are significantly underrepresented in low-income and minority urban communities. Using census tract-level data on banking services and neighborhood characteristics in Alameda County, California, Figlio and Genshlea (1999) explore the neighborhood racial composition and the level of banking services and concludes that minority neighborhoods, particularly black neighborhoods, are less likely to have access to banking services and are more likely to lose services during bank consolidations. Chang, Chaudhuri, and Jayaratne (1997) show that bank branches in New York City tend to be spatially clustered, with new branches opening near at least one other existing branch. A recent study by Ding and Reid (2020) presents evidence for a greater decline in bank branches in lower-income neighborhoods than in more affluent communities between 2009 and 2017.

The decrease in bank branches is mainly attributable to fewer newly opened branches and more branch closures since 2009 (Figures 2.1 and 2.2). This drop imposes difficulties for older, poorer, and less-educated individuals who depend on local access to banking branches and may be unable to do their banking online. The situation is worse for communities of color, given the scarcity of census tracts with minority-owned banks. Ergungor (2010) shows that the presence of a bank branch in a low-income community improves the access to mortgage loans. Nguyen (2019) demonstrates that branch closures lead to a persistent decline in local small-business lending.

Many studies have documented the importance of community banks to small businesses and the local economy, and their impact on the availability and cost of funding. Community banks tend to locate close to their customers (Petersen and Rajan,1994; Degryse and Ongena, 2005; Jagtiani, 2008). Kowalik (2014) finds support for the view that community banks have a special role in supporting small, local, opaque businesses. Hakenes, Hasan, Molyneux, and Xie (2015) find that small regional banks are more important funding providers than big interregional banks in regions with low access to finance.

Researchers have also explored how banks respond to changes in competition. Flechsig (1965) shows that short-term business loan rates correlate with differences in local and regional loan markets. Corvoisier and Gropp (2002) find that increased concentration in the euro area may result in less competitive pricing. Unite and Sullivan (2003) find evidence that foreign bank entry is associated with a reduction in interest rate spreads in some domestic banks in the Philippines. Degryse and Ongena (2005) look at bank loans in Belgium and conclude that loan rates decrease with the distance between the firm and the lending bank, and increase with the distance between the firm and competing banks. Looking at US data, Ergungor (2010) shows mortgage originations increase and interest spreads decline when there is a bank branch located in a low- or moderate-income neighborhood. Liebersohn (2017) finds that greater bank competition causes an increase in deposit rates and the quantity and composition of commercial loans. Azar, Raina, and Schmalz (2019) provide evidence at the county level that variation in bank concentration correlates with variation in deposit account interest rates, maintenance fees, and fee thresholds. In general, increased competition increases credit supply and decreases its price. Most of the literature pertaining to competition and credit is conducted at the bank level and focuses on either commercial loans or deposit rates. As far as we know, this paper is the first one to investigate the relation between competition and consumer loans at an even smaller market level, the census tract level.

This paper use bank branch openings and closures to measure whether competition in the local credit market increase or decrease. Based upon the theory of supply, we hypothesize that bank branch openings will increase the competition and further decrease the loan price, or loan rate. On the other hand, branch closures will decrease the competition and then increase the price of the loan in the local credit market.

Based on the empirical results, a negative relationship exists between bank branch openings and local loan rates which is consistent with the hypothesis, while the relationship between bank branch closures and local loan rates is insignificant, which is inconsistent with the hypothesis. In particular, an occurrence of a new bank branch in a census tract will cause the average loan rates decreased by 21.6 basis points. I obtain similar results for each of three types of auto loans, and use number variables instead of dummy variables[38]. The results generally indicate that there is a negative effect of branch openings on loan rates while no effect of branch closures. The paper distinguishes between CB and non-CB branches and finds that the negative effects of branch openings on loan rates are mainly driven by CBs. An occurrence of new community bank branch in a census tract will cause the average loan rates to decrease by 27.6 basis points.

The results also indicate that the likelihood that a new bank branch opens in a particular community increases when a small percentage of the total population is the minority, and when the community shows a relatively high household income and a high education level. In addition, there is evidence of herd behavior of branch openings, which reinforces Chang, Chaudhuri, and Jayaratne (1997), who conclude that bank branches in New York City tend to be spatially clustered.

This paper aims to help close the research gap and stimulate further research on the effects of local credit market competition. It offers three contributions to the literature: First, its census tract-level dataset complements other studies that seek to identify factors that affect branch presence. Second, it sheds more light on the role of community banks in their neighborhoods. Third, it tackles the overlooked area of consumer loans and tests the relationship between changes in competition and consumer loan rates, based on nationwide data at the census tract level.

The remainder of the paper proceeds as follows. Section 2.3 describes the data used in the empirical analysis. Section 2.4 introduces the models and associated variables used in assessing the herd behavior of branch openings as well as the effects of openings and closures on local loan

[38] In the robustness test, the variables used are number of bank branch openings and number of closures instead of dummy variables of openings and closures.

rates. Section 2.5 discusses the empirical results. Section 2.6 presents the results of robustness tests. Finally, the conclusion in Section 2.7 provides suggestions for future research.

2.3. Data

The data for estimating the empirical specifications come primarily from four sources: the US Census Bureau, the Federal Financial Institutions Examination Council (FFIEC), the Federal Deposit Insurance Corporation (FDIC), and RateWatch. The US Census Bureau provides the boundary shapefiles of census tracts, which are used here to map demographic data and bank branches in their corresponding census tracts. The shapefiles are available from 1990, 2000, and then for every year from 2010 to 2019. The tract boundaries are occasionally "split," but as noted, they are designated with the expectation of relative permanence "so that statistical comparisons can be made from census to census." To ensure the accuracy of bank locations within census tracts for this paper, office addresses in a given year are matched with the latest available file for the same year. For example, addresses from 2001 to 2009 are all matched to those in the 2000 shapefile; addresses from 2010 to 2019 are matched with each corresponding year's shapefile.

Demographic data are available from the 2000 census and the 2009−2018 five-year estimates from the American Community Survey (ACS) provided by the Census Bureau. Linear interpolation allows data to be obtained for the years 2001 to 2008. Additionally, the FFIEC Census Report is useful for determining whether a census tract is a low- or moderate-income tract.[39]

Bank branch data are obtained from the FDIC's Summary of Deposits (SOD) dataset, an annual collection of information on the branches of all insured institutions, as of June 30 of each year, and available from 1994 to 2019. The SOD data provide the latitude and longitude of each

[39] See www.ffiec.gov/censusapp.htm, accessed August 31, 2020.

branch and help confirm the accuracy of the match of a branch's location to the latitude and longitude. The coordinates are used to map the branches in census tracts by using ArcGIS. This step produces census tract numbers for each bank branch. Based on the census tract number, we know in what county the branch is located. If the county code from the census tract number differs from that in the SOD, the branch address can be geocoded again by using a Google API, and the branch can then be mapped in the census tract using the new latitude and longitude.

The Events & Changes dataset (E&C), which is publicly available from the FDIC, offers information on bank branch closures and openings starting from August 1999. The month of branch closures and openings is identified by the effective date. To find the census tracts where branch closures and openings occur, E&C data are matched with SOD data according to certificate numbers and branch numbers. The FDIC also provides a list of community banks to help flag them. Quarterly Standard Industry Reports (SDIs) provide bank balance sheets and income statement information. The SDIs are available from Q4 1992 to Q2 2020.

A series of local loan rates are built with data provided by RateWatch, which takes surveys of consumer loans from loan institutions across the country. It covers many types of loan institutions, among them, banks, credit unions, trust companies, loan production offices, and mortgage companies. Because of the limited address information, this paper focuses on banks. The data are available monthly since 2001. RateWatch dataset is widely used in literature involving depsit rates and loan rates.[40] Our analyses focus on several most commonly tracked loan products among U.S. branches, 36-month term on a two-year used auto loan with a loan size of $15,000, and 48-month term on a new auto with a loan size of $25,000. The key results hold when using

[40] Levine, et al (2020) collects certificate deposit rates from RateWatch dataset, and Auerbach, et al (2020) collects loan rates from the dataset.

new auto loan with different features, such as 60-month term on a new auto loan with a loan size of $25,000, in the robustness test. In this paper, we construct tract-level series of loan rates. For each loan rate series, we take the average rate across surveyed institutions in a census-month.

2.4. Models and Variables

Based on the question we will explore in this paper, three basic models are proposed. The first one is a panel logit regression that explores whether there is herd behavior in branch openings. The detailed regression is as followed,

$$\text{Openings}_{i,t} = \beta_0 + \beta_1 \text{Number of Branches}_{i,t-1} + \sum_{j=2}^{7} \beta_j \text{Demographic Variables}_{j,i,t-1} + \text{Year Fixed Effects} + \varepsilon_{i,t}, \quad (1)$$

where $\text{Openings}_{i,t}$ is a binary variable which indicates whether there are bank branch openings in census tract *i* at time *t*, 1 implies the presence of bank branch openings, 0 otherwise. The variable of interest is $\text{Number of Branches}_{i,t-1}$, i.e., the number of bank branches in census tract i at time t-1. A positive coefficient of $\text{Number of Branches}_{i,t-1}$ indicates new bank branches are more likely to be located in census tracts that contained more bank branches in the previous year. Year fixed effects are included to control for variables that are constant across the country but vary over time. The standard errors are robust to heteroskedasticity.

The specific demographic variables for each of the census tracts include the following: total population; the percentage of the minority population; the percentage of the population with high school as the highest education level; the percentage of the population over 65; log median household income; and vacancy rate. All the explanatory variables are lagged one period. It is generally expected that the census tracts with a higher population, a lower fraction of minority population, higher education level, higher median household income, and lower vacancy rate in the previous year are more likely to have more new bank branches located in them this year.

The second and the third models test whether bank branch openings and closures affect the average rates of specific loan products in a census tract. The specific models are as follows:

$$\text{Loan Rate}_{i,t} = \beta_0 + \sum_{j=1}^{12} \beta_j Openings_{i,t-j} + \sum_{j=13}^{18} \beta_j \text{Demographic Variables}_{j,i,t-1} + \lambda_i + \sigma_t + \varepsilon_{it}, \tag{2}$$

$$\text{Loan Rate}_{i,t} = \beta_0 + \sum_{j=1}^{12} \beta_j Closures_{i,t-j} + \sum_{j=13}^{18} \beta_j \text{Demographic Variables}_{j,i,t-1} + \lambda_i + \sigma_t + \varepsilon_{it}, \tag{3}$$

where Loan $\text{Rate}_{i,t}$ is the average loan rate on specific loan products of census tract i at time t. A tract-level series of rates is constructed on specific loan products with the most number of observations available (e.g., new auto loans with a maturity of 48 months). For each loan rate series, the average rate is taken across surveyed institutions in a census tract by month. The variables of interest are $Openings_{i,t-j}$ ($Closures_{i,t-j}$) which are dummy variables indicating whether there is a branch opening (closure) or not in census tract i at time t-j, j ranges from 1 to 12. A significant negative coefficient on these variables indicates that the average loan price of a census tract that has newly opened (closed) bank branches in the last j period is higher than those that do not have a branch opening (closure). The models also include λ_i, which are tract-level fixed effects that control for the time-invariant variables, and σ_t, which are time-fixed effects that control for macroeconomic variables that are constant across the country but vary over time.

In addition to focusing on the marginal effect of a single lagged period, the sum of the point estimates of coefficients is studied on all lagged interested variables. This helps to understand the cumulative effect of branch openings (closures) on the loan rates. The sum of coefficients with t-test value in parentheses is provided at the end of each regression.

One period lagged demographic variables, such as minority population percentages, poverty rates, and median household income are included in the model. Additionally, the time and census tract fixed effects are included as explanatory variables.

Table 2.1 reports the summary statistics of variables for analysis at the tract level. The sample includes more than 1 million tract-year observations from 2001 to 2019. On average, there are 0.7 bank branches, 0.03 branch openings, and 0.02 branch closures per census tract per year. The population in a census tract has a mean of 4,375. The average percentage of minority populations is 35.8; the average percentage of the population with a lower than high school degree is 16.5.

[Insert Table 2.1 About Here]

2.5. Empirical Results

The empirical analysis begins with the correlations between tract-level branch numbers and demographic features. Table 2.2 contains the correlation matrix for all the demographic variables and the number of bank branches for census tracts. The table indicates that in census tracts with bank branch openings, the openings are positively correlated with the number of bank branches, the number of population, and higher median household income; and negatively correlated with a higher percentage of the minority population, higher percentage of high school education (only) attainment, and percentage of the population over 65 in the previous year. All the correlations, moreover, are statistically significant at one percent level. The table includes correlations among the other variables, all of which have expected signs and are statistically significant.

[Insert Table 2.2 About Here]

Of the three main empirical models, Table 2.3 presents the results based on the panel logit regression assessing the likelihood of a new branch being in a census tract, based on the number of branches in this census tract last year. I first run the test for total openings of banks, then follow with individual tests for CB and non-CB branch openings. The findings in Column (1) indicate a herd behavior of bank branches. New bank branches are more likely to be located in census tracts that contained more bank branches in the previous year. On average, a unit increase in bank branches in a census tract in the previous year increase the probability of an occurrence of a new bank branch next year by 0.17 percentage points.[41] This result is consistent with Chang, Chaudhuri, and Jayaratne (1997). Column (1) also shows that the new bank branches are more likely to be located in census tracts with a higher population, a lower percentage of the minority population, higher education attainment level, and higher median household income in the previous year. Two variables, the population percentage over 65 and the vacancy rate are not significant. The signs of the coefficients of these variables show the same as we expect.

[Insert Table 2.3 About Here]

Turning to the individual results for CBs and non-CBs, Column (2) shows that new CB branches are more likely to be located in census tracts with more bank branches, higher population, a lower minority population percentage, lower education attainment level, higher vacancy rate, and higher median household income in the previous year. Column (3) shows that new non-CB branches are more likely to be located in census tracts with more bank branches, higher population, higher minority population percentage, higher education attainment level, lower vacancy rate, and higher median household income the previous year. All three regressions show a herd behavior of bank branches, i.e., new bank branches are more likely to operate in a census tract that already has

[41] Average marginal effects (AMEs) is reported. The AMEs for both CB and non-CBs are 0.09 percentage points.

at least one other branch. Of interest, some tract-level demographic variables have a different sign for new CB branches and new non-CB branches. Unlike what we expect, compared with CBs, non-CBs are more likely to open new branches in census tracts with a higher percentage of the minority population.

Table 2.4 presents the panel regression results for the effect of all bank branch openings as well as the two individual types (CBs and non-CBs) on tract-level loan rates. The regression results for all banks are in the first column, while those for CBs and non-CBs are in the second and the third columns. All the columns contain results, including one-period lagged demographic control variables, and both tract and time fixed effects. The significance level of the cumulative effects is reported at the end of each regression. The t-test is utilized to test the significance of the cumulative effect. In the first column, the Open (t-j) variables are dummy variables (equal to 1), if there is at least one newly opened bank branch in a census tract in the j periods ago, and 0 otherwise. For example, the Open (t-1) variable equals 1, if there is at least one newly opened bank branch in a census tract one month ago; the remaining variables are done in the same manner. In the other two columns, the Open (t-j) dummies equal 1, if there is at least one newly opened CB branch or non-CB branch in a census tract in the j periods ago, and 0 otherwise. The dependent variable in Panel A is the average loan rates on a 36-month term on a two-year used automobile of a census tract at time t. The dependent variable in Panel B is the average loan rates on a 48-month term on a new automobile.

[Insert Table 2.4 About Here]

The regressions include 12 lagged periods. On average, the effect starts to be significant immediately following a branch opening and becomes insignificant after six months. The coefficients of the first six months after the presence of the new branches are reported, as well as the cumulative effect of these periods. The loan rates used in Panel A are 36-month term on a two-year used auto and 48-month term loan rates on a new auto in Panel B.

The regression results from Column (1) in Panel A indicate that an occurrence of a new branch in a census tract will cause a decrease in the average loan rates in the ensuing six months. The cumulative effect for these six months is -21.6 basis points on average, and statistically significant at the 1% level. This number indicates that an occurrence of a new branch in a census tract will cause the average loan rates to decrease by 21.6 basis points over 6 months. Columns (2) and (3) deliver results when CB branch openings are distinguished from non-CB branch openings. Column (2) tells a similar story as Column (1) that an occurrence of a new community bank branch in a census tract will cause the average loan rates to decrease by 27.6 basis points over 6 months. Column (3) shows that the effect is not significant for the census tracts with new non-CB branches. Only when CBs change the competition do loan rates decrease. Panel B shows similar results with Panel A. This evidence provides a new view of the role of community banks in the neighborhoods in which they operate.

Table 2.5 shows the panel regression results for the effect of all bank branch closures as well as the two types of bank branch (CBs and non-CBs) closures. The dependent variables in Panel A and Panel B are the same as in Table 2.4. The regression results for all banks are in the first column, while those for CBs and non-CBs are in the second and the third columns. In the first column, the Close (t-j) variables are dummy variables (equal to 1) if there is at least one closed branch in a census tract in the j periods ago, and 0 otherwise. For example, the Close (t-7) variable equals 1 if there is at least one bank branch closed in a census tract seven months ago, the remaining variables are done in the same manner. In the other two columns, the Close (t-j) dummies equal 1 if there is at least one CB branch or non-CB branch closed in a census tract j periods ago, and 0 otherwise.

[Insert Table 2.5 About Here]

The regressions include 12 lagged periods; however, unlike what we expect, the effect is not significant over the period. The coefficients of the last six months are reported as well as the last six-month and three-month cumulative effects. The regression results from Column (1) indicate that an occurrence of a branch closure in a census tract does not affect the local loan rates. Columns (2) and (3) deliver results CB branch openings are distinguished from non-CB branch openings. Similar with the results in Column (1), results from Column (2) and (3) show that an occurrence of a CB or a non-CB branch closure in a census tract will not cause changes in the average loan rates. Either the cumulative effect for the last six months or the cumulative effect for the last quarter is not significant.

2.6. Robustness Tests

As a further check on our empirical results, we re-estimate our basic equation for CBs and non-CBs using a third loan product, 60-month term loans on a new auto. I also use the number of openings and closures, instead of dummy variables, as explanatory variables and run the same regression with Tables 2.4 and 2.5. All robustness tests show results similar to the main results. There is a negative and significant coefficient on the lagged open variables as well as the summation of these coefficients. This indicates that an occurrence of a new branch in a census tract will cause a decrease in the average loan rates in the following six months, cumulatively, or 20.6 basis points lower. As regards branch closures, the results indicate that closures of bank branches will not cause changes in the local loan rate.

[Insert Table 2.6 About Here]

2.7. Conclusions

The existing literature provides evidence that an increase in competition will increase credit supply and decrease its price. However, most of the competition-credit literature is conducted at the bank level and focuses on either commercial loans or deposit rates. As far as we know, this is the first study to investigate the relation between competition and consumer loans at the much smaller market level of a census tract. This paper examines the relationship between change in credit market competition and loan rates using nationwide census tract-level data and a large sample of branch-level loan rates nationwide.

The results show that an increase in competition of the banking industry, measured by the presence of a new bank branch in a census tract, will cause a decrease in the average loan rates in the ensuing six months. The effects are driven by the competition changes made by CBs. Unexpectedly, a decrease in competition, measured by the occurrence of a branch closure in a census tract, will not cause an increase in the average loan rates. In addition, evidence supports the herd behavior of new bank branches documented by previous literature, where new bank branches are more likely to be sited in census tracts with more bank branches in the previous year.

Figures

Figure 2.1 Monthly bank branch openings and closures (January 2001 to December 2019)

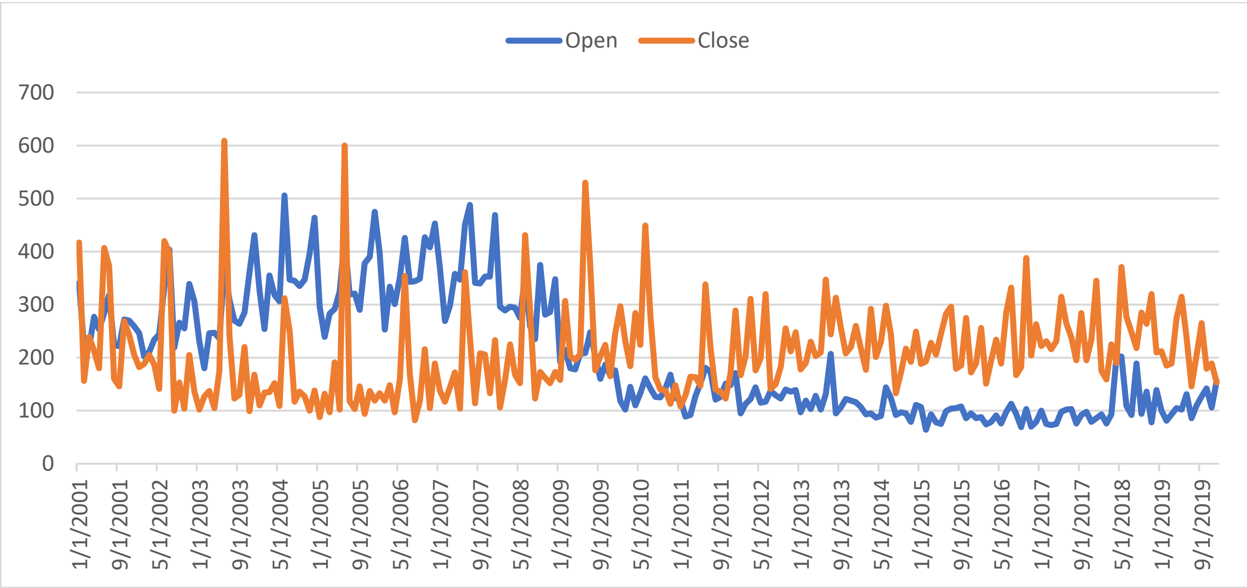


Figure 2.2 Monthly net change, all bank branches (January 2001 to December 2019)

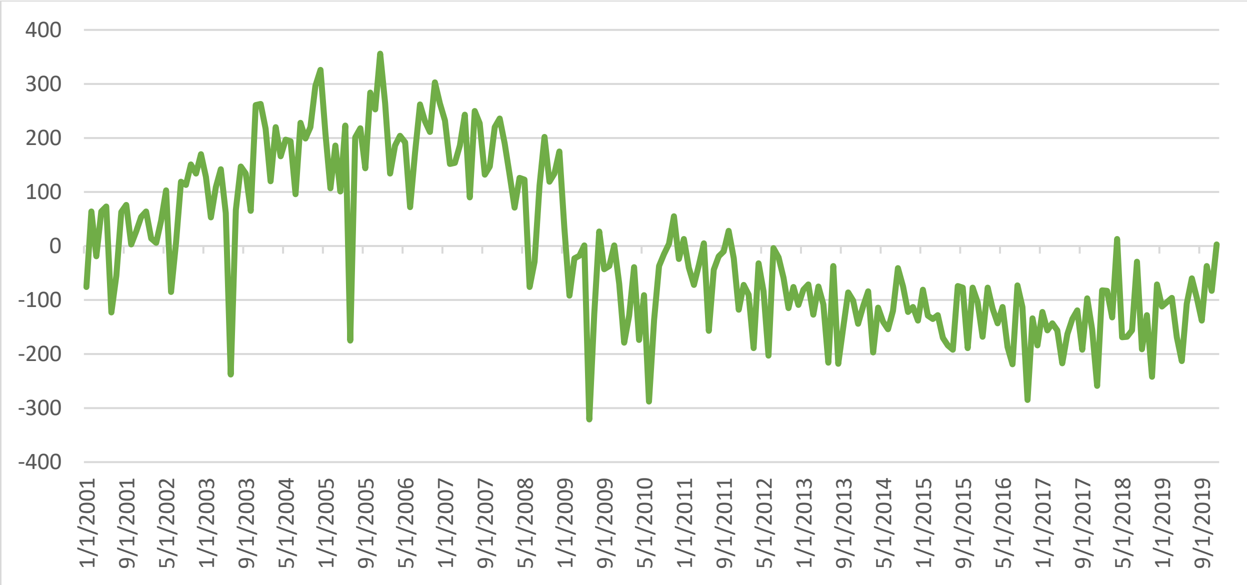


Figure 2.3 Monthly branch openings, CBs and non-CBs (January 2001 to December 2019)

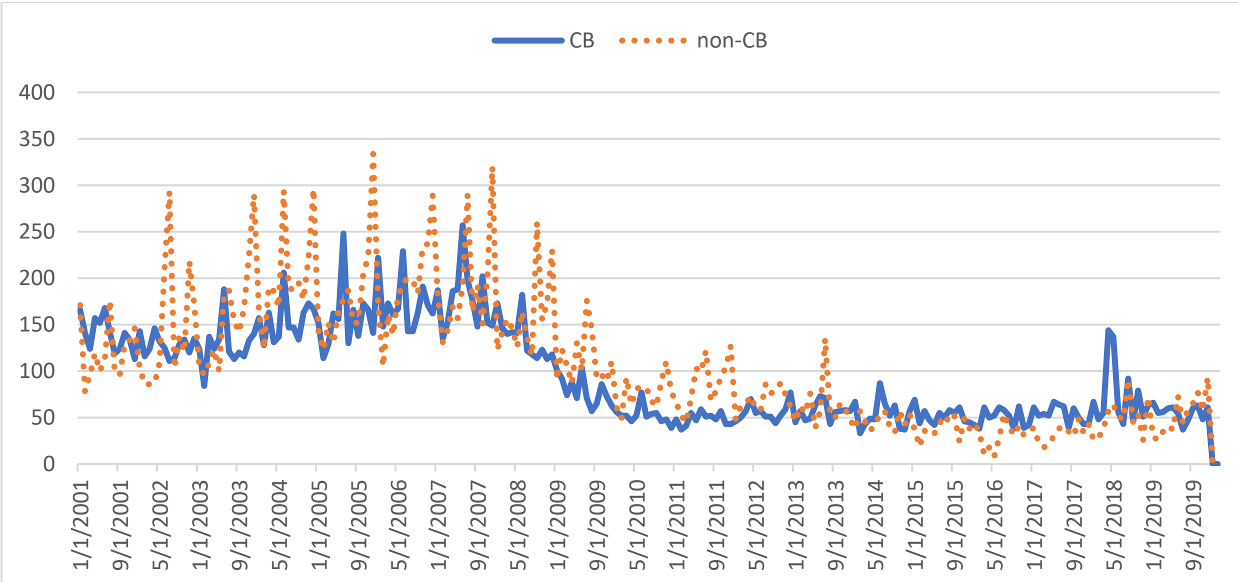
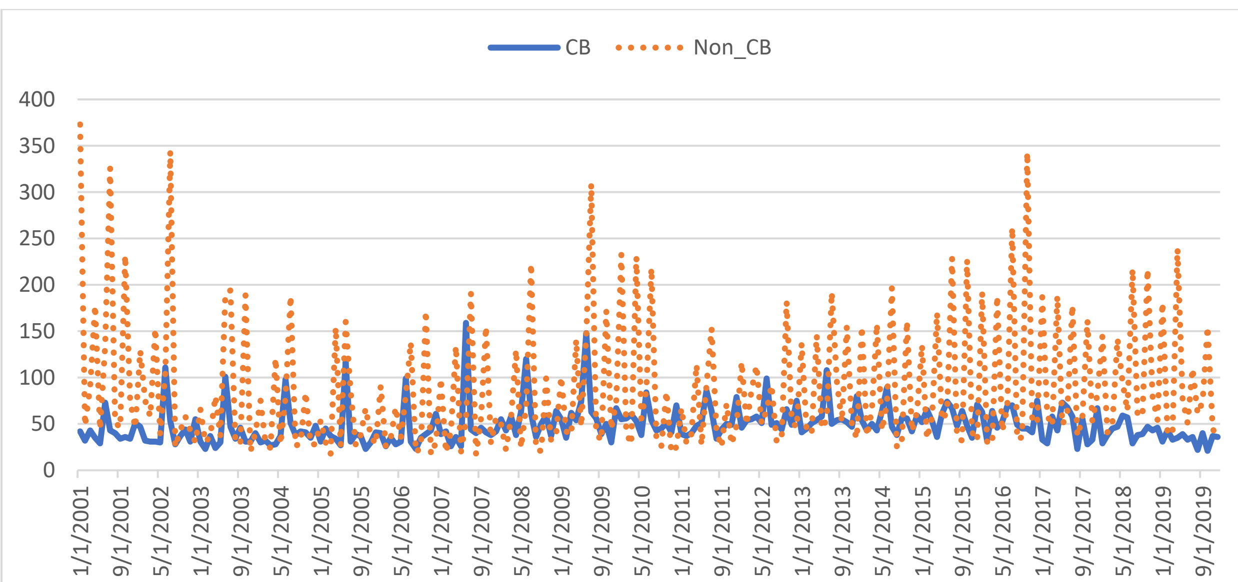


Figure 2.4 Monthly branch closures, CBs and non-CBs (January 2001 to December 2019)



Tables

Table 2.1 Summary Statistics of Census Tracts, 2001–2019

This table presents the summary statistics of selected branch information and demographic variables at the census tract level. The sample uses annual data from 2001 to 2019. The branch variables include the total number of branches [No. Branches], the number of newly opened branches [No. Openings], and the number of branch closures [No. Closures]. The demographic variables include total population; the percentage of the minority population [Minority Rate]; the percentage of the population with high school as the highest education level [Education Attainment]; the percentage of the population over 65 [Over 65 (%)]; vacancy rate; and log median household income [Ln Median Income].

	No. Observations	Mean	Std. Deviation	Min	Max
No. Branches	1,262,585	0.705	1.677	0	67
No. Openings	1,262,585	0.027	0.182	0	11
No. Closures	1,262,585	0.024	0.165	0	27
Total Population	1,262,585	4,375	2,287	0	70,271
Minority Rate	1,262,585	0.358	0.310	0	1
Education Attainment	1,262,585	0.165	0.125	0	1
Over 65 (%)	1,262,585	1.060	4.189	0	100
Vacancy rate	1,262,585	0.113	0.103	0	1
Ln Median Income	1,262,585	10.772	0.494	6.032	12.429

Table 2.2 Correlation Matrix

This table presents a correlation matrix for all the demographic variables and the number of bank branches per census tract. The demographic variables include total population [Population]; the percentage of the minority population [Minority]; the percentage of the population with high school as the highest education level [Education]; the percentage of the population over 65 [Over 65]; vacancy rate [Vacancy]; and log median household income [Median Income]. The sample uses annual data from 2001 to 2019. *, **, and *** indicates statistically significant at the 10%, 5%, and 1% levels, respectively.

	Openings	Branches	Population	Minority	Education	Over 65	Vacancy
Branches	0.048***						
Population	0.047***	0.161***					
Minority	-0.038***	-0.162***	-0.001				
Education	-0.031***	-0.084***	-0.070***	0.579***			
Over 65	-0.019***	-0.117***	-0.016***	-0.023***	-0.079***		
Vacancy	-0.019***	-0.013***	-0.248***	0.037***	0.127***	0.057***	
Median Income	-0.040***	0.023***	0.213***	-0.376***	-0.642***	0.092***	-0.322***

Table 2.3 Herd Behavior of Branch Openings

This table presents results for the panel logit regression. The dependent variable a binary variable which indicates whether there are newly opened bank branches or not, coding 1 if there are new branches, 0 otherwise. The t-values and p-values are based on robust standard errors. Standard errors are reported in the parentheses below the estimates. Symbols ***, **, and * denote statistical significance at 1%, 5%, and 10% levels, respectively. The sample uses annual data from 2001 to 2019. The demographic variables include total population; the percentage of the minority population [Minority]; the percentage of the population with high school as the highest education level [Education]; the percentage of the population over 65 [Population over 65]; vacancy rate [Vacancy]; and log median household income [Median Household Income]. All the explanatory variables are lagged one period.

	All Banks	Community Banks	Non-Community Banks
	(1)	(2)	(3)
Branch Number (t-1)	0.077***	0.081***	0.077***
	(0.004)	(0.005)	(0.005)
Total population (t-1)	0.059***	0.061***	0.059***
	(0.003)	(0.004)	(0.004)
Minority (t-1)	-0.612***	-1.777***	0.614***
	(0.039)	(0.051)	(0.049)
Education (t-1)	-0.197*	0.737***	-1.274***
	(0.112)	(0.142)	(0.146)
Population over 65 (t-1)	0.007	-0.003	0.010
	(0.004)	(0.006)	(0.007)
Vacancy (t-1)	-0.030	0.336***	-0.489***
	(0.094)	(0.111)	(0.131)
Median Household Income (t-1)	0.517***	0.226***	0.825***
	(0.026)	(0.033)	(0.034)
Observations	1,262,197	1,262,197	1,262,197

Table 2.4 Loan Rates and Branch Openings

Panel A. Loan rate of a 36-month term on a two-year used auto

This table presents results for panel regressions with dummy variables indicating whether there is (code 1) a newly opened bank branch in a census tract or not (coded 0) in previous months. For example, if there is a new bank branch three months ago, Open (t-3) is coded 1, otherwise 0. The regressions use monthly data from 2001 to 2019. The dependent variable is the average loan rate of a 36-month term on a two-year used automobile in a census tract. All regressions include tract-level demographic variables as control variables. All the explanatory variables are lagged one period. Tract and time fixed effects are included, and robust standard errors are clustered at the tract level. T statistics are reported in the parentheses below the estimates. *, **, and *** indicates statistically significant at the 10%, 5%, and 1% levels, respectively.

	All Banks	Community Banks	Non-Community Banks
	(1)	(2)	(3)
Open (t-1)	-0.035**	-0.054**	-0.010
	(-2.16)	(-2.47)	(-0.42)
Open (t-2)	-0.049***	-0.063***	-0.029
	(-3.04)	(-2.94)	(-1.22)
Open (t-3)	-0.031**	-0.044**	-0.016
	(-1.98)	(-2.13)	(-0.65)
Open (t-4)	-0.039**	-0.054**	-0.024
	(-2.33)	(-2.53)	(-0.93)
Open (t-5)	-0.022	-0.018	-0.031
	(-1.31)	(-0.83)	(-1.24)
Open (t-6)	-0.041**	-0.043**	-0.044*
	(-2.52)	(-2.05)	(-1.77)
Cumulative	-0.216***	-0.276***	-0.154
	(-2.68)	(-2.62)	(-1.24)
Demographic Controls	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes
Obs. Number	365,624	365,624	365,624
R-squares	0.59	0.59	0.59

Panel B. Loan rate of a 48-month term on a new auto

This table presents results for panel regressions with dummy variables indicating whether there is (code 1) a newly opened bank branch in a census tract or not (coded 0) in previous months. For example, if there is a new bank branch three months ago, Open (t-3) is coded 1, otherwise 0. The regressions use monthly data from 2001 to 2019. The dependent variable is the average loan rate of a 48-month term on a new automobile in a census tract. All regressions include tract-level demographic variables as control variables. Tract and time fixed effects are included, and robust standard errors are clustered at the tract level. T statistics are reported in the parentheses below the estimates. *, **, and *** indicates statistically significant at the 10%, 5%, and 1% levels, respectively.

	All Banks	Community Banks	Non-Community Banks
	(1)	(2)	(3)
Open (t-1)	-0.031**	-0.053***	-0.002
	(-2.30)	(-3.12)	(-0.09)
Open (t-2)	-0.043***	-0.052***	-0.031
	(-3.29)	(-3.15)	(-1.49)
Open (t-3)	-0.033**	-0.038**	-0.027
	(-2.46)	(-2.26)	(-1.27)
Open (t-4)	-0.034**	-0.042**	-0.027
	(-2.48)	(-2.40)	(-1.26)
Open (t-5)	-0.024*	-0.023	-0.030
	(-1.83)	(-1.33)	(-1.41)
Open (t-6)	-0.024*	-0.019	-0.037*
	(-1.80)	(-1.11)	(-1.77)
Cumulative	-0.189***	-0.226***	-0.154
	(-2.81)	(-2.69)	(-1.43)
Demographic Controls	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes
Obs. Number	412,129	412,129	412,129
R-squares	0.64	0.64	0.64

Table 2.5 Loan Rates and Branch Closures

Panel A. Loan rate of a 36-month term on a two-year used auto

This table presents results for panel regressions with dummy variables indicating whether there is (code 1) a bank branch closed in a census tract or not (coded 0) in previous months. For example, if there is a new bank branch seven months ago, Close (t-7) is coded 1, otherwise 0. The regressions use monthly data from 2001 to 2019. The dependent variable is the average loan rate of a 36-month term on a two-year used automobile in a census tract. All regressions include tract-level demographic variables as control variables. Tract and time fixed effects are included, and robust standard errors are clustered at the tract level. T statistics are reported in the parentheses below the estimates. *, **, and *** indicates statistical significant at the 10%, 5%, and 1% levels, respectively.

	All Banks	Community Banks	Non-Community Banks
	(1)	(2)	(3)
Close (t-7)	0.003	0.011	0.000
	(0.17)	(0.40)	(0.01)
Close (t-8)	0.003	0.012	0.000
	(0.16)	(0.45)	(0.01)
Close (t-9)	-0.006	0.031	-0.021
	(-0.41)	(1.15)	(-1.15)
Close (t-10)	-0.006	0.032	-0.022
	(-0.40)	(1.18)	(-1.17)
Close (t-11)	-0.018	0.009	-0.029
	(-1.12)	(0.33)	(-1.52)
Close (t-12)	-0.006	0.018	-0.016
	(-0.39)	(0.67)	(-0.89)
Cumulative last 6	-0.031	0.112	-0.088
	(-0.40)	(0.83)	(-0.96)
Cumulative last 3	-0.030	0.058	-0.067
	(-0.71)	(0.81)	(-1.32)
Demographic Controls	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes
Obs. Number	365,624	365,624	365,624
R-squares	0.59	0.59	0.59

Panel B. Loan rate of a 48-month term on a new auto

This table presents results for panel regressions with dummy variables indicating whether there is (code 1) a bank branch closed in a census tract or not (coded 0) in previous months. For example, if there is a new bank branch seven months ago, Close (t-7) is coded 1, otherwise 0. The regressions use monthly data from 2001 to 2019. The dependent variable is the average loan rate of a 48-month term on a new automobile in a census tract. All regressions include tract-level demographic variables as control variables. Tract and time fixed effects are included, and robust standard errors are clustered at the tract level. T statistics are reported in the parentheses below the estimates. *, **, and *** indicates statistical significant at the 10%, 5%, and 1% levels, respectively.

	All Banks	Community Banks	Non-Community Banks
	(1)	(2)	(3)
Close (t-7)	-0.005	-0.013	-0.001
	(-0.36)	(-0.62)	(-0.06)
Close (t-8)	-0.003	-0.010	0.001
	(-0.20)	(-0.44)	(0.05)
Close (t-9)	-0.004	0.012	-0.011
	(-0.30)	(0.53)	(-0.71)
Close (t-10)	-0.006	0.002	-0.010
	(-0.47)	(0.08)	(-0.61)
Close (t-11)	-0.010	0.010	-0.019
	(-0.77)	(0.45)	(-1.28)
Close (t-12)	0.003	0.031	-0.012
	(0.26)	(1.41)	(-0.82)
Cumulative last 6	-0.023	0.031	-0.052
	(-0.37)	(0.28)	(-0.69)
Cumulative last 3	-0.013	0.042	-0.041
	(-0.36)	(0.70)	(-0.99)
Demographic Controls	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes
Obs. Number	412,129	412,129	412,129
R-squares	0.64	0.64	0.64

Table 2.6 Robustness Check

Panel A. Openings and Loan Rate

This table presents results for panel regressions with dummy variables indicating whether there is (code 1) a newly opened bank branch in a census tract or not (coded 0) in previous months. For example, if there is a new bank branch three months ago, Open (t-3) is coded 1, otherwise 0. The regressions use monthly data from 2001 to 2019. The dependent variable is the average loan rate of a 60-month term on a new automobile in a census tract. All regressions include tract-level demographic variables as control variables. Tract and time fixed effects are included, and robust standard errors are clustered at the tract level. T statistics are reported in the parentheses below the estimates. *, **, and *** indicates statistical significant at the 10%, 5%, and 1% levels, respectively.

	All Banks	Community Banks	Non-Community Banks
	(1)	(2)	(3)
Open (t-1)	-0.029**	-0.046***	-0.006
	(-2.17)	(-2.72)	(-0.29)
Open (t-2)	-0.042***	-0.046***	-0.035*
	(-3.20)	(-2.81)	(-1.69)
Open (t-3)	-0.033**	-0.035**	-0.032
	(-2.54)	(-2.08)	(-1.53)
Open (t-4)	-0.033**	-0.041**	-0.027
	(-2.40)	(-2.35)	(-1.20)
Open (t-5)	-0.024*	-0.023	-0.029
	(-1.78)	(-1.35)	(-1.35)
Open (t-6)	-0.025*	-0.018	-0.041*
	(-1.86)	(-1.04)	(-1.93)
Cumulative	-0.186***	-0.209**	-0.169
	(-2.78)	(-2.50)	(-1.56)
Demographic Controls	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes
Obs. Number	409,989	409,989	409,989
R-squares	0.64	0.64	0.64

Panel B. Closures and Loan Rate

This table presents results for panel regressions with dummy variables indicating whether there is (code 1) a bank branch closed in a census tract or not (coded 0) in previous months. For example, if there is a new bank branch seven months ago, Close (t-7) is coded 1, otherwise 0. The regressions use monthly data from 2001 to 2019. The dependent variable is the average loan rate of a 60-month term on a new automobile in a census tract. All regressions include tract-level demographic variables as control variables. Tract and time fixed effects are included, and robust standard errors are clustered at the tract level. T statistics are reported in the parentheses below the estimates. *, **, and *** indicates statistical significant at the 10%, 5%, and 1% levels, respectively.

	All Banks	Community Banks	Non-Community Banks
	(1)	(2)	(3)
Close (t-7)	0.001	-0.006	0.004
	(0.05)	(-0.27)	(0.25)
Close (t-8)	0.002	0.000	0.002
	(0.113)	(0.02)	(0.14)
Close (t-9)	-0.002	0.020	-0.012
	(-0.18)	(0.92)	(-0.77)
Close (t-10)	-0.003	0.011	-0.010
	(-0.23)	(0.50)	(-0.58)
Close (t-11)	-0.008	0.017	-0.020
	(-0.65)	(0.76)	(-1.31)
Close (t-12)	0.002	0.031	-0.014
	(0.17)	(1.42)	(-0.93)
Cumulative last 6	-0.009	0.073	-0.048
	(-0.14)	(0.66)	(-0.64)
Cumulative last 3	-0.009	0.058	-0.043
	(-0.26)	(0.97)	(-1.04)
Demographic Controls	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes
Obs. Number	409,989	409,989	409,989
R-squares	0.64	0.64	0.64

Panel C. Number of Branch Openings and Loan Rate

This table presents results for panel regressions with variables indicating how many new bank branches in a census tract in previous months. The regressions use monthly data from 2001 to 2019. The dependent variable is the average loan rate of a 48-month term on a new auto in a census tract. All regressions include tract-level demographic variables as control variables. Tract and time fixed effects are included, and robust standard errors are clustered at the tract level. T statistics are reported in the parentheses below the estimates. *, **, and *** indicates statistical significant at the 10%, 5%, and 1% levels, respectively.

	All Banks	Community Banks	Non-Community Banks
	(1)	(2)	(3)
Open (t-1)	-0.029**	-0.052***	-0.001
	(-2.22)	(-3.15)	(-0.04)
Open (t-2)	-0.042***	-0.052***	-0.031
	(-3.34)	(-3.19)	(-1.58)
Open (t-3)	-0.033**	-0.037**	-0.029
	(-2.58)	(-2.26)	(-1.43)
Open (t-4)	-0.034***	-0.040**	-0.028
	(-2.63)	(-2.38)	(-1.35)
Open (t-5)	-0.025**	-0.023	-0.028
	(-1.98)	(-1.41)	(-1.41)
Open (t-6)	-0.027**	-0.019	-0.038*

	(-2.14)	(-1.18)	(-1.92)
Cumulative	-0.190***	-0.225***	-0.155
	(-2.95)	(-2.73)	(-1.49)
Demographic Controls	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes
Obs. Number	412,129	412,129	412,129
R-squares	0.64	0.64	0.64

Panel D. Number of Branch Closures and Loan Rate

This table presents results for panel regressions with variables indicating how many bank branches closed in a census tract in previous months. The regressions use monthly data from 2001 to 2019. The dependent variable is the average loan rate of a 48-month term on a new auto in a census tract. All regressions include tract-level demographic variables as control variables. Tract and time fixed effects are included, and robust standard errors are clustered at the tract level. T statistics are reported in the parentheses below the estimates. *, **, and *** indicates statistical significant at the 10%, 5%, and 1% levels, respectively.

	All Banks	Community Banks	Non-Community Banks
	(1)	(2)	(3)
Close (t-7)	-0.001	-0.010	0.003
	(-0.10)	(-0.48)	(0.20)
Close (t-8)	-0.00	-0.007	0.003
	(-0.00)	(-0.32)	(0.20)
Close (t-9)	-0.001	0.016	-0.008
	(-0.05)	(0.72)	(-0.54)
Close (t-10)	-0.001	0.005	-0.003
	(-0.05)	(0.25)	(-0.22)
Close (t-11)	-0.008	0.008	-0.015
	(-0.66)	(0.36)	(-1.02)
Close (t-12)	0.005	0.031	-0.007
	(0.38)	(1.49)	(-0.49)
Cumulative last 6	-0.006	0.044	-0.027
	(-0.10)	(0.40)	(-0.37)
Cumulative last 3	-0.004	0.045	-0.025
	(-0.12)	(0.76)	(-0.63)
Demographic Controls	Yes	Yes	Yes
Fixed Effects	Yes	Yes	Yes
Obs. Number	412,129	412,129	412,129
R-squares	0.64	0.64	0.64

Chapter 3

Large Plant Openings and Rural Housing Market Impacts

3.1. Introduction

The opening of a large plant bringing new jobs to a rural area is the quintessential example of a positive shock to local housing markets discussed in real estate valuation and appraisal courses across the United States. The converse is true, as well. That is, the closing of a large plant in a rural area is often given as an example of a negative shock to local housing markets. These are most often discussed in the context of an externality that results in "economic obsolescence" to be estimated in the cost approach to real estate valuation. The link between the new jobs offered at the plant and increased local housing values is assumed to be true, but very little research has been done to actually support this widely accepted notion.

The decades-long out-migration from rural areas to more urban areas is well documented and discussed *ad nauseum* across fields ranging from industry and politics to retirement planning and healthcare. The oft-overlooked story within the rural out-migration story is that regional variation exists across rural areas in terms of migration patterns. In short, some rural areas are (and have been) experiencing significant net in-migration, despite the fact that the overall statistics support the net out-migration interpretation of the data. People are indeed leaving rural areas *overall*, but they are moving in significant numbers to *certain* rural areas. Clearly, then, it is not productive to think of Americans as having a bias against any and all rural living. This is supported by data from a recent Gallup survey indicating that the largest percentage of Americans express a desire to live in a rural area.[42] The percentage of Americans who wish to live in a rural area is

[42] Newport, Frank. "Americans Big on Idea of Living in the Country." Gallup Organization, December 7, 2018. Available at https://news.gallup.com/poll/245249/americans-big-idea-living-country.aspx. It is worth noting that this survey was from well before the global pandemic. It is quite likely that the preference for non-metropolitan areas is even more pronounced now.

considerably larger than the percentage of Americans who currently live in a rural area, so there are a significant number of people who would move to a rural area if given the chance.

The link between new plant openings and regional economic performance is better understood than the link between plant openings and housing markets. In fact, investigations of this related topic typically ignore housing market impacts.[43] In pursuit of this associated improvement in regional economic performance, cities, counties, and states offer ever-increasing incentive packages aimed at attracting plant openings to a particular area. The cost to state and/or local taxpayers of these incentive packages is typically justified by pointing to the new jobs that will be created and all of the perceived expected benefits that flow from those new jobs.[44] As the competition among regions for new plant sitings increases in frequency and size, the question of proper and complete measurement of expected regional benefits becomes ever more pressing.

Even commercially available regional economic forecasting software has only recently realized the importance of accounting for housing price impacts from regionally important policy decisions.[45] For example, Regional Economic Models, Inc. (REMI) only made metropolitan statistical area (MSA) and state-specific housing price elasticities of population and employment available in their product beginning in 2017. It is laudable that the housing price elasticities are now included, but it is very telling that they were overlooked for so long.

This topic is particularly relevant right now, given the recent global pandemic. A number of industries have decided to "re-shore" certain parts of their manufacturing operations due to supply chain vulnerabilities highlighted by the pandemic. At the same time, the pandemic has

[43] Greenstone and Moretti (2003) is one exception, but it is a National Bureau of Economic Research Working Paper that has never moved on to publication. Parts of the analysis are included in Greenstone, Hornbeck, and Moretti (2010), but the attention paid to housing markets in the earlier version is not present in the later study.
[44] This issue is addressed in further detail in Greenstone, Hornbeck, and Moretti (2010).
[45] See the information from Regional Economic Models, Inc. (REMI) on house price inclusion.

made the traditionally safe decision to locate a new plant near other plants somewhat less safe. In a time when population density is not our friend, perhaps adding to industrial concentration in an already crowded region is not ideal. Add to this stew the fact that a larger percentage of Americans still express a desire to live in rural areas than actually live in rural areas, as discussed above. We may be entering a time when profit-maximizing siting decisions and locational preferences of potential workers align very closely. It seems reasonable, then, to investigate the impact that a new manufacturing plant might have on the largest asset class held by most homeowners: their house.

As a preview of the findings, the study highlights significant evidence of a positive impact on rural housing markets defined at the county level in counties where large plant openings take place. The positive house price effects are estimated relative to the trajectory of housing values in otherwise similar control counties within the same state as the subject county. The remainder of the study is organized as follows: Section 3.2 presents some relevant background literature and the motivation for the study, Section 3.3 describes the data used in the empirical analysis, and Section 3.4 details methodology used in the analysis, Section 3.5 provides a discussion of the results, and Section 3.6 concludes.

3.2. Literature Review and Motivation

This study touches on several somewhat inter-related strands of literature: the continued urbanization of America, the redistribution of the rural population and rural development economics, regional housing markets, manufacturing facilities' impact on housing values, and so-called "million-dollar plants." A quick glance at Table 3.1 provides confirmation of the story heard so often in the popular media: a focus on the top line number for percent of the U.S. population living in urban areas shows an increase in this percentage over time. There is substantial literature on the continued urbanization of America, wherein the population continues to cluster more and more densely in metropolitan and metropolitan-adjacent areas. Similar data from other developed and developing countries shows that this is a global phenomenon, not simply an American aberration.

[Insert Table 3.1 About Here]

A thorough review of this literature is well beyond the scope of the current study, but the literature can be broadly categorized into two main streams: studies looking at agglomeration and resultant spillover effects and studies looking at amenities offered by urban living. Recent examples from the agglomeration line of literature include Faggio, Silva, and Strange (2020), Dong (2020), Engelberg, Ozoguz, and Wang (2018), and Ellison, Glaeser, and Kerr (2010). Giuliano, Kang, and Yuan (2019) provide an excellent review of this line of literature as it relates to how cities grow and evolve over time. Melo, Graham, and Noland (2009), in conducting a meta-analysis of agglomeration economies, also provide a review of agglomeration literature.[46] Recent examples from the urban amenities line of literature include Couture and Handbury (2020) Carlino and Saiz (2019), Lee, Lee, and Shubho (2019), and Lee and Lin (2017). Helsey and Strange (2014) provide a formal model of coagglomeration, which leaves room for something other than agglomeration economies to drive industry siting decisions. There are also occasional studies that do not necessarily lend themselves to categorization as easily. These studies generally take urbanization as given and seek to explain some facet of its existence. For example, Mantegazzi, McCann, and Venhorst (2019) use areas of language discontinuity to explain how spillovers in agglomeration economies are shaped and how far their influence can spread.

[46] See also Duranton and Puga (2004), Head and Mayer (2004), and Rosenthal and Strange (2004) for additional, slightly more dated, reviews.

Closely related to the literature on increasing urbanization is the line of studies investigating firms' choices regarding location. Again, this is a long and well-developed line of literature, and full coverage of it is beyond the scope of the current research endeavor. One stream of this literature looks at how firms choose to locate or re-locate. Examples would include the managerial culture argument in Doeringer, Evans-Klock, and Terkla (2004), the cost of worker training argument in Almazan, De Motta, and Titman (2007), or the agglomeration economies in rural areas argument of Artz, Kim, and Orazem (2016). Another stream of this literature analyzes firm selection as firm survival, meaning firms in more competitive agglomerative markets must compete harder to continue operations. This stream of literature builds largely on the work of Melitz (2003) regarding product differentiation and includes a number of possible explanations for the more intense competition observed in large, agglomerative markets: cost markups (Melitz and Ottaviano (2008), Holmes, Hsu, and Lee (2014)), entrepreneurial talent (Nocke (2006)), linkage utilization (Baldwin and Okubo (2006)), or product specialization (Holmes and Stevens (2014)). Combes et al. (2012) provide evidence on the relative importance of the firm selection argument and find the effect is swamped by agglomeration economies.

Barca, McCann, and Rodriguez-Pose (2012) discuss the choice between regional development policies that focus on "place-neutral" versus "place-based" approaches. The former set of development policies, they argue, has been the status quo for decades, and it is this set of policies that have resulted in much of the rural-to-urban migration. These policies explicitly encourage agglomeration and mobility of labor as the way to spread the wealth among people, not places. That is, the rural poor are expected to move to richer urban areas to avail themselves of better opportunities. Given how ubiquitous these policies have been over a long period of time, it is to be expected that the academic literature has focused primarily on causes and outcomes related to this set of policy prescriptions, as has been noted in the review of the literature thus far.

The second, and much newer, set of policies aim to disrupt that thinking by encouraging development that is more spatially diverse. Viewed in that light, the American (and global) trend toward increasing urbanization is an understandable welfare-maximizing decision in response to the incentives put in place by decades of agglomerative policy-making. Limited forays into place-based policies existed prior to this relatively recent focus on such policies. For example, enterprise zones would be considered a place-based development policy in the United States. Early research on these policies typically found little to no benefit for employment or property values. Wilder and Rubin (1996) provide a review of the very early studies of outcomes associated with this place-based policy. The interested reader is also referred to Boarnet and Bogart (1996), Dowall (1996), Engberg and Greenbaum (1999), Bondonio and Engberg (2000), Lambert and Coomes (2001), Greenbaum and Bondonio (2004), and Neumark and Kolko (2010). Greenbaum and Engberg (2004) and Bondonio and Greenbaum (2007) provide some later, more nuanced analysis revealing enterprise zones have positive effects on new business creation, but the new businesses hasten the closure and exit of incumbent firms. O'Keefe (2004) also provides evidence of positive, but short-lived benefits from enterprise zone designation. Komarek and Loveridge (2015) provide additional support for positive impacts from place-based policies in the form of employment growth among small firms, which comes at the expense of income growth overall.

Barca, McCann, and Rodriguez-Pose (2012) point to several governmental, quasi-governmental, and private reports published around the 2008-2010 timeframe that, on balance, provide support for a more place-based development strategy.[47] Perhaps not surprisingly, the

[47] The World Bank's 2009 report and Glaeser and Gottlieb (2008) both argue and present evidence in favor of traditional place-neutral development policies, although the World Bank report does allow that such policies may be

arrival of these institutional arguments mostly in favor of more place-based development policies was also accompanied by renewed academic interest in the set of place-based policies advocated for in the reports. Kline (2010) shows that it is possible for such place-based policies to leave at least some communities in a permanent "poverty trap." Busso, Gregory, and Kline (2013) find that targeted place-based policies achieve the desired outcome of spurring growth in the targeted area without inducing population and cost of living increases. Kline and Moretti (2014) find evidence of agglomeration economies in manufacturing related to the Tennessee Valley Authority that persists over a long period, but that is offset by manufacturing losses elsewhere. Gaubert (2018) finds that place-based policies will lead to worse societal (i.e. aggregate) outcomes.

Also intertwined with literature on people migrating to urban centers and firms choosing to cluster or build further away is distinct literature on rural development economics. While this has always been a secondary consideration of the literature on agglomeration, even stretching as far back as Marshall (1890), it is relatively rarely the focus of analysis. Lewis and Prescott (1972) are one notable exception. They offer a prescient warning about the continued focus on place-neutral development policies that favor the growth of urban areas. The concerns they highlight, particularly regarding the deleterious impacts on commercial activity in rural areas, are issues that still plague many of those same rural areas. Renkow and Hoover (2000) discuss possible causes of the shifting fortunes of urban versus rural areas in terms of population growth. Irwin et al. (2010) and Kilkenny (2010) provide reviews of the most relevant academic literature on this topic, as well as additional context for the relative dearth of academic focus on the rural areas that serve as feeders for agglomeration economies in the more traditional literature.

necessary to address specific issues. The remaining six reports, Barca (2009), two 2009 reports from the Organization for Economic Co-Operation and Development, and an additional report from Corporacion Andina de Fomento (a Latin American development bank), all argue for place-based development policies to supplement and/or supplant the place-neutral policies that have failed to produce desired results over a long timeframe.

Absent from the Kilkenny (2010) review are two additional studies directly related to rural development. Reeder and Robinson (1992), which is also not mentioned in the Irwin et al. (2010) review, analyze enterprise zones with an eye specifically toward their use as a rural development tool. They find that rural enterprise zones are equally as cost-effective as urban enterprise zones at creating jobs, which they argue support the designation of more rural enterprise zones. Henry, Barkley, and Bao (1997) present evidence on the spillover effects from urban cores to their rural neighbors, and suggest several place-based policies that rural areas can utilize to attract more of the growth that spills out of their nearest urban neighbors.

Recent rural development literature has celebrated the renewed interest in place-based development policies, which tend to favor rural areas more than place-neutral policies. Much of this research has been focused on recommended actions to enable rural areas to take full advantage of place-based policy-making. Examples of these policy-focused studies include Olfert and Partridge (2010), Castle, Wu, and Weber (2011), Partridge and Olfert (2011), and Mulligan, Partridge, and Carruthers (2012). Additional recent entrants in this line of literature look at firm entry and survival in rural areas (Renski (2008), Yu, Orazem, and Jolly (2011), and Renski and Wallace (2012)), population shifts among rural areas (Rickman and Rickman (2011)), and the housing bubble in rural areas (Rickman and Guettabi (2015)). Using data from Japan, Okubo and Tomiura (2012) find that policies aimed at encouraging development in rural areas end up attracting relatively less productive plants that re-locate away from urban areas with higher levels of competition.

This study is also very closely related to the literature dealing with the effects of manufacturing facilities on housing prices. Boyle and Kiel (2001), Jackson (2001), and Simons and Saginor (2006) review this literature in more detail. The Boyle and Kiel (2001) review are notable in that it also reviews the broader, but intimately related, topics of air pollution and water pollution. A significant portion of the literature dealing with manufacturing facilities and house prices focuses on local housing price impacts from proximity to power plants of various types. Proximity here is typically defined by concentric rings centered on the manufacturing facility being studied. The frequency with which power plants of some type are studied is likely due to these facilities being located near residential areas more often than other types of manufacturing plants. Studies typically find a negative impact on surrounding housing values from the power plant in question, as seen in Blomquist (1974), Davis (2011), and Hodge (2011).[48] Davis (2011), in particular, finds that areas near power plants also experience decreases in household income, educational attainment, and percentage of owner-occupied properties. The result with regard to household income echoes Hanna (2007), who also finds decreased house prices near the source of the pollution. Kahn (2009) finds that population growth is occurring at significantly higher levels in counties with less exposure to the highest-emitting type of power plants. Muehlenbachs, Spiller, and Timmins (2015) find a similar, and substantially larger, negative price impact related to proximity to shale oil extraction facilities for nearby homes reliant on groundwater. They report a positive price impact for homes receiving water from a municipal source, which they attribute to lease payments for shale oil access.

Another area of frequent focus within the literature on manufacturing impacts on housing markets is hazardous waste sites (so-called "Superfund" sites). This stream of literature is also covered in the three reviews mentioned above: Boyle and Kiel (2001), Jackson (2001), and Simons and Saginor (2006). The negative impact of Superfund sites on surrounding housing values has

[48] Interestingly, Gamble, Downing, and Sauerlender (1980), Nelson (1981), and Gamble and Downing (1982) fail to find any significant housing price impact related to proximity to nuclear power plants, including the plant located at Three Mile Island, even though they are studying the issue relatively soon after the Three Mile Island incident.

been documented in numerous studies over the years, dating back at least to Greenberg and Hughes (1992). The cleaning of those sites has also been a topic of considerable interest in the literature. Gayer, Hamilton, and Viscusi (2000) find that housing markets are able to correctly interpret new information about Superfund sites, as the classification of hazard level by the Environmental Protection Agency (EPA) has an impact on the depressed prices for homes nearest hazardous sites. Kiel and Williams (2007) add evidence in support of that notion, finding that not all Superfund sites have negative property value impacts and that not all impacts are ameliorated by site cleanup. Greenstone and Gallagher (2008) find that the positive housing market impacts from site cleanup do not outweigh the costs of such cleanup, which again points to the possibility of reduced housing market impacts for Superfund sites deemed less hazardous. Gamper-Rabindran and Timmins (2013) find that houses nearest a hazardous site are also the lowest-priced houses, on average. They show that these closest neighbors experience the largest percentage price increases from site cleanup, but the low starting values depress cleanup value estimates viewed at an aggregate level.

A small subset of this literature has developed that analyzes the closing of manufacturing facilities. Deng, Hernandez, and Xu (2020) find that the closing of two large coal-fired power plants is associated with significant increases in both nearby residential property prices and transaction volume in previously affected areas. Currie et al. (2015) study both openings and closings of toxic plants as related to housing values and health risks. Presumably, it is the health risks that are being capitalized into house prices, so it seems logical to extend the analysis to include such risks. They find that plant openings lead to a decline of approximately 11% of house price and about a 3% increase in the probability of low infant birth weight within one mile of the plant.

The last stream of literature related to the topic at hand is the set of studies dealing specifically with large plant openings (so-called "Million Dollar Plants," or MDP). Greenstone and Moretti (2003) find that the county that wins the bidding process for a new large manufacturing plant experiences a significantly higher growth rate for incomes and property values, with no identifiable deterioration of municipal-level finances versus the county that was the runner-up in the industrial siting decision-making process. Greenstone, Hornbeck, and Moretti (2010) find that MDPs increase the total factor productivity of existing plants in winning counties, and they also increase labor costs in those counties. In other words, incomes go up in winning counties at incumbent plants, too, rather than solely at the newly constructed plant. Bloom et al. (2019) find that new MDPs create learning spillovers in winning counties, which improves management practices at incumbent firms, perhaps explaining the increase in total factor productivity noted by Greenstone, Hornbeck, and Moretti (2010). Finally, Kim (2020) finds that the opening of an MDP leads to a significant increase in the debt-to-capital ratio of incumbent firms in the winning county, relative to the runner-up county. This is interpreted as evidence that the cost of financial distress for firms in winning counties has been reduced.

As can be seen throughout the literature review, the current study touches on a variety of strands of literature. The idea of siting a manufacturing facility in a rural area, presumably without the benefit of local agglomeration economies, cuts against the bulk of the agglomeration economies literature. However, it dovetails with the emerging literature on the desirability of urban amenities nicely, as "urban" amenities likely are not the only amenities that workers value. In fact, if the Gallup poll cited above is to be believed, the largest percentage of Americans value rural amenities quite highly. This research also provides another look at the use of place-based development policies, as the municipal and state incentives provided to attract the plant opening are a type of place-based policy and one that is used with increasing frequency. Given that the literature is mixed at best on the question of impact from place-based development policies, an additional look at a specific policy is a welcome addition to the literature.

The current study also contributes to the rural development literature more broadly, in that it analyzes firm entry decision-making in a rural area and housing market impacts specific to a rural area, both topics of recent interest in that literature. The current study also advances the literature on manufacturing and housing markets. Most of this literature focuses on housing market impacts from older plants subject to older pollution regulations. In fact, the entire Superfund literature focuses on now-defunct plants that were most definitely subject to different emission standards. This study focuses on newer plants that would be built to meet modern pollution standards. If homeowners are indeed able to properly assess risk information about nearby plant locations (*a la* Gayer, Hamilton, and Viscusi (2000)), then we should expect to see different reactions in the property market to plants built to a different emission standard. Finally, this study adds to the growing literature on the special importance of Million Dollar Plants as an engine for local and regional growth.

A note here is warranted about differences from Greenstone and Moretti (2003). They report an increase in county property values of about 1.1% for the winning county versus the runner-up county. First, their study does not focus solely on plant sitings in rural areas. Second, of the 82 winning counties and 129 losing counties in their sample, they could only develop property value data for 92 of those 211 counties. Third, the property value data they used relied mainly on tax assessor data, which is subject to severe reporting issues. Data availability since their study has increased dramatically.

3.3. Data

Data on newly opened large plants are collected from AreaDevelopment.com. Their website claims that they are "considered the leading executive magazine covering corporate site selection and relocation." Indeed, their data on plant siting decisions have been used in prior published academic studies.[49] Unfortunately, the feature they formerly published on runners-up in corporate site selection tournaments is no longer a regularly reported item. Nonetheless, data on winning locations are still provided, and counties for comparison can be reliably identified as described below.

The housing data utilized in the study is obtained from Zillow. Specifically, the median list price at the county level and a monthly frequency are used as the housing price index for comparison across winner and comparison counties. While a more recognized housing price index would be preferable, the published housing price indices focus understandably on larger market areas. This has been a historical impediment to rural housing market research when something other than individual transaction data was needed. Fortunately, the wide coverage offered by Zillow will now begin to help fill in such lingering gaps in the literature.

Demographic variables are collected from the American Community Survey (ACS), which is available from the U.S. Census Bureau. The demographic data are collected at the county level include total households, total population, total housing units, age distribution, distribution of educational attainment, percent of the working-age population (i.e. population aged 16 or higher) in the labor force, median household income, median family income, per capita income, and percent of the population below poverty level in the past 12 months. The urban versus rural classification used to determine where rural plant openings are taking place also comes from the

[49] See Greenstone, Hornbeck, and Moretti (2010), for example.

U.S. Census Bureau.[50] In short, counties with less than 50 percent of the population living in rural areas are classified as mostly urban, counties with 50 to 99.9 percent of the population living in rural areas are classified as mostly rural, and counties with 100 percent of the population living in rural areas are classified as completely rural. In this study, mostly rural and completely rural as both classified as rural. This fits with similar definitions used in prior related literature, and it makes intuitive sense. For many types of plants under consideration, a minimum level of infrastructure would be necessary to allow a county to be considered at all. Table 3.2 provides a list of variable abbreviations and definitions, while Table 3.3 provides some selected summary statistics for the sample data.

[Insert Tables 3.2 and 3.3 Panel A About Here]

3.4. Methodology

3.4.1. Forecasting Framework

The Prophet is a time series forecasting library introduced by Facebook. Taylor and Letham (2018) use a decomposable times series model with three main model components: trend, seasonality, and holidays. In detail, these three features are combined in the following equation:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t.$$

Where $g(t)$ is the trend function which models nonperiodic changes in the value of the time series, $s(t)$ represents periodic changes, say seasonality, and $h(t)$ represents the effects of holidays that occur on potentially irregular schedules over one or more days. The error term ε_t represents any idiosyncratic changes that are not accommodated by the model. It works best with time series that have strong seasonal effects and several seasons of historical data. Prophet provides

[50] Further detail on the classification scheme used by the Census Bureau for categorizing rural versus urban counties is available at: https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural.html

some practical advantages, it can easily accommodate seasonality with multiple periods, and it also handles outliers well.

In this paper, the input variable is monthly MHP before the event month, we forecast monthly MHP from event month ($Time = 0$) until December 2019. The summary statistics of MHP by county are presented in Table 3.3 Panel A. The time periods for fifteen counties in our sample are available from January 2010 to December 2019, which result in 120 mouth MHPs for each of them. The other five counties have relatively shorter time periods.

3.4.2. Regression Model

In this paper, we are interested in the question asking whether the MHP will increase after a new large plant opening. We propose a regression model as follows to test the question:

$$MHP = \beta_0 + \beta_1 \times Real + \beta_2 \times Time + \beta_3 \times Real \times Time + \varepsilon,$$

where MHP is median housing price; $Real$ is whether the data is real MHP or predicted MHP, 1 indicates real, 0 otherwise; $Time$ is time relative to the month when the plant opened, thereafter event month. For example, 0 is event month, 1 is one month after the large plant opening, 2 is two months after the large plant opening, and so on.

When $Real = 1$, the regression model becomes

$$MHP = \beta_0 + \beta_1 + \beta_2 \times Time + \beta_3 \times Time + \varepsilon.$$

When $Real = 0$, the regression model becomes

$$MHP = \beta_0 + \beta_2 \times Time + \varepsilon.$$

Therefore, the null hypothesis of the regression is $H_0: \beta_3 = 0$. When the null hypothesis is rejected, it indicates that the MHP of a county changes after a new large plant opening, *vice versa*. Further, a positive β_3 indicates the price trend is upward and a negative β_3 indicates the price trend is downward.

3.4.3. Comparable Counties

We use propensity-score matching (PSM), a causal inference approach, to find comparable counties that associate with similar demographic characteristics with the winner counties. This procedure reduces potential bias in estimating treatment effects. To avoid the misspecification of the PSM issue, we employ PSM with random forest to conduct the matching procedure. We matched winner counties with comparable counties based on demographic variables of one year before the new plant opening and the comparable counties are required to be located in the same state where the winner counties are located.

The specific county-level demographic variables include the following: total households; total population; population 25 years and over; percent less than 9th grade; percent 9th to 12th grade no diploma; percent high school graduate; percent some college degree; percent associate's degree; percent bachelor's degree; percent graduate or professional degree; population 16 years and over; population 16 years and over in labor force; median household income; median family income; per capita income; percent in the past 12 months is below the poverty level; percent 65 years and over; total housing units.

Table 3.3 Panel B provides summary statistics of these variables across counties. The first column is variable names. The second column presents the selected statistics of each variable of the full sample, the first number in each cell is mean with the median in parentheses and standard deviation in bracket followed. The third and fourth columns present the same statistics of each variable with the second column of the winner counties and the comparable counties generated by PSM correspondingly. The last column presents the t-statistics of differences in means of each variable between the winner counties and the comparables counties. Based on the last column, we find that there is no difference in means of the demographic variables between winner counties and comparables counties.

[Insert Table 3.3 Panel B About Here]

3.4.4. Meta-analysis

Meta-analysis is a statistical analysis that allows us to combine the results of multiple studies. It is normally conducted when multiple studies are addressing the same question. The purpose of a meta-analysis is to use statistical approaches to derive a pooled estimate which is closest to the unknown common truth based on the error associate with each study. It has been widely used in epidemiology and evidence-based medicine. It also draws attention from Economics in recent years.

In this paper, each pair of regression results of winner counties (treatment group) and comparable counties (control group) is treated as an individual study. We collect estimated coefficients and corresponding standard errors from each of the regression and undertake an inverse-variance-weighted random-effects meta-analysis of the means. Meta-analysis can quantify the effect of large plant openings on the rural housing market. We will discuss more details of the reasons to choose a random-effects model rather than a fixed-effects model in the empirical results section.

Under the random-effects model, the goal is not to estimate one true effect but to estimate the mean of a distribution of effects. The summary effect, therefore, represents the mean of the population of true effects.

3.5. Empirical Results

Before presenting the regression results, Table 3.4 contains Spearman correlations information for each county's real MHPs and predicted ones. To support the correlation analysis, the relationships between MHPs, both predicted MHPs and real MHPs, and times are visually presented in form of scatter plots (see Appendix 3). The correlations show that real MHPs are negatively correlated with predicted prices for seven (seventy percent) winner counties, Greene County, Tishomingo County, Harrison County, Marshall County, Tioga County, Russell County, and Gulf County. Correlation estimators of Greene County, Tishomingo County, and Harrison County are highly significant. On the other hand, real MHPs are positively correlated with predicted prices for seven (seventy percent) comparable counties, Schuyler County, George County, Hardin County, Pulaski County, Anderson County, Seneca County, Jackson County, and York County. Correlation estimators of Schuyler County, Hardin County, and York County are significant. Scatter plots in Appendix 3 display similar relationships between forecast MHPs and real MHPs for each county.

Turning to the regression results for both winner counties and comparable counties, since we observe potential autocorrelations from scatter plots of some counties in our sample, we estimated the sample autocorrelation function (ACF) for residuals. Appendix 2 shows ACFs of residuals of each regression and confirms that some of the regressions have autocorrelation between the errors. We then correct the correlation in our formal regressions.

[Insert Table 3.4 About Here]

Table 3.5 presents baseline regression results using generalized least squares (GLS) with a corrected correlation for each county. There are eight columns in total in Table 3.5, the first four columns contain regression results of winner counties, and the other four columns contain results of comparable counties. The columns named β_3 show coefficients of our variable of interest. In general, the table indicates that MHPs increase in winner counties after an occurrence of large plant opening compared with comparable counties. The coefficients of β_3s are positive for all winner counties and are highly significant for eight (eighty percent) winner counties. The coefficients of four out of ten comparable counties are positively significant, though the magnitudes are much smaller compared to the ones of winner counties; two have negatively significant coefficients, and four coefficients are positive but not significant. The columns named S.E. and N report standard errors and the number of observations correspondingly. Based on this information, we are able to conduct a meta-analysis to generate the summary effect.

[Insert Table 3.5 About Here]

We estimate the summary effect across regressions based on a meta-analysis using random effects for two reasons. First, a large amount of the variance across regressions is attributable to heterogeneity ($I^2 = 99.81\%$; $P < 0.01$, test of heterogeneity). Second, the random-effects model is more appropriate as we expect the effect sizes are not identical across the regressions.

We also conduct a leave-one-out case diagnostic to determine which one study has the highest influence on heterogeneity. Based on the leave-one-out I^2 values, no one study overly contributes to the heterogeneity.

Using our proposed regression, β_3s are estimated separately for each county, including winner counties and comparable counties. Figure 3.1 represents the results for each study, the difference in change of MHP between the winner county and comparable county, on one plot. Random-effects meta-analysis shown in the diamond. The dashed vertical line denotes a mean difference of 0, which represents no difference in MHP change between winner counties and comparable counties. The horizontal bars represent 95% CIs. The results of the statistical tests of the overall effect, the z-test statistics is 4.5, which is statistically significant at 1 percent level.

[Insert Figure 3.1 About Here]

The random-effects meta-analysis indicates the median housing prices in a rural county increase after an occurrence of a new large plant opening. The average effect size is 1,385.26 which is statistically significant at one percent level ($95\%\ CI, 782.43 - 1{,}988.09$). This indicates that on average, the median housing price will increase by $1,385 per month after a new manufacturing plant takes place.

3.6. Conclusion

The purpose of this paper is to motivate and contribute to a better understanding of the role of establishments of manufacturing plants in the rural area. We investigate the impact that a new manufacturing plant might have on the largest asset class held by most homeowners: their house.

The findings based on the empirical results highlight significant evidence of a positive impact on rural housing markets defined at the county level in counties where large plant openings take place. The positive house price effects are estimated relative to the trajectory of housing values in otherwise similar control counties within the same state as the subject county. As regards the average effect, based on a random-effects meta-analysis, we conclude that on average, the median housing price will increase by $1,385 per month after a new manufacturing plant takes place.

We hope our study will stimulate further research on the role of new manufacturing plants in rural areas. Future research on related topics could utilize a more comprehensive dataset with more housing price index and a more complete list of new manufacturing plants. Future research can also be extended to include different types of plants other than manufacturing plants and examine whether the impacts on the housing market caused by a different category of the plant are different. Future research can also examine the impact of a new plant on the local economy, social welfare, and the interaction between the new plant and other businesses.

Tables and Figures

Table 3.1 Urban Percentage of State Populations

Area Name	1950	1960	1970	1980	1990	2000	2010
United States	64.0	69.9	73.6	73.7	75.2	79.0	80.7
Alabama	43.8	54.8	58.6	60.0	60.4	55.4	59.0
Alaska	26.6	37.9	56.9	64.3	67.5	65.6	66.0
Arizona	55.5	74.5	79.6	83.8	87.5	88.2	89.8
Arkansas	33.0	42.8	50.0	51.6	53.5	52.5	56.2
California	80.7	86.4	90.9	91.3	92.6	94.4	95.0
Colorado	62.7	73.7	78.5	80.6	82.4	84.5	86.2
Connecticut	77.6	78.3	78.4	78.8	79.1	87.7	88.0
Delaware	62.6	65.6	72.2	70.6	73.0	80.1	83.3
Florida	65.5	73.9	81.7	84.3	84.8	89.3	91.2
Georgia	45.3	55.3	60.3	62.4	63.2	71.6	75.1
Hawaii	69.0	76.5	83.1	86.5	89.0	91.5	91.9
Idaho	42.9	47.5	54.1	54.0	57.4	66.4	70.6
Illinois	77.6	80.7	83.2	83.3	84.6	87.8	88.5
Indiana	59.9	62.4	64.9	64.2	64.9	70.8	72.4
Iowa	47.7	53.0	57.2	58.6	60.6	61.1	64.0
Kansas	52.1	61.0	66.1	66.7	69.1	71.4	74.2
Kentucky	36.8	44.5	52.3	50.9	51.8	55.8	58.4
Louisiana	54.8	63.3	66.5	68.6	68.1	72.6	73.2
Maine	51.7	51.3	50.8	47.5	44.6	40.2	38.7
Maryland	69.0	72.7	76.6	80.3	81.3	86.1	87.2
Massachusetts	84.4	83.6	84.6	83.8	84.3	91.4	92.0
Michigan	70.7	73.4	74.0	70.7	70.5	74.7	74.6

Minnesota	54.5	62.2	66.5	66.9	69.9	70.9	73.3
Mississippi	27.9	37.7	44.5	47.3	47.1	48.8	49.4
Missouri	61.5	66.6	70.1	68.1	68.7	69.4	70.4
Montana	43.7	50.2	53.4	52.9	52.5	54.1	55.9
Nebraska	46.9	54.3	61.5	62.9	66.1	69.8	73.1
Nevada	57.2	70.4	80.9	85.3	88.3	91.5	94.2
New Hampshire	57.5	58.3	56.4	52.2	51.0	59.3	60.3
New Jersey	86.6	88.6	88.9	89.0	89.4	94.4	94.7
New Mexico	50.2	65.9	69.8	72.1	73.0	75.0	77.4
New York	85.5	85.4	85.7	84.6	84.3	87.5	87.9
North Carolina	33.7	39.5	45.5	48.0	50.4	60.2	66.1
North Dakota	26.6	35.2	44.3	48.8	53.3	55.9	59.9
Ohio	70.2	73.4	75.3	73.3	74.1	77.4	77.9
Oklahoma	51.0	62.9	68.0	67.3	67.7	65.3	66.2
Oregon	53.9	62.2	67.1	67.9	70.5	78.7	81.0
Pennsylvania	70.5	71.6	71.5	69.3	68.9	77.1	78.7
Rhode Island	84.3	86.4	87.1	87.0	86.0	90.9	90.7
South Carolina	36.7	41.2	48.3	54.1	54.6	60.5	66.3
South Dakota	33.2	39.3	44.6	46.4	50.0	51.9	56.7
Tennessee	44.1	52.3	59.1	60.4	60.9	63.6	66.4
Texas	62.7	75.0	79.7	79.6	80.3	82.5	84.7
Utah	65.3	74.9	80.4	84.4	87.0	88.2	90.6
Vermont	36.4	38.5	32.2	33.8	32.2	38.2	38.9
Virginia	47.0	55.6	63.2	66.0	69.4	73.0	75.5
Washington	63.2	68.1	73.4	73.5	76.4	82.0	84.1
West Virginia	34.6	38.2	39.1	36.2	36.1	46.1	48.7
Wisconsin	57.9	63.8	65.9	64.2	65.7	68.3	70.2

Wyoming	49.8	56.8	60.5	62.7	65.0	65.1	64.8

Note: The definition of urban population changed beginning with the 2000 Decennial Census. For the years 1950-1990, urban included all population in "urbanized areas" and incorporated places or Census Designated Places with population 2,500 or greater located outside of urbanized areas. For years 2000 and 2010, urban included all population in urbanized areas and urban clusters. Urbanized areas are defined as densely settled territory with specific population thresholds, while urban clusters can each have their own density and population thresholds.

Source: U.S. Census Bureau

Table 3.2 Variable Legend

Variable	Definition
MHP	Median housing price
Real	One if the housing price is the real price, zero otherwise
Time	Time relative to the month when the plant opened. 0 is the month when the plant opened, 1 is one month after month 0, 2 is two months after month 0, and so on.
TP	Total population
P16	The population that are 16 years and over
L16	The population that are 16 years and over in the labor force
P25	The population that are 25 years and over
P65	Percent of population that are 65 years and over
E9	Percent of population' education attainment is less than 9^{th} grade
E12	Percent of population' education attainment is 9^{th} to 12^{th} grade no diploma
EHigh	Percent of population' education attainment is high school graduate
EAss	Percent of population' education attainment is associate degree
ECol	Percent of population' education attainment is some college without a degree
EBac	Percent of population' education attainment is bachelor's degree
EGra	Percent of population' education attainment is graduate or professional degree
TH	Total households
MHI	Median household income
MFI	Median family income
PCI	Per capita income
POV	Percent of population in the past 12 months is below the poverty level

THU	Total housing units

Table 3.3 Summary Statistics

Panel A. Monthly Median Housing Price by County

Winner County	MHP	N	Time Range	Comparable County	MHP	N	Time Range
Greene, NY	226,540.8 (223,675.0) [12,563.5]	120	2010/01-2019/12	Schuyler, NY	168,777.8 (169,900.0) [13,760.8]	120	2010/01-2019/12
Tishomingo, MS	149,664.0 (149,000.0) [13,723.8]	84	2013/01-2019/12	George, MS	128,696.8 (127,500.0) [11,899.7]	120	2010/01-2019/12
Harrison, TX	185,755.4 (185,000.0) [11,638.0]	120	2010/01-2019/12	Hardin, TX	178,538.9 (169,900.0) [19,301.8]	120	2010/01-2019/12
Polk, MO	138,908.7 (134,900.0) [13,222.9]	120	2010/01-2019/12	Pulaski, MO	148,974.6 (149,900.0) [6,305.9]	120	2010/01-2019/12
Marshall, KY	158,789.6 (159,900.0) [10,689.2]	120	2010/01-2019/12	Anderson, KY	144,466.5 (144,500.0) [12,883.6]	120	2010/01-2019/12
Tioga, NY	141,168.8 (140,000.0) [7,563.2]	120	2010/01-2019/12	Seneca, NY	145,542.4 (142,500.0) [14,257.9]	120	2010/01-2019/12
Jones, MS	120,776.9 (122,600.0) [14,410.4]	120	2010/01-2019/12	Jackson, MS	148,091.2 (148,300.0) [10,866.0]	120	2010/01-2019/12
Botetourt, VA	256,657.8 (258,750.0) [13,466.5]	120	2010/01-2019/12	York, VA	317,803.9 (319,900.0) [13,326.9]	120	2010/01-2019/12
Russell, KY	141,933.6 (143,900.0) [10,045.6]	67	2014/06-2019/12	Wayne, KY	155,346.2 (154,900.0) [18,708.5]	106	2011/03-2019/12
Gulf, FL	322,488.1 (329,900.0) [32,922.0]	45	2016/04-2019/12	Glades, FL	159,553.6 (162,500.0) [12,712.0]	47	2016/02-2019/12

Note: Numbers reported are means, with medians in parentheses and standard deviations in brackets.

Panel B. Summary Statistics of Demographic Variables

Variable Name	Full Sample	Winner Counties	Comparable Counties	t-test of differences
TP	41,553.85 (32,165) [29,795.27]	38,226.90 (32,165.00) [19,318.31]	44,880.80 (28,998.00) [38,423.45]	0.63
P16	33,099.65 (26,558.00) [23,153.54]	30,733.40 (26,558.00) [14,856.90]	35,465.90 (23,126.50) [29,976.12]	0.66
L16	19,476.65 (15,342.50) [14,453.90]	17,414.20 (15,567.00) [9,096.73]	21,539.10 (13,871.50) [18,677.26]	0.54
P25	27,971.35 (23,375.00) [19,579.58]	26,290.70 (23,375.00) [12,461.43]	29,652.00 (19,844.50) [25,450.96]	0.71
P65	16.90 (17.55) [3.72]	18.13 (18.80) [2.06]	15.66 (14.55) [4.65]	0.14
E9	5.60 (4.30) [3.51]	5.44 (4.95) [2.94]	5.76 (4.30) [4.17]	0.84
E12	9.73 (10.05) [3.12]	10.09 (10.25) [3.27]	9.37 (9.25) [3.09]	0.62
EHigh	35.80 (37.00) [5.29]	36.14 (36.70) [3.36]	35.45 (38.30) [6.89]	0.78
EAss	9.08 (8.60) [1.96]	8.81 (8.35) [1.73]	9.35 (9.40) [2.23]	0.99
ECol	20.68 (21.20) [2.63]	20.67 (21.70) [2.02]	20.69 (20.45) [3.24]	0.55
EBac	11.63 (11.40) [4.20]	11.66 (11.00) [3.52]	11.60 (11.40) [4.98]	0.98
EGra	7.50 (7.80) [3.64]	7.20 (7.40) [1.92]	7.80 (7.80) [4.91]	0.72
TH	15,228.40 (12,920.00) [10,554.83]	14,393.00 (12,920.00) [6,890.85]	16,063.80 (11,059.00) [13,643.78]	0.73
MHI	48,041.35 (47,450.00) [12,204.73]	45,767.50 (44,929.50) [9,812.42]	50,315.20 (49,191.50) [14,376.64]	0.42
MFI	58,396.35 (56,770.00) [13,672.47]	56,016.20 (56,227.00) [11,702.51]	60,776.50 (60,131.50) [15,655.86]	0.45
PCI	23,877.70 (23,579.50) [5,289.57]	23,647.30 (22,268.50) [5,188.82]	24,108.10 (24,351.50) [5,659.16]	0.85
POV	11.35 (11.35) [4.46]	11.83 (11.75) [4.78]	10.86 (10.85) [4.32]	0.64
THU	18,693.55 (15,333.50) [12,468.77]	18,161.20 (15,333.50) [8,082.93]	19,225.90 (13,572.50) [16,194.18]	0.85

Note: Numbers reported are means, with medians in parentheses and standard deviations in brackets. Medians and standard deviations are not reported for the indicator variables.

Table 3.4 Spearman Correlation between Real MHP and Predicted MHP by County

Winner County	Spearman's rho	Comparable County	Spearman's rho
Greene, NY	-0.46***	Schuyler, NY	0.41***
Tishomingo, MS	-0.62***	George, MS	0.21
Harrison, TX	-0.49***	Hardin, TX	0.78***
Polk, MO	0.19	Pulaski, MO	-0.29*
Marshall, KY	-0.07	Anderson, KY	0.20
Tioga, NY	-0.17	Seneca, NY	0.13
Jones, MS	0.38*	Jackson, MS	0.08
Botetourt, VA	0.17	York, VA	0.56**
Russell, KY	-0.43	Wayne, KY	-0.43
Gulf, FL	-0.32	Glades, FL	-0.41

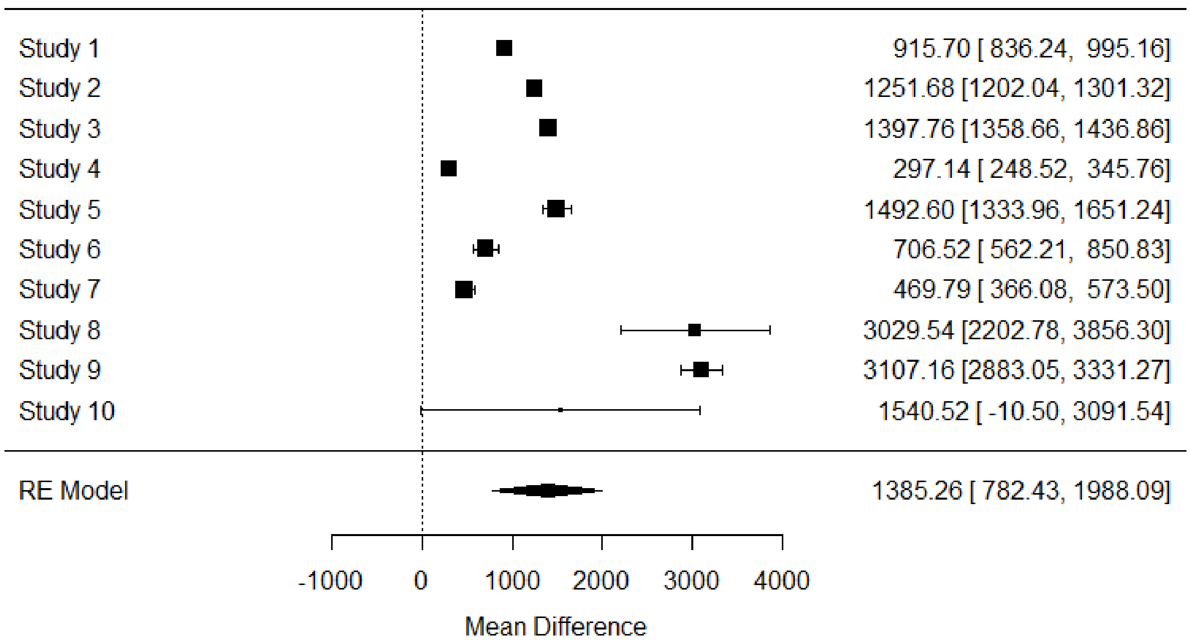
Note: ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, respectively.

Table 3.5 Baseline Regression Results

Winner County	β_3	S.E.	N	Comparable County	β_3	S.E.	N
Greene, NY	1,088.20***	128.82	100	Schuyler, NY	172.50	384.41	100
Tishomingo, MS	1,759.23***	167.82	102	George, MS	507.55***	193.06	102
Harrison, TX	1,092.67***	101.362	110	Hardin, TX	-305.09*	183.05	110
Polk, MO	836.37***	160.76	82	Pulaski, MO	539.23***	156.87	82
Marshall, KY	1,483.37***	128.85	80	Anderson, KY	-9.23	712.39	80
Tioga, NY	1,225.51***	404.21	64	Seneca, NY	518.99	428.48	64
Jones, MS	850.13***	291.65	50	Jackson, MS	380.34	234.39	50
Botetourt, VA	2974.60	2,150.37	30	York, VA	-54.94	845.00	30
Russell, KY	2,281.18***	359.36	40	Wayne, KY	-825.98	627.56	40
Gulf, FL	2,511.97	2,725.46	32	Glades, FL	971.45	3551.26	32

Note: ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, respectively.

Figure 3.1 Random-Effects Meta-Analysis



Reference

Ammons, Lila. (1996) "The evolution of Black-owned banks in the United States between the 1880s and 1990s." *Journal of Black Studies,* 26.4: 467-489.

Baradaran, Mehrsa. (2017). **The Color of Money: Black Banks and the Racial Wealth Gap**, Harvard University Press, Cambridge, MA.

Barth, James R., and Aron Betru. (2019). "Turbocharging Opportunity Zones." *Milken Institute Review: A Journal of Economic Policy,* 21.3: 14-17.

Bates, Timothy, and William Bradford. (1980). "An analysis of the portfolio behavior of black-owned commercial banks." *The Journal of Finance,* 35.3: 753-768.

Black, Harold A., M. Cary Collins, and Ken B. Cyree. (1997). "Do black-owned banks discriminate against black borrowers?" *Journal of Financial Services Research*, 11.1-2: 189-204.

Boorman, John T., and Myron L. Kwast. (1974). "The start-up experience of minority-owned commercial banks: A comparative analysis." *The Journal of Finance,* 29.4: 1123-1141.

Breitenstein, Eric C., Karyen Chu, Kathy Kalser, and Eric W. Robbins. (2014). "Minority depository institutions: Structure, performance, and social impact." *FDIC Quarterly,* 8, no. 3.

Breitenstein, Eric C., Karyen Chu, Richard D. Cofer, Jr., Jeffrey A. DeLuca and Shawn E. Schreier. (2019). "2019 Minority depository institutions: Structure, performance, and social impact." FDIC.

Brimmer, Andrew F. (1971). "The black banks: An assessment of performance and prospects." *The Journal of Finance,* 26.2: 379-405.

Clair, Robert T. (1988). "The performance of black-owned banks in their primary market areas." *Economic and Financial Policy Review*, Federal Reserve Bank of Dallas, November, 11-20.

Elyasiani, Elyas, and Seyed Mehdian. (1992). "Productive efficiency performance of minority and nonminority-owned banks: A nonparametric approach." *Journal of Banking and Finance,* 16.5, 933-948.

Fairchild, Gregory B., Young Kim, Megan E. Juelfs, and Aron Betru. (2020) "Good money after bad? The comparative efficiency of minority depository institutions." *Journal of Developmental Entrepreneurship*, 25.01: 2050002.

Federal Deposit Insurance Corporation (2017). "Preservation and Promotion of Minority Depository Institutions." Report to Congress.

Henderson, Christopher C. (1999). "The economic performance of African-American-owned banks: The role of loan loss provisions." *American Economic Review,* 89.2: 372-376.

Hoque, Monzurul, William Lepley, Robert Nagy, and Mussie Teclezion. (2015). "Impact of the 2008-2009 recession on African-American owned US banks." *Managerial Finance*.

Iqbal, Zahid, Kizhanathan V. Ramaswamy, and Aigbe Akhigbe. (1999). "The output efficiency of minority-owned banks in the United States." *International Review of Economics & Finance,* 8.1: 105-114.

Kashian, Russell, and Robert Drago. (2017). "Minority-Owned Banks and Bank Failures After the Financial Collapse." *Economic Notes: Review of Banking, Finance and Monetary Economics,* 46.1: 5-36.

Kashian, Russ, Richard McGregory Jr, and Neil Lockwood. (2014). "Do minority-owned banks pay higher interest rates on CDs?" *The Review of Black Political Economy,* 41.1: 13-24.

Kashian, Russell, Richard McGregory, and Robert Drago. (2017). "Minority owned banks and efficiency revisited." *Journal of Productivity Analysis,* 48.2-3: 97-116.

Kwast, Myron L., and Harold Black. (1983). "An analysis of the behavior of mature black-owned commercial banks." *Journal of Economics and Business,* 35.1: 41-54.

Lawrence, Edward C. (1997). "The viability of minority-owned banks." *The Quarterly Review of Economics and Finance,* 37.1: 1-21.

Meinster, David, and Elyas Elyasiani. (1988). "The performance of foreign owned, minority owned, and holding company owned banks in the US." *Journal of Banking and Finance,* 12.2, 293-313.

Price, Douglas A. (1990). "Minority-owned banks: History and trends." *Economic Commentary*, Federal Reserve Bank of Cleveland, July.

Price, Gregory N. (1994). "The cost of government deposits for black-owned commercial banks." *The Review of Black Political Economy,* 23.1: 9-24.

Price, Gregory N. (1995). "The determinants of entry for black-owned commercial banks." *The Quarterly Review of Economics and Finance,* 35.3: 289-303.

Ricks, Anthony W. (2018). "Access to Credit in African American Communities: The Case for the Development of Black Owned Banks and Financial Institutions." Consumer Financial Protection Bureau.

Rosenbaum, Paul R., and Donald B. Rubin. (1983). "The central role of the propensity score in observational studies for causal effects." *Biometrika*, 70.1: 41-55.

Toussaint-Comeau, Maude, and Robin Newberger. (2017). "Minority-owned banks and their primary local market areas." *Economic Perspectives,* Federal Reserve Bank of Chicago, 4: 1-31.

Azar, José, Sahil Raina, and Martin C. Schmalz. “Ultimate ownership and bank competition.” Available at SSRN 2710252 (2019).

Barth, James R, and Jiayi Xu. “US Minority Banks: Why Still So Few—After 150 Years?” (2021). Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3438390.

Caskey, John P. “Bank representation in low-income and minority urban communities.” *Urban Affairs Quarterly* 29.4 (1994): 617–638.

Chang, Angela, Shubham Chaudhuri, and Jith Jayaratne. “Rational herding and the spatial clustering of bank branches: An empirical analysis.” (1997). *Federal Reserve Bank of New York Research Paper*, no. 9724.

Corvoisier, Sandrine, and Reint Gropp. “Bank concentration and retail interest rates.” *Journal of Banking & Finance* 26.11 (2002): 2155–2189.

Degryse, Hans, and Steven Ongena. “Distance, lending relationships, and competition.” *The Journal of Finance* 60.1 (2005): 231–266.

Ding, Lei, and Carolina K. Reid. “The Community Reinvestment Act (CRA) and bank branching patterns.” *Housing Policy Debate* 30.1 (2020): 27–45.

Ergungor, Ozgur Emre. “Bank branch presence and access to credit in low- to moderate-income neighborhoods.” *Journal of Money, Credit and Banking* 42.7 (2010): 1321–1349.

Figlio, David N., and Joseph W. Genshlea. “Bank consolidations and minority neighborhoods.” *Journal of Urban Economics* 45.3 (1999): 474–489.

Flechsig, Theodore G. “The effect of concentration on bank loan rates.” *The Journal of Finance* 20.2 (1965): 298–311.

Hakenes, Hendrik, Iftekhar Hasan, Philip Molyneux, and Ru Xie. “Small banks and local economic development.” *Review of Finance* 19.2 (2015): 653–683.

Jagtiani, Julapa. “Understanding the effects of the merger boom on community banks.” *Economic Review*, 93.3 (2008): 28–48.

Kowalik, Michal. “Can Small Banks Survive Competition from Large Banks?” Working paper. Federal Reserve (2014).

Liebersohn, Jack. “How does competition affect bank lending? Quasi-experimental evidence from bank mergers.” *Working paper* (2017).

Nguyen, Hoai-Luu Q. “Are credit markets still local? Evidence from bank branch closings.” *American Economic Journal: Applied Economics* 11.1 (2019): 1–32.

Petersen, Mitchell, and Raghuram Rajan. “The benefits of lending relationships: Evidence from small business data,” *Journal of Finance* 49.1 (1994): 3–37.

Unite, Angelo A., and Michael J. Sullivan. “The effect of foreign entry and ownership structure on the Philippine domestic banking market.” *Journal of Banking & Finance* 27.12 (2003): 2323–2345.

Almazan, Andres, Adolfo De Motta, and Sheridan Titman. "Firm location and the creation and utilization of human capital." *The Review of Economic Studies* 74.4 (2007): 1305-1327.

Artz, Georgeanne M., Younjun Kim, and Peter F. Orazem. "Does agglomeration matter everywhere?: new firm location decisions in rural and urban markets." *Journal of Regional Science* 56.1 (2016): 72-95.

Baldwin, Richard E., and Toshihiro Okubo. "Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting." *Journal of economic geography* 6.3 (2006): 323-346.

Barca, Fabrizio. **Agenda for a reformed cohesion policy**. Brussels: European Communities, 2009.

Barca, Fabrizio, Philip McCann, and Andrés Rodríguez-Pose. "The case for regional development intervention: place-based versus place-neutral approaches." *Journal of regional science* 52.1 (2012): 134-152.

Bloom, Nicholas, et al. "What drives differences in management practices?." *American Economic Review* 109.5 (2019): 1648-83.

Blomquist, Glenn. "The effect of electric utility power plant location on area property value." *Land Economics* 50.1 (1974): 97-100.

Boarnet, Marlon G., and William T. Bogart. "Enterprise zones and employment: evidence from New Jersey." *Journal of urban economics* 40.2 (1996): 198-215.

Bondonio, Daniele, and John Engberg. "Enterprise zones and local employment: evidence from the states' programs." *Regional Science and Urban Economics* 30.5 (2000): 519-549.

Bondonio, Daniele, and Robert T. Greenbaum. "Do local tax incentives affect economic growth? What mean impacts miss in the analysis of enterprise zone policies." *Regional Science and Urban Economics* 37.1 (2007): 121-136.

Boyle, Melissa, and Katherine Kiel. "A survey of house price hedonic studies of the impact of environmental externalities." *Journal of real estate literature* 9.2 (2001): 117-144.

Busso, Matias, Jesse Gregory, and Patrick Kline. "Assessing the incidence and efficiency of a prominent place based policy." *American Economic Review* 103.2 (2013): 897-947.

Castle, Emery N., JunJie Wu, and Bruce A. Weber. "Place orientation and rural–urban interdependence." *Applied Economic Perspectives and Policy* 33.2 (2011): 179-204.

Carlino, Gerald A., and Albert Saiz. "Beautiful city: Leisure amenities and urban growth." *Journal of Regional Science* 59.3 (2019): 369-408.

Combes, Pierre-Philippe, et al. "The productivity advantages of large cities: Distinguishing agglomeration from firm selection." *Econometrica* 80.6 (2012): 2543-2594.

Couture, Victor, and Jessie Handbury. "Urban revival in America." *Journal of Urban Economics* 119 (2020): 103267.

Currie, Janet, et al. "Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings." *American Economic Review* 105.2 (2015): 678-709.

Davis, Lucas W. "The effect of power plants on local housing values and rents." *Review of Economics and Statistics* 93.4 (2011): 1391-1402.

Deng, Guoying, Manuel A. Hernandez, and Shu Xu. "When Power Plants Leave Town: Environmental Quality and the Housing Market in China." *Environmental and Resource Economics* 77.4 (2020): 751-780.

Doeringer, Peter, Christine Evans-Klock, and David Terkla. "What attracts high performance factories? Management culture and regional advantage." *Regional Science and Urban Economics* 34.5 (2004): 591-618.

Dong, Yilin. "Determinants of entry: Evidence from new manufacturing firms in the US." *Growth and Change* 51.4 (2020): 1542-1561.

Dowall, David E. "An evaluation of California's enterprise zone programs." *Economic Development Quarterly* 10.4 (1996): 352-368.

Duranton, Gilles, and Diego Puga. "Micro-foundations of urban agglomeration economies." *Handbook of regional and urban economics*. Vol. 4. Elsevier (2004): 2063-2117.

Ellison, Glenn, Edward L. Glaeser, and William R. Kerr. "What causes industry agglomeration? Evidence from coagglomeration patterns." *American Economic Review* 100.3 (2010): 1195-1213.

Engberg, John, and Robert Greenbaum. "State enterprise zones and local housing markets." *Journal of Housing Research* 10.2 (1999): 163-187.

Engelberg, Joseph, Arzu Ozoguz, and Sean Wang. "Know thy neighbor: Industry clusters, information spillovers, and market efficiency." *Journal of Financial and Quantitative Analysis* 53.5 (2018): 1937-1961

Faggio, Giulia, Olmo Silva, and William C. Strange. "Tales of the city: what do agglomeration cases tell us about agglomeration in general?." *Journal of Economic Geography* 20.5 (2020): 1117-1143.

Fomento, C. A. "Dessarollo Local: Haciaunnuevo protagonismo de las ciudades y regiones." Caracas: Corporacion Andina de Fomento (2010).

Gamble, Hays B., Roger H. Downing, and Owen H. Sauerlender. "Community growth around nuclear power plants." *Real Estate Economics* 8.3 (1980): 268-280.

Gamble, Hays B., and Roger H. Downing. "Effects of nuclear power plants on residential property values." *Journal of Regional Science* 22.4 (1982): 457-478.

Gamper-Rabindran, Shanti, and Christopher Timmins. "Does cleanup of hazardous waste sites raise housing values? Evidence of spatially localized benefits." *Journal of Environmental Economics and Management* 65.3 (2013): 345-360.

Gaubert, Cecile. "Firm sorting and agglomeration." *American Economic Review* 108.11 (2018): 3117-53.

Gayer, Ted, James T. Hamilton, and W. Kip Viscusi. "Private values of risk tradeoffs at superfund sites: housing market evidence on learning about risk." *Review of Economics and Statistics* 82.3 (2000): 439-451.

Giuliano, Genevieve, Sanggyun Kang, and Quan Yuan. "Agglomeration economies and evolving urban form." *The Annals of Regional Science* 63 (2019): 377-398.

Glaeser, Edward L., and Joshua D. Gottlieb. The economics of place-making policies. No. w14373. *National Bureau of Economic Research*, 2008.

Greenberg, Michael, and John Hughes. "The impact of hazardous waste Superfund sites on the value of houses sold in New Jersey." *The Annals of Regional Science* 26.2 (1992): 147-153.

Greenbaum, Robert, and Daniele Bondonio. "Losing focus: A comparative evaluation of spatially targeted economic revitalization programmes in the US and the EU." *Regional Studies* 38.3 (2004): 319-334.

Greenbaum, Robert T., and John B. Engberg. "The impact of state enterprise zones on urban manufacturing establishments." *Journal of Policy Analysis and Management* 23.2 (2004): 315-339.

Greenstone, Michael, and Enrico Moretti. Bidding for industrial plants: Does winning a'million dollar plant'increase welfare?. No. w9844. *National Bureau of Economic Research*, 2003.

Greenstone, Michael, and Justin Gallagher. "Does hazardous waste matter? Evidence from the housing market and the superfund program." *The Quarterly Journal of Economics* 123.3 (2008): 951-1003.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings." *Journal of Political Economy* 118.3 (2010): 536-598.

Hanna, Brid Gleeson. "House values, incomes, and industrial pollution." *Journal of Environmental Economics and Management* 54.1 (2007): 100-112.

Head, Keith, and Thierry Mayer. "Market potential and the location of Japanese investment in the European Union." *Review of Economics and Statistics* 86.4 (2004): 959-972.

Helsley, Robert W., and William C. Strange. "Coagglomeration, clusters, and the scale and composition of cities." *Journal of Political Economy* 122.5 (2014): 1064-1093.

Henry, Mark S., David L. Barkley, and Shuming Bao. "The hinterland's stake in metropolitan growth: Evidence from selected southern regions." *Journal of Regional Science* 37.3 (1997): 479-501.

Hodge, Timothy R. "The effect of ethanol plants on residential property values: Evidence from Michigan." *Journal of Regional Analysis and Policy* 41.1100-2016-89865 (2011): 148-167.

Holmes, Thomas J., Wen-Tai Hsu, and Sanghoon Lee. "Allocative efficiency, mark-ups, and the welfare gains from trade." *Journal of International Economics* 94.2 (2014): 195-206.

Holmes, Thomas J., and John J. Stevens. "An alternative theory of the plant size distribution, with geography and intra-and international trade." *Journal of Political Economy* 122.2 (2014): 369-421.

Irwin, Elena G., et al. "A century of research on rural development and regional issues." *American Journal of Agricultural Economics* 92.2 (2010): 522-553.

Jackson, Thomas. "The effects of environmental contamination on real estate: A literature review." *Journal of Real Estate Literature* 9.2 (2001): 91-116.

Kahn, Matthew E. "Regional growth and exposure to nearby coal fired power plant emissions." *Regional Science and Urban Economics* 39.1 (2009): 15-22.

Kiel, Katherine A., and Michael Williams. "The impact of Superfund sites on local property values: Are all sites the same?." *Journal of urban Economics* 61.1 (2007): 170-192.

Kilkenny, Maureen. "Urban/regional economics and rural development." *Journal of Regional Science* 50.1 (2010): 449-470.

Kim, Hyunseob. "How does labor market size affect firm capital structure? Evidence from large plant openings." *Journal of Financial Economics* 138.1 (2020): 277-294.

Kline, Patrick. "Place based policies, heterogeneity, and agglomeration." *American Economic Review* 100.2 (2010): 383-87.

Kline, Patrick, and Enrico Moretti. "Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority." *The Quarterly journal of economics* 129.1 (2014): 275-331.

Komarek, Timothy, and Scott Loveridge. "Firm sizes and economic development: Estimating long-term effects on US county growth, 1990–2000." *Journal of Regional Science* 55.2 (2015): 262-279.

Lambert, Thomas E., and Paul A. Coomes. "An evaluation of the effectiveness of Louisville's enterprise zone." *Economic Development Quarterly* 15.2 (2001): 168-180.

Lee, Sanghoon, and Jeffrey Lin. "Natural amenities, neighborhood dynamics, and persistence in the spatial distribution of income." *The Review of Economic Studies* 85.1 (2018): 663-694.

Lee, Yongsung, Bumsoo Lee, and Md Tanvir Hossain Shubho. "Urban revival by millennials? Intraurban net migration patterns of young adults, 1980–2010." *Journal of Regional Science* 59.3 (2019): 538-566.

Lewis, William C., and James R. Prescott. "Urban-Regional Development and Growth Centers: An Econometric Study." *Journal of Regional Science* 12.1 (1972): 57-70.

Mantegazzi, Daniele, Philip McCann, and Viktor Venhorst. "The impact of language borders on the spatial decay of agglomeration and competition spillovers." *Journal of Regional Science* 60.3 (2020): 558-577.

Marshall, Alfred. "" Some Aspects of Competition." The Address of the President of Section F--Economic Science and Statistics--of the British Association, at the Sixtiet Meeting, held at Leeds, in September, 1890." *Journal of the Royal Statistical Society* 53.4 (1890): 612-643.

Melitz, Marc J. "The impact of trade on intra-industry reallocations and aggregate industry productivity." *Econometrica* 71.6 (2003): 1695-1725.

Melitz, Marc J., and Gianmarco IP Ottaviano. "Market size, trade, and productivity." *The review of economic studies* 75.1 (2008): 295-316.

Melo, Patricia C., Daniel J. Graham, and Robert B. Noland. "A meta-analysis of estimates of urban agglomeration economies." *Regional science and urban Economics* 39.3 (2009): 332-342.

Muehlenbachs, Lucija, Elisheba Spiller, and Christopher Timmins. "The housing market impacts of shale gas development." *American Economic Review* 105.12 (2015): 3633-59.

Mulligan, Gordon F., Mark D. Partridge, and John I. Carruthers. "Central place theory and its reemergence in regional science." *The Annals of Regional Science* 48.2 (2012): 405-431.

Nelson, Jon P. "Three Mile Island and residential property values: Empirical analysis and policy implications." *Land Economics* 57.3 (1981): 363-372.

Neumark, David, and Jed Kolko. "Do enterprise zones create jobs? Evidence from California's enterprise zone program." *Journal of Urban Economics* 68.1 (2010): 1-19.

Newport, Frank. "Americans Big on Idea of Living in the Country." Gallup Organization, December 7, 2018. Available at https://news.gallup.com/poll/245249/americans-big-idea-living-country.aspx.

Nocke, Volker. "A gap for me: Entrepreneurs and entry." *Journal of the European Economic Association* 4.5 (2006): 929-956.

OECD. 2009. How Regions Grow. Paris : *Organisation for Economic Growth and Development*.

OECD. 2009. Regions Matter: Economic Recovery, Innovation and Sustainable Growth. Paris : *Organisation for Economic Growth and Development*.

Okubo, Toshihiro, and Eiichi Tomiura. "Industrial relocation policy, productivity and heterogeneous plants: Evidence from Japan." *Regional Science and Urban Economics* 42.1-2 (2012): 230-239.

Olfert, M. Rose, and Mark D. Partridge. "Best practices in twenty-first-century rural development and policy." *Growth and Change* 41.2 (2010): 147-164.

O'Keefe, Suzanne. "Job creation in California's enterprise zones: a comparison using a propensity score matching model." *Journal of Urban Economics* 55.1 (2004): 131-150.

Partridge, Mark D., and M. Rose Olfert. "The winners' choice: Sustainable economic strategies for successful 21st-century regions." *Applied Economic Perspectives and Policy* 33.2 (2011): 143-178.

Renski, Henry. "New firm entry, survival, and growth in the United States: A comparison of urban, suburban, and rural areas." *Journal of the American Planning Association* 75.1 (2008): 60-77.

Renski, Henry, and Ryan Wallace. "Entrepreneurship in rural America." *Financing economic development in the 21st century* 245 (2012).

Rickman, Dan S., and Shane D. Rickman. "Population growth in high-amenity nonmetropolitan areas: What's the prognosis?." *Journal of Regional Science* 51.5 (2011): 863-879.

Rickman, Dan S., and Mouhcine Guettabi. "The great recession and nonmetropolitan America." *Journal of Regional science* 55.1 (2015): 93-112.

Rosenthal, Stuart S., and William C. Strange. "Evidence on the nature and sources of agglomeration economies." *Handbook of regional and urban economics*. Vol. 4. Elsevier, 2004. 2119-2171.

Renkow, Mitch, and Dale Hoover. "Commuting, migration, and rural-urban population dynamics." *Journal of regional science* 40.2 (2000): 261-287.

Reeder, Richard J., and Kenneth L. Robinson. "Enterprise zones: Assessing their rural development potential." *Policy Studies Journal* 20.2 (1992): 264.

Simons, Robert, and Jesse Saginor. "A meta-analysis of the effect of environmental contamination and positive amenities on residential real estate values." *Journal of Real Estate Research* 28.1 (2006): 71-104.

Taylor, Sean J., and Benjamin Letham. "Forecasting at scale." *The American Statistician* 72.1 (2018): 37-45.

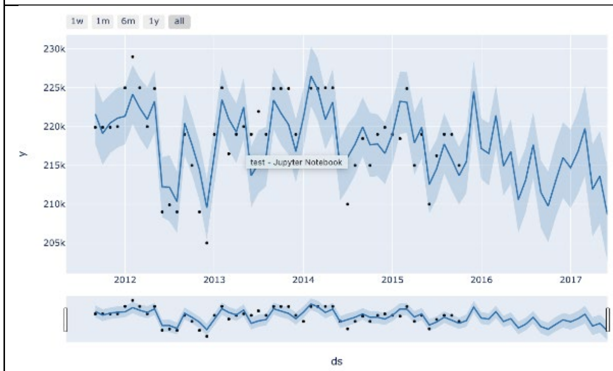
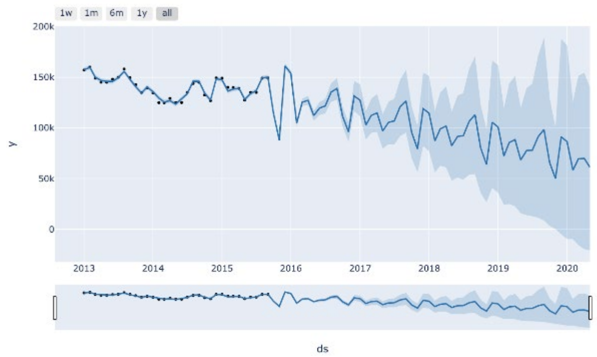
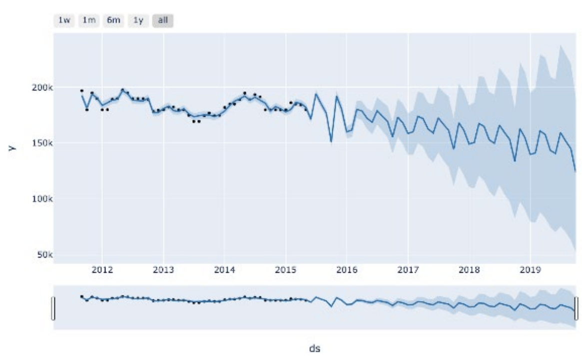
Wilder, Margaret G., and Barry M. Rubin. "Rhetoric versus reality: A review of studies on state enterprise zone programs." *Journal of the American Planning Association* 62.4 (1996): 473-491.

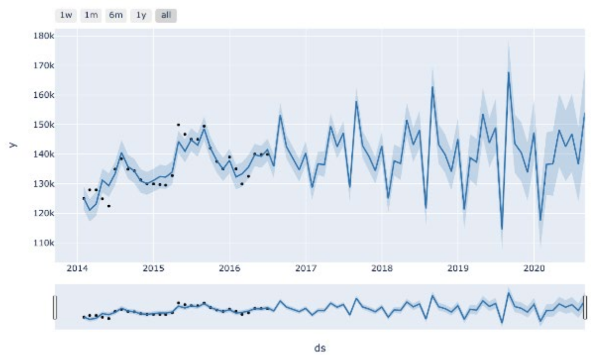
World Bank. World development report 2009: Reshaping economic geography. The World Bank, 2008.

Yu, Li, Peter F. Orazem, and Robert W. Jolly. "Why do rural firms live longer?." *American Journal of Agricultural Economics* 93.3 (2011): 673-692.

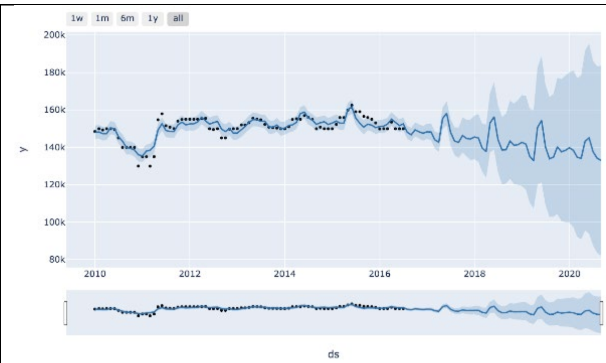
Appendix

A1. Predictions by using Prophet

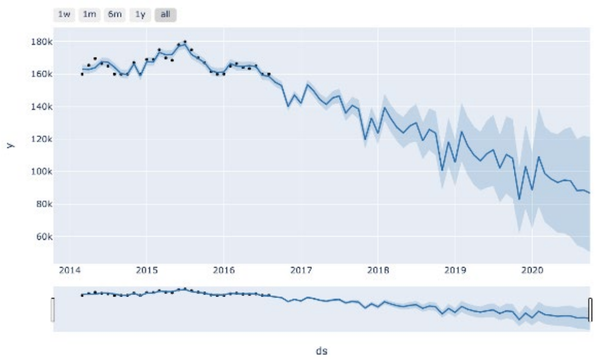
Winner County	**Comparable County**
Greene, NY	Schuyler, NY
	
Tishomingo, MS	George, MS
	
Harrison, TX	Hardin, TX
	
Polk, MO	Pulaski, MO



Marshall, KY



Anderson, KY

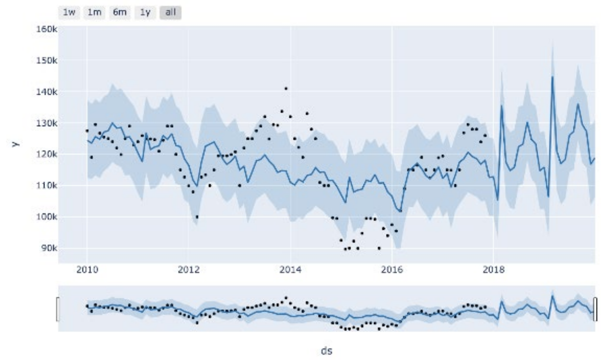


Tioga, NY

Seneca, NY

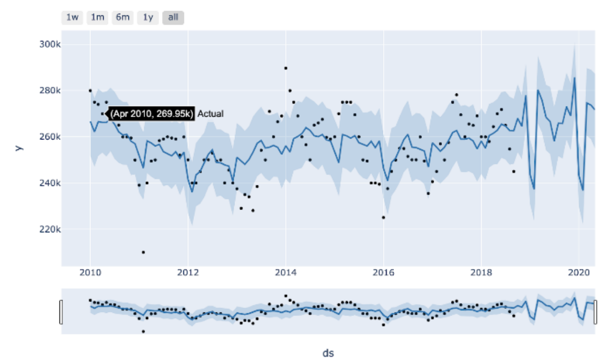
Jones, MS

Jackson, MS

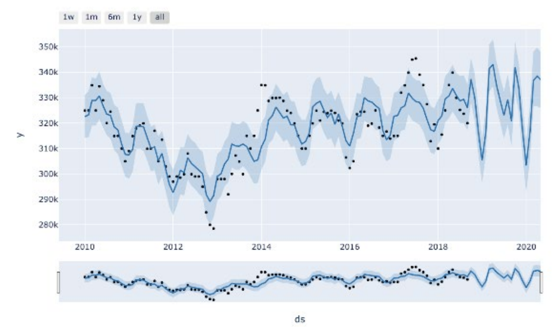


Botetourt, VA

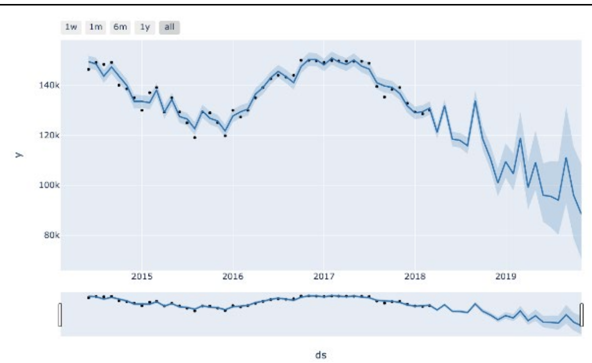
York, VA



Russell, KY

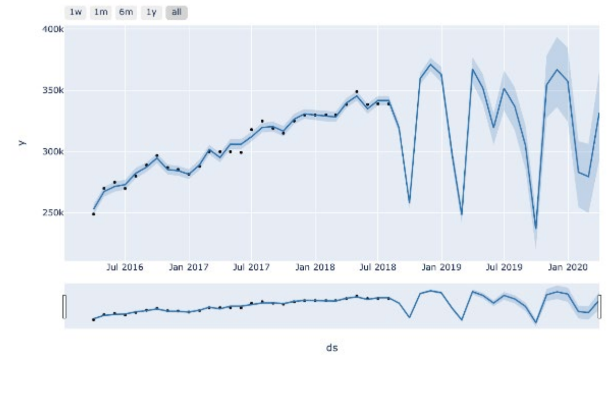
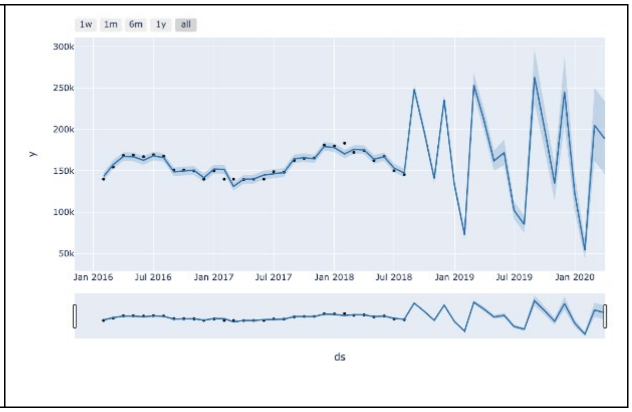


Wayne, KY

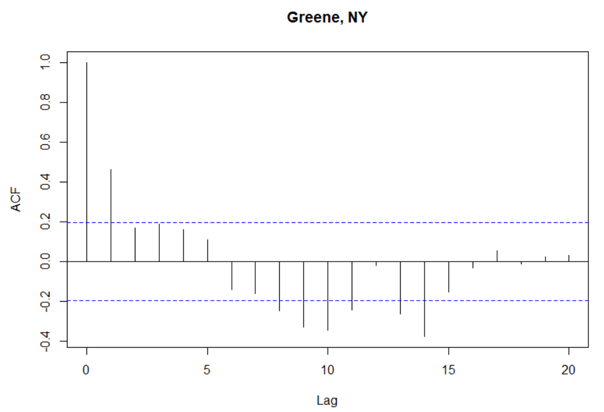
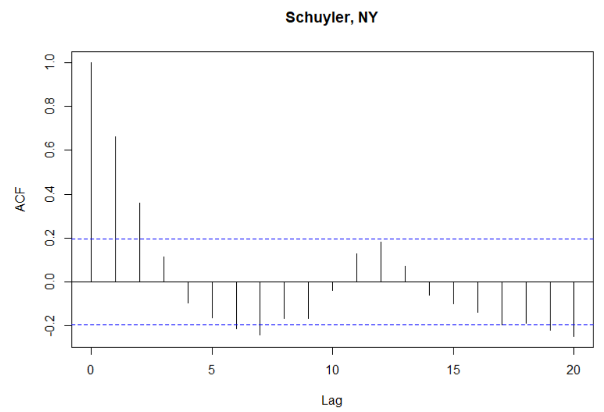
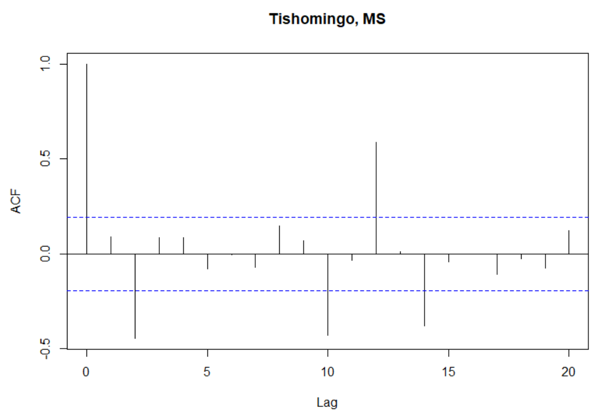
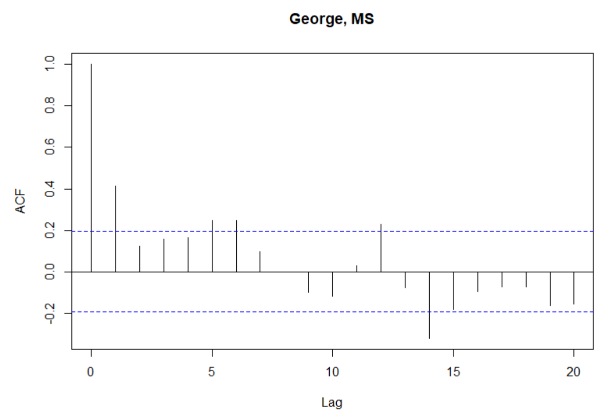


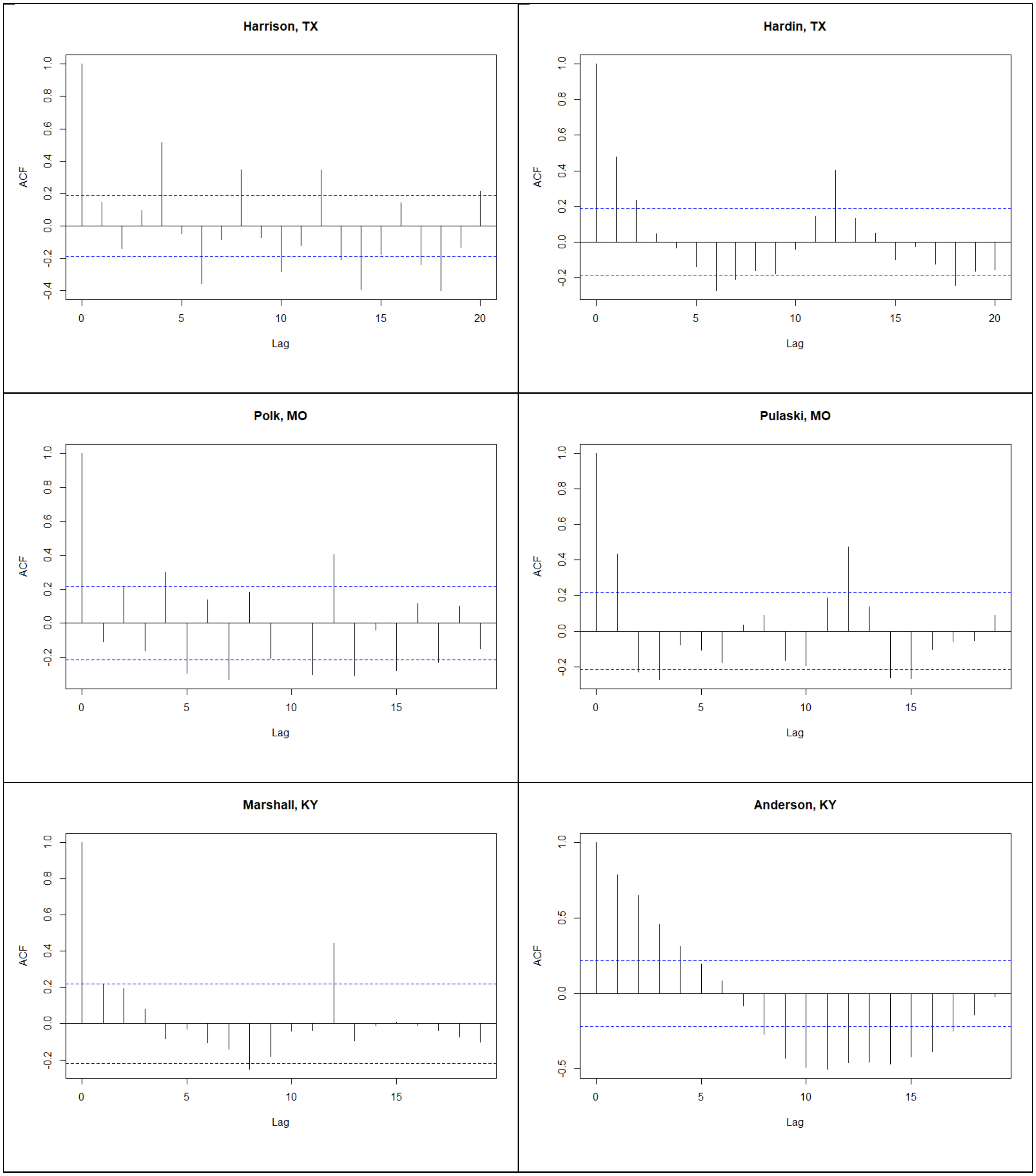
Gulf, FL

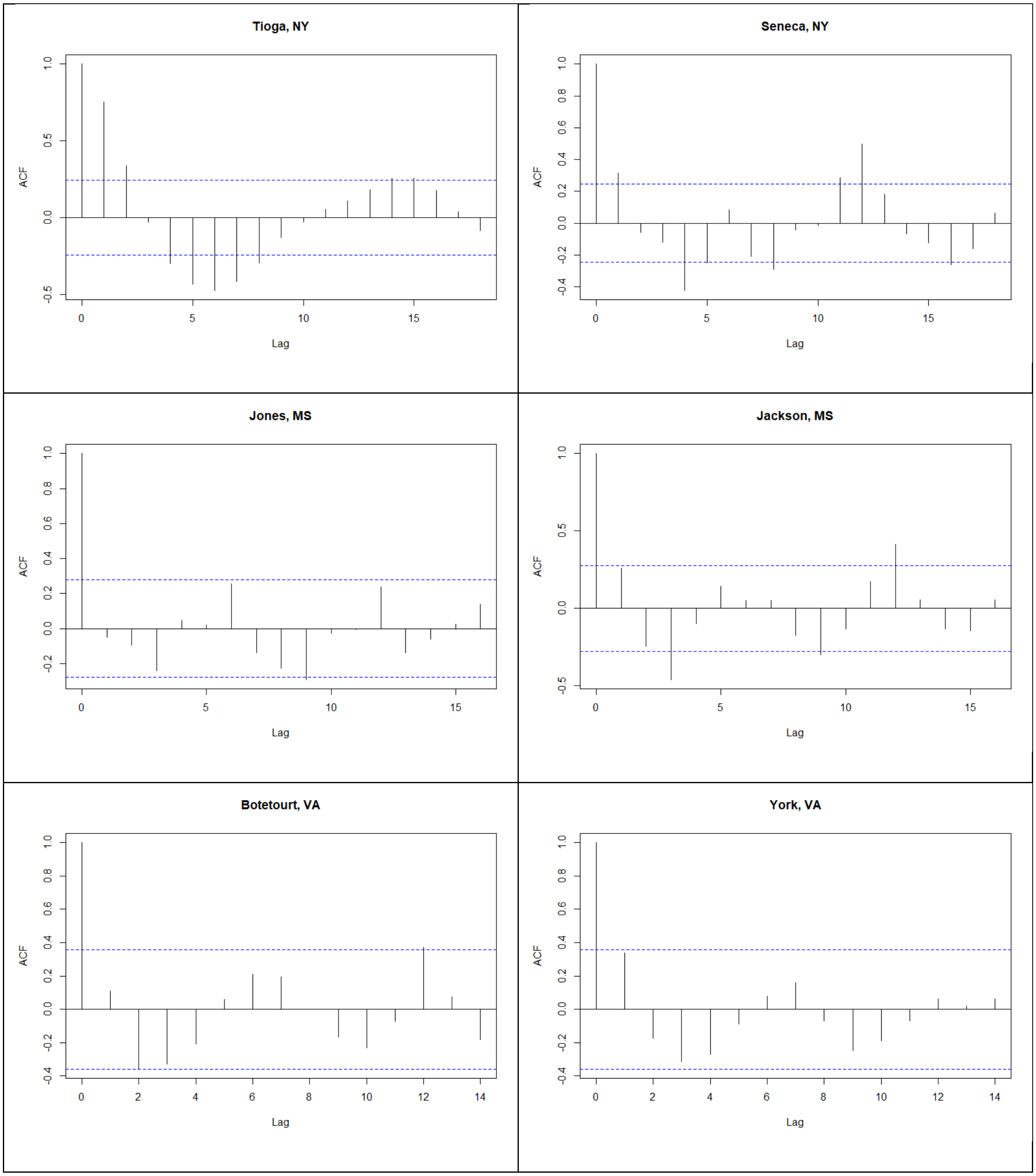
Glades, FL

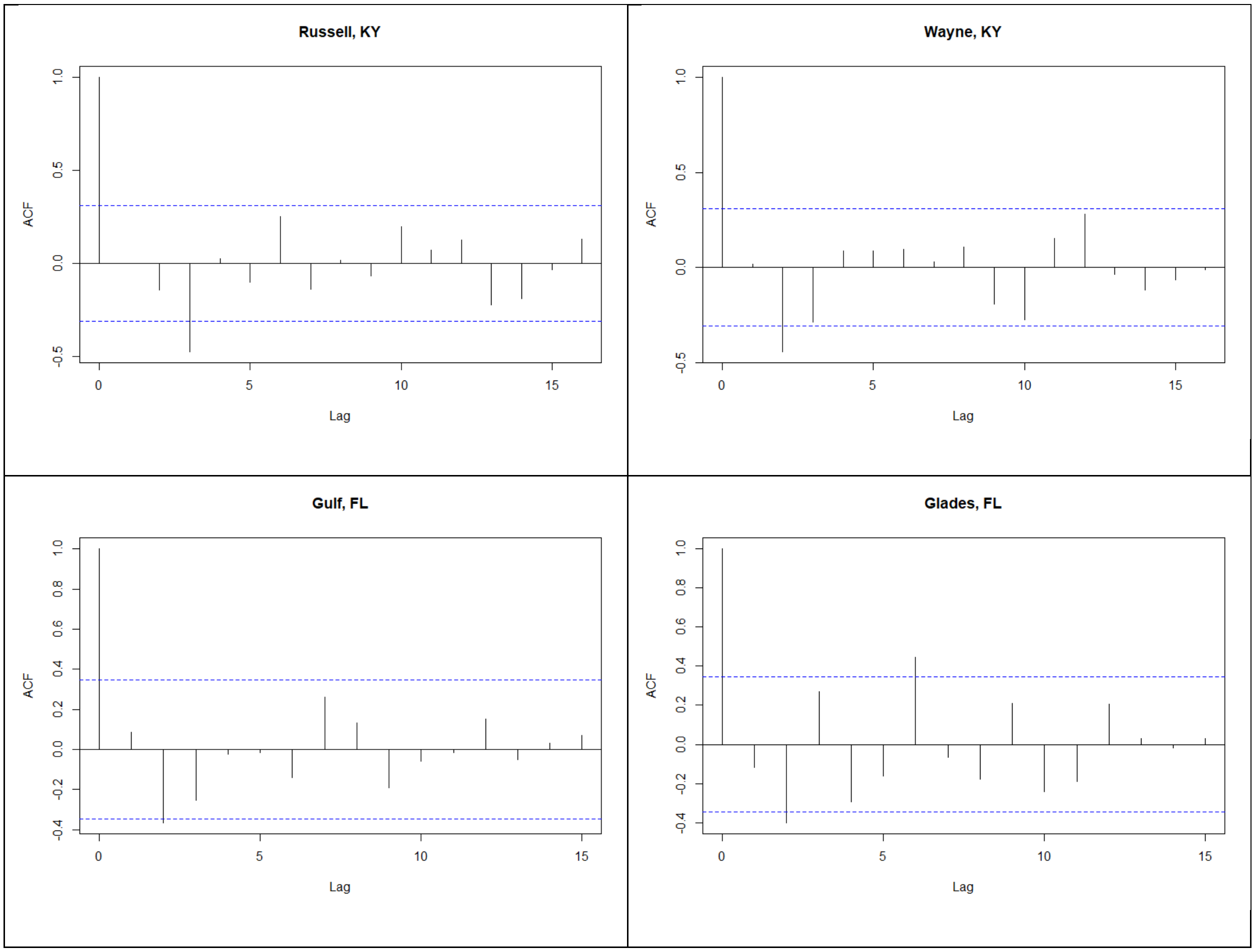
	

A2. Autocorrelation function (ACF)

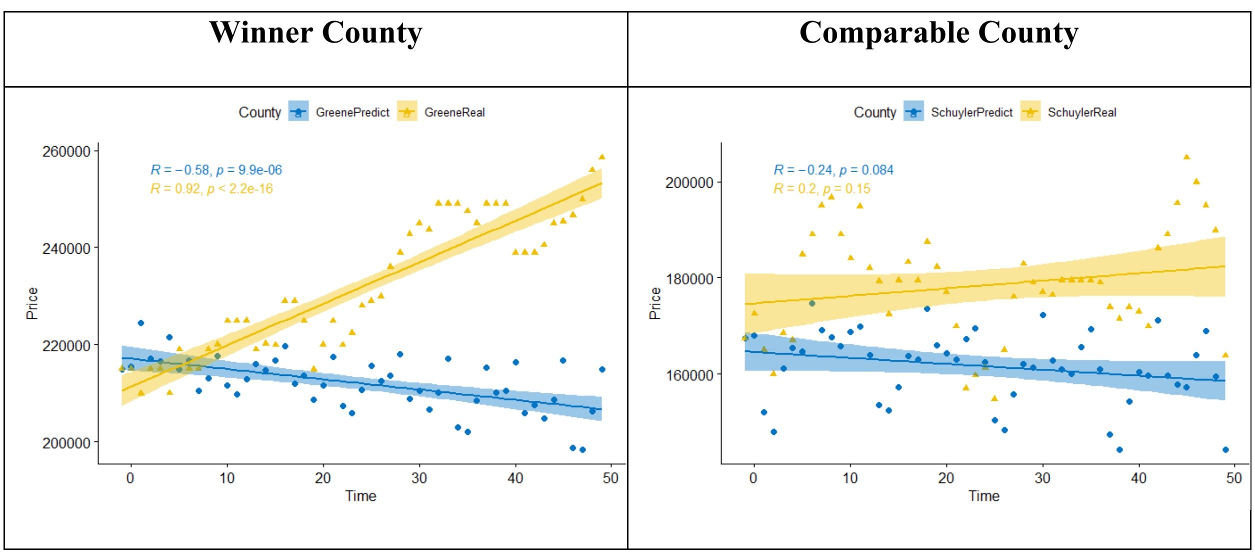
Winner County	**Comparable County**
	
	

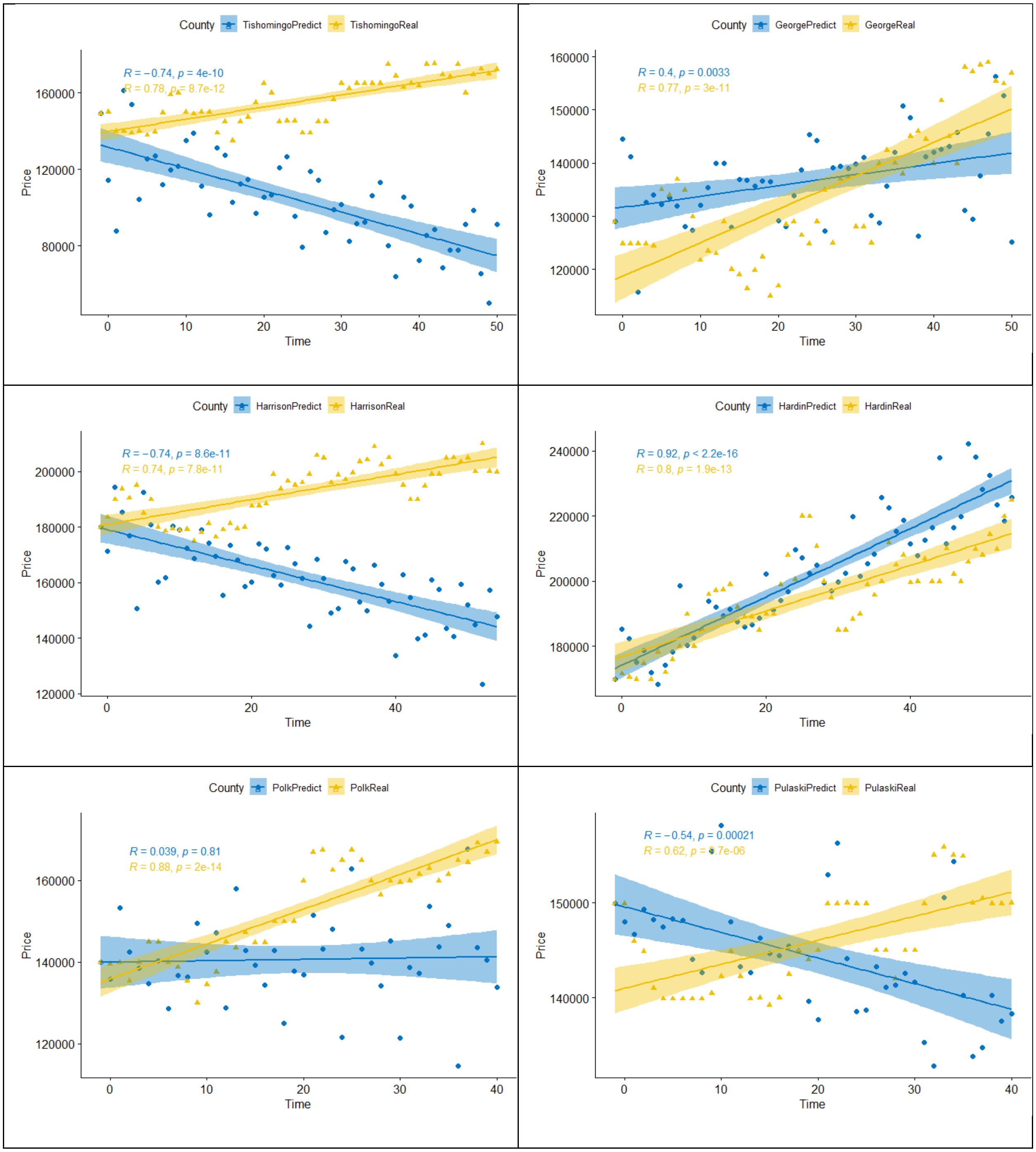

Harrison, TX
Hardin, TX
Polk, MO
Pulaski, MO
Marshall, KY
Anderson, KY
ACF
Lag

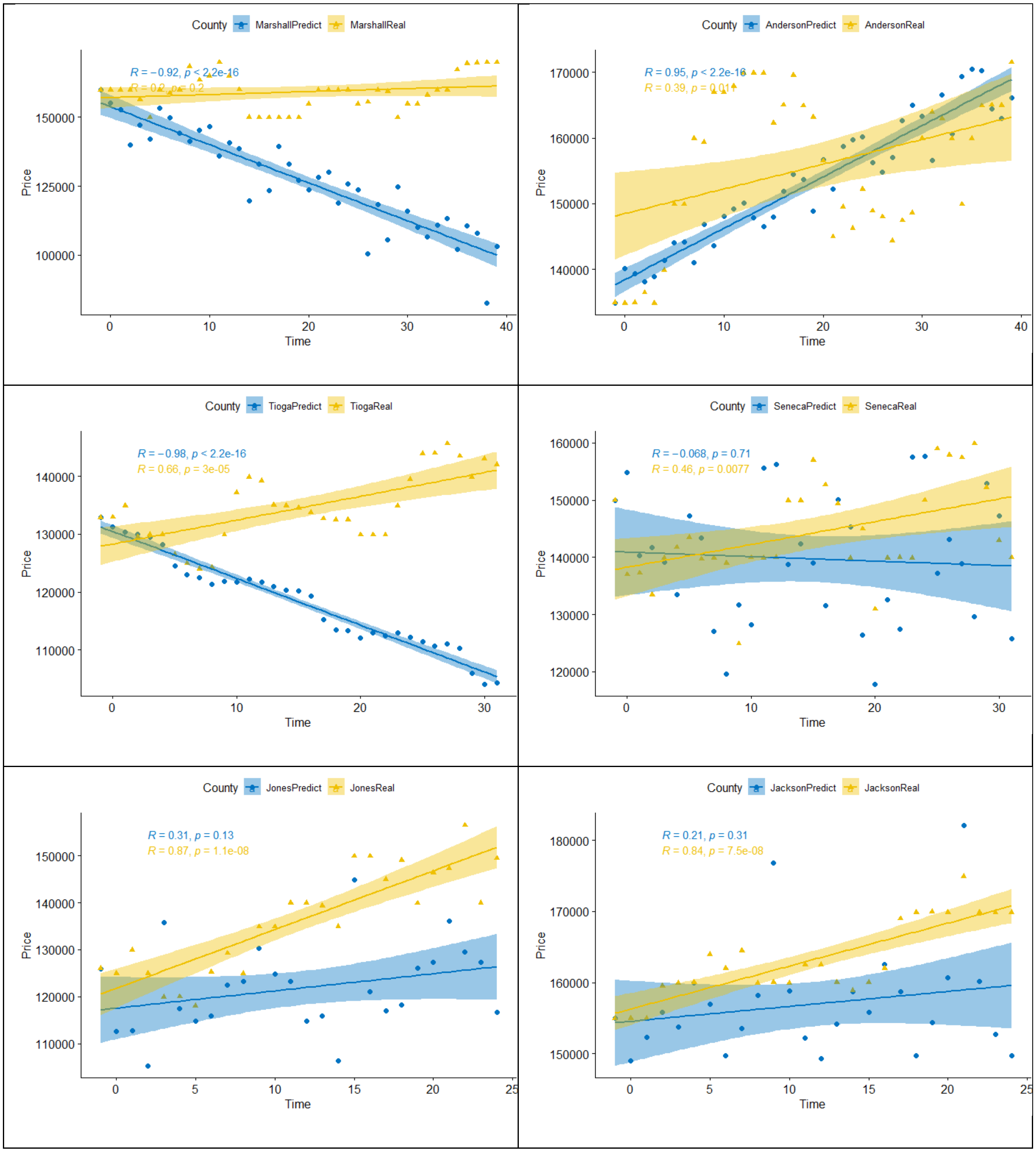

Tioga, NY
Seneca, NY
Jones, MS
Jackson, MS
Botetourt, VA
York, VA
ACF
Lag

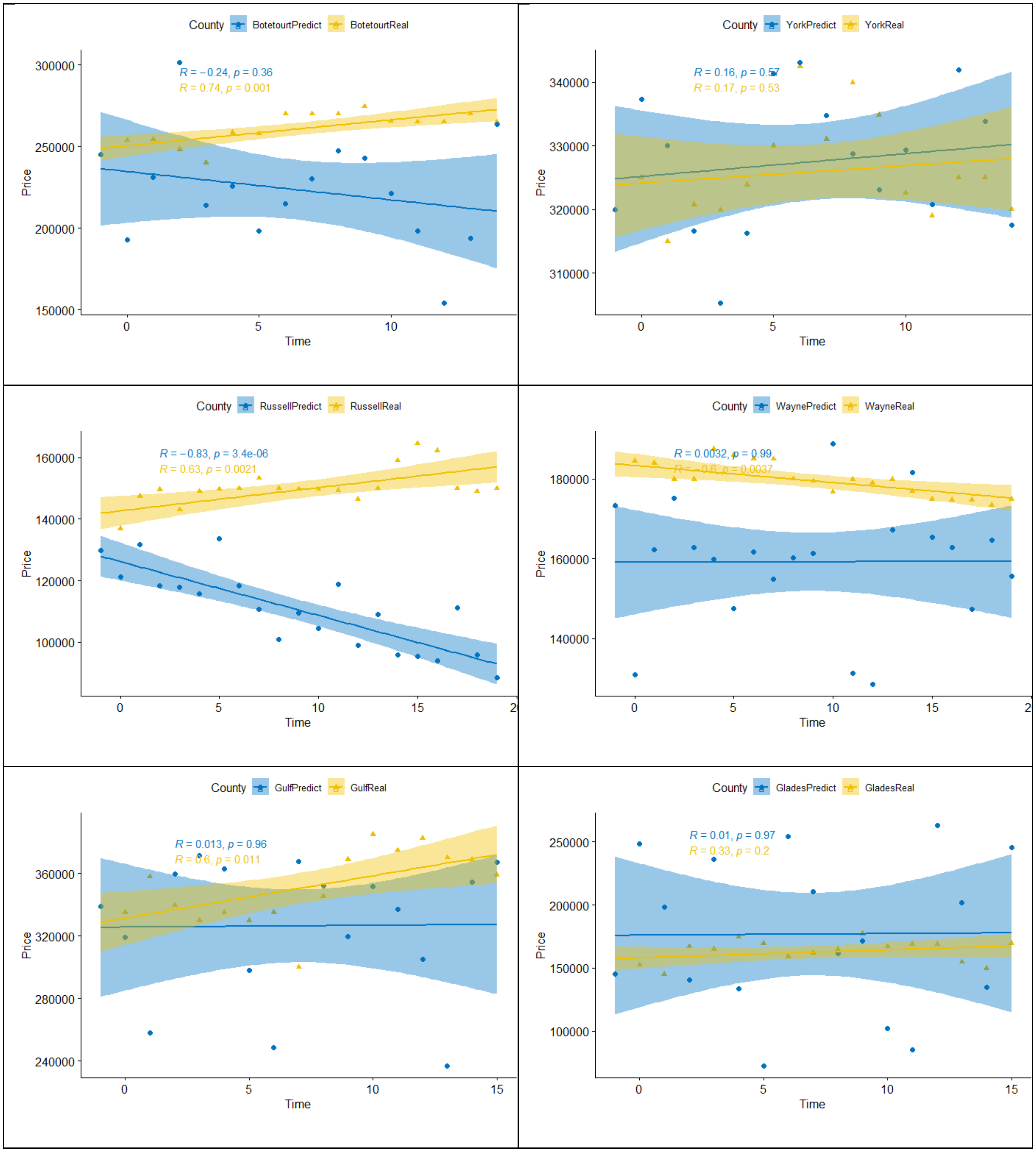


A3. Scatter plot of median housing prices and time




County TishomingoPredict TishomingoReal
R = −0.74, p = 4e-10
R = 0.78, p = 8.7e-12
Price
Time
County GeorgePredict GeorgeReal
R = 0.4, p = 0.0033
R = 0.77, p = 3e-11
Price
Time
County HarrisonPredict HarrisonReal
R = −0.74, p = 8.6e-11
R = 0.74, p = 7.8e-11
Price
Time
County HardinPredict HardinReal
R = 0.92, p < 2.2e-16
R = 0.8, p = 1.9e-13
Price
Time
County PolkPredict PolkReal
R = 0.039, p = 0.81
R = 0.88, p = 2e-14
Price
Time
County PulaskiPredict PulaskiReal
R = −0.54, p = 0.00021
R = 0.62, p = 1.7e-06
Price
Time


County MarshallPredict MarshallReal
R = −0.92, p < 2.2e-16
R = 0.2, p = 0.2
Price
Time
County AndersonPredict AndersonReal
R = 0.95, p < 2.2e-16
R = 0.39, p = 0.011
Price
Time
County TiogaPredict TiogaReal
R = −0.98, p < 2.2e-16
R = 0.66, p = 3e-05
Price
Time
County SenecaPredict SenecaReal
R = −0.068, p = 0.71
R = 0.46, p = 0.0077
Price
Time
County JonesPredict JonesReal
R = 0.31, p = 0.13
R = 0.87, p = 1.1e-08
Price
Time
County JacksonPredict JacksonReal
R = 0.21, p = 0.31
R = 0.84, p = 7.5e-08
Price
Time


County BotetourtPredict BotetourtReal
R = −0.24, p = 0.36
R = 0.74, p = 0.001
County YorkPredict YorkReal
R = 0.16, p = 0.57
R = 0.17, p = 0.53
County RussellPredict RussellReal
R = −0.83, p = 3.4e-06
R = 0.63, p = 0.0021
County WaynePredict WayneReal
R = 0.0032, p = 0.99
County GulfPredict GulfReal
R = 0.013, p = 0.96
R = 0.6, p = 0.011
County GladesPredict GladesReal
R = 0.01, p = 0.97
R = 0.33, p = 0.2
Price
Time