THE ECONOMICS OF GEOGRAPHIC AND DEMOGRAPHIC HETEROGENEITY IN DIGITALLY TRANSFORMED MARKETS

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Mohammed S. Alyakoob

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2019

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. Mohammad Rahman, Chair
Krannert School of Management
Dr. Prabuddha De
Krannert School of Management
Dr. Justin Tobias
Krannert School of Management
Dr. Kevin Mumford
Krannert School of Management

Approved by:
Yanjun Li
Head of Management Department, Krannert School of Management

To my amazing wife, Zainab.

Thank you for everything.

ACKNOWLEDGMENTS

I would like to acknowledge the dedication, support, and love provided by my wonderful wife, Zainab. Put simply, you are my inspiration. I am forever indebted to you for your constant and unconditional support. While being my primary source for editing and feedback, you somehow managed to find time to be an amazing mother and wife. All this while you completed your own Ph.D. You are truly incredible.

I also want to recognize the role my parents have had in shaping who I am today. My father is, and continues to be, my academic and intellectual inspiration. I thank you wholeheartedly for continuing to push me and always believing in my potential. The generosity, kindness, and unwavering support my mother has always shown me is a constant in my life that I have always depended upon.

I want to extend my profound gratitude to my committee members—Mohammad Rahman, Prabuddha De, Justin Tobias, and Kevin Mumford. Their guidance throughout this whole process has been invaluable. I will forever be in the debt of my advisor and committee chair, Mohammad Rahman. Working with Mohammad has been one of the great opportunities of my life. His insights, dedication, and passion for research are an inspiration to me and I hope that I will be able to live up the lofty standards he has set. I consider Mohammad not only an advisor and a mentor, but a dear friend and colleague whom I am privileged to have worked closely with. I will always be indebted to Prabuddha De for his wise council and guidance. Interacting constantly with him has provided me with a sincere example of service, intellectual curiosity, dedication to academic rigor, and honesty. I will always fondly remember his constant support and many enthralling conversations which have heavily shaped how I approach my academic career. Many thanks go to Justin Tobias, who provided me with the essential econometric training and insights that greatly improved the quality of my work. Kevin Mumford generously provided constant feedback, comments, and advice. I learned a great deal from his rigorously and methodologically examining approach to thinking through problems.

I would also like to thank my friends and colleagues who have always offered a helping hand and a friendly ear. I am grateful for all the time we have spent together and will always remember it fondly. A special thanks goes to my area colleagues: Abhishek, Vandith, Yipu, and Sameer. To my office mate, Anpu, I want to thank you for constantly engaging with me on my research ideas and on all the topics which we have discussed in the office.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Alyakoob, Mohammed S. PhD, Purdue University, August 2019. The Economics of Geographic and Demographic Heterogeneity in Digitally Transformed Markets. Major Professor: Mohammad Rahman.

The digital transformation of markets can remove traditional geographic restrictions, democratizing access to previously unattainable products, and enable individuals to extract rent from their personal assets. However, these digital innovations often have competitors and complementors that are not immune to the impact of local factors such as the local market structure, economic condition, and even demographics. This dissertation examines the geographic and demographic heterogeneity driven disparities in two digitally transformed markets, the financial and accommodations sectors respectively.

First, we study the impact of local financial market competition in managing online peer-to-peer loans. With the boom of financial technologies (FinTech), a critical question is whether the local financial market structure still matters. Unlike traditional retail financial institutions that are predominantly territorial, FinTech-based platforms, in particular peer-to-peer (P2P) lending, provide individuals equal access to funds by removing typical geographic restrictions. Combined with other benefits such as ease-of-use and lower interest rates, P2P lenders are increasingly threatening the traditional local lenders. A largely unanswered question in the literature is whether the local retail financial institutions strategically respond to the rise of such P2P platforms. Moreover, if the strategic reaction of traditional institutions continues the legacy of being territorial, borrowers will ultimately gain unevenly from the competition. That is, where a borrower lives may still matter. In this chapter, we devise multiple strategies to empirically analyze the extent and nature of the

strategic response of traditional institutions to P2P lending. This includes: (1) utilization of a Probit model that leverages the richness of our local market data and (2) exploitation of bank mergers as exogenous shocks to local market structure. We find consistently that a borrower from a more competitive market is more likely to prepay, suggesting that local market structure plays a pivotal role in P2P borrowers' debt management. We validate the underlying mechanism by studying the improving credit profiles of borrowers and platforms' (exogenous) changes in pricing in moderating the main effect. This mechanism reveals that traditional banks, especially when their local market conditions support, credibly responds to the growth of P2P and are successful in attracting consumers back to traditional financial products. Relatedly, we document heterogeneity in the benefits that borrowers gain from the local market structure (using a machine learning algorithm) and verify the robustness of our main findings. We discuss the implications for P2P lending, other crowd-based markets, and local retail financial markets.

Second, we examines the heterogeneous economic spillover effects of a home sharing platform—Airbnb—on the growth of a complimentary local service—restaurants. By circumventing traditional land-use regulation and providing access to underutilized inventory, Airbnb is attracting visitors of a city to vicinities that are not traditional tourist destinations. Although visitors generally bring significant spending power, it is, however, not clear if the visitors use Airbnb primarily for lodging, thus, not contributing to the adjacent vicinity economy. To evaluate this, we focus on the impact of Airbnb on the restaurant employment growth across vicinities in New York City (NYC). Our results indicate that if the intensity of Airbnb activity (Airbnb reviews per household) increases by 1%, the restaurant employment in an average area grows by approximately 1.03%. We also investigate the role of demographics and market concentration in driving the variation. Notably, restaurants in areas with a relatively high number of Black residents do not benefit from the economic spillover of Airbnb activity. Also, restaurants in more competitive areas reap the benefit from this spillover most. We validate the underlying mechanism behind the main result by evaluating the impact of Airbnb on Yelp visitor reviews – areas with increasing Airbnb activity experience a surge in their share of NYC visitor reviews. This result is further validated by evaluating the impact of a unique Airbnb neighborhood level policy recently implemented in New Orleans.

1. INTRODUCTION

The widespread adoption of Internet services has reduced the traditional search costs that prevented wide scale consumer to consumer transactions (Bakos, 1997). This reduction in search costs has spawned innovate companies that have challenged traditional business models in various industries, including accommodations, financial, and transportation (Einav et al., 2016; Zervas et al., 2017). These digital platforms can remove traditional geographic restrictions, democratizing access to previously unattainable products, and enable individuals to extract rent from their personal assets. However, they also have competitors and complementors that are often not immune from the impact of local factors such as the local market structure, economic condition, and even demographics. This dissertation consists of two essays which examine the geographic and demographic heterogeneity driven disparities in digitally transformed markets.

In Chapter 2, we examine the competitive dynamics between online P2P lending platforms and traditional, locally embedded, financial institutions in the US. We study the impact of local financial market structure on borrowers personal loan management decisions. By studying these debt management choices, we are able to examine the extent and nature of substitution between traditional banks and online lending/funding platforms. We find that P2P borrowers from more competitive local markets are more likely to prepay on their online loan. This indicates that the banking competition in a locality dictates the competitive/strategic responses of banks in that area to the online alternative provided by P2P lending. Particularly, P2P borrowers from more competitive banking markets are more likely to substitute back to traditional banking entities. We provide an extensive empirical evaluation that leverages bank mergers as an exogenous shock to local banking competition. We delve into the underlying mechanism of our results by exploiting another natural experiment relating to changes in platform assigned interest rates for otherwise equivalent borrowers. We use these differences in assigned interest rates to establish that borrowers with higher interest rates are more likely to substitute their loans if they live in a competitive banking market. To further understand the intricacies driving the mechanism, we use machine learning tools (specifically, Causal Forests) to examine the heterogeneity of these results as they relate to a borrowers credit worthiness and loan amount.

In Chapter 3, we focus on the heterogeneous economic spillover effects of home sharing platforms on the growth of complimentary local services. Since home sharing platforms circumvent traditional land use regulations, they allow visitors to a city to locate in areas that would otherwise have been inaccessible. Visitors that choose to locate in these sharing economy enabled areas have two options. On the one hand, they may exploit the area in which they are lodging strictly for accommodation purposes and commute to more traditional tourist locations. As a result, they will spend their non-accommodation based tourism dollars in the traditional tourist locations. On the other hand, they may go beyond staying in an area and spend their tourism dollars locally. To evaluate the economic spillover eect of this spending, we focus on restaurants in New York City (NYC). Our results indicate that if the intensity of Airbnb activity (Airbnb reviews per household) increases by 1%, the restaurant employment in that neighborhood grows by approximately 1.03%. Notably, this beneficial economic spillover effect does not extend to majority Black or Hispanic neighborhoods in NYC. To empirically identify this effect, we collected Airbnb, Yelp, and a plethora of local variables from various sources to create an expansive data set. We employ a difference-in-difference (DID) specification that leverages the temporal and spatial distribution of Airbnb entry and intensity in NYC neighborhoods that are not traditionally considered tourist locations. We corroborate our main findings by implementing various algorithmic matching techniques and an extensive study of the necessary underlying mechanism. Specifically, this result requires that Airbnb redistributed visitors are frequenting restaurants located in the vicinity of their Airbnb lodgings. Therefore, using our comprehensive Yelp data collection, we assess the impact of Airbnb activity on Yelp visitors restaurant review behavior. The extensive data collection process enables us to identify the location of Yelp reviewers to establish whether they are NYC residents or visitors.

This result is further validated by evaluating the impact of a unique Airbnb neighborhood level policy driven natural experiment in New Orleans. In 2017, New Orleans implemented regulation whereby Airbnb was deemed illegal in one neighborhood while it was ocially legalized in adjacent neighborhoods. The policy shift caused an increase in the Airbnb activity in the legalized neighborhood and a decrease in the now banned neighborhood. An examination of the Yelp restaurant reviews written by visitors in the overlapping time indicates a clear shift in restaurant activity towards the areas where Airbnb activity was legalized. To delineate the intricacies of Airbnbs eect across localities, we also investigated the role of demographics and market concentration in driving the variation. Our results suggest that both demographics and market structure have an important role in determining the areas that benet from the economic spillover of Airbnb. Notably, restaurants in areas with a relatively high number of Black residents or a relatively high number of Hispanic residents do not benet from the economic spillover of Airbnb activity. This trend continues when expanding the results to cities beyond NYC, especially as it pertains to the lack of spillover aect in majority Black localities. An exception to this trend is Los Angeles, which has 49% Hispanic population; the impact of home sharing on restaurant employment does extend to majority Hispanic areas. For the market structure heterogeneity analysis, we use local Yelp reviews to identify the concentration of restaurant activity in certain areas. We nd that in areas where a few restaurants capture the majority of local Yelp reviewshigh concentration areasthe impact of Airbnb on restaurant employment is diminished.

2. WHERE YOU LIVE MATTERS: THE IMPACT OF LOCAL FINANCIAL MARKET COMPETITION IN MANAGING PEER-TO-PEER LOANS

2.1 Introduction

Retail financial institutions predominately operate within specific geographic territories and, consequently, offer heterogeneous services and opportunities to consumers (Amel and Starr-McCluer, 2002; Cyrnak and Hannan, 1999). Even the regulatory changes that removed intra-state banking restrictions did not significantly change the importance of boundaries, especially in reaping the benefits of retail bank competition (Amel and Starr-McCluer, 2002; Cyrnak and Hannan, 1999).[1] The recent advent of financial technologies (FinTech), e.g., peer-to-peer (P2P) lending (Lending Club and Prosper), social mobile payments (Venmo and Zelle), and brokerage (Robinhood), challenges these notions. In particular, these platforms transcend specific localities and provide consumers from across the country with equal access. However, if the strategic reaction of traditional institutions continues the legacy of being territorial, borrowers will ultimately gain unevenly from the competition between their local financial institutions and FinTech platforms.

In this study, we examine the competitive dynamics between online P2P lending platforms and traditional, locally embedded, financial institutions in the US.[2] Specifically, we focus on *the role of local competition in driving strategic responses of the traditional banking markets to the growth of online P2P lending.* Online P2P lending platforms compete with local commercial banks and credit unions by facilitating the

[1] In the hallmark case, *United States vs. Philadelphia National Bank (1963)*, the Supreme Court opined that banks operate in geographic "clusters" and any evaluation of the impact of mergers must focus on the competition within these clusters.

[2] We focus on the largest P2P lending platform in the US, Lending Club, which has, according to Lending Club's 10-K filings, originated $10.9 billion worth of loans in 2018.

funding of unsecured personal loans.[3] The majority of borrowers (over 80%) who have utilized P2P platforms describe their loan purpose as a form of debt reconciliation, often specifically focused on reconciling credit card debts. As such, the bulk of P2P borrowers have migrated from traditional financial offerings—being credit card debtors who usually pay high interest rates—to P2P platforms. This migration is occurring at an unprecedented rate—the share of the personal loan market that can be attributed to FinTech lenders skyrocketed from 1% in 2010 to 34% in 2016 (see Figure 2.1).

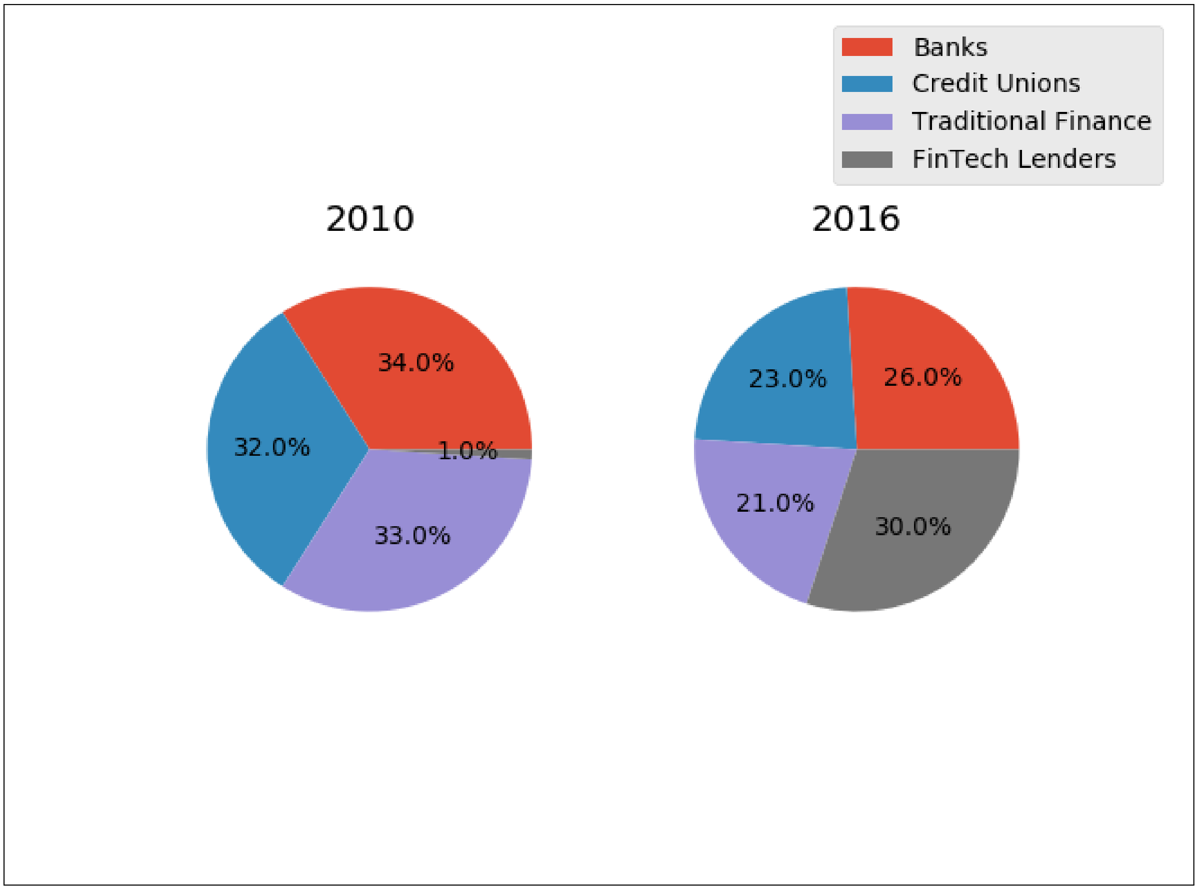


Fig. 2.1. The Change in Unsecured Personal Loan Market Shares

[3]We use “P2P lending platforms,” “P2P platforms,” and “P2P lenders” interchangeably throughout the paper.

P2P lending platforms are widely believed to provide certain advantages over local banking institutions for borrowers. According to Lending Club, their borrowers utilize the platform because of its convenience and potentially lower interest rates. First, regarding the convenience of obtaining personal loans, Lending Club states in their 10-k filings that they "offer a fast and easy-to-use online and mobile application process and provide borrowers with access to live support and online tools." Second, the filings state that their "proprietary lending marketplace model, easily accessible online delivery and process automation enable us to offer a wide range of borrowers interest rates that are lower on average than the rates charged by banks for credit cards." While it is conceivable that borrowers may also have access to the funds through their local financial institutions, the advantages provided by P2P platforms (i.e., convenience and low costs) generally circumvent traditional barriers to credit switching, such as effort and information asymmetry. Also, traditional financial institutions conventionally do not consolidate debts with personal loans.

Notably, unlike traditional lenders, P2P platforms transcend geographic boundaries. Specifically, the platforms assign borrowers with similar credit profiles with the same interest rate, regardless of their locations. For example, Figure 2.1 draws the interest rate Lending Club assigned to borrowers with a credit grade "D1" in June of 2011, 2012, 2013, and 2014 respectively.[4] The graph shows that in all four months, borrowers all received the same interest rate regardless of the number of banking institutions in their locality. In sharp contrast, the interest rates offered by traditional lenders largely depend on the competitive environment in a specific locality. Borrowers from more competitive markets have access to funds with more favorable terms (Hannan and Prager, 2004; Dick, 2007; Degryse and Ongena, 2005).

The growth and maturity of P2P lending—combined with its structural necessity to offer homogeneous rates across regions as well as the market opportunity it creates for the personal loan divisions of traditional institutions—would arguably lead to strategic reactions from traditional retail financial institutions. By losing potential

[4]Lending Club uses "sub-grades" to categorize borrowers based on their credit profiles and borrowing history. Section 2.4 provides more details of the grading system.

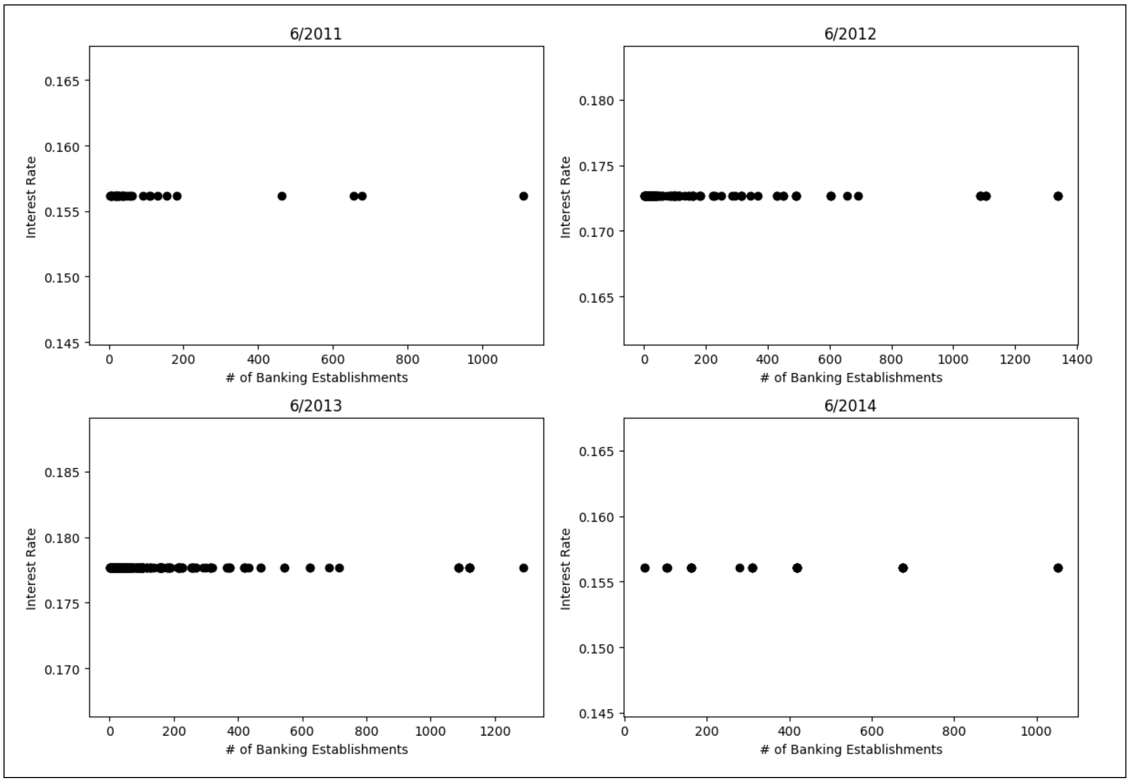


Fig. 2.2. Interest Rates Assigned to "D1" Borrowers on Lending Club

consumers to P2P lenders and thus leaving money on the table, the local lenders may, first, compete directly by adjusting their interest rates and other terms. More importantly, local lenders may seek to convert P2P borrowers back to local markets by, e.g., refinancing P2P loans with more favorable terms. This is particularly lucrative once a borrower shifts to a personal loan product, eliminating any prior cannibalization threats between different lending arms of a traditional retail bank.

These strategic reactions would naturally not be homogeneous across geographies. Local lenders in more competitive markets are poised to react more quickly and more extensively to the rise of P2P lenders. As such, considering the other side of the market, P2P borrowers benefit disproportionately from the local market competition in their debt management—those who reside in more competitive local financial markets benefit more. In other words, if P2P platforms are not immune to the competition from local retail financial institutions, borrowers may ultimately gain from the competition (disproportionately). That is, where a borrower lives may still matter. P2P borrowers from more competitive local markets will find them in better positions because they may find more superior terms in refinancing their P2P loans (or other debts). This advantage from local competition translates to, ceteris paribus, higher prepayment rates of P2P loans in more competitive localities. Hence, P2P borrowers' prepayment behavior is an ideal proxy for analyzing the local financial institution's competitive response to FinTech alternatives. Consequently, we seek answers to the question of *whether borrowers of online P2P lending platforms systematically vary in their propensity to pay off early based on their local market structure*. We also *isolate the factors related to local market structure that dictates the variations in the prepayment of P2P loans*.

We combine rich data on individual loan transactions and payment histories from Lending Club with an extensive set of variables of local market structure, economic conditions, and demographic information.[5] The Lending Club data provides a comprehensive collection of information on each borrowers' characteristics including, among others, credit score, credit history, debt-to-income ratio, past delinquencies, loan purpose, and income. The depth of the individual level controls, in combination with a broad set of local economic and demographic information, enable us to identify the impact of market structure on the prepayment decisions of P2P borrowers. Furthermore, we complement this analysis by exploiting exogenous shocks to an area's banking market structure induced by bank mergers to identify the causal impact of these changes on P2P borrowers' debt management behavior (Nguyen, 2019; Liebersohn, 2017).

Our main results suggest that the number of local lending institutions—as a proxy for the local market competition—has a non-trivial impact on the prepayment behavior of P2P borrowers. In particular, a borrower from a more competitive market, i.e., more lending institutions in the borrower's locality, is more likely to prepay his or her P2P loan. That is to say, the benefits accrued from this alternative source

[5]We verify our main findings using a similar data set obtained from Prosper.com—the second largest P2P platform in the US. We find qualitatively similar results across the two P2P platforms.

of financing remain dependent on geography-based competitive dynamics. We find qualitatively similar results using bank mergers as exogenous shocks to the local market competition (or structure). Specifically, we find that bank mergers, which often lead to a reduction in local market competition, are negatively associated with P2P borrowers' likelihood of prepayment. Reassuringly, these findings remain unchanged across different specifications and robustness checks.

We delve further into the main findings by assessing the heterogeneity in the impact of local market structure by lending institution types and borrower characteristics. First, we examine the role of the type of local institutions prevalent in a borrower's locality and its moderation effects on our results. Because different lending institutions capitalize on different tools, the dispersion in bank type distribution in a locality will also determine the reaction to online alternatives. On the one hand, localized (and often smaller) banks incorporate "soft information" in their lending decision processes (Cole et al., 2004; Berger et al., 2005), such as the characteristics of a borrower that are observable only through on-site interviews. This informational advantage is location specific. But these localized banks' competitive advantage from relationship lending may not be useful to P2P borrowers, because these online borrowers are unlikely to be relationship-based. This is evident by the fact that they initially choose P2P platforms to fulfill their financial needs. On the other hand, larger institutions that cover more geographic markets often rely more on their credit models, e.g., algorithmic decision making and marketing campaigns, other than those relationship-based factors.

In our analysis, we find that the presence of large non-community banks is more beneficial to P2P borrowers than the presence of community banks in obtaining funds to prepay on their P2P loans, which suggests that the large non-community banks are more likely to respond to the rise of P2P alternatives than the smaller community banks. Besides, we examine the heterogeneity in outcomes in terms of borrower characteristics. We find that the impact of local market structure on P2P borrowers' prepayment behavior is amplified when borrowers have significantly improved their credit profiles (e.g., increase in their FICO scores). These findings indicate that P2P platforms may initially be able to entice borrowers that are unattractive to local markets, but that these borrowers are more likely to return to the local market offerings once their profiles improve.

The remainder of the chapter is organized as follows. We summarize the literature and highlight our contributions in Section 2.2. Section 2.3 introduces our research context—P2P lending and Lending Club—and the data set we obtained. Then we present our main empirical findings in Section 2.4 and further evidence in Section 2.5. Section 2.6 concludes and discusses the managerial implications of our findings.

2.2 Literature and Contributions

We first contribute to the burgeoning literature on FinTech in general and P2P lending more specifically. The studies on P2P lending thus far have primarily focused on investor behaviors and, subsequently, funding outcomes and market efficiency (*e.g.*, Pope and Sydnor, 2011; Iyer et al., 2015; Berkovich, 2011; Lin et al., 2013; Lin and Viswanathan, 2016).[6] Tang (2019) find that P2P platforms are substitutes for traditional banks, especially for borrowers with lower credit grades. They utilize changes in bank regulation in 2010 that caused traditional banks to tighten credit access to identify this impact. We extend the literature by focusing on P2P borrower's debt management behavior. To the best of our knowledge, we provide the first evidence that local retail financial institutions strategically respond to the rise of P2P lend-

[6]Pope and Sydnor (2011) find evidence of a disparity in funding due to the race and attractiveness of the borrower. Iyer et al. (2015) find that lenders are able to ascertain more specific credit scores than those provided by the P2P lending platform (the platform generally provides a range as opposed to a specific number) by utilizing factors that are not usually used in traditional financial analysis (they do not consider a borrower's local lending options). Berkovich (2011) and Zhang and Liu (2012) find evidence of herding behavior by lenders on the P2P lending platforms. Lin et al. (2013) investigate the role of friendships (using group memberships on Prosper) on the probability of funding a loan, obtaining lower interest rates, and also lowering default rates. Lin and Viswanathan (2016) find evidence of home bias in the online lending setting implying that lenders are more inclined to lend money to borrowers that reside in the lender's state of residence.

ing (see, e.g., Morse (2015), for a more comprehensive review of the literature).[7] A related study, Butler et al. (2016), explores the association between local access to finance and P2P loan applications. In particular, they examine how the local market structure affects the desired term of individual borrowers' P2P loan applications on a short-lived auction mechanism of Prosper.com. They, however, do not provide any insights related to whether or not the local lending institutions react strategically to the growth of P2P lending and how such reaction may vary in exploiting the structural limitations of P2P platforms. Also, our findings complement their initial understanding of the role of local market structure by identifying the types of borrowers who stand to gain most from traditional financial institutions as well as how the composition of the traditional local market structure plays a role in strategically reacting to the rise of P2P.

Our study also contributes to the extensive literature on banking competition across disciplines. We provide evidence that recent technological advancements, specifically crowd-based P2P lending, have not removed the boundaries within which retail lending institutions compete. Regulations over the past three decades have lifted restrictions on larger banks that prevented them from opening branches in specific locations. Much of the literature focuses on whether this would remove the geographic boundaries within which competition was defined. Elliehausen and Wolken (1992), Kwast et al. (1997), and Amel and Starr-McCluer (2002) use periodic household surveys on consumer finance usage and find that consumers continue to rely on their local banks for their financial needs. Cyrnak and Hannan (1999) reinforce these finding with evidence from the market of small business loans. In contrast, Petersen and Rajan (2002) find that the distance between small businesses and their credit providers is increasing. Thus, while the local retail banking market's importance for depository services is well defined, its role regarding credit offering services is not as clear. We complement these studies by providing detailed transaction level evidence that these boundaries remain valid.

[7] Li et al. (2016) focus on predicting both prepayment and default risks using individual borrower characteristics, which is akin to the risk measures currently provided by the platforms.

2.3 Research Context and Data

2.3.1 Peer-to-Peer Lending

P2P lending is generally considered a revolutionary form of consumer loans and, more importantly, act as an alternative to local lending institutions. P2P platforms primarily facilitate the online exchange of money between individual lenders and borrowers. The first step of a typical funding process is that a potential borrower places a loan request on the platform. The borrower specifies an amount of money and provides information such as a description of the loan, debt, income, home ownership status, and other credit information. Then the platform examines the borrower's credentials and assigns an interest rate. As soon as the loan request is listed on the platform, individual lenders will decide whether and how much to lend to the borrower. Lenders can fund a portion of the loan (most lenders invest less than $500 in a specific loan). If the listing is funded within a certain period, it converts to a fully amortized personal loan, and the borrower receives the money. Typically, P2P platforms facilitate the transfer of funds by working with a third-party bank. The borrower then starts making monthly payments to the platform, which then distributes the money to the appropriate lenders.

During the entire payment process, which is typically 36 months, the borrower has the option to pay off the remaining debt at any given time. This form of prepayment is of interest to us.[8] The borrower generally does not pay any penalty for prepayment on P2P lending platforms. A borrower may also default, which occurs when the borrower fails to make the required monthly payments after a certain period. The platform will then assign an agency to collect the money from the borrower. Regardless of the outcome of the collection process, this loan is declared as default

[8] If the borrower prepays by paying an amount greater than the monthly required amount but does not prepay all the remaining principal, the monthly payment remains the same, but the length of the loan will be adjusted. The personal loans are amortized, which means that if the borrower prepays a portion of the loan, the total interest they pay over the life of the loan is reduced. We define a prepayment as a loan that is fully paid off after the contracted date of the 6th payment and before half the contracted payments have expired. Furthermore, the amount of the prepayment should be greater than 35% of the loan amount. Later, we perform robustness checks regarding this definition.

2.3.2 Data and Samples

We obtain our main data from the largest P2P lending platform in the US: LendingClub.com.[9] Lending Club (LC) has facilitated the funding of over 3 million personal loans by June 2018. In our analysis, we use transactions data of loans on LendingClub.com that originated from January 2011 to March 2014.[10] All listings on LC were successfully funded during this period. Furthermore, loans that are listed as "current" (not yet completed) in our data set are removed as their final payment results are not available. We also limit our sample to only those loans with a predetermined maturity date before our data collection date. Therefore, loans with a late maturity date but having been prepaid are also dropped for consistency. After sanitizing, we have 187,122 loans and their associated payment outcomes. Table 2.1 reports the summary statistics of key variables.

Table 2.1.: Variables and Summary Statistics

			Mean	**SD**
Prepaid(1/0)			0.20	0.40
	Sub Grade			
		A1	0.19	0.40
		B1	0.19	0.40
		C1	0.19	0.39
		D1	0.18	0.38
Default(1/0)			0.12	0.33
	Sub Grade			
		A1	0.03	0.16
		B1	0.08	0.26
		C1	0.13	0.34

continued on next page

[9]We complement our main analysis using data from the other major P2P lending platform in the US—Prosper.com.

[10]On March 31, 2014, Lending Club stopped providing the borrower's city.

Table 2.1.: *continued*

		Mean	**SD**
	D1	0.18	0.39
Lending Institutions		149.92	252.60
Business Degree		0.20	0.05
Race: Black		0.16	0.15
Origin: Hispanic		0.21	0.19
Unemployment Momentum		-0.09	0.04
Poverty		0.17	0.08
Lender-Population Ratio		0.74	1.08
Rent Adjusted Income		0.19	0.05
Amount Requested		12.86	7.73
IR(%)		0.13	0.04
Income		70.76	55.40
DTI(%)		0.16	0.08
Home-ownership		0.55	0.50
Bank Card Utilization		0.56	0.33
Current Credit Lines		10.78	4.69
Delinquencies		0.24	0.72
Inquiries Last 6 Months		0.79	1.02
Years Employed		5.65	3.60
Public Records in Last 10 Years		0.13	0.48
Revolving Credit Balance		15.05	18.01

For each loan, we observe the borrower's city, state, loan amount, debt to income ratio, income, home-ownership status, and other information related to the borrower's credit rating such as the number of delinquencies and current credit accounts. Lending

Club also assigns a credit grade for each borrower based on the credit profile. Besides, we obtain the current status for each loan. It can be one of the two states—default or paid. We also gather details on the amount of each payment made during the life cycle of a loan and the changes in FICO score associated with each borrower.

We complement the LC data with information about the market structure of the local lending industry, demographics, and economic conditions. We obtain the local lending market structure from the Business Pattern Data maintained by the US Census Bureau. Specifically, we use the number of credit offering institutions in each zip code as the measure of local market structure. Credit offering institutions include commercial banks, savings institutions, and credit unions. Also from the Census Bureau, we obtain demographic and educational attainment data. From the Bureau of Labor Statistics (BLS), we obtain historical annual unemployment data. Lastly, we get fair market rent data from the U.S. Department of Housing and Urban Development (HUD) as a proxy for the local cost of living.

To better reflect the access to local finance opportunities and other local area characteristics, it is desirable to construct area specific measurements at as granular a level as possible. Lending Club reports the city where a borrower resides in, which corresponds to the "Places" from U.S. Census Bureau, and also 3-digit zip code. We use the finer of these two different aggregation levels in calculating all local area specific measurements. In the rest of the paper, we refer to this as the location of the borrower. Because not all data are pre-processed at the desired location level, we streamline the measurement issues by employing a granular matching system using ArcGIS (Version 10.4). Briefly, we obtain the census tract population data from the U.S. Census Bureau. Census tracts are small subdivisions, with a population of generally around 4,000 residents. For each geographic level, we split each observation of that geographic level into census tracts. Then we use the matched census tracts to calculate population weighted variables. Governed by the granularity of our P2P borrower information, we aggregate other information from various sources at the borrower's location level using this method.

2.4 Empirical Analysis

We present our empirical findings in this section. Section 2.4.1 introduces our main empirical specification and reports the main results (based on various measures of local market structures). In Section 2.4.2, we report further evidence from bank mergers as exogenous shocks to local market structures. We present evidence demonstrating the underlying mechanisms in Section 2.4.3. Robustness checks are presented in Section 2.4.4.

2.4.1 Main Specifications

We define a loan as *prepaid* when the borrower repays the full amount before half of the contracted payments have been made and after the contracted date of the 6th payment. In addition, we require that the total amount of the last three payments must be greater than 35% of the total loan amount. This definition restricts that a prepaid loan is not paid off in small increments.[11]

Based on the definition of prepayment, We model the latent variable form as follows:

$$\begin{aligned} Y_{ijt}^{*} &= \beta_0 + \beta_1 \cdot \textit{Lending Institutions}_{jt} + \beta_2 \cdot \textit{Demographics}_{jt} + \beta_3 \cdot \textit{Econ Condition}_{jt} \\ &\quad + \beta_4^{'} \mathbf{X}_{ijt} + \mu_k + \nu_t + \epsilon_{ijt}, \qquad (2.1) \\ Y_{ijt} &= 1[Y_{ijt}^{*} > 0], \end{aligned}$$

Y_{ijt}^{*} is the latent underlying utility of the borrower prepaying. Our primary variable of interest is *Lending Institutions*$_{jt}$, which is the natural log of the number of lending institutions in area j at time t. *Demographics*$_{jt}$ includes all demographic characteristics in area j at time t. *Econ Condition*$_{jt}$ contains all variables related to local economic conditions in area j at time t. $\mathbf{X}_{ijt}$ are all observed characteristics of individual borrower i; μ_k are state dummies, and ν_t are time fixed effects. Lastly, ϵ_{ijt} is the unobserved error term that captures idiosyncratic shocks to individual borrowers'

[11] We use alternative definitions of "being prepaid" in section 2.4.4 as a robustness check.

utility of prepaying. In regressions, we cluster the standard error at the area level (j) to account for correlations among local level unobservables.[12]

We include the following demographic variables, with each variable representing a specific area j: the percentage of residents who have an undergraduate degree in a business major (*Business Degree*), the proportion of residents who identify their race as Black (*Race: Black*), and the proportion of Hispanic residents (*Origin: Hispanic*). We include the proportion of four-year college degree holders with a business degree as a measure of a resident's exposure to individuals with higher levels of financial literacy. Borrowers from areas with a larger proportion of such population may have a more comprehensive understanding of the risk and return trade-off associated with debt management. We also include data on the percentage of minority residents in an area—the race and origin—to alleviate concerns regarding their role in determining a borrower's debt management behavior (e.g., see Anderson and VanderHoff (1999); Ambrose and Capone (1998); Pedersen and Delgadillo (2007) for examples of the role of race in debt management).

In addition, we include the following economic controls that are associated with a specific area: *Unemployment Momentum*, *Poverty*, *Rent Adjusted Earnings*, and *Lender-to-Population Ratio*.[13] *Unemployment momentum* is calculated as the monthly unemployment rate of area j for a moving window that spans from 6 months before time t to 18 months after time t. So, there are 25 observations associated with each area. The *Unemployment momentum* is the slope of the fitted line (obtained by minimizing the sum of squared errors between the 25 observations and the fitted line). This variable captures the momentum of the overall economic environment—either improving or deteriorating. The economic momentum may have an impact on a borrower's debt management because borrowers from areas with improving local con-

[12]Not all our socio-economic variables are available at the month level. Therefore, when necessary, we use the measures obtained for a particular year for all months in that year.

[13]In a robustness check, we include the population information in each borrower's area. The size of the location is arguably relevant to debt management. Specifically, we calculate quartiles for the population distribution and construct dummy variables corresponding to each quartile (25%, 50%, 75%, and 100%). We do not include the population variable as it is, because it is highly correlated with other variables. The results from this robustness check are qualitatively consistent.

ditions are likely to have different local opportunities compared with borrowers from areas in deteriorating economic conditions (holding everything else constant). We also control for *Rent Adjusted Earnings*, which is the ratio of the median household income and the fair market rental value of a 3-bedroom residence in that area. We include this variable to control for the overall purchasing power in an area, which should capture a borrowers' ability to utilize their income for repayment. *Lenders-to-Population Ratio* is included to account for potential disparities in the number of people served by each lending institutions.

Lastly, the rich individual-level data provided by Lending Club allows us to include a multitude of borrower and loan characteristics. Loan characteristics such as *Amount Requested* and *Interest Rate* provide a measure of the relative difficulty of paying back a specific loan. For example, larger amounts should arguably be more difficult to pay back on average (holding other characteristics the same). Borrower characteristics provide an in-depth overview of each borrower's personal characteristics that may influence their financial management behavior. We also include state fixed effects and year fixed effects in our main specification.

Evidence based on the Number of Lending Institutions

Table 2.2 reports the results with the dependent variable *Prepaid*. Column 1 reports the results of specification 2.1 using a Probit model.[14] The *Lending Institutions* coefficient is positive and statistically significant. The coefficient estimates suggest that when the number of institutions increases from the 25^{th} percentile to the 75^{th} percentile, the conditional probability of prepayment increases by about 0.5 percentage point.[15] The likelihood of prepaying is 14.18% at the 25^{th} percentile of the *Lending Institutions* distribution. Our estimate translates to a 3.5% increase in the prepaying likelihood (= 0.5/14.18). Borrowers would require an increase in income of approxi-

[14]We use Probit regressions as our main specification. Results from Logit regressions and linear probability models are qualitatively the same.
[15]All other variables are evaluated at their median values.

mately $75,000 to achieve a similar increase in prepayment probability. These results indicate that the competition among lending institutions in the local area positively affects the borrower's propensity to prepay their online P2P loans.

Table 2.2.: Main Results on Prepayment

Dep. Variable: Prepaid(1/0)	(1)	(2)	(3)	(4)
Lending Institutions	**0.009*** **	**0.010*** **	**0.006*** **	**-0.005** **
	(0.002)	(0.003)	(0.002)	(0.002)
Demographics:				
Business Degree	0.073	0.076	0.080	0.068
	(0.079)	(0.078)	(0.078)	(0.079)
Race: Black	0.010	0.013	0.013	0.035
	(0.033)	(0.033)	(0.032)	(0.032)
Origin: Hispanic	0.022	0.023	0.024	0.028
	(0.028)	(0.028)	(0.028)	(0.028)
Economic Conditions:				
Unemployment Momentum	-0.003	0.000	-0.008	0.006
	(0.100)	(0.100)	(0.100)	(0.100)
Poverty	-0.188***	-0.188***	-0.184***	-0.158**
	(0.065)	(0.065)	(0.064)	(0.064)
Lender-Population Ratio	-0.002	-0.002	-0.002	-0.003
	(0.003)	(0.003)	(0.003)	(0.003)
Rent Adjusted Income	0.002***	0.002***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)
Borrower/Listing Characteristics:				

continued on next page

Table 2.2.: *continued*

Dep. Variable: Prepaid(1/0)	(1)	(2)	(3)	(4)
Amount Requested	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
IR(%)	0.012***	0.012***	0.012***	0.011***
	(0.003)	(0.003)	(0.003)	(0.003)
Income	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
DTI(%)	-0.821***	-0.820***	-0.819***	-0.831***
	(0.054)	(0.054)	(0.054)	(0.055)
Home-ownership	0.082***	0.082***	0.082***	0.079***
	(0.008)	(0.008)	(0.008)	(0.008)
Bank Card Utilization	-0.205***	-0.205***	-0.205***	-0.205***
	(0.014)	(0.014)	(0.014)	(0.014)
Current Credit Lines	0.003***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Delinquencies	-0.020***	-0.020***	-0.020***	-0.019***
	(0.005)	(0.005)	(0.005)	(0.005)
Inquiries Last 6 Months	0.040***	0.040***	0.040***	0.040***
	(0.003)	(0.003)	(0.003)	(0.003)
Years Employed	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Public Records in Last 10 Years	0.026***	0.026***	0.026***	0.025***
	(0.008)	(0.008)	(0.008)	(0.008)
Revolving Credit Balance	-0.002***	-0.002***	-0.002***	-0.002***

continued on next page

Table 2.2.: *continued*

Dep. Variable: Prepaid(1/0)	(1)	(2)	(3)	(4)
	(0.000)	(0.000)	(0.000)	(0.000)
Grade FE	Yes	Yes	Yes	Yes
Loan Category FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Constant	-0.872***	-0.876***	-0.935***	-0.829***
	(0.114)	(0.114)	(0.116)	(0.114)
Observations	187,122	187,122	187,122	186,778

Alternative Measures of Local Market Structure

Thus far, we have utilized the number of lending institutions as defined in the Census Business Pattern Data as our measure of local financial market structure. The method of using the number of lending institutions is novel in studies of financial markets.[16] In particular, the measure includes the number of credit unions in an area. Including credit union information is necessary given their large share in personal loan markets. In this section, we explore other measures of local lending market structure to check the robustness of our main findings.

FDIC Insured Bank Branches and Deposits

An alternative measure of the local lending institution market structure that has been widely adopted in the literature utilizes branch and deposit information for

[16]Similar measures have been used extensively in studies of other markets (e.g., Brynjolfsson et al. (2009), Chevalier (1995), Mazzeo et al. (2014)).

each Federal Deposit Insurance Corporation (FDIC) insured bank branch in the US (Becker, 2007; Butler and Cornaggia, 2011). A potential concern associated with this measure in our context is that credit unions are not insured by FDIC and, therefore, are not included in the measure. Regardless, we first evaluate the robustness of our results by incorporating the FDIC data on bank branches and deposits. Column 2 of Tables 2.2 report the results from this alternative specification. The coefficient estimates are qualitatively similar to the main findings—online borrowers from areas with more branches are more likely to prepay their online loans.

FDIC also provides details regarding the dollar amount of deposits held by each FDIC insured bank branch. Deposits are the total amount of funds in the form of savings accounts, money market accounts, and checking accounts. They have been used as a proxy for the capacity of lending institutions in a local area (Becker, 2007). To examine whether our results are robust to changes in the nature of the proxy for market size, we replace the number of institutions with deposits (summed across all branches in a borrower's local area). Column 3 of Table 2.2 show the coefficients for deposits in a Probit specification. Consistent with our results, the total deposits contained in a borrower's local area are positively associated with prepayment likelihood.

Market Concentration: Herfindahl-Hirschman Index

As a third measure of the local market structure, we calculate the Herfindahl-Hirschman Index (HHI) of the local banking market. HHI is widely adopted by government agencies such as FDIC, the Department of Justice (DOJ), and the Federal Trade Commission (FTC) to monitor the potential impacts of banking mergers.[17] Specifically, we first calculate the market share of each bank in a given area based on deposits. Then, the HHI is the sum of the squared market shares in a given area. FDIC notes that an HHI below 1,000 is considered "low concentration," which implies

[17] See for example https://www5.fdic.gov/sod/definitions.asp?systemForm=SOD&helpItem=HHI&swhichPg=DMH.

that the deposits are more evenly distributed among the local lending institutions. In contrast, an HHI over 1,800 implies that the area is in "high concentration."

FDIC specifies geographic boundaries for banking markets. These geographic boundaries are collected from the Competitive Analysis and Structure Source Instrument for Depository Institutions (CASSIDI). The Federal Reserve System and the Department of Justice use this tool to evaluate competition among banks.[18] The evaluation is conducted by determining the changes in HHI caused by the merger of banking institutions. The authorities are specifically concerned with mergers that affect the overall competitiveness in a specific market. Similarly, we evaluate the HHI as an indicator of the market structure at the location level governed by our main empirical specification.

To examine the role of market concentration, in regressions we use the HHI of the local lending market instead of the lending institution count as in the main specification Equation 2.1. Column 4 in Table 2.2 reports Probit estimates. The coefficients indicate that higher market concentration leads to a lower likelihood of prepayment, which is consistent with our main findings because higher market concentration suggests a lower level of competition in a local market.

2.4.2 Evidence from Bank Mergers

Bank Mergers

Thus far, we have incorporated a rich set of individual and area-level controls to establish the role of local market structure in determining P2P borrowers' debt management outcomes. The results are consistent across multiple specifications. An ancillary approach to determining the role of market structure is to assess identical borrowers from statistically similar markets and exogenously alter the local market structure in one area. Evaluating the impact of this alteration and the subsequent debt management differences between individuals from the affected and non-affected

[18]More information can be found at https://cassidi.stlouisfed.org.

markets allows us to address the research question through the lenses of natural experimentation. Specifically, bank mergers are ideal “natural experiments” to establish the impact of market structure *changes* on online borrowers’ debt management behavior.

On the one hand, bank mergers have significant impacts on the market structure of the local lending industry. Fraisse et al. (2018) find that, in areas with physical branches of two merging banks, overall bank lending declines by approximately 2.7%. Liebersohn (2017) finds that bank mergers increase the deposit rates by 0.13 percentage point. Nguyen (2019) uses bank mergers as an instrumental variable for branch closures and finds that bank closures lead to a decline in credit supply. On the other hand, bank mergers are exogenous relative to P2P borrowers’ debt management. Considering the complexities and stakes involved in a merger, it is unreasonable to suggest that merger decisions are driven by the need to alter the debt management behavior of P2P borrowers.

To test how bank mergers lead to changes in P2P borrowers’ prepayment outcomes, we collect bank merger data from FDIC. Operationally, we assign borrowers as “treated” if their local market was affected by a merger among banks. Specifically, for each bank merger, we identify areas where both merging banks had branches open before the merger. We refer to these areas as bank merger affected areas. We consider two criteria when determining if a P2P borrower is treated: (1) whether the borrower is from one of the determined bank merger impacted areas and (2) whether the borrower obtained their loan in the 12 months before the merger. Arguably, the merger-induced changes to the local financial markets will only affect treated borrowers in the period after they obtained their P2P loans. Therefore, the merger was not related to the decision to acquire a P2P loan. We assign a value of one to borrowers that satisfy the two criteria and zero to all other borrowers. We replace the number of lending institutions with this treatment dummy in Equation 2.1 and study its impact on prepayment.

The results are presented in column 1 of Table 2.3. The binary treatment variable (*Merger*) has a negative impact on the probability of prepaying. This suggests that the merger-induced reduction in market competition affects P2P borrowers' debt management behavior.

Table 2.3.
Prepayment Results from Bank Mergers

Dep. Variable: Prepaid(1/0)	(1)	(2)	(3)
Merger(1/0)	**-0.022***	**-0.061***	**-0.070***
	(0.012)	(0.025)	(0.040)
Demographic Variables	Yes	Yes	Yes
Economic Cond. Variables	Yes	Yes	Yes
Borrower Char. Variables	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Constant	-0.796***	-0.796***	-1.470**
	(0.079)	(0.079)	(0.613)
Observations	179,857	179,857	7,552

As mentioned in previous discussions, the HHI of a banking market is used by FDIC to determine the impact of mergers on bank markets. Specifically, the FDIC uses a benchmark of HHI change in a value of 200 as a minimum threshold to determine whether potential intervention is necessary. In conjunction, we replicate the process used to obtain the results in column 1 of Table 2.3, except we only use the mergers in markets where the HHI increased by at least 200 in the post-merger year. Column 2 of Table 2.3 presents the results of this specification. The coefficient for the binary treatment variable (*Merger*) is negative and statistically significant. More importantly, the magnitude of the coefficient is larger than that in column 1, which means that the disadvantages of the merger-induced reduction in the competition are more pronounced in locations where the merger caused a substantial shift in competitive dynamics.

Borrower Matching

The merger results, especially those with the increased coefficient magnitude when setting the minimum HHI criteria at 200, provide significant evidence towards supporting the posited mechanism. One underlying assumption is that mergers are not endogenous to the debt management behavior of P2P borrowers. This assumption will hold unless there exists a causal relationship between P2P borrower's debt management choices and bank decision to merge. To assuage any lingering doubts regarding the comparability of the borrowers in the merger impacted areas and the other borrowers, we perform analysis based on a nearest neighbor matching between "treated" borrowers and untreated borrowers (based on a one-to-one matching) to create a directly comparable sample.

To match borrowers, we first identify the borrowers who were impacted by a merger in the same manner as mentioned previously. We focus on areas where the HHI increased by at least 200. For each treated borrower, we identify areas that have a comparable HHI (within 5%) to the pre-merger HHI in the focal borrower's locality.[19] We keep only the borrowers from the matched areas with the same sub-grade as the treated borrower and those who obtained their loans during the specified period for that merger (12 months preceding the merger date). By keeping only borrowers with the same sub-grade, we ensure that the credit profile of the borrowers is similar. Finally, we use the treated borrower's debt-to-income ratio, loan amount, and the total value of deposits in that borrower's locality as the matching variables. We use debt-to-income ratio and loan amount to represent the subtle differences within sub-grades and use the total value of deposits to ensure similar access to funds across the treatment group and the control group. The results using the specification that only

[19]We only include borrowers from areas that have never gone through a merger to identify control borrowers.

uses the subset of matched borrowers is shown in column 3 of Table 2.3. The results are consistent with our main findings, indicating that the reduction in competitive dynamics causes prepayment to decrease. The magnitude of the coefficient is slightly larger than the coefficient in column 2, meaning that the non matched sample does not suffer from a bias that may inflate our results.

2.4.3 The Underlying Mechanism

Evidence from Borrower FICO Score Improvements

We have shown evidence that the debt management behavior of P2P borrowers is affected by the competitive dynamics in a local area. In particular, borrowers are more likely to prepay their loans (possibly) due to strategic local banking reactions if they reside in more competitive markets. We continue to explore evidence that supports the mechanism behind our findings. If the competition in a local area prompts banks to react to the rising P2P lending by offering individual borrowers favorable terms, then arguably these borrowers who have improved their credit rating through P2P lending will benefit more from the local lending institutions' reactions—more likely to prepay. In addition, this benefit from improved credit rating should be more salient for those who reside in more competitive localities. Governed by this argument, we utilize the richness of the Lending Club data to examine the impact of local market structure on borrowers with improving credit profiles to examine further the validity of the underlying mechanism driving our findings.

Operationally, for each borrower, we have data on their monthly FICO score changes throughout the loan's life cycle. To utilize this information, we estimate a specification similar to Equation 2.1 by adding a binary term that indicates whether the borrower's FICO score increased by more than 40 points. This variable, *FICO Increase (1/0)*, has a value of 1 if the borrower's FICO score at the time when the payment period ends is at least 40 points greater than the FICO score at the beginning of the payment period. Moreover, we interact our primary independent variable, *Lending Institutions*, with this binary term. Table 2.4 shows the results of this specification. Column 1 reports the results on the subsample of borrowers with an initial credit score of less than 700. These are borrowers who would most likely benefit from an improvement in credit rating. The results indicate that, in general, borrowers who have increased their FICO scores by at least 40 points are more likely to prepay their online loans. These borrowers have improved their credit standing and, as such, can obtain superior interest rates from local lending institutions. Arguably, these borrowers are prime candidates for loan replacement. Therefore, if the local market competition drives the strategic reaction of local financial institutions, then we would expect this effect to be more significant in areas with a higher level of competition. The positive and statistically significant coefficient associated with the interaction term, between the FICO score increase indicator and the number of lending institutions, provides evidence supporting this mechanism. Column 2 replicates the analysis for borrowers with initial credit scores higher than 700. The results indicate that these borrowers do not benefit from the local market as their FICO score increases do not result in significant changes to their local market options. This finding is comforting because borrowers with already competitive scores are least likely beneficiaries of score improvements.

Evidence from Lending Club Interest Rate Changes

To provide further insights regarding the underlying mechanism, we exploit Lending Club's pricing mechanism. LC sub-grades are assigned to borrowers by the LC proprietary algorithm and used to assign interest rates to borrowers—all borrowers with the same sub-grade are charged the same interest rate at any given point in time. In addition, the interest rate assigned to each sub-grade is assigned by the LC Interest Rate Committee.[20] The committee assign interest rates based on four

[20]This process is reported in the 2018 LC 10-K form found at https://ir.lendingclub.com. The Interest Rate Committee is made up of Chief Executive Officer, Chief Financial Officer, Chief Risk Officer, Chief Operating Officer, and General Council.

Table 2.4.
Examining FICO Score Changes and Interest Rate Changes

Dep. Variable: Prepaid(1/0)	FICO $\leq$ 700 (1)	FICO >700 (2)	Interest Rate Experiment (3)
FICO Increase (1/0) x Lending Inst.	0.015**	-0.007	
	(0.007)	(0.010)	
FICO Increase (1/0)	0.072***	0.023	
	(0.028)	(0.042)	
Lending Inst.	0.007**	0.008*	0.009**
	(0.003)	(0.004)	(0.004)
Interest Rate (1/0) x Lending Inst.			0.008***
			(0.002)
Demographic Variables	Yes	Yes	Yes
Economic Cond. Variables	Yes	Yes	Yes
Borrower Char. Variables	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Constant	-0.783***	-1.318***	-0.885***
	(0.137)	(0.508)	(0.041)
Observations	121,148	65,974	89,649

criteria: (1) general economic conditions; (2) the balance of funds and demand for credit through the LC platform, taking into account whether borrowing requests exceed investor commitments; (3) the estimated default likelihood per loan; and (4) consumer credit rates by other lending platforms and major financial institutions. Among the criteria, the second implies that the interest rate changes are potentially due to platform-specific differences that are not driven by local market conditions. In other words, the interest rate for a specific sub-grade can change over time.

We use a hypothetical example to illustrate the relationship between the underlying mechanism and the interest rate changes. Assume two P2P borrowers reside in a highly competitive market. Further assume that these two borrowers are statistically identical except for the fact that LC assigns one borrower a rate of δ_h and the other a rate of δ_l, where $\delta_h > \delta_l$. The potential local rate for these borrowers is γ_{comp}. Since both borrowers are P2P borrowers, we can plausibly assume that $\delta_l < \delta_h < \gamma_{comp}$. Suppose there are two further analogous borrowers who reside in less competitive markets and who are also assigned rates of δ_h and δ_l by the platform. Their assumed local rate is $\gamma_{not-comp}$, where $\delta_l < \delta_h < \gamma_{comp} < \gamma_{non-comp}$. In this scenario, the borrower in the competitive market with the higher platform assigned interest rate has the smallest difference when considering offline to online interest rates $(\delta_h - \gamma_{comp})$. This implies that if the borrowers' credit profile was to improve post-loan or the interest rates decreased overall, this borrower would be the most likely to prepay due to superior local rates. This is based on the assumption that competitive markets react to interest rate changes *at-least* as fast as less competitive markets. Thus, if the platform were to assign a higher interest rate to one of two identical borrowers, then the role of local market structure in determining their prepayment would be stronger for the borrower with the higher interest rate.

We leverage the interest rate changes to test the underlying mechanism. We identify all interest rate changes that increased the prevailing rates for a specific sub-grade. For each change, we examine a three-month window around the event. We determine all the borrowers within the sub-grade that was impacted by the change. We create a variable called *Interest Rate (1/0)* as a binary treatment term. Borrowers with the higher interest rate are assigned a value of 1 and others are assigned a value of 0. We identify 89,660 borrowers who obtained loans when the interest rates of their associated sub-grades were altered. Since the interest rate differences of these borrowers are plausibly exogenous to each borrower's specific local market conditions, we can evaluate the mitigating role of this treatment on our variable of interest in specification 2.1. Specifically, we add an interaction term between *Lending Institutions* and *Interest Rate (1/0)* in the model. The results are reported in column 3 of Table 2.4. Unsurprisingly, the interaction term is positive and statistically significant, indicating that the impact of the local financial market structure is stronger for the borrower with the higher platform assigned rate.

Using Causal Forests to Estimate Heterogeneous Treatment Effects

We next supplement our analysis by estimating heterogeneous treatment effects following the method proposed by Wager and Athey (2018). Specifically, we identify the heterogeneity of the treatment effect with regard to the size of a borrower's local market (# of lending institutions) and the loan amount of the loan. This analysis enables us to delve further into the underlying mechanism. With regard to market size, it reveals whether the marginal benefits of a more competitive local market diminish after a market reaches a certain size. With regard to the loan amount, it allows us to identify the loan sizes that are most vulnerable to local market reactions.

Recall that we have a sample of 89,660 borrowers who obtained loans in periods when the platform altered the interest rate. Furthermore, recall that borrowers are considered treated if they are charged higher interest rates than similar borrowers in a comparable period. To estimate the heterogeneous treatment effects, we randomly split the sample into two groups: a training sample (65% of the entire sample) and a test sample (the remaining 35%).

With the training sample, we use Causal Forests (Wager and Athey, 2018) to estimate the decision tree that is used to assign treatment effects to each point. We then apply this decision tree on the test sample and obtain the estimated individual treatment effect for each borrower in the test sample. Figure 2.4.3 shows the average treatment effect of the number of lending institutions. The results support our findings in section 2.4.3, suggesting that the treatment effect is more pronounced for borrowers with a more competitive local financial market structure. That is, the causal impact of a borrower charged an exogenously higher interest rate on their prepayment rate is stronger for borrowers with more local banking competition. Moreover, this also indicates that the marginal benefits of more competitive local markets are not diminishing. For example, borrowers from medium sized markets would also benefit from an increase in local banking competition. Figure 2.4.3 emulates this analysis and outlines the differences in treatment effect across borrower loan amounts. The figure indicates that larger loan amounts are associated with higher treatment effects, which further corroborates the underlying mechanism as borrowers with larger loan amounts will have more savings from exchanging loans for lower local interest rate offerings. The average treatment effects in Figure 2.4.3 also indicate that after a specific loan size (approximately \$12,000), there is an increase in treatment effect which is homogeneous for all loan amounts. That is, local markets do not significantly vary in their strategic reactions once a loan amount crosses a specific minimum threshold.

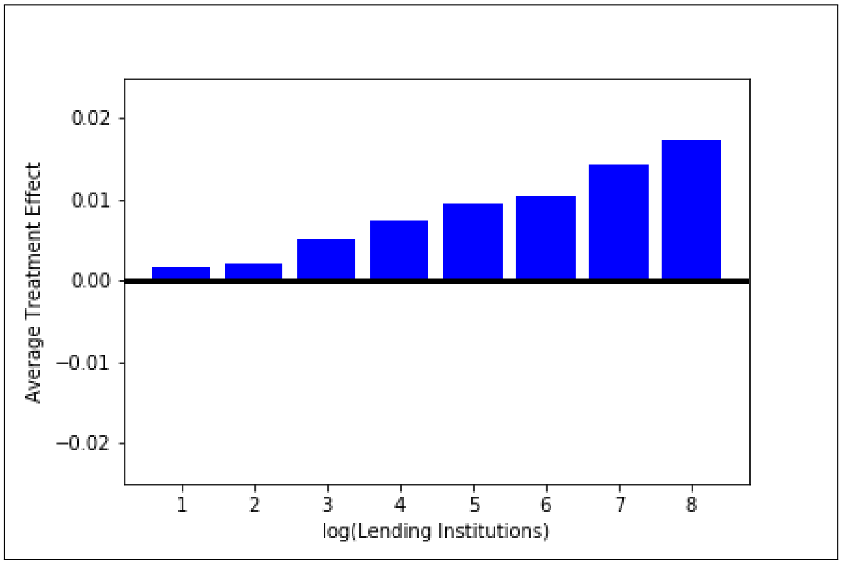


Fig. 2.3. Heterogeneous Treatment Effects on Prepayment by the Number of Lending Institutions

2.4.4 Robustness Checks

Alternative P2P Lending Platform

To assess the robustness of our results to potentially unobserved platform specific factors, we reproduce our primary analysis (reported in Table 2.2) for data collected from Prosper.com, the second largest P2P platform in the US. We obtain 42,822 loans from Prosper aggregated at the location level. These loans cover the same origination period as our main analysis—January 2011 to March 2014. We present the results

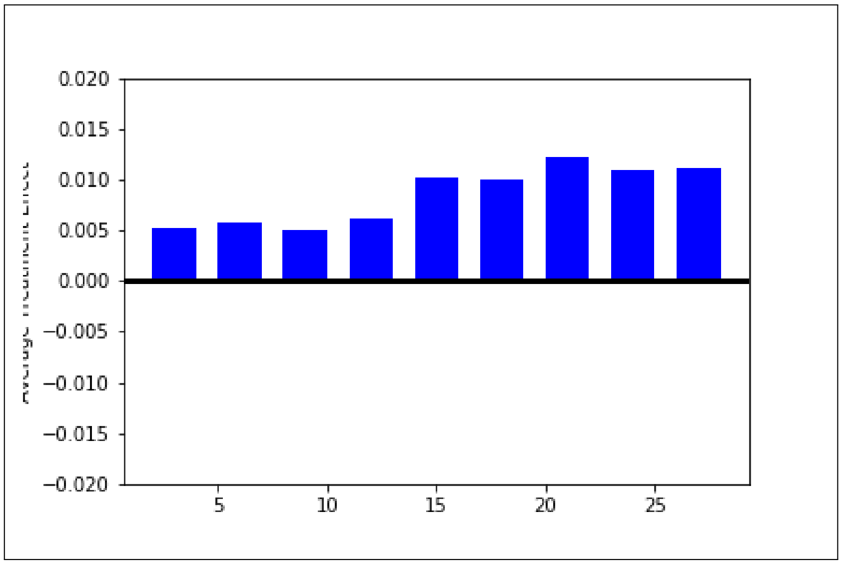


Fig. 2.4. Heterogeneous Treatment Effects on Prepayment by Loan Amount

for the Prosper data to alleviate concerns regarding the generalizability of our results to borrowers from other platforms. Table 2.5 shows the results of our main Probit specification (Equation 2.1) using Prosper data. The findings are consistent with our main results using Lending Club data.

Table 2.5.: Robustness Check: An Alternative P2P Lending Platform—Prosper.com

Dep. Variable: Prepaid(1/0)	(1)	(2)	(3)	(4)
Lending Institutions	**0.011****	**0.011****	**-0.013*****	**-0.014*****
	(0.005)	(0.005)	(0.003)	(0.003)
Demographics:				
Business Degree	-0.003	-0.003	0.426**	0.429**
	(0.156)	(0.156)	(0.196)	(0.196)
Race: Black	-0.177***	-0.179***	0.069	0.074
	(0.064)	(0.064)	(0.078)	(0.078)

continued on next page

Table 2.5.: *continued*

Dep. Variable: Prepaid(1/0)	(1)	(2)	(3)	(4)
Origin: Hispanic	-0.042	-0.043	0.038	0.040
	(0.060)	(0.060)	(0.077)	(0.077)
Economic Conditions:				
Unemployment Momentum	0.400**	0.400**	0.489*	0.490*
	(0.197)	(0.197)	(0.255)	(0.255)
Poverty	0.035	0.034	0.218	0.223
	(0.127)	(0.127)	(0.157)	(0.157)
Lender-Population Ratio	-0.009	-0.009	-0.015	-0.016
	(0.009)	(0.009)	(0.012)	(0.013)
Rent Adjusted Income	0.002	0.002	0.002	0.002
	(0.001)	(0.001)	(0.002)	(0.002)
Borrower/Listing Characteristics:				
Amount Requested	-0.005***	-0.005***	0.016***	0.016***
	(0.001)	(0.001)	(0.002)	(0.002)
IR(%)	0.041***	0.041***	0.048***	0.048***
	(0.004)	(0.004)	(0.006)	(0.006)
Income	0.000**	0.000**	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
DTI(%)	-0.000***	-0.000***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Home-ownership	0.047***	0.047***	-0.051***	-0.052***
	(0.015)	(0.015)	(0.020)	(0.020)
Bank Card Utilization	-5.608**	-5.609**	-1.983	-1.991
	(2.648)	(2.648)	(3.405)	(3.405)

continued on next page

Table 2.5.: *continued*

Dep. Variable: Prepaid(1/0)	(1)	(2)	(3)	(4)
Current Credit Lines	-0.004**	-0.004**	0.009***	0.009***
	(0.001)	(0.001)	(0.002)	(0.002)
Delinquencies	0.000	0.000	-0.004***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Inquiries Last 6 Months	0.044***	0.044***	0.031***	0.031***
	(0.005)	(0.005)	(0.006)	(0.006)
Years Employed	-0.002***	-0.002***	-0.003**	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
Public Records in Last 10 Years	0.023**	0.023**	0.002	0.002
	(0.010)	(0.010)	(0.014)	(0.014)
Revolving Credit Balance	-0.001***	-0.001***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Grade FE	Yes	Yes	Yes	Yes
Loan Category FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Constant	-2.042***	-2.043***	-2.756***	-2.754***
	(0.163)	(0.163)	(0.228)	(0.228)
Observations	42,808	42,808	42,808	42,808

Prepayment

A potential concern is how we define a loan being prepaid. Our main model defines a prepayment as any loan that is paid in full after the 6^{th} contracted payment date but before the contracted date of the middle payment[21]. Moreover, the sum of the last three payments must be greater than 35% of the initial loan value. To mitigate concerns regarding the definition, we use three alternative methods to define prepayment and report the results in Table 2.6. The first method (column 1 of Table 2.6) defines a prepaid loan being paid in full after the 3^{rd} contracted payment date and before the contracted date of the middle payment. The second method (Column 2 of Table 2.6) defines a prepayment as a loan where the sum of the last three payments was greater than 45% of the initial loan value. The third method (Column 3 of Table 2.6) replicates Column 2 except where the sum of the last three payments is 55%. The results are qualitatively consistent with our main findings. Moreover, the coefficient size of the 55% benchmark (Column 3 of Table 2.6 represents an approximate 50% increase relative to the 45% benchmark (Column 2). This is reassuring, as it indicates that the local market is significantly more important for borrowers seeking to replace more than half their loans at the time of prepayment.

Size of Local Institutions

In our main specification, we measure the number of lending institutions using business pattern data from the U.S. Census Bureau (we sum all the credit offering institutions in a borrower's local area). This relies on the number of credit offering institutions in a given zip code collected annually. However, it may be that some lending institutions are significantly larger than others, potentially representing more competition. For example, an area with ten small institutions may not be comparable to an area with ten large institutions. To alleviate this concern, we obtain the

[21] The middle payment is the 18^{th} payment for a 36-month loan and 30^{th} payment for a 60-month loan.

Table 2.6.
Robustness Check: Definitions of Prepayment

Prepayment Definition:	Prepaid Between 3rd and Mid Payment	Prepaid Last Payment 45%	Prepaid Last Payment 55%
	(1)	(2)	(3)
Lending Institutions	**0.009***** (0.002)	**0.008***** (0.002)	**0.012***** (0.002)
Demographic Variables	Yes	Yes	Yes
Economic Cond. Variables	Yes	Yes	Yes
Borrower Char. Variables	Yes	Yes	Yes
Grade FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Constant	-0.884*** (0.111)	-0.886*** (0.114)	-0.801*** (0.110)
Observations	187,122	187,122	187,122

number of institutions in each zip code split into subcategories based on the number of employees at the institution (*Source:* U.S. Census Business Pattern Data), which is not an exact number of employees in each institution, but rather a range of the number of employees. For example, at each zip code level, we have the number of institutions with 1 - 4 employees, 5 - 10 employees, and so on. We multiply the number of institutions in each area by the midpoint of the associated employee range and then sum for across levels within a zip code. For example, a zip code with three institutions of 1 - 4 employees and three institutions of 5 - 10 employees would have a weighted lending institution measure of $3*2.5+3*7.5 = 30$. The results indicate that when accounting for the size of the institutions, a borrower's propensity to prepay continues to be affected by the number of institutions.

2.5 Further Analysis

In this section, we delve further into the competitive structure of local lending institutions and understand the influence of different structures on the strategic reactions of traditional financial institutions to P2P lending.

2.5.1 Large Banking Institutions

We seek to augment our previous results by evaluating the potentially different roles of large and small/community banks in providing access to funds. Arguably, branches associated with large banks can offer better alternatives for P2P Lending Platform borrowers in a specific location. Smaller/community banks are often dependent upon the relationships they build in the community and the continual management of these relationships. P2P borrowers are generally seeking a short term alternative and, therefore, may not benefit from the community banks (these are likely to be regional banks) as much as large banks. Also, a local bank is often constrained by limited resources (DeYoung et al. (2007), Elyasiani and Goldberg (2004)). In essence, P2P borrowers are unlikely to establish close relationships with community banks, which means that they may benefit more from the presence of a larger number of multi-state bank branches in their locations.

To test the conjecture, we define a large bank as the top 1% of all banking institutions ranked by the total assets.[22] We use the FDIC branch level data to determine the total assets of each branch's parent institution. For each locality, we calculate the ratio of large banks to total banks. We add this ratio as an additional variable in our main model specification (Equation 2.1). We provide results in Table 2.7. Our findings lend support to our prediction and show that borrowers with access to a larger ratio of large banks have a higher propensity to prepay.

[22] We assess the robustness of this measure to other percentiles (2%,5%) as well as absolute asset figures. Our results are consistent and suggest that the impact of large banks increases when the criteria are more strict.

Table 2.7.
Further Analysis: Large Banks

Dep. Variable:	Prepaid(1/0) (1)
Large Banks / Total Banks	**0.033*** (0.019)
log(# of FDIC Branches)	0.009*** (0.003)
Demographic Variables	Yes
Economic Cond. Variables	Yes
Borrower Char. Variables	Yes
Grade FE	Yes
Year FE	Yes
Constant	-0.871*** (0.117)
Observations	186,751

2.5.2 Examining Default Behavior

We have thus far examined the strategic reaction of local financial institutions to the online P2P alternative by studying borrower prepayment behavior. Another debt management behavior that potentially relates to offline strategic reactions is the P2P borrowers' likelihood of default. Local financial markets are important in determining the availability of funds to P2P borrowers, in particular to those in relatively weaker financial situations. We can see this in three aspects. First, borrowers residing in the more competitive banking environments have greater access to bank finance (Cetorelli and Strahan, 2006; Beck et al., 2004). Second, Keeley (1990) predicts and tests that banking competition can lead to an increase in risk-taking behavior. That is to say, increased bank competition "forces" financial institutions to offer more non-traditional services and engage in risky loan provisioning (Liangliang Jiang, 2017; Jiménez et al., 2013). Literature has also documented a positive relationship between bank competition and losses from banks' provision of loans (Dick, 2006; Berger et al., 2004). Similarly, Marquez (2002) finds that an increase in the number of banks in an area results in greater access to credit for low-quality borrowers. Lastly, banks, specifically large institutions, often have debt consolidation programs that can assist at-risk borrowers. Thus, borrowers with access to more institutions are more likely to find a local bank offering such a program. Therefore, P2P borrowers in a more competitive local market, ceteris paribus, will be arguably less likely to default.

Table 2.8 reports the results of estimating Equation 2.1 with the dependent variable *Default*, which equals 1 if the borrower defaulted on their loan and 0 otherwise. The results indicate that borrowers from more competitive markets are indeed more likely to obtain local funding (to avoid defaults) and thus less likely to default. Our previous results about prepayment outcomes provide evidence of local banks' strategic reactions to the growing market share of the P2P alternative. The findings about default outcomes reported in this section further suggest that local competition can also benefit struggling P2P borrowers.

Table 2.8.: Results on Default

Dep. Variable: Default(1/0)	(1)	(2)	(3)	(4)
Lending Institutions	**-0.015*****	**-0.020*****	**-0.010*****	**0.002**
	(0.003)	(0.003)	(0.002)	(0.002)
Demographics:				
Business Degree	0.762***	0.765***	0.737***	0.722***
	(0.174)	(0.173)	(0.173)	(0.174)
Race: Black	-0.112**	-0.101*	-0.115**	-0.172***
	(0.053)	(0.053)	(0.053)	(0.052)
Origin: Hispanic	-0.122***	-0.120***	-0.120***	-0.132***
	(0.043)	(0.043)	(0.043)	(0.044)
Economic Conditions:				

continued on next page

Table 2.8.: *continued*

Dep. Variable: Default(1/0)	(1)	(2)	(3)	(4)
Unemployment Momentum	-0.183	-0.180	-0.182	-0.207*
	(0.114)	(0.114)	(0.114)	(0.114)
Poverty	0.588***	0.618***	0.549***	0.480***
	(0.129)	(0.129)	(0.128)	(0.128)
Lender-Population Ratio	-0.009	-0.009	-0.009	-0.015
	(0.009)	(0.009)	(0.009)	(0.011)
Rent Adjusted Income	-0.004***	-0.004***	-0.004***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Borrower/Listing Characteristics:				
Amount Requested	0.007***	0.007***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
IR(%)	0.060***	0.060***	0.060***	0.060***
	(0.004)	(0.004)	(0.004)	(0.004)
Income	-0.003***	-0.003***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
DTI(%)	0.806***	0.802***	0.807***	0.827***
	(0.061)	(0.061)	(0.061)	(0.062)
Home-ownership	-0.065***	-0.066***	-0.064***	-0.057***
	(0.009)	(0.009)	(0.009)	(0.009)
Bank Card Utilization	-0.035**	-0.035**	-0.035**	-0.035**
	(0.016)	(0.016)	(0.016)	(0.016)
Current Credit Lines	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)

continued on next page

Table 2.8.: *continued*

Dep. Variable: Default(1/0)	(1)	(2)	(3)	(4)
Delinquencies	-0.009*	-0.009*	-0.009*	-0.009
	(0.006)	(0.006)	(0.006)	(0.006)
Inquiries Last 6 Months	0.043***	0.043***	0.043***	0.043***
	(0.004)	(0.004)	(0.004)	(0.004)
Years Employed	-0.004***	-0.004***	-0.004***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Public Records in Last 10 Years	-0.008	-0.008	-0.008	-0.007
	(0.008)	(0.008)	(0.008)	(0.008)
Revolving Credit Balance	-0.001*	-0.001*	-0.001*	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)
Grade FE	Yes	Yes	Yes	Yes
Loan Category FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Constant	-2.340***	-2.331***	-2.249***	-2.383***
	(0.128)	(0.128)	(0.130)	(0.128)
Observations	187,093	187,093	187,093	186,749

2.6 Concluding Remarks

2.6.1 Summary and Discussions

In this paper, we present empirical evidence that the growing P2P lending market is not immune to the competition from traditional retail lending institutions. In particular, our findings suggest that local competition positively affects the likelihood of strategic reactions from local lending institutions. The potential for improved local interest rates has economically sound impacts on payment outcomes of P2P loans. Numerically, our estimates indicate that if an LC borrower moved from an area at the 25^{th} percentile of local competition to an area at the 75th percentile, the probability of prepayment would have increased by approximately 4.49%. To achieve the same increase in the likelihood of prepayment, a borrower would require an increase of at least $75,000 in annual income. These findings withstand a battery of robustness checks. In addition, we exploit bank mergers as an exogenous shock to a borrower's local banking market and find that mergers (which are associated with a reduction in competition) cause a decrease in prepayment. To validate the mechanism, we find that P2P borrowers with improving credit profiles are most susceptible to offline competition. We also leverage platform-level interest rate changes to isolate the mechanism attributed to the role of competition in driving local lending institutions' reactions. Specifically, borrowers from more competitive markets are more likely to access improved rates and replace their loans accordingly.

In addition, research analyzing competition and substitution that occur between online and offline channels has mainly focused on retail products (e.g., Forman et al. (2009), Brynjolfsson et al. (2009)). Our findings indicate that this competition persists when the product category is financial. The arguments often cited for the persistence of local retail establishments (e.g., delivery hurdles and a customer's need for instant gratification) are not relevant for financial products. However, financial products, unlike retail products, are often tailored to the needs of individual consumers. Therefore, a local bank's intimate knowledge of the borrower heterogeneity in their area allows them to offer more personalized services. A deeper understanding of the relationship between local market structure and P2P borrowers' debt management may shed light on the factors that are relevant to online and offline channel substitution for financial products. Specifically, we find that local financial market structures with a relatively higher ratio of large banks have a higher prepayment propensity.CV

2.6.2 Managerial Implications

As more consumers of financial services become aware and comfortable with the use of FinTech offerings to fulfill their needs, businesses must acquire a deeper understanding of the interaction between alternative and traditional financial products. As shown in Figure 2.1, the market share of P2P lenders is rapidly expanding. Furthermore, a large portion of online P2P borrowers utilizes the platform to consolidate their credit card debt. These are not new borrowers. They are simply offsetting their current debt with a different type of debt. Moreover, as an example, Lending Club recently expanded its offerings and now provides automobile refinancing services. The implication is that online P2P lending platforms are not necessarily creating new demand, but are shifting the debt (and, as a result, future interest payments) away from the more traditional providers. However, these same borrowers are utilizing local markets to further facilitate and manage that debt. As such, traditional banks are not losing customers to FinTech for good so long as their local presence supports a favorable strategy. Moreover, they are able to take advantage of improving credits of P2P borrowers for whom the platforms bear the initial unattractive risks. Interestingly, based on the online banking behavior of customers of a large bank, Abhishek et al. (2018) show that even online channel utilization of a bank is significantly impacted by local branch opening and/or closing. Thus, the physical presence of banks can be a powerful lever in facing increasing competition from FinTech and digitization.

This dynamic, undeniably, is quite challenging for managers of financial service providers because they often have competing product services. For example, a bank may offer a credit card to a consumer. That consumer may be tempted to consolidate that credit later on through the online P2P service. However, the local affiliate of the bank may intervene and offer a better rate than the one provided by the P2P platform (specifically in the case where the local bank can utilize its understanding of the heterogeneity of local borrowers). In this case, the local bank is competing with both the P2P platform and its own larger parent institution (the one that initially offered the credit card). Coordination between regional branches and their parent institutions can help isolate the geographical locations where the bank should intervene (places where borrowers are more susceptible to alternative FinTech products). However, as in the retail sector, this coordination is often difficult to efficiently accomplish.

From the perspective of the providers of FinTech products, the operational restriction to offer homogeneous terms across markets is a critical weakness in fending off competition from localized traditional financial institutions. Our study provides evidence that the local area characteristics—specifically market structure, economic conditions, and borrower heterogeneity—need to be considered when assessing the sustainability of the FinTech product in a specific geographic area. These factors should be incorporated into the marketing strategies of FinTech firms to allocate marketing expenditures more optimally. Since local market structure impacts P2P borrowers' debt management behavior, FinTech providers would need to devise strategies to tailor services to local market attributes. Studies that assess the role of local market characteristics on the receptiveness of local residents to FinTech products would provide more insights into this vital issue.

Furthermore, given that the P2P lending market is a two-sided network, our study provides insights into the borrower side behavior, which is under-represented in the existing literature. Our results suggest that disregarding the borrower side of the P2P lending market results in an incomplete understanding of the factors that affect market equilibrium in this setting. Further research on borrower side behavior in other P2P financial markets, such as P2P Mortgage markets, can provide valuable insights for practitioners and regulators attempting to understand the nuances of these innovative markets.

3. SHARED PROSPERITY (OR LACK THEREOF) IN THE SHARING ECONOMY

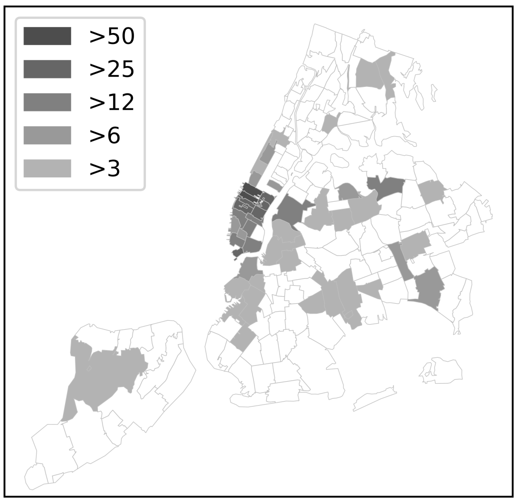
3.1 Introduction

The rapid growth of the sharing economy—44% of Americans have participated in the sharing economy as of 2016—and its impact on local economies is a topic of discussion among practitioners, regulators, and researchers.[1] Much of this attention has focused on the sharing economy's impact on traditional economic activity that directly competes with a platform—e.g., Uber on the taxi industry (Cramer and Krueger, 2016; Wallsten, 2015) and Airbnb on the hotel industry (Zervas et al., 2017). Other work has focused on the potentially negative spillover effects of home sharing platforms on the housing markets in large cities (Barron et al., 2018; Sheppard et al., 2016). Further studies have found a spillover effect on entrepreneurial activity (Burtch et al., 2018), durable goods market (Gong et al., 2018), and interactions between different sharing economy products (Zhang et al., 2018). In contrast, we focus on the economic spillover effects of home sharing platforms, specifically Airbnb, on local non-competing complimentary economic establishments.

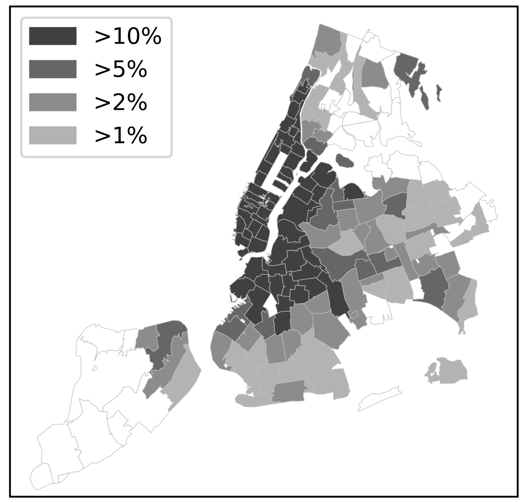
Home sharing platforms such as Airbnb connect residents of a city with potential visitors/tourists. Through the platform, visitors find acceptable hosts to stay with during their visit and the hosts receive compensation for allowing the visitors to stay in their home. By providing access to underutilized inventory (Einav et al., 2016), home sharing platforms have the potential to attract the visitors of a city to vicinities without a significant hotel presence. Hosts who use these platforms are not restricted by land-use regulations and large fixed costs, allowing them to provide accommodations to visitors in areas that would otherwise have been infeasible

[1]http://time.com/4169532/sharing-economy-poll/

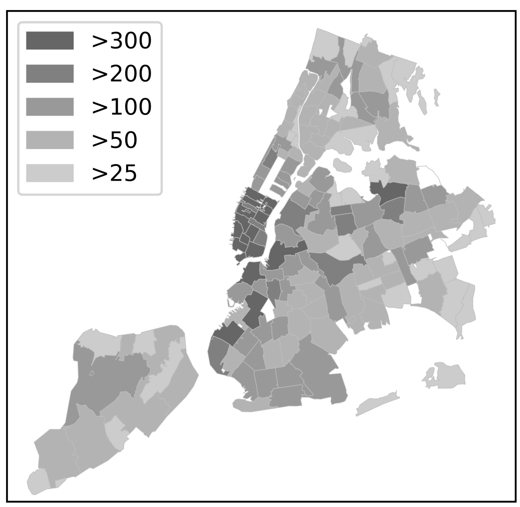
(Coles et al., 2018). Figure 3.1 shows the distribution of hotels, Airbnb activity, and restaurants in New York City (by zipcode) in 2016. The majority of hotels are centrally located while Airbnb stays are more geographically distributed. In other words, Airbnb visitors have the opportunity to, and do, locate in areas without a significant hotel presence.



(a) Hotels.



(b) Airbnb.



(c) Restaurants.

Fig. 3.1. Distribution of Hotels, Airbnb Reviews, and Restaurants in NYC in 2016.

Visitors that choose to locate in these sharing economy enabled areas have two options. On the one hand, they may exploit the area in which they are lodging strictly for accommodation purposes and commute to more traditional tourist locations. As a result, they will spend their non-accommodation based tourism dollars in the traditional tourist locations. On the other hand, they may go beyond utilizing the area simply for lodging and spend their tourism dollars at establishments near their Airbnb listings. Establishments that would not have access to significant amounts of tourism expenditure without the presence of the home-sharing platform. To evaluate the economic spillover effect of this spending, we focus on restaurants in New York City (NYC). NYC is the most visited city in the United States and, in 2012, 21% of tourist spending in NYC, or $7.4 billion, was spent at restaurants.[2] Only accommodation ($10 billion) and shopping ($8 billion) expenditures accounted for a higher proportion of tourist spending.[3] As shown in Figure 3.1, the geographic distribution of restaurants is dispersed across the whole city. This implies that while some areas may not have a large hotel presence, all areas, for the most part, have a significant restaurant presence. Therefore, if a home-sharing visitor chooses to locate in areas without a hotel presence, they would still have access to a substantial number of local restaurants. If the home-sharing visitors in these areas do utilize local restaurants, then these restaurants will improve their financial performance. This improvement would be reflected in the aggregate area level restaurant employment. Consequently, due to both the significance of restaurants with regard to tourist spending and the dispersed geographic distribution of restaurants across NYC, we ask the following research question: *What is the impact of home-sharing activity on local restaurant employment growth? Does this effect vary across areas and, if so, what are the local drivers of this heterogeneous impact?*

Importantly, even in the cases where visitors using home-sharing platform choose not to commute to the more traditional tourist areas, the new visitors' impact on restaurants near their Airbnb listing is still uncertain. There are a multitude of factors

[2]http://www.thisisinsider.com/most-visited-us-cities-2017-12#2-los-angeles-california-9
[3]https://skift.com/2013/07/09/how-tourists-to-new-york-city-spend-their-money/

that will influence the magnitude of this effect. For example, unlike the majority of traditional accommodation alternatives in the hospitality industry (hotels/motels), home-sharing services often provide access to the host's kitchen. This option enables visitors to forgo restaurants and prepare their meals in their homes, which would reduce the potential impact of home-sharing visitors on local restaurants.

Another factor that will determine the impact of the home-sharing platform on restaurants near Airbnb listings is the potential impact of home-sharing visitors on the dynamics of local restaurant demand and its subsequent impact on local residents' behavior. Depending on the agreement, the Airbnb visitor may occupy the home without the host being present or the visitor may share the home with the host during their stay. These two alternatives present different ramifications for the potential demand for local restaurants. The visitors that do not share the homes with the hosts during their stay are consequently replacing the host for the duration of their visit. This temporarily alters the dynamics of local restaurant demand but does not necessarily impact the size of the market. Conversely, visitors that share the homes with the hosts are potentially affecting both the size and dynamics of the restaurant market. Also, since home-sharing listings are often in areas that do not traditionally cater to visitors/tourists, there is a potential for these new home-sharing enabled visitors to affect local residents' behavior as well. For example, a resident's utility to frequent a local restaurant may decrease as more visitors frequent that restaurant. Moreover, even in the case where the visitor displaces the host, the host has obtained access to additional financing through the short term rental income. Therefore, while the host is temporarily unavailable, they also have more income to potentially spend on restaurants when they are not renting out their homes.

To empirically identify the impact of home-sharing visitors on restaurant employment at restaurants in the vicinity of Airbnb listings, we employ a difference-in-difference (DID) specification which utilizes the fact that restaurant employment data is available for the years prior to Airbnb entry into NYC as well as the spatial and temporal variation of Airbnb intensity across zipcodes. We conduct this analysis on a sample of NYC zipcodes that were not significant tourist destinations prior to Airbnb entry into NYC and that did not have significant changes in the educational and income profiles of their residents post Airbnb entry. This is discussed in detail in section 3.3.1. The DID approach extracts the difference in area level restaurant employment before and after Airbnb entry for the first difference. The second difference in the DID framework compares this difference in high intensity Airbnb areas with the analogous difference in low intensity Airbnb areas.

The specification incorporates fixed effects and local variables. These variables create a framework whereby Airbnb intensity is conditionally exogenous to the local factors that may impact restaurant employment. Specifically, the panel structure of this approach incorporates area level fixed effects which rule out the effect of potentially endogenous unobserved time invariant local attributes. We also include a time effect which captures city and/or national factors which may impact local economic activity across NYC during a specific time period. We augment these controls with time varying local area characteristics such as retail employment, hotel employment, and local restaurant popularity. As further validation, we utilize matching algorithms to pair areas with a higher intensity of Airbnb activity with comparable areas with lower intensity of Airbnb activity. We conduct both static and dynamic matching and find consistent results.

Our results indicate that, for an average zipcode, if the intensity of Airbnb activity (Airbnb reviews per household) increases by 1%, then the restaurant employment in that zipcode grows by approximately 1.03%. This result is validated across multiple specifications matching, dynamic matching, and various examinations of the parallel trends assumption necessary for the DID specification (section 3.3.4). We also conduct various robustness checks to assess the definitions of Airbnb intensity, restaurant employment, aggregation level (NYC neighborhoods instead of zipcodes), and matching metrics. To examine the generalizability of this result to other cities, we expand our analysis to an additional 5 major U.S. cities and find similar results.

The mechanism behind these findings is evaluated by assessing the impact of Airbnb activity on Yelp visitor reviews. We find that, on average, if the intensity of Airbnb activity (Airbnb reviews per household) increases by 1%, then the proportion of NYC Yelp visitor reviews that occur in that zipcode increases by approximately 3.4%.[4] The mechanism is corroborated by an examination of a neighborhood level policy shift that occurred in New Orleans in 2017. New Orleans implemented a policy whereby Airbnb was deemed illegal in one neighborhood while it was officially legalizing in adjacent neighborhoods. The policy shift caused the proportion of Yelp visitor reviews in the newly legalized neighborhood to increase and simultaneously decrease in the illegal neighborhood. In summary, our results indicate a three stage process: 1) Airbnb brings visitors to areas that would not otherwise have had access to visitor spending 2) The Airbnb visitors frequent local restaurants and 3) The Airbnb visitors that frequent local restaurants have a tangible economic impact on the performance of the restaurants in these neighborhoods.

To delineate the intricacies of Airbnb's effect across localities, we investigate the role of demographics and market concentration in driving the benefits garnered by specific areas. Our results indicate that both demographics and market structure have an important role in determining the areas that benefit from the economic spillover of Airbnb. We find that the complimentary spillover effect of Airbnb activity on restaurant employment is more pronounced in areas with higher levels of market competition. We also find that restaurants in areas with a relatively high number of Black residents or a relatively high number of Hispanic residents do not benefit from the economic spillover of Airbnb activity.

The sharing economy has altered the landscape of many traditional industries. As regulators struggle with ways to frame legislative discussions surrounding its impact, it is imperative to also assess the economic spillover these alternatives create. This is crucial as regulators seek to obtain a holistic picture of the sharing economy's impact. We provide evidence to the importance of this discussion by establishing a

[4] This interpretation is obtained by evaluating the impact of a 1% increase in Airbnb intensity for a zipcode with the median proportion of NYC Yelp visitor reviews in 2014.

positive economic spillover effect of home sharing on restaurant employment. Perhaps most critically to the regulatory discussion, this benefit is not homogeneous. This implies that a focus on the purely negative direct effects of these platforms may be shortsighted. Furthermore, any discussion surrounding positive spillover benefits must be tempered by an understanding that these benefits are not homogeneously benefiting all localities.

3.2 Primary Empirical Context

Our empirical context is NYC Airbnb activity, restaurant employment, and restaurant reviews on Yelp written by visitors to NYC from the year 2007 to 2016. We aggregate the data at the zipcode level and obtain restaurant employment data from the Bureau of Labor Statistics (BLS) Business Pattern Data.[5] This data provides the number of employees in the restaurant industry by zipcode for a specific year.[6] Our data ends in 2016 as this is the last year of publicly available Business Pattern Data at the time of writing. We also obtain local demographic data from the U.S. Census Bureau. This includes race, origin, median income, and number of households.

3.2.1 Home Sharing Platform Data from Airbnb

We obtain consumer facing data from Airbnb, the most prominent home sharing platform in the world.[7] Specifically, we gather Airbnb listings and review data for NYC. We periodically gather this data and combine it with data from insid-

[5]The most granular level of data provided by the BLS is the zipcode level. We include the following institution categories and their associated NAICS codes in defining the restaurant sector: full-service restaurants (722511), limited-service restaurants (722513), drinking places (722410), cafeteria and grill buffets and buffets (722514), and snack and nonalcoholic beverage bars (722515). URL: https://www.naics.com/six-digit- naics/?code=72

[6]This is not an exact number of employees in each restaurant, but rather a range of the number of employees. For example, at each zipcode level, we have the number of restaurants with 1 - 4 employees, 5- 10 employees, and so on. We multiply the number of institutions in each area by the midpoint of the associated employee range, and then sum for across levels within a zipcode. For example, a zip code with 3 institutions of 1-4 employees and 3 institutions of 5-10 employees would have a total number of employees of 3*2.5+ 3*7.5 = 30.

[7]www.airbnb.com

eairbnb.com, which is a website providing access to periodic snapshots of Airbnb listings and reviews.[8] To validate the accuracy of the data collected, we reproduce results from Coles et al. (2018), who study Airbnb usage and growth patterns in NYC and have access to proprietary data obtained from Airbnb. Our analysis indicates that our data patterns are similar to the patterns found in the proprietary Airbnb data obtained by the authors.

Airbnb allows hosts to list their properties, either whole homes or just rooms in their homes, on their online platform and potential visitors can choose from the selection of listings. Visitors that utilize the platform and, subsequently, stay at a listing are asked to review the hosts/listings after their stays and have 14 days to submit their reviews.[9] We use the total number of reviews written for hosts with listings in a specific area and specific time period as a proxy for Airbnb demand. This method has been used in prior research (Barron et al., 2018; Horn and Merante, 2017; Zervas et al., 2017) and, furthermore, the variation in reviews follows similar patterns to Coles et al. (2018), who use proprietary data from Airbnb in NYC.

3.2.2 Restaurant Review Data from Yelp

We also obtain consumer facing data from restaurant reviews on Yelp. We obtain this information in September 2017. We collect all public facing reviews from restaurants in NYC. Yelp reviews are used as a proxy for local restaurant economic activity (Glaeser et al., 2017, 2018). We also collect every review ever written by each reviewer in our sample. These include reviews written for restaurants in NYC and out of NYC. By doing so, we have access to two important data features. First, we are able capture the reviews for the NYC restaurants that were closed at the time of our collection.[10] Therefore, if a restaurant is open in a previous year, but was closed

[8] http://insideairbnb.com/get-the-data.html

[9] https://www.airbnb.com/help/article/13/how-do-reviews-work

[10] Yelp does not remove closed restaurant review pages from their directory, however they are removed from the main search page. The URL for these closed restaurants can be obtained through the review pages of reviewers that had previously reviewed these restaurants when they were open. Therefore, if a reviewer in our sample has previously reviewed a restaurant that is now closed, then we can

at the time of our data collection, we would have review data for that restaurant for the period when it was open. Second, by obtaining all the reviewers' review histories, we can separate the reviewers into two categories: 1) reviews written by visitors to NYC and 2) reviews written by residents of NYC. We refer to these as visitor and local reviewers, respectively.

Each reviewer on Yelp lists their location. However, these self-reported locations can be erroneous, especially for a city like NYC where the reviewers could be reporting their original locations. For example, a person originally from Seattle, WA who lives and works in NYC may list their location as Seattle. The reverse may apply for someone living in Seattle but that is originally from NYC. To resolve this issue, we use the review history of all the reviewers in our sample to identify their locations. To err on the side of caution, we consider a reviewer a "local" (resident of NYC) if their stated location is within NYC and 75% or more of their reviews are for restaurants located in NYC. We consider a reviewer a "visitor" if their stated location is not NYC and less than 75% of their reviews are for restaurants in NYC.

By separating the reviews into these two categories (visitors and locals), we can focus on the category of Yelp reviews that would be impacted by Airbnb visitors. Specifically, if Airbnb users have introduced a significant amount of new restaurant demand to an area, then that demand would be reflected in the number of visitor Yelp reviews. We calculate the proportion of all NYC Yelp restaurant reviews written by visitors in a specific area and year to measure time varying visitor restaurant activity. We obtain more than 3.5 million Yelp reviews corresponding to 34,331 restaurants in NYC (these include both open and closed restaurants).

access that restaurant's URL from the reviewer's review history page. As such, our process is as follows: (1) collect the reviews for the restaurants in our initial sample of restaurants obtained from crawling the Yelp main search page for NYC restaurants, (2) combine all the restaurant reviews and create a list of all unique reviewers, (3) collect all the reviews from each reviewer's review history page, (4) combine all the reviews from step 3 and create a list of NYC restaurants that were not scraped in step 1 (these are likely closed but could also be restaurants that were not obtained from the initial Yelp search (5) collect the reviews from these restaurants and repeat the process. The process is finished when the next set of restaurants obtained from the reviewers' review histories provides no new restaurants.

We also utilize the Yelp reviews to calculate a restaurant popularity index for each area and year combination. First, for each area (i) and quarter (t) combination, we calculate the proportion of reviews written by residents of NYC that occurred in restaurants in area i.[11] Specifically, this is calculated as follows: $\frac{\text{\# Yelp Local Reviews in area i}}{\text{\# of Yelp Local Reviews in NYC}}$. Second, we obtain the yearly average of this proportion for each area i. We refer to this as the *Prop. of NYC Local Yelp Reviews in Area i*. The area with the highest *Prop. of NYC Local Yelp Reviews in Area i* in the base year (2007) is used as a benchmark. Third, for a given year and area i, we calculate the restaurant popularity index as follows: $\frac{\text{Prop. of NYC Local Yelp Reviews in Area i}}{\text{Prop. of Local Yelp Reviews in base area}}$. The restaurant popularity index is incorporated as a control in our model.

The restaurant popularity index is a powerful control as it captures the time variant area level characteristics that are associated with the popularity of restaurants in an area. For example, if a very popular restaurant was to open in an area in time period t+1, then that area would likely see an increase in Yelp reviews. This new popular restaurant may attract patrons that are both locals and visitors. The visitors that frequent this new popular restaurant may not necessarily be lodging in the local area, but may simply be attracted to the restaurant due to its growing popularity. Therefore, this new restaurant will cause an increase in both local and visitor restaurant activity from period t to period t+1. In this case, attributing the increase in visitor reviews to Airbnb activity would be incorrect as the increase in popularity may also attract Airbnb users to the same area. By incorporating the restaurant popularity index we have controlled for the economic variation that causes this new popular restaurant to open in area i. Therefore, this rich control allows us to better isolate Airbnb visitors' impact from the more general local economic factors that drive restaurant employment growth. A univariate regression of *log(Restaurant Employment)* on the restaurant popularity index yields an R-squared of 0.59. As such, 59% of the variation in restaurant employment is explained by the variation in the restaurant popularity index.

[11]We average quarterly observations to mitigate the impact of one time shocks in local popularity that may be due to singular events.

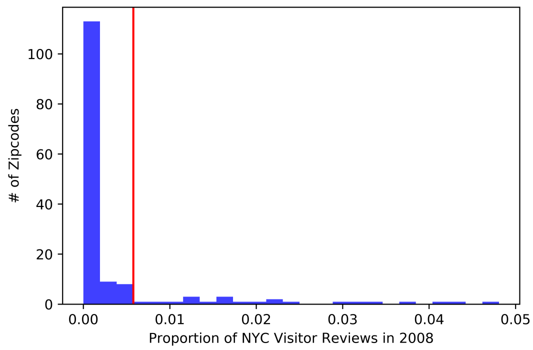
3.3 Impact of Airbnb on Restaurant Employment

3.3.1 Sample Construction

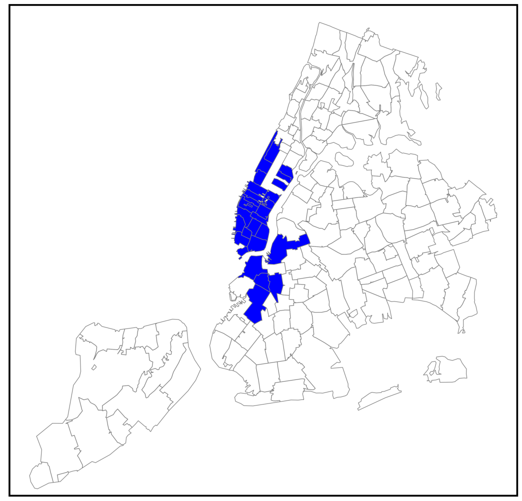
NYC is the largest tourist destination in the United States, and, as a result, contains areas that were and continue to be established tourist destinations. These areas attract a considerable amount of visitors with or without the presence of Airbnb. To identify the areas with endogenous local characteristics that attracted a large number of visitors with and without Airbnb, we examine the distribution of Yelp visitor reviews in 2008. Since Airbnb had negligible presence in 2008, Yelp visitor review activity in that year cannot be attributed to Airbnb visitors. Figure 3.2, panel A, shows the distribution of the proportion of the Yelp visitor reviews for zipcodes in NYC in 2008. We exclude the zipcodes where the proportion of Yelp visitor reviews captured in 2008 is greater than 0.5% (all the zipcodes to the right of the vertical line) and refer to the remaining areas as *NYC Sample 1*. We henceforth refer to areas excluded from *NYC Sample 1* as *traditionally tourist areas*. Thus, areas included in *NYC Sample 1* are considered *traditionally residential areas*. Figure 3.2, panel B, displays a map outlining the zipcodes that comprise the *traditionally tourist areas*. As expected, these zipcodes are centrally located in southern Manhattan and northwest Brooklyn.

Figure 3.3.1 plots the 2008 Yelp Visitor reviews and 2016 Airbnb reviews in NYC. The plots show that there is very little variation in Airbnb activity among the *traditionally tourist areas*. This indicates that the pre-Airbnb factors that attracted tourists to these areas have a significant role in attracting Airbnb visitors and justifies pruning these areas in our analysis.

By design, the areas in *NYC Sample 1* did not attract a significant amount of visitors in the periods preceding Airbnb's launch. Therefore, for zipcodes in *NYC Sample 1*, the post-2008 visitor activity can not be related to time invariant area level factors that attract tourists. To account for local changes in the economic environment of an area that may impact restaurant employment, we also identify and



(a)



(b)

Fig. 3.2. Sample Construction: Traditionally Tourist Areas

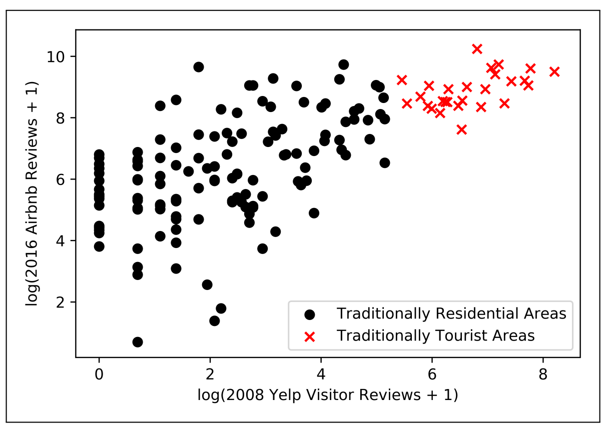


Fig. 3.3. 2008 Yelp Visitor Reviews / 2016 Airbnb Reviews

exclude areas in *NYC Sample 1* with significant increases in housing construction, income (based on IRS filings), or educational attainment post-2008. We utilize these factors as they have been used previously in research studying gentrification and its impacts (Freeman, 2005). We refer to the remaining areas as *NYC Sample 2*, which is the main sample used in our analysis.

3.3.2 Identification Strategy

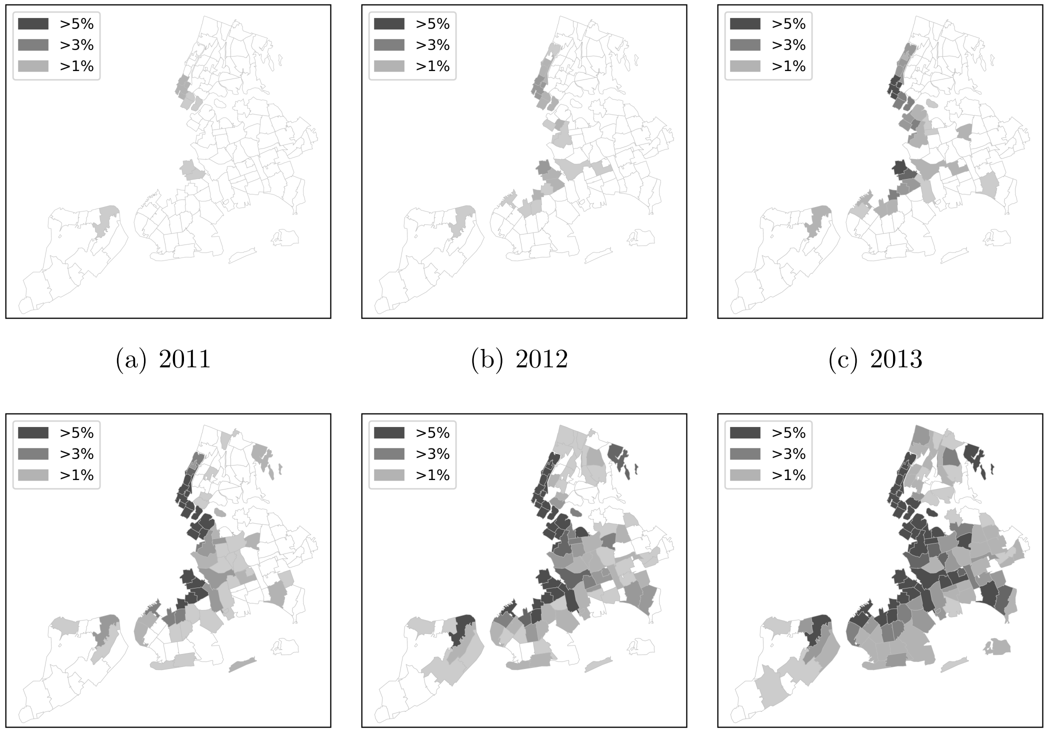
Figure 3.4 outlines Airbnb entry times and intensity levels across the zipcodes in *NYC Sample 2*. Airbnb intensity is defined as the Airbnb reviews per housing unit in a specific year.[12] We utilize the spatial and temporal variation in Airbnb intensity levels across NYC zipcodes to construct the following difference-in-difference (DID) specification which examines the role of Airbnb on NYC restaurant employment:

$$log(Restaurant\ Employment)_{i,t} = \beta_0 + \alpha_i + \delta_t + \beta_1 \cdot \frac{Airbnb\ Reviews_{i,t}}{Households_i} + X_{i,t} + \epsilon_{i,t} \tag{3.1}$$

where i represents the zipcode and t represents the year. Our variable of interest is $\frac{Airbnb\ Reviews_{i,t}}{Households_i}$, which is the ratio of the number of Airbnb reviews written for Airbnb listings to the number of households in zipcode i during year t.[13] We use this ratio as a proxy for Airbnb intensity in a zipcode. In section 3.5, we conduct various robustness checks for this measure as a proxy for Airbnb entry and intensity. α_i is a fixed effect for zipcode i which captures time invariant unobserved local characteristics for each zipcode. These include demographic and economic variables without significant year to year changes. For example, if two zipcodes have large differences in the number of households, then this will impact the number of restaurants in the zipcode and, as a result, the restaurant employment. As such, the zipcode fixed effects capture the average level of employment in an area. Furthermore, δ_t is a fixed effect for the year. This captures unobserved factors that impact restaurant employment in NYC as a whole for a specific time period. For example, there may be an event which increases the number of tourists that come to NYC which will positively impact all NYC zipcodes in a year.

[12]This measure defines Airbnb intensity throughout this paper. In section 3.5.3 we examine the robustness of our results to different measures of Airbnb intensity and find consistent results.

[13]The denominator in $\frac{Airbnb\ Reviews_{i,t}}{Households_i}$ is time invariant. It is simply to cross-sectionally normalize the size of each area i.



(a) 2011 (b) 2012 (c) 2013

(d) 2014 (e) 2015 (f) 2016

Fig. 3.4. NYC Zipcodes and Airbnb Intensity for *NYC Sample 2*

$X_{i,t}$ is a vector of local time varying controls. This includes the local restaurant popularity index calculated from the Yelp reviews for restaurants in zipcode i. As previously detailed, this variable helps control for variation in restaurant employment that is hard to account for with other time invariant observables. We also include local market structure variables from the BLS Business Pattern Data. Specifically, we include the number of employees in the hotel industry and the number of employees in the retail industry. The number of hotel industry employees is correlated with both the number, size, and performance of hotels in an area. Therefore, this variable controls for the impact of an increase in visitors that are using hotels as opposed to Airbnb and are utilizing local restaurants. This is particularly important given that the number of hotels in NYC increased by 35% between 2004 and 2013 and that many of these new hotels were located outside of the central tourist areas.[14] The number of retail employees controls for the number, size, and performance of retail stores in a local area. This is correlated with improving retail establishment conditions and relates to overall improving economic conditions in an area. Table 3.1 shows the summary statistics for the variables used in Equation 3.1.

Table 3.1.
Summary Statistics

		2011	**2012**	**2013**	**2014**	**2015**	**2016**
Restaurant Employment	Total	72,317	75,230	80,764	86,204	90,830	96,143
	Median	452	508	494	550	600	620
	Mean	598	622	667	712	751	795
	St. Dev	420	430	486	521	539	571
Airbnb Intensity							
	Median	0.0%	0.0%	0.1%	0.5%	1.1%	2.3%
	Mean	0.1%	0.3%	0.8%	2.0%	4.2%	6.4%
	St. Dev	0.3%	0.6%	1.4%	3.6%	7.0%	9.7%
Airbnb Reviews	Total	3,333	8,291	20,599	54,028	113,728	176,893
	Median	0	7	23	82	214	417
	Mean	28	69	170	447	940	1,462
	St. Dev	64	137	327	892	1,857	2,709
Local Popularity Index							
	Median	0.006	0.009	0.011	0.014	0.020	0.022
	Mean	0.018	0.021	0.024	0.029	0.036	0.041
	St. Dev	0.029	0.030	0.033	0.037	0.043	0.048
Hotel Employment	Total	3,206	3,466	3,558	3,340	3,462	3,542
	Median	2	2	2	2	2	2
	Mean	26	29	29	28	29	29
	St. Dev	83	81	83	74	78	76
Retail Employment	Total	167,670	171,805	174,062	180,249	187,324	187,831
	Median	1,086	1,100	1,085	1,158	1,146	1,203
	Mean	1,386	1,420	1,439	1,490	1,548	1,552
	St. Dev	1,164	1,198	1,197	1,213	1,236	1,215

To further control for time varying local economic and social factors that may impact restaurant employment, it would be useful to include the number of restaurant establishments in area i at time t in $X_{i,t}$. However, since we are examining the potential for Airbnb intensity to impact restaurant employment, Airbnb intensity

[14]http://prattcenter.net/sites/default/files/hotel_development_in_nyc_report-pratt_center-march_2015.pdf.

may also have an impact on the number of restaurants. This will bias our estimates. To alleviate this concern we include the residuals of the following specification in $X_{i,t}$:

$$log(Restaurant\ Establishments)_{i,t} = \beta_0 + \alpha_i + \delta_t + \beta_1 \cdot \frac{Airbnb\ Reviews_{i,t}}{Households_i} + X^*_{i,t} + \epsilon^*_{i,t} \tag{3.2}$$

$X^*_{i,t}$ contains the restaurant popularity index, the log of the number of employees in the hotel industry, and the log of the number of employees in the retail industry. As a result, the residuals, $\epsilon^*_{i,t}$, captures the unobserved variation in restaurant establishments, after the impact of Airbnb intensity, local popularity, retail, and hotel employment have been controlled for. We refer to $\epsilon^*_{i,t}$ as *Adjusted Restaurants Count* and include it in $X_{i,t}$ in Equation 3.1. Finally, the error term, $\epsilon_{i,t}$, in Equation 3.1 is the unobserved random shock associated with a area (i) during a specific time (t). We calculate robust standard errors that allow $\epsilon_{i,t}$ to be correlated for a specific zipcode i across time t (Moulton, 1990).

This identification strategy requires plausible conditional exogenous variation in Airbnb intensity. A violation of this requirement could manifest in two ways: 1) a violation of the parallel trends assumption or 2) omitted variables and simultaneity issues. With regard to point 1, in section 3.3.4 we present various tests examining the validity of this assumption. We find strong and consistent evidence that the pre-Airbnb restaurant employment trends are parallel across all areas in our sample. With regard to point 2, we explore the drivers of Airbnb intensity, given our constructed sample and specification, to examine the plausibility of this concern.

The sample construction procedure utilized to assemble *NYC Sample 2* excludes zipcodes with significant changes in income, educational attainment, and residential construction. Therefore, the residents in the zipcodes used in our analysis had consistent levels of income and educational attainment for the period considered. Furthermore, these zipcodes did not experience significant increases in new housing construction. As such, our specification would be susceptible to simultaneity issues under the following circumstance: (i) There exists unobserved variation that simultaneously impacts restaurant employment and Airbnb intensity. (ii) This simultaneous impact occurs in areas that do not experience significant changes in income, educational attainment, and residential construction. (iii) This unobserved simultaneous factor must be time variant (given the area level fixed effects) and systematic. It is not enough for this factor to sporadically and/or randomly occur in certain areas (i.e., one area has a concert in a specific year). (iv) It cannot be accounted for by the time-varying controls included in the main specification, $X_{i,t}$.

To examine the plausibility of such a scenario, we consider the potential sources of Airbnb intensity variation. In our view, there are two potential systematic sources driving the temporal variation in Airbnb intensity: (i) An area's economic condition improves, making it more attractive for home-sharing visitors or (ii) Home-sharing popularity among both hosts and visitors is increasing due to the increased levels of comfort and awareness consumers have with home-sharing platforms. Given the rich set of controls and the fact that *NYC Sample 2* excludes areas with changes in income, educational attainment, and residential construction, is is significantly more plausible that the systematic variation in Airbnb intensity is driven by the latter. Given the outlined conditions necessary for simultaneity to be a concern, in conjunction with the potential drivers of Airbnb variation, we are unable to identify a factor that meets all the conditions. As a result, we are confident that omitted variables and simultaneity do not systematically pose an endogeneity concern to specification 3.1. Nevertheless, we revisit this issue in Section 3.3.3 for robustness.

3.3.3 Main Results

Table 3.2 displays the results for the specification in Equation 3.1 at the zipcode level of analysis. Column 1 shows the results for the analysis which includes all 157 zipcodes in NYC.[15] Column 2 of the table reports the results of the speci-

[15] This includes all NYC zipcodes with more than 20 restaurants and more than 1,000 households.

fication in Equation 3.1 using only the zipcodes identified in *NYC Sample 2* with only Airbnb intensity $\left(\frac{Airbnb\ Reviews_{i,t}}{Households_{i,t}}\right)$, year effects (δ_t), and zipcode fixed effects (α_t) estimated. Columns 3 adds controls for the local restaurant popularity index and *Adjusted Restaurants Count.* Column 4 incorporates local employment controls, specifically *log(Hotel Employees)* and *log(Retail Employees).* Across all specifications, our results indicate that Airbnb has a positive and salient impact on restaurant employment in a zipcode. The coefficient for the specification estimated with *NYC Sample 2* with all the covariates (column 4) is 1.026. This result indicates that if Airbnb intensity in a zipcode increases by 1%, then restaurant employment will increase by approximately 1.03%.

Table 3.2.
Airbnb Zipcode Level Impact on Restaurant Employment

	(1)	(2)	(3)	(4)
Dep. Variable:	**log(Rest. Employment)**	**log(Restaurant Employment)**	**log(Restaurant Employment)**	**log(Restaurant Employment)**
Airbnb Reviews per Household	**0.481*****	**1.206*****	**1.065*****	**1.026*****
	(0.091)	(0.271)	(0.188)	(0.191)
Local Rest. Popularity	1.265***		1.569***	1.653***
	(0.251)		(0.421)	(0.417)
Adjusted Restaurants Count	0.767***		0.736***	0.736***
	(0.061)		(0.073)	(0.074)
log(Hotel Employees)	0.030***			0.020*
	(0.010)			(0.012)
log(Retail Employees)	0.178***			0.033
	(0.066)			(0.069)
Zipcode Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Constant	4.982***	5.993***	5.982***	5.725***
	(0.477)	(0.016)	(0.013)	(0.473)
Observations	1,570	1,210	1,210	1,210
R-squared	0.652	0.488	0.588	0.591
Number of zipcodes	157	121	121	121

Matching

Our results have thus far shown that Airbnb has a positive impact on restaurant employment in NYC. We have identified the effect of Airbnb by utilizing a DID framework that incorporates a rich set of controls including fixed effects and time varying observables. We now incorporate matching to further evaluate the robustness of these finding with regard to any lingering endogeneity concerns associated with the conditional exogeneity assumption in the DID framework. That is, given the controls in our specification, does there remain an unobserved time varying factor that impacts both Airbnb and restaurant employment simultaneously. While this concern is unlikely given the aforementioned controls, we utilize algorithmic matching to examine the robustness of this claim and reduce model dependency biases (Ho et al., 2007; Imai et al., 2008).

Using the zipcodes in *NYC Sample 2* (121 zipcodes), we use matching to define a subset of the data where zipcodes with significant Airbnb activity are matched with zipcodes without significant Airbnb activity. By comparing areas with similar characteristics—except for Airbnb intensity—we remove biases associated with zipcodes that are not comparable to any other zipcodes in the sample in terms of their capacity to attract Airbnb visitors (Heckman et al., 1998). Before matching the zipcodes, we must select criteria under which we define each zipcode as either a treated or control zipcode. We use 2016 Airbnb activity to determine whether a zipcode is treated. Based on the distribution of 2016 Airbnb reviews per household (Airbnb intensity), we establish upper and lower treatment criteria. If a zipcode has more Airbnb intensity in 2016 than the upper treatment criterion, then that zipcode is considered treated. If a zipcode has less Airbnb intensity in 2016 than the lower treatment criterion, then that zipcode is considered a control. Zipcodes where the number of 2016 Airbnb reviews falls between the upper and lower treatment criteria are removed. Unavoidably, this measure of treatment is subjective. Therefore, to assuage doubts regarding model dependency, we present a complete sensitivity analysis for all matching results based on various treatment definitions.

To match treated and control zipcodes, we compare a set of pretreatment local variables. The pretreatment variables are selected so as to predict the probability of an area obtaining an amount of Airbnb reviews that is higher than the upper criteria mentioned above. Since pretreatment variables are designed to predict Airbnb activity, we focus on the year 2008, which is the year prior to significant Airbnb activity. First, we include the number of households in a zipcode. Zipcodes with more households have a larger pool of potential Airbnb supply and, as such, will likely have more Airbnb activity. We also include the number of retail establishments and hotels in the zipcode. Areas that have underlying foundational structures that attracted visitors before Airbnb are likely to attract a larger number of Airbnb visitors. Furthermore, drawing on findings from Quattrone et al. (2016), we include the ratio of homes that are rented and the median income of residents in a zipcode. Quattrone et al. (2016) find that these factors have a persistent effect on Airbnb intensity in London.

To conduct the pretreatment matching, we use the aforementioned pretreatment matching variables as the predictors in a logistic regression to obtain the conditional probability of treatment (Propensity Score) for each zipcode (Rosenbaum and Rubin, 1983; Caliendo and Kopeinig, 2008). For each treated zipcode, we find the nearest neighbor by comparing the conditional probability of treatment to a set of control zipcodes. We remove the zipcodes that are not matched. This leaves us with a reduced subsample of zipcodes where each zipcode with a high level of Airbnb activity is matched with a zipcode with low Airbnb activity. Specifically, we first identify the Airbnb intensity (Airbnb reviews to households) in 2016 of each zipcode. We then calculate the 70$^{\text{th}}$ percentile of the distribution of 2016 Airbnb intensity levels (5.03%). Any zipcode with a level of Airbnb intensity that is greater than the 70$^{\text{th}}$ percentile is defined as a treated zipcode. To identify control zipcodes we calculate the 35$^{\text{th}}$ percentile of the distribution of 2016 Airbnb intensity (1.35%). A zipcode with 2016

Airbnb intensity that is lower than the 35^{th} percentile is defined as a control zipcode. All other zipcodes are discarded. Using the aforementioned Propensity Score, each treated zipcode is matched with a control, and the zipcodes that are not matched are discarded. This process results in a subset of 52 zipcodes.

Column 1 of Table 3.3 shows the results of Equation 3.1 on only the matched subsample. Once again, the results indicate a positive and salient impact of Airbnb activity on restaurant employment. The coefficient size for Airbnb Reviews per Household indicates that if the Airbnb intensity in a zipcode increases by 1%, then restaurant employment would increase by approximately 0.95%.

Table 3.3.
Airbnb Zipcode Level Impact on Restaurant Employment with Matching

	(1)	(2)	(3)
Dep. Variable:	**log(Restaurant Employment)**	**log(Restaurant Employment)**	**log(Restaurant Employment)**
Airbnb Reviews per Household	**0.951*****	**0.786*****	**1.072*****
	(0.228)	(0.282)	(0.279)
Local Rest. Popularity	1.982**	2.160***	2.548***
	(0.943)	(0.749)	(0.776)
Adjusted Restaurants Count	0.746***	0.633***	0.650***
	(0.103)	(0.079)	(0.083)
log(Hotel Employees)	-0.012	0.014	-0.007
	(0.011)	(0.014)	(0.012)
log(Retail Employees)	0.078	0.076	0.063
	(0.117)	(0.086)	(0.092)
Zipcode Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Constant	5.415***	5.448***	5.558***
	(0.804)	(0.589)	(0.627)
Observations	520	592	528
R-squared	0.656	0.604	0.626
Number of zipcodes	52	97	89

Matching: Sensitivity Analysis

Since the selection of treatment in our design is subjective, we conduct a sensitivity analysis on the upper and lower treatment criteria. Table 3.4 displays the estimated coefficient values for the *Airbnb Reviews per Household* variable from Equation 3.1 for different specifications of upper and lower treatment criteria based on the distribution of Airbnb intensity in 2016. Specifically, each row represents the minimum boundary for treatment based on whether the Airbnb intensity indicator (ratio of Airbnb activity to households) was greater than the respective treatment criteria for that zipcode in 2016. For example, the coefficient in the first row and first column represents a minimum treatment threshold corresponding to the 70$^{\text{th}}$ percentile and a maximum untreated threshold corresponding to the 35$^{\text{th}}$ percentile. This entails that only the upper 30$^{\text{th}}$ percentile and lower 35$^{\text{th}}$ percentile of observations (according to 2016 Airbnb intensity) will be considered in the matching phase. The matching phase will then match each of the treated zipcodes with an untreated zipcode for the specific treatment criteria. The results indicate that for all specifications of treatment criteria Airbnb has a positive impact on restaurant employment. The findings indicate that a 1% increase in Airbnb intensity results in an increase in restaurant employment between 0.9% and 1.1% at the zipcode level.

Table 3.4.
Matching Sensitivity Analysis for Airbnb Treatment (Dep. Var.: Restaurant Employment)

		Lower Treatment Percentile				
		35$^{\text{th}}$%	40$^{\text{th}}$%	45$^{\text{th}}$%	50$^{\text{th}}$%	55$^{\text{th}}$%
Upper	70$^{\text{th}}$%	0.951***	1.014***	0.923***	0.935***	1.069***
Treatment	65$^{\text{th}}$%	0.958***	0.893***	0.981***	1.004***	0.982***
Percentile	60$^{\text{th}}$%	1.031***	1.058***	0.979***	1.028***	1.098***

Dynamic Matching

The matching approach utilized thus far is dependent upon the following assumption: conditional on the time varying controls incorporated in the model and year fixed effects, an area's attractiveness to Airbnb activity does not change. This implies that two matched areas retain the same level of attractiveness (conditional on controls). To remove the dependency of this assumption and examine the robustness of our results, we also perform dynamic matching to examine the role of time varying economic changes in driving our results. Specifically, we conduct the following procedure:

1. Identify each zipcode/year combination with an Airbnb intensity greater than 2%. We consider these zipcode/year combinations as treated.

2. For each identified zipcode/year combination, using all other zipcodes in the same year as potential controls, conduct nearest neighbor matching (based on propensity score) to match the treated zipcode/year combination with the most similar control zipcode based on the following criteria:

    - Restaurant employment in the previous year.
    - The number of Starbucks locations in the current year.
    - The local popularity index in the current year.
    - The number of hotels in the current year.

We match on the restaurant employment to find an area with comparable employment in the previous year. We include the number of Starbucks locations as it is often seen as an indicator of gentrification and improving economic conditions (Glaeser et al., 2018). We include the local popularity index and the number of hotels to give a measure of the relative attractiveness of an area's restaurants to locals and visitors. Therefore, the comparison is between areas where the restaurant employment in the previous year is similar and the current year variables are also comparable.

We estimate specification 3.1 and include all the dynamically matched zipcode/year combinations. We additionally include all observations for any zipcode that was matched for the period between 2007-2010.[16] The results are presented in columns 2 (dynamic matching on zipcodes in *NYC Sample 1*) and 3 (dynamic matching on zipcodes in *NYC Sample 2* of Table 3.3 and both indicate a positive and salient impact of Airbnb on restaurant employment.

3.3.4 Examining the Parallel Trends Assumption

A key assumption of DID specification 3.1 is that the restaurant employment trends in the periods preceding significant Airbnb activity in an area must be parallel, conditional on controls. To examine the validity of this assumption, we employ two strategies. First, we examine the impact of Airbnb intensity on alternative industries that should not be impacted by visitor activity but should increase with improving economic conditions in an area. Second, we directly test this assumption using the leads and lags model (Autor, 2003).

Falsification Industries

We identify four industries that should not be impacted by increases in Airbnb intensity: banking services, pet care, fitness, and hair/beauty salons. All these industries would be expected to experience growth during periods of improving economic conditions that result from changes such as gentrification. However, Airbnb visitors' expenditures should not have an impact on these industries. Columns 1-4 of Table 3.5 present the results where the dependent variable is the log of employment in these

[16]By including the observations for the periods preceding Airbnb entry into NYC, we are able to replicate the DID structure from Specification 3.1.

sectors respectively.[17] Reassuringly, the results indicate that Airbnb intensity does not have an impact on any of these industries.

Table 3.5.
Falsification Tests

Dep. Variable:	(1) Comm. Banking Employment	(2) Pet Care Employment	(3) Fitness Employment	(4) Beauty Salon Employment
Airbnb Reviews per Household	**0.157** (0.207)	**0.561** (0.515)	**-0.303** (1.083)	**0.089** (0.171)
Local Rest. Popularity	-0.591 (0.708)	0.258 (0.814)	-1.325 (2.250)	-0.358 (0.423)
log(Falsification Industry Establishments)	1.093*** (0.088)	1.833*** (0.051)	2.127*** (0.123)	1.103*** (0.063)
log(Hotel Employees)	0.004 (0.021)	-0.001 (0.010)	0.033 (0.043)	0.005 (0.010)
log(Retail Employees)	0.208** (0.094)	0.132 (0.091)	-0.106 (0.224)	-0.033 (0.053)
Zipcode Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Constant	0.655 (0.646)	-0.900 (0.626)	1.016 (1.493)	1.052*** (0.355)
Observations	1,210	1,210	1,210	1,210
R-squared	0.475	0.857	0.534	0.652
Number of Zipcodes	121	121	121	121

Leads and Lags Model

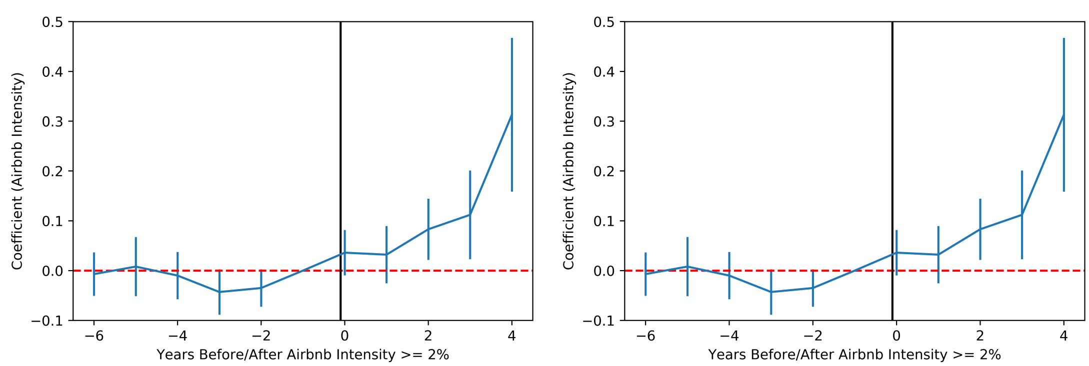
The falsification tests in section 3.3.4 provide support that the parallel trends assumption is not violated. However, we directly test this assumption using the relative time model (leads and lags model). Specifically, we incorporate leads and lags dummies that indicate the temporal difference (in years) between an observation and the time of treatment. Since this test requires a binary treatment variable, we estimate two versions of this model, one with the treatment criterion set at 2%

[17] Since we do not have to worry about the impact of Airbnb on the growth of falsification industry establishments, we do not need to derive an adjusted establishment counts for each category. Consequently, we include the log of the number of establishments in the specific falsification industry analyzed.

intensity of Airbnb and another set at 3% Airbnb intensity. The model is specified as follows:

$$\begin{aligned} log(Restaurant\ Employment)_{i,t} &= \alpha_i + \delta_t + \sum_j \tau_j \cdot PreAirbnb(j) \\ &+ \cdot\delta AirbnbBinary_{i,t} + \sum_k \omega_k \cdot PostAirbnb(k) + X_{i,t} + \epsilon_{i,t} \quad (3.3) \end{aligned}$$

Figure 3.5 provides a graphical representation of the leads and lags coefficients and their estimated confidence intervals. In both versions (2% and 3%) the results indicate that, in the periods proceeding Airbnb activity, there is not a statistically significant difference in the restaurant employment between the Airbnb and non-Airbnb intense areas. However, in the periods after Airbnb intensity reaches 2% or 3%, the impact of Airbnb intensity on restaurant employment is positive, significant, and growing.



(a) (b)

Fig. 3.5. Leads and Lags Coefficients

3.4 Evidence Supporting the Validity of the Underlying Mechanism

3.4.1 The Impact of Airbnb on Yelp Visitor Reviews

Thus far, our results indicate that Airbnb has a positive impact on restaurant employment. The necessary underlying mechanism is that Airbnb visitors' are frequenting local restaurants. Therefore, to evaluate the validity of this mechanism, we use the Yelp visitor data that we collected to assess the impact of Airbnb activity on Yelp visitors' restaurant review behavior. We utilize the following DID specification:

$$\frac{\textit{Yelp Visitor Reviews}_{i,t}}{\textit{Yelp Visitor Reviews}_{t}} = \beta_0 + \alpha_i + \delta_t + \beta_1 \cdot \frac{\textit{Airbnb Reviews}_{i,t}}{\textit{Households}_{i,t}} + X_{i,t} + \epsilon_{i,t} \quad (3.4)$$

This is the same specification as Equation 3.1 except that the dependent variable is the proportion of NYC Yelp visitor reviews that were written for restaurants in zipcode i. This captures the spatial distribution of restaurant visitor activity across NYC. If Airbnb activity is impacting restaurant employment, then the areas with increasing Airbnb activity should capture an increasing proportion of the NYC restaurant visitor activity.

In parallel with our restaurant employment analysis, we control for zipcode fixed effects (α_i), year fixed effects(δ_t), and a vector of local time varying controls $X_{i,t}$, which includes retail employment, hotel employment, restaurant local popularity index, and *Adjusted Restaurants Count*. The error term, $\epsilon_{i,t}$, is the unobserved random shock associated with a zipcode (i) during a specific time (t).

Table 3.6 presents the results of Equation 3.4. Column 1 of the table reports the results of the specification in Equation 3.4 for the zipcodes in *NYC Sample 2*. Column 2 reports the results of the matched sample of zipcode. The results indicates that Airbnb has a positive and salient impact on the proportion of NYC Yelp visitor reviews written in a zipcode. To provide economic interpretation for the coefficient (Airbnb Reviews per Household) we evaluate the effect on a hypothetical zipcode where the proportion of NYC visitor reviews is 0.057%—the median value in 2012. If

Airbnb intensity, as measured by *Airbnb Reviews per Household*, increased by 1% then the proportion of Yelp visitor restaurant reviews in this zipcode would increase by approximately 0.0023%.[18] Given that the current proportion of NYC visitor reviews is 0.057%, this represents a 4.04% increase in the proportion of Yelp visitor reviews. Furthermore, given that NYC tourist restaurant spending was $7.4 billion in 2012, the extra 0.0023% that would be captured by the median zipcode would translate to approximately $170,000 of extra tourist restaurant expenditure.

Table 3.6.
Airbnb Impact on Proportion of NYC Visitor Yelp Reviews

	(1)	(2)	(3)
Dep. Variable:	**Prop. NYC Yelp Visitor Reviews**	**Prop. NYC Yelp Visitor Reviews**	**Prop. NYC Yelp Local Reviews**
Airbnb Reviews per Household	**0.150****	**0.229*****	**-0.063**
	(0.064)	(0.054)	(0.056)
Local Rest. Popularity	0.759***	1.098***	5.480***
	(0.198)	(0.329)	(0.231)
Adjusted Restaurants Count	0.037**	0.051**	0.020*
	(0.018)	(0.025)	(0.011)
log(Hotel Employees)	0.000	-0.002	-0.001
	(0.002)	(0.003)	(0.002)
log(Retail Employees)	0.001	-0.006	0.010
	(0.008)	(0.011)	(0.013)
Year Fixed Effects	Yes	Yes	Yes
Constant	0.048	0.085	-0.022
	(0.059)	(0.074)	(0.085)
Observations	1,210	520	520
R-squared	0.363	0.517	0.929
Number of zipcodes	121	52	52

We also estimate equation 2 for the various matching treatment criteria that were explained in section 3.3.3. Table 3.7 shows the coefficient for Airbnb Reviews per Household from Equation 3.4 for the various treatment criteria. Across all specifi-

[18]This is based on results from matched zipcodes, column 2 of Table 3.6.

cations, the results indicate that, Airbnb has a positive and salient impact on the proportion of Yelp visitor reviews in a locality.

Table 3.7.
Matching Sensitivity Analysis for Airbnb Treatment (Dep. Var.: Proportion Yelp Visitor Reviews)

		Lower Treatment Percentile				
		35^{th}%	40^{th}%	45^{th}%	50^{th}%	55^{th}%
Upper	70^{th}%	0.231***	0.234***	0.217***	0.222***	0.237***
Treatment	65^{th}%	0.233***	0.234***	0.221***	0.230***	0.241***
Percentile	60^{th}%	0.237***	0.242***	0.225***	0.227***	0.234***

3.4.2 The Impact of Airbnb on Yelp Local Reviews

To isolate the impact of Airbnb on the proportion of NYC Yelp visitor reviews, we controlled for time invariant zipcode effects, year effects, as well as a multitude of time varying local factors. While our results provide evidence regarding the impact of Airbnb activity on visitors' restaurant activity, its effect on the restaurant activity of local residents remains uncertain. If the aforementioned controls have adequately captured the local activity that relates to the attractiveness of an area to both Airbnb visitors and restaurants, then Airbnb activity should not have a positive impact on the proportion of NYC local Yelp reviews. A positive effect would imply that local residents are frequenting restaurants more often because of Airbnb, which is unrealistic. However, it may be the case that Airbnb has a negative impact on residents' restaurant behavior, as opposed to no impact. First, Airbnb visitors that do not share their rented homes with the hosts are temporarily displacing local residents and reducing the potential market size of resident restaurant demand. Second, a local resident's utility associated with frequenting a local restaurant may be reduced if that restaurant is now becoming more popular among visitors.

To assess the impact of Airbnb on local demand, as well as further validating the effectiveness of the controls in specification 3.1, we replace the proportion of Yelp visitor reviews with the proportion of Yelp local reviews in Equation 3.4. Effectively, this also serves as a falsification test for the main mechanism. The following outlines this specification:

$$\frac{\textit{Yelp Local Reviews}_{i,t}}{\textit{Yelp Local Reviews}_{t}} = \beta_0 + \alpha_i + \delta_t + \beta_1 \cdot \frac{\textit{Airbnb Reviews}_{i,t}}{\textit{Households}_{i,t}} + X_{i,t} + \epsilon_{i,t} \quad (3.5)$$

The results are presented in Columns 3 of Table 3.6. The impact of Airbnb intensity is negative and insignificant. This indicates that while Airbnb has a positive and extremely salient impact on visitor activity, its impact on local resident activity is not as clear. Table 3.8 shows the sensitivity of this result to various treatment criteria for the matching analysis. The results support the findings of a negative leaning effect of Airbnb on local demand.

Table 3.8.
Matching Sensitivity Analysis for Airbnb Treatment (Dep. Var.: Prop. Yelp Local Reviews)

		Lower Treatment Percentile				
		$35^{th}\%$	$40^{th}\%$	$45^{th}\%$	$50^{th}\%$	$55^{th}\%$
Upper	$70^{th}\%$	-0.061	-0.057	-0.074	-0.076	-0.063
Treatment	$65^{th}\%$	-0.067	-0.063	-0.072	-0.066	-0.064
Percentile	$60^{th}\%$	-0.065	-0.060	-0.075	-0.075	-0.073

3.4.3 Evidence from a Policy Shift in New Orleans

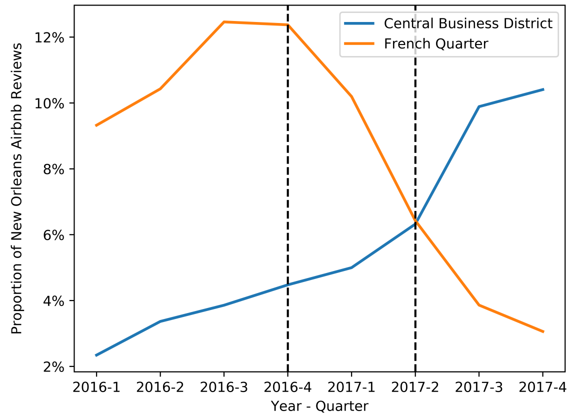
Thus far, we have relied upon the time varying entry and intensity of Airbnb into various areas to identify its impact on local restaurant activity. We have incorporated various controls, including area and time fixed effects as well as time varying local factors, that impact restaurants. This DID specification, specifically given the rich set of controls, has provided a framework that utilizes the conditional exogeneity of Airbnb intensity to identify the causal impact. However, access to a setting with a regulatory based policy shift would relieve the need for conditional exogeneity and further validate the consistency of the results. This policy shift is not available in NYC, however, a neighborhood level policy on Airbnb legality was implemented recently in New Orleans. In 2017, after various discussions between New Orlean's officials and Airbnb, the New Orlean's City Council voted to legalize short-term rentals in the city (these had previously been illegal but Airbnb was active in the city regardless). However, to legally run a short term rental property the hosts are required to register with the city. The city offered various types of short term rental registration options. This new policy also included a ban on short term rentals in the French Quarter neighborhood, which is an extremely popular tourist destination.[19] New Orleans officials also claimed that they would fine those hosts that were not in compliance with the new regulation.

As a result of this new policy, Airbnb supply shifted heavily away from the French Quarter neighborhood, which had attracted a significant number of Airbnb listings and visitors due to the popularity of the location. At the same time, the Central Business District, a bordering neighborhood, experienced significant increases in Airbnb demand. Figure 3.6 shows the temporal impact of the policy on the proportion of Airbnb reviews associated with locations in the French Quarter and Central Business District neighborhoods, respectively. The dashed lines represent the timing of the policy. The first dashed lines indicates the announcement of the upcoming policy (Q4:2016) and second is the actual implementation of the policy (Q2:2017). The graph indicates a clear impact of the policy on Airbnb activity in the two neighborhoods.

This policy provides a unique opportunity to further assess the underlying mechanism behind our results. An exogenous policy shock shifts Airbnb from one neighborhood to a nearby neighborhood. We can then evaluate the impact on Yelp visitor

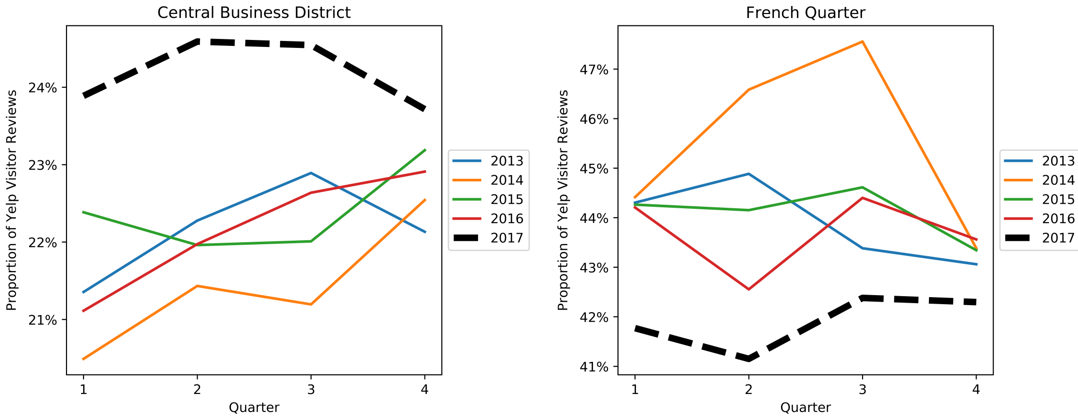
[19]This ban did not extend to all of the French Quarter as some streets were exempted from the ban due to zoning regulations.

Fig. 3.6. New Orleans Airbnb Policy and Airbnb Activity



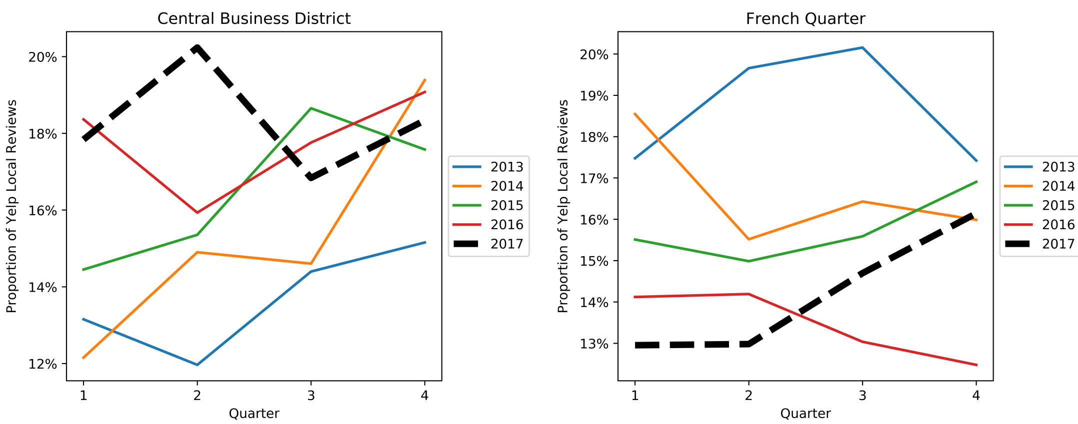
and local activity in the same time period. We collect Yelp restaurant reviews and Airbnb review data for New Orleans. We aggregate this data to the New Orleans neighborhood level based on the neighborhood definitions used by Airbnb.[20] Figure 3.7 plots the proportion of New Orleans Yelp visitor reviews that are captured by each neighborhood respectively. The x-axis represents the quarter and each year is plotted separately. This is to account for seasonal factors of demand. We are interested in the dashed line which represents 2017, the year in which the policy was implemented. In the Central Business District, Yelp visitor review activity is mixed for all the years prior to 2017. However, in 2017, it has a significant increase across all quarters. Conversely, in the French Quarter, Yelp visitor reviews are decreasing. Consistent with our results, Airbnb activity appears to drive visitor restaurant behavior. We also assess the impact on local reviews in Figure 3.8 and find that the impact of the policy does not significantly vary the behavior of local residents.

[20] Insideairbnb.com provides shapefiles they obtain from the Airbnb website which outline the boundaries for each neighborhood in a city. We use these boundaries to identify whether a restaurant is located in that neighborhood by mapping the geocoded location of a restaurant and the neighborhood boundaries.



(a) (b)

Fig. 3.7. Effect of New Orleans Airbnb Policy on the Proportion of Yelp Visitor Reviews



(a) (b)

Fig. 3.8. Effect of New Orleans Airbnb Policy on the Proportion of Yelp Local Reviews

3.5 Robustness Checks

3.5.1 Neighborhood Level Analysis

To examine the robustness of our aggregation level (zipcodes), we replicate our analysis at the NYC neighborhood level. Neighborhoods are large enough to be selfsufficient if a visitor wishes to remain locally, but are also not so large that it would be impossible for a visitor to choose to commute to other neighborhoods if they desire. Furthermore, neighborhoods are organized in a manner so as to represent similar economic, demographic, and market structure characteristics. We determine neighborhoods based on the boundaries indicated by the New York State Department of Health.[21] The report splits NYC into 42 neighborhoods and allocates each zipcode in NYC to a neighborhood.

We estimate specification 3.1 at the neighborhood aggregation level and present the result in Table 3.9. Column 1 reports the results estimated using all 42 neighborhoods in New York City. Column 2 of the table reports the results of the specification in Equation 3.1 using only the neighborhoods identified in *NYC Sample 2* with only Airbnb intensity $\left(\frac{\textit{Airbnb Reviews}_{i,t}}{\textit{Households}_{i,t}}\right)$, year effects (δ_t), and neighborhood fixed effects (α_t) estimated. Columns 3 adds controls for the local restaurant popularity index and *Adjusted Restaurant Counts*. Column 4 incorporates local employment controls, specifically *log(Hotel Employees)* and *log(Retail Employees)*. Across all specifications, our results indicate that Airbnb has a positive and salient impact on restaurant employment in a neighborhood.

3.5.2 Neighborhood Level Matching

Our motivation for removing the *traditionally tourist areas* was due to the endogenous characteristics associated with their capacity to attract visitors. However, to further assuage any remaining doubts regarding this effect, we can utilize matching. Specifically, we can identify neighborhoods with significant Airbnb activity and find similar neighborhoods without Airbnb presence. For example, assume there are two neighborhoods where each neighborhood is able to attract x amount of visitors in the period before Airbnb. The value of x is indicative of each neighborhood's capacity to attract visitors. We assume that, without Airbnb, each neighborhood would

[21]https://www.health.ny.gov/statistics/cancer/registry/appendix/neighborhoods.htm.

Table 3.9.
Robustness Check: Airbnb Neighborhood Level Impact on Restaurant Employment

	(1)	(2)	(3)	(4)
Dep. Variable:	**log(Rest. Empl.)**	**log(Rest. Empl.)**	**log(Rest. Empl.)**	**log(Rest. Empl.)**
Airbnb Reviews per Household	**0.796*****	**1.359*****	**1.235*****	**1.106*****
	(0.133)	(0.414)	(0.322)	(0.258)
Local Rest. Popularity	0.756***		1.267**	1.121***
	(0.158)		(0.511)	(0.387)
Adjusted Restaurants Count	0.655***		0.714***	0.714***
	(0.079)		(0.127)	(0.105)
log(Hotel Employees)	0.058***			0.048***
	(0.017)			(0.014)
log(Retail Employees)	0.371***			0.217*
	(0.107)			(0.111)
Neighborhood Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Constant	4.313***	7.462***	7.442***	5.467***
	(0.931)	(0.028)	(0.018)	(0.928)
Observations	420	290	290	290
R-squared	0.855	0.767	0.835	0.863
Number of neighborhoods	42	29	29	29

attract a relatively similar proportion of visitors in the future. However, if Airbnb became popular in only one of the neighborhoods, then the difference in restaurant performance is attributable to Airbnb's impact. Therefore, in our sample, we pair neighborhoods based on the number of Yelp reviews in 2008 and Airbnb intensity in 2015. Our goal is to pair neighborhoods that have similar Yelp visitor activity in 2008 but differing Airbnb intensity in 2015. Table **??** shows the number of 2008 Yelp reviews for each neighborhood in the pruned sample. Each neighborhood pair is assigned manually and displayed in column 3. Column 5 of Table 3.9 shows the results of specification 3.1 on only the 24 matched neighborhoods. The coefficient for Airbnb Reviews per Household is positive and statistically significant providing further evidence of the impact of Airbnb on restaurant employment.

3.5.3 Measure of Airbnb Demand

In our analysis, we have used the ratio of Airbnb reviews to households as a proxy for Airbnb intensity. The denominator (households) is used to normalize the Airbnb activity by the number of potential hosts in a zipcode. An alternative method to normalize Airbnb activity is to use the log of the number of Airbnb reviews to represent Airbnb intensity. A log-log framework would allow us to capture the impact of percent changes in Airbnb activity on our dependent variables of interest. However, log transformations are problematic when the number of Airbnb reviews is small (i.e., using a log transformation, an increase in Airbnb reviews from 4 to 8 would be given greater importance than it warrants). While the ratio of Airbnb reviews to households does not suffer from this problem, we replace the independent variable in Equation 3.1 with a variable that takes on the following values: 0 if the ratio of Airbnb to households is less than 2%, otherwise its value is equal to *log(Airbnb Reviews)*. Column 1 of Table 3.10 shows the results for this specification and indates that the results are robust to this alternative defintion.

We also examine the robustness of our results to including a binary treatment definition for Airbnb activity instead of the continuous measure used thus far. Specifically, we create a binary variable that has a value if 0 if the Airbnb reviews per household for a specific zipcode/year combination is less than 1%, otherwise the variable has a value of one. We also construct a secondary binary variable that has a cutoff point of 2%. Columns 2 and 3 of Table 3.10 show the results for these specifications. The results indicate the our findings are robust to differing Airbnb intensity measures, including strict binary measures of treatment.

3.5.4 Measure of Restaurant Employment

To measure restaurant employment, we obtain data from the U.S. Bureau of Labor Statistics. Specifically, we include the following institution categories and their associated NAICS codes in defining the restaurant sector: full-service restaurants

Table 3.10.
Robustness Check: Alternative Definitions for Airbnb Intensity

	(1)	(2)	(3)
Dep. Variable:	**Airbnb Measure (1)**	**Airbnb Measure (2)**	**Airbnb Measure (3)**
Airbnb Intensity Measure	**0.012*****	**0.061*****	**0.092*****
	(0.003)	(0.022)	(0.023)
Local Rest. Popularity	1.830***	1.983***	1.840***
	(0.454)	(0.467)	(0.449)
Adjusted Restaurants Count	0.765***	0.772***	0.765***
	(0.078)	(0.078)	(0.078)
log(Hotel Employees)	0.022*	0.023**	0.023*
	(0.011)	(0.011)	(0.012)
log(Retail Employees)	0.055	0.062	0.059
	(0.070)	(0.070)	(0.069)
Year Fixed Effects	Yes	Yes	Yes
Constant	5.566***	5.521***	5.540***
	(0.478)	(0.478)	(0.475)
Observations	1,210	1,210	1,210
R-squared	0.587	0.586	0.587
Number of zipcodes	121	121	121

(722511), limited-service restaurants (722513), drinking places (722410), cafeteria and grill buffets and buffets (722514), and snack and nonalcoholic beverage bars (722515). To determine whether our results are robust to the selection of NAICS codes to measure restaurant employment, we create two new definitions. The first alternative definition includes only full-service restaurants and limited services restaurants. The second alternative definition adds drinking places to the first alternative. Table 3.11 shows the results of Equation 3.1 for the two alternative definitions of restaurant employment. The results are consistent with our main findings, providing evidence that our analysis is robust to alternative restaurant employment measures.

Table 3.11.
Robustness Checks: Alternative Definitions for Restaurant Employment

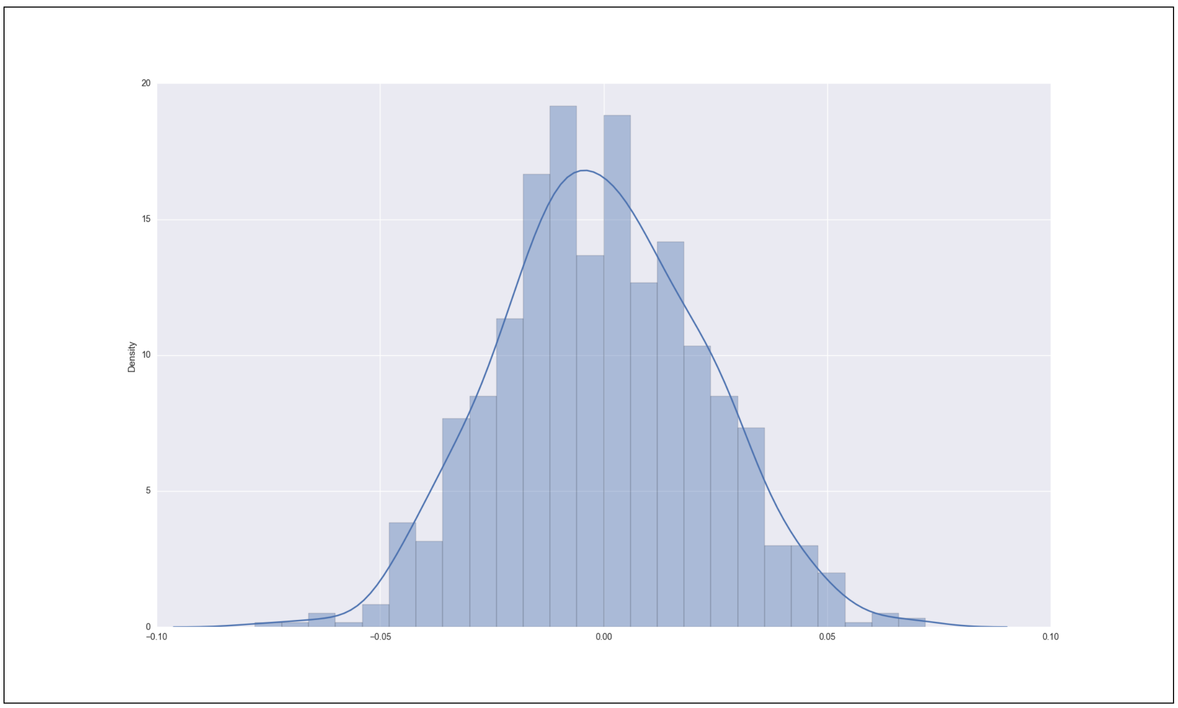
	(1)	(2)
Dep. Variable:	**log(Restaurant Employment)**	**log(Restaurant Employment)**
Airbnb Reviews per Household	**0.946*** **	**1.070*** **
	(0.200)	(0.201)
Local Rest. Popularity	1.746***	1.730***
	(0.483)	(0.460)
Adjusted Restaurants Count	0.728***	0.755***
	(0.081)	(0.081)
log(Hotel Employees)	0.024*	0.023*
	(0.014)	(0.013)
log(Retail Employees)	0.059	0.064
	(0.083)	(0.077)
Year Fixed Effects	Yes	Yes
Constant	5.374***	5.376***
	(0.568)	(0.527)
Observations	1,210	1,210
R-squared	0.519	0.541
Number of Zipcodes	121	121

3.5.5 Placebo Test

To determine the robustness of our results to a potential spurious effect driven by serial correlation of restaurant employment with unobserved activity, we implement a randomized treatment test (Bertrand et al., 2004). Specifically, we use the binary treatment allocation of Airbnb demand described in section 3.5.3 based on a 2% threshold for treatment. That is, if the number of Airbnb reviews by households for an area in a given year is greater than 2%, that area is considered treated. To perform the randomization, we randomly assign a treatment year to each zipcode and estimate specification 3.1 with the Airbnb intensity defined as the randomly assigned treatment variable. We repeat this process for 1,000 iterations.

Figure 3.9 displays the distribution of the resulting coefficient values. The distribution is centered around zero which indicates that, if the treatment was randomly assigned, the resulting impact of Airbnb intensity on restaurant employment would be non-existent. Furthermore, we conduct a t-test that evaluates whether the mean of the distribution is statistically different than zero and obtain a p-value of 0.51. This indicates that we fail to reject the hypothesis that the mean of the Airbnb intensity coefficients from randomly assigned treatments is different from zero. Finally, the coefficient value for the same specification that we estimated in section 3.5.3 is 0.092 (column 3 of Table 3.10). Based on the distribution in Figure 3.9, the probability associated with obtaining this coefficient in a randomized framework is less than 0.001.

Fig. 3.9. Distribution of Airbnb Binary Coefficient for Randomly Assigned Treatments



3.5.6 Robustness of Matching Method

In section 3.3.3, we utilized the propensity score of each matched unit as the distance metric to finding a matching zipcode. While propensity score as a matching metric is widely used in the literature (Dehejia and Wahba, 2002), to further alleviate concerns regarding choice dependency, we repeat the analysis of 3.3.3 for two other distance metrics: Mahalanobis Distance and Coarsend Exact Matching (CEM) (Iacus et al., 2012). We present the sensitivity analysis of a Mahalanobis distance based matching analysis and CEM in Tables 3.12 and 3.13 respectively. The results indicate a similarly positive and salient impact of Airbnb on restaurant employment, alleviating potential concerns regarding the selection of distance metric in our matching analysis.

Table 3.12.
Robustness Check: Matching (using Mahalanobis distance) Sensitivity Analysis for Airbnb Treatment (Dep. Var.: Restaurant Employment)

		Lower Treatment Percentile				
		35^{th}%	40^{th}%	45^{th}%	50^{th}%	55^{th}%
Upper	70^{th}%	0.867***	0.839***	0.860***	0.769***	0.897***
Treatment	65^{th}%	0.904***	0.904***	0.838***	0.884***	0.853***
Percentile	60^{th}%	0.891***	0.978***	0.993***	0.938***	0.933***

Table 3.13.
Robustness Check: Matching (using CEM) Sensitivity Analysis for Airbnb Treatment (Dep. Var.: Restaurant Employment)

		Lower Treatment Percentile				
		35^{th}%	40^{th}%	45^{th}%	50^{th}%	55^{th}%
Upper	70^{th}%	0.989***	1.023***	0.824***	0.855***	0.902***
Treatment	65^{th}%	0.984***	1.016***	0.816***	0.844***	0.891***
Percentile	60^{th}%	1.100***	1.123***	0.919***	0.934***	0.993***

3.6 Heterogeneity Analysis

Our results indicate that home-sharing platforms have a salient spillover effect on the economic performance of local complimentary services, specifically restaurants. This effect captures the average impact of Airbnb across NYC areas. We extend this analysis by evaluating the role of area level heterogeneity in driving the relationship between Airbnb and restaurant employment. We focus specifically on two categories: demographics and market structure.

3.6.1 Heterogeneous Impact of Airbnb Due to Local Demographics

Ideally, the economic spillover effect that Airbnb provides to restaurants would be independent of the demographic characteristics of the locality. However, this may not be the case. Evidence already indicates that Airbnb hosts and visitors may be incorporating race into their decision making process when finding a match on the platform (Edelman et al., 2017).

To examine the role of demographics on the relationship between Airbnb and restaurant employment, we partition the zipcodes in our data based on demographics. For each zipcode, we determine the proportion of residents that identify as White, Black, and/or Hispanic.[22] Table 3.14 summarizes the distribution of each demographic for the zipcodes in our sample. We create three subsamples of zipcodes, one for each demographic group. Each subsample contains only zipcodes where the proportion of residents that identify with the related demographic is greater than 50%. For example, in the White subsample, only zipcodes where 50% or more of the residents are White is included. The same is done for Black and Hispanic, respectively.

For each subsample, we estimate Equation 3.1 to determine the impact of Airbnb on restaurant employment in areas with a high presence of a certain demographic. To identify this impact, it is necessary that each subsample have areas with high and

[22] We use the 2011 American Community Survey from the U.S. Census Bureau to obtain zipcode level data on race/origin. White refers to White and not Hispanic and Black refers to Black and not Hispanic.

Table 3.14.
Demographic Statistics

Demographic	**Zipcode Average**	**Zipcode Median**	**Zipcode 75$^{\mathbf{th}}$ Percentile**	**Airbnb Intensity** $< 3\%$	**Airbnb Intensity** $\geq 3\%$ **Average**
White	31.2%	21.8%	51.8%	75.8%	24.2%
Black	22.9%	11.3%	33.6%	33.3%	66.7%
Hispanic	29.9%	24.0%	43.8%	50.0%	50.0%

low Airbnb intensity respectively. In other words, each subsample should have areas with and without Airbnb activity. Table 3.14 also shows the percentage of zipcodes within each subsample where the Airbnb intensity is high (Airbnb per household is greater than 3% in 2016). The percentages indicate that, while there is variation between the subsamples, they all independently have a distribution of Airbnb active and inactive zipcodes, which enables us to identify the impact of Airbnb activity in each subsample.

Table 3.15 shows the results of Equation 1 for all the subsamples. Columns 1, 2, and 3 show the results for the sample of zipcodes with a high proportion of White, Black, and Hispanic residents. Column 4 shows the result for the subsample that have a majority of either Black or Hispanic residents respectively. The results indicate that not all the subsamples are benefiting from the spillover effect of Airbnb. Specifically, the results indicate that, among the selected demographics, only areas with a high proportion of White residents benefit from the home-sharing platform facilitated visitors. Restaurant in areas with a high proportion of Black or Hispanic resident do not appear to benefit from the spillover effect.

One potential rationale for this result relates to the perception of crime and lack of safety often associated with areas with a higher proportion of minorities. To examine this potential among Airbnb visitors, we identify Airbnb reviews that contain phrases and words suggesting that the Airbnb visitors felt that the area was unsafe.[23].

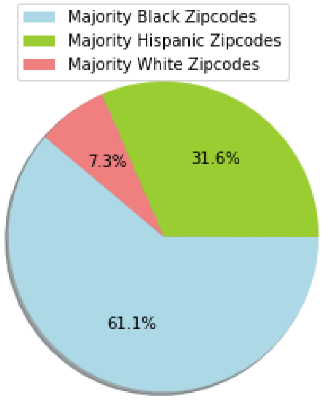
[23]Specifically, we identify Airbnb reviews that mention an area as being one of the following: unsafe, dangerous, shady, crime, risky, or seedy

Table 3.15.
Heterogeneity of Airbnb Impact (Subsample by Demographics)

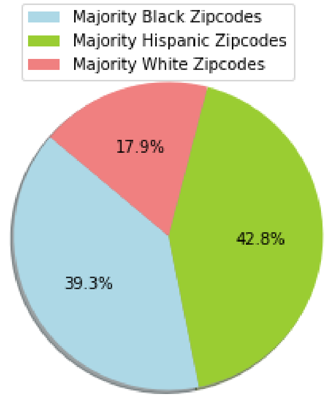
Dep. Variable: log(Rest. Empl.)	(1) High Ratio of White Residents	(2) High Ratio of Black Residents	(3) High Ratio of Hispanic Residents	(4) High Ratio of Hispanic/Black Residents
Airbnb Reviews per Household	**2.398***** (0.794)	**0.330** (0.418)	**0.239** (0.522)	**0.440** (0.516)
Local Rest. Popularity	1.261** (0.503)	10.910*** (2.408)	2.033 (1.958)	2.680* (1.399)
Adjusted Restaurants Count	0.646*** (0.170)	0.494*** (0.135)	0.837*** (0.206)	0.670*** (0.100)
log(Hotel Employees)	0.002 (0.021)	0.025 (0.023)	0.008 (0.026)	0.024 (0.019)
log(Retail Employees)	0.079 (0.134)	0.009 (0.097)	-0.314** (0.130)	-0.043 (0.089)
Zipcode Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Constant	5.725*** (0.930)	5.548*** (0.636)	8.019*** (0.897)	6.079*** (0.618)
Observations	280	200	220	566
R-squared	0.569	0.709	0.488	0.541
Number of zipcodes	28	20	22	59

The upper left chart in Figure 3.10 shows the distribution of these reviews across the NYC zipcodes whose residents are either majority Black, White, or Hispanic. There is a clear disparity in the proportion of these reviews that are associated with Airbnb listings in zipcodes where the majority of residents are Black. We also obtain data from NYC open data initiative on reported crimes and felonies in a zipcode. The upper right chart in Figure 3.10 shows the proportion of crimes reported in the zipcodes where the residents are predominately of one demographic and the bottom chart shows the felonies. The distribution of crimes indicates that the Airbnb guests' perceptions are significantly worse than actual reported crime statistics. Moreover, the discrepancy is particularly stark for the predominately Black areas.

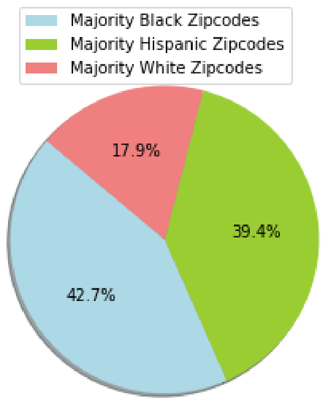
The review contents reveal that Airbnb visitors are more likely to discuss negative aspects of a local area that relate to safety if they are staying in a predominately Black zipcode. This may be partly attributed to the ignorance of the Airbnb visitors



(a) Airbnb Reviews Mentioning Safety Concerns



(b) NYC Crimes Reported



(c) NYC Felonies Reported

Fig. 3.10. Airbnb Reviews and Crime Distribution

to the demographic makeup of a specific location while making their reservation.[24] To corroborate this, we examine Airbnb visitors that stayed in NYC more than once. We identify those visitors that stayed in a predominately Black or White zipcode during their first stay. We then calculate the proportion of these visitors that stayed in the same zipcode the second time they visited NYC and used Airbnb. Among these visitors, we find that 13% of visitors that stayed in a predominately Black zipcode stayed in the same zipcode for their second trip. This compares to a 21% retention rate in the majority White zipcodes. This suggests that a larger portion of the visitors

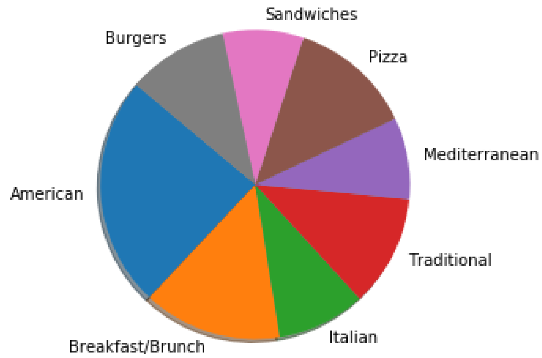
[24]Over 60% of Airbnb hosts in the predominately Black areas are not Black.

that stayed in the predominately Black zipcodes were dissatisfied with their location and chose to change their location for thier subsequent visit.

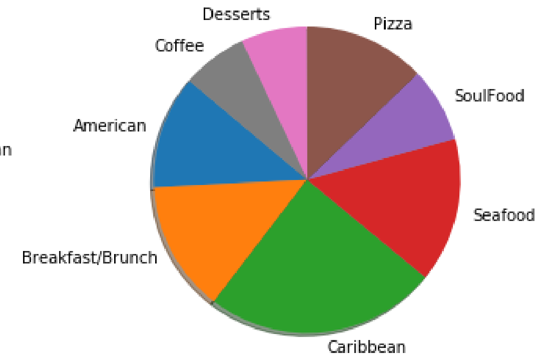
While these numbers are potentially troublesome, they should be taken with caution when seen as indicative of a racial bias in behavior by Airbnb visitors. The demographics of a location may be partially affecting the Airbnb visitor behavior due to less auspicious reasons that relate to food preferences. To ascertain the differences in restaurant offerings, we use the category label for each restaurant from its Yelp.com page. Figure 3.11 displays the 8 most common categories in the majority Black, majority White, and *Traditionally Tourist Areas*. A Comparison of the majority White and the *Traditionally Tourist Areas* shows that the categories have significant overlap. The categories are all the same except Mediterranean replaces bars in the majority White areas. However, the *Traditionally Tourist Areas* and the majority Black areas are significantly different in their offerings. Specifically, Carribean, Seafood, and Soulfood make up almost half of the proportions in the majority Black areas, but are not in the top 8 in the *Traditionally Tourist Areas*.

3.6.2 Heterogeneous Impact of Airbnb due to Restaurant Market Structure

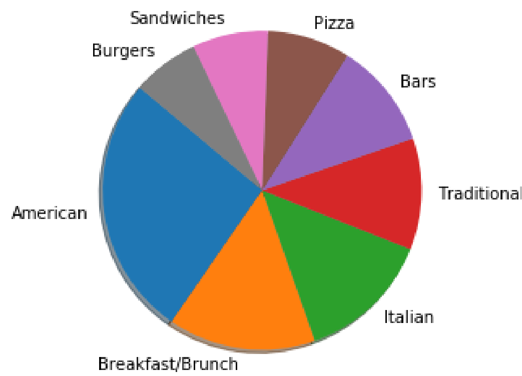
The competitive nature of the restaurants in a locality may also impact the spillover effect of Airbnb on local restaurant performance. On the one hand, more competition may improve the quality and variety of restaurant availability. This would entice visitors of an area to frequent the restaurants in the locality. On the other hand, areas with a more competitive restaurant dynamic have a multitude of restaurants that are popular. This implies that visitors may distribute to these restaurant and, as a result, all the restaurants obtain small increases in activity. Given that these restaurants were already popular and likely had significant local activity, it is unclear whether areas with more competition and, as a result, greater diversity in vis-



(a) Majority White Zipcodes



(b) Majority Black Zipcodes



(c) Traditionally Tourist Areas

Fig. 3.11. Yelp Restaurant Category Distribution for NYC Zipcodes

itor restaurant activity, would necessarily increase restaurant employment to appease this new demand.

To determine the competitive dynamics among the restaurants in a specific zipcode, we calculate the Herfindahl-Hirschman Index (HHI).[25] The market share of each restaurant is calculated using the share of the local Yelp reviews written in 2011 (Gutt et al., 2019). If the local reviews in 2011 are distributed across many restaurants—indicating that there is high competition among the restaurants—then the HHI will be low and indicates a more competitive local area. In contrast, if a few restaurants dominate the majority of the reviews than the competition among the restaurants would be low and the HHI would be large. We use the local Yelp restaurant reviews

[25] The HHI index ranges between 0-10,000. It is calculated as the sum of the square of the market share of each restaurant in the zipcode.

in 2011 as there is still relatively little Airbnb activity in 2011. Based on the distribution of HHI across the zipcodes we identify each zipcode as having low (below the 33^{rd} percentile), medium (between the 33^{rd} and 66^{th} percentile), or high (greater than the 66^{th} percentile) concentration.

Table 3.16 shows the results of Equation 1 on the three subsamples of varying local restaurant competition. The results indicate that Airbnb does not have an impact on restaurants in zipcodes that have a low level of restaurant competition (column 3). The impact is strongest in the areas with high competition (column 1). The implication is that in areas where a few restaurants dominate the local markets, the benefit from the spillover effects of Airbnb is diminished. Since the restaurants have finite physical capacity, the dominant restaurants in areas without significant competition do not benefit from the visitors as their capacity is perhaps already reached. The finite capacity issue is likely less problematic in areas with more competitive restaurants as the demand in those areas is distributed among the restaurants. In these areas, restaurants can hire more employees to service the greater demand without necessarily being constrained by physical capacity.

3.7 Generalizing the Findings to Other Cities

Thus far, we have identified the impact of Airbnb on restaurant employment in NYC. We have focused on NYC as it is the most active Airbnb city in the United States and is the most visited city overall. To evaluate the extent that our results are generalizable to other cities, we assess the impact of Airbnb intensity on restaurant employment in 5 other cities. Specifically, we obtain Airbnb, Yelp, and local employment data for Austin, TX; Chicago, IL; Los Angeles, CA; Portland, OR; and San Francisco, CA.[26] We aggregate the data at the zipcode level for each city. We replicate the pre-analysis that was done for NYC by removing the zipcodes where the number of Yelp visitor reviews is significantly higher than other zipcodes in the city.

[26]We select these cities based on the availability of Airbnb data from insideairbnb.com.

Table 3.16.
Heterogeneity of Airbnb Impact (Subsample by Market Structure)

Dep. Variable: log(Restaurant Employment)	(1) **High Competition**	(2) **Medium Competition**	(3) **Low Competition**
Airbnb Reviews per Household	**0.952*** **	**0.550**	**0.481**
	(0.322)	(0.333)	(0.298)
Local Rest. Popularity	1.883***	10.969***	1.248
	(0.507)	(1.529)	(2.438)
Adjusted Restaurants Count	0.396***	0.521***	0.871***
	(0.139)	(0.110)	(0.115)
log(Hotel Employees)	0.007	-0.019	0.042**
	(0.015)	(0.012)	(0.019)
log(Retail Employees)	0.163*	0.020	0.000
	(0.093)	(0.102)	(0.112)
Year Fixed Effects	Yes	Yes	Yes
Constant	5.379***	5.739***	5.415***
	(0.691)	(0.674)	(0.733)
Observations	400	400	410
R-squared	0.687	0.609	0.603
Number of zipcodes	40	40	41

This is to remove those zipcodes that are attractive to visitors regardless of Airbnb availability.

We combine the zipcodes from all the cities and run Equation 3.1 except that we replace the year fixed effect (δ_t) with a year/city fixed effect. A year/city fixed effect captures city specific events that are time variant such as seasonal festivals. Column 1 of Table 3.17 presents the results for this analysis and indicates that Airbnb has an impact on restaurant employment beyond NYC. We also conduct Equation 3.1 on each city individually to evaluate whether the effect holds for all cities. Columns 2-6 of Table 3.17 present the results. They show that the impact of home sharing and restaurant employment is consistent for all cities.

To further delineate the drivers of the heterogeneity between cities, we assess the role of demographic differences across cities. In section 3.6.1 we found that, in NYC,

Table 3.17.
Airbnb Impact on Restaurant Employment for Cities Beyond NYC

	(1) *All Cities*	(2) *Austin,TX*	(3) *Chicago,IL*	(4) *LA,CA*	(5) *Portland,OR*	(6) *San Fran.,CA*
Dep. Variable:	**log(Rest. Empl.**	**log(Rest. Empl.**	**log(Rest. Empl.**	**log(Rest. Empl.**	**log(Rest. Empl.**	**log(Rest. Empl.**
Airbnb Reviews per Household	**0.460***** (0.057)	**0.471***** (0.066)	**0.267***** (0.061)	**0.543***** (0.118)	**0.494***** (0.068)	**0.411*** (0.221)
Local Rest. Popularity	0.482*** (0.078)	-0.054 (0.231)	0.560*** (0.088)	0.727*** (0.148)	0.569** (0.210)	0.059 (0.099)
Adjusted Rest. Count	0.925*** (0.063)	1.072*** (0.197)	0.968*** (0.087)	0.802*** (0.116)	0.735*** (0.103)	1.251*** (0.271)
log(Hotel Employees)	-0.011* (0.006)	0.000 (0.013)	0.005 (0.011)	-0.036** (0.018)	-0.022** (0.010)	-0.002 (0.026)
log(Retail Employees)	0.145*** (0.026)	0.191*** (0.065)	0.184*** (0.041)	0.153*** (0.052)	0.089 (0.055)	-0.078 (0.083)
City-Year Fixed Effects	Yes					
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes
Constant	6.010*** (0.189)	5.593*** (0.473)	5.502*** (0.291)	5.651*** (0.393)	6.325*** (0.389)	7.864*** (0.664)
Observations	1,650	240	460	430	290	230
R-squared	0.664	0.719	0.593	0.677	0.710	0.704
Number of zipcodes	165	24	46	43	29	23

restaurants in areas with majority Black and/or Hispanic residents did not benefit from the spillover effect of home sharing. We replicate the analysis in Table 3.15 for the zipcodes in the 5 new cities we introduced in this section. Column 1 of Table 3.18 shows the results of Equation 3.1 on the full sample of zipcodes in the 5 additional cities. Since we are unable to use city-year fixed effects in the subsample analysis due to sample size limitations, we use year fixed effects and we include the effect on the full sample with year fixed effects for comparability. Columns 2,3, and 4 show the results for the subsamples with majority White, Black, and Hispanic residents. The results for the predominately White and Black zipcodes mimic the results in NYC with the effect only present in the predominately White areas. Interestingly, the results for the Hispanic areas indicate that there is a home-sharing spillover effect in this areas.

Table 3.18.
Heterogeneity of Airbnb Impact (Subsample by Demographics) for Additional Cities Beyond NYC

	(1)	(2)	(3)	(4)
Dep. Variable: log(Rest. Empl.)	**Full Sample**	**High Ratio of White Residents**	**High Ratio of Black Residents**	**High Ratio of Hispanic Residents**
Airbnb Reviews per Household	**0.465*****	**0.396*****	**0.368**	**0.471*****
	(0.057)	(0.062)	(0.839)	(0.065)
Local Rest. Popularity	0.486***	0.320***	-0.663	0.457*
	(0.106)	(0.104)	(1.158)	(0.250)
Adjusted Restaurants Count	-0.026	-0.024	-0.049	-0.053
	(0.024)	(0.030)	(0.067)	(0.045)
log(Hotel Employees)	-0.010	-0.004	-0.044	-0.013
	(0.008)	(0.008)	(0.061)	(0.012)
log(Retail Employees)	0.167***	0.182***	0.289**	0.115*
	(0.038)	(0.048)	(0.118)	(0.065)
Year Fixed Effects	Yes	Yes	Yes	Yes
Constant	5.740***	5.749***	4.486***	6.022***
	(0.272)	(0.328)	(0.878)	(0.479)
Observations	1,650	1,040	170	350
R-squared	0.475	0.503	0.180	0.587
Number of zipcodes	165	104	17	35

3.8 Conclusions and Discussion

As home sharing platforms have gained popularity, they have been met with resistance from local regulators and other stakeholders concerned about their negative impact on local communities. Researchers have studied the impact of home sharing platforms on the hotel industry (Zervas et al., 2017), rental prices (Barron et al., 2018), and even its potential for racial discrimination (Edelman et al., 2017). Home sharing platforms are unique in the context of the sharing economy because, on the surface, the negative local externalities (rental prices, housing prices, and negative impact on communities) are directed towards local residents while the positive local externalities are constrained to the Airbnb hosts themselves (Filippas and Horton, 2017). In essence, hosts are micro-entrepreneurs who are monetizing inventory that would otherwise have remained stagnant (rooms in their homes or whole homes when they are traveling) and visitors have a larger supply of potential short term rental accommodations to choose from. The advantage for the visitor may be realized through a lower fee, a more organic/localized experience, or potentially both. However, the negative economic impact is limited to the residents of the local area.

Regulators in many major cities have focused on these negative aspects to motivate regulatory frameworks designed to limit the impact of home sharing platforms. However, we find that Airbnb, the most prominent home sharing platform in the world, has a positive and salient economic spillover effect on local restaurants. The platforms capacity to attract visitors to areas that would otherwise not have had access to these visitor dollars can act as a local economic engine supporting these local restaurants. Our results indicate that if the Airbnb intensity in a zipcode increases by 1%, then restaurant employment would increase by approximately 1%.

Our employment growth estimates compare favorably with a back-of-the-envelope calculation. Specifically, there were 113,728 Airbnb reviews in *NYC Sample 2* in 2015. Considering that 75% of visitors wrote reviews (Zervas et al., 2017), this amounts to approximately 150,000 visits. If we assume that, on average, each visit consists of two individuals and the visitors stay for two weeks (the average stay indicated by Airbnb), then this results in approximately 300,000 Airbnb visitors and 8,400,000 meals per year.[27] A typical restaurant has 100 seats, requires 22 employees, and expects the patrons to spend considerable time in eating their meals (Batt et al., 2014). As such, we assume 5 meals per day per seat for each restaurant which translates to approximately 23 meals serviced per employee. Therefore, the number of additional full-time employees needed to serve the Airbnb visitors in 2015 is approximately 1,000 (8,400,000 meals per year / 365 days / 23 meals per emlpoyee). Given that there were 86,204 restaurant employees in 2014, this suggests a 1.16% increase in restaurant

[27] 300,000 visitors x 2 meals per day * 14 days = 8,400,000.

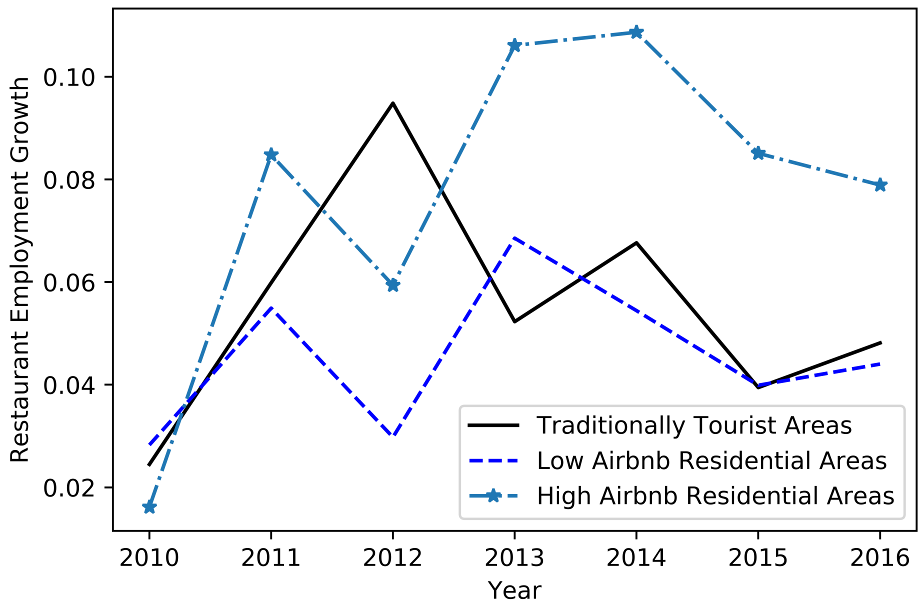
employment in 2015. Reassuringly, this is remarkably close to the estimates from our model.[28]

A question that is important in this context is whether this increase in restaurant employment in the *traditionally residential areas* is coming at the cost of employment in the *traditionally tourist areas*. There are two main possibilities: 1) The home-sharing visitors who are frequenting *traditionally residential areas* would have stayed in the *traditionally tourist areas* without the home-sharing platform or 2) The home-sharing visitors would not have come to NYC if not for the home-sharing platform. The most plausible reality is that home-sharing visitors are a combination of both possibilities. In 2016, *traditionally tourist areas* accounted for approximately 51% of the Airbnb reviews in our data. Figure 3.12 displays the restaurant employment growth in these areas during the period of our study and indicates positive restaurant employment growth throughout the years we study. Figure 3.12 also displays the growth rates in *traditionally residential areas* with and without high levels of Airbnb intensity (high Airbnb areas correspond to areas with higher than 3% Airbnb intensity in 2016). The *traditionally residential areas* with significant Airbnb presence have the highest growth rate (especially after 2012, which is when Airbnb activity significantly increases). However, the restaurant employment growth rate in the *traditionally tourist areas* remains positive.

We also find that the impact of Airbnb on restaurant employment is not homogeneously benefiting all areas. Specifically, demographics and market structure have an important role in determining the value extracted by local restaurants from Airbnb activity. Spillover effects of Airbnb on restaurants are diminished in areas with a relatively high number of residents who identify their race as Black. We find a similar result for areas with a relatively high proportion of residents that identify their origin as Hispanic. In contrast, restaurant in areas with a high proportion of White residents benefit from the economic spillover of Airbnb activity. Similar analysis was conducted for 5 additional cities. The results indicate that the lack of spillover for predominately

[28]The median Airbnb intensity was 1.1% in 2015. Considering the estimated elasticity of 1.03%, this results in a job growth of 1.13%.

Fig. 3.12. Restaurant Employment Growth



Black areas is consistent for the extended sample. However, interestingly, the result associated with predominately Hispanic areas does not hold. Most of the Hispanic zipcodes in the extended sample are located in Los Angeles. The fact that Los Angeles is approximately 48% Hispanic may change perceptions of potential interactions of people visiting the city. For the market competition heterogeneity analysis, we find that in areas where a few restaurants capture the majority of local Yelp reviews—high concentration areas—the impact of Airbnb on restaurant employment is diminished.

These findings contribute to the growing stream of literature on the direct and indirect impacts of Internet-enabled alternatives on traditional local establishments. The literature on the direct effect has covered the retail market (Brynjolfsson et al., 2009; Forman et al., 2009), local print market (Seamans and Zhu, 2013), taxi industry (Cramer and Krueger, 2016; Wallsten, 2015), and hotel industry (Zervas et al., 2017). We contribute to the growing stream of literature on the spillover effect of these Internet-enabled platforms, with a specific focus on sharing economy platforms (Burtch et al., 2018; Sheppard et al., 2016; Quattrone et al., 2016; Filippas and Horton, 2017; Gong et al., 2018; Barron et al., 2018). Our work is novel in that it focuses on *complimentary* spillover effects. Specifically, we are able to ascertain the affect of an Internet-enabled phenomena—sharing platform induced visitor redistribution—on the actualized economic impact of complimentary services—restaurants.

Our results are also useful for the discussion surrounding the purpose of home-sharing platforms. Airbnb and other home-sharing platforms argue that the majority of their users simply use the platform to augment their income and not as a means of creating investment properties. This would seem to indicate a preference for hosts that share their properties. Our NYC sample focused on areas that are not traditional tourist locations, and descriptive statistics point to the fact that the shared listings are over-represented in our sample. Specifically, in the tourist locations that were removed, 37% of the reviews were attributable to shared listings. This is compared to 60% in our sample of non-tourist locations. As such, our results indicate that areas with a relatively higher proportion of shared listings benefit from the spillover impact of home-sharing visitors. Further analysis regarding the differences in behavior of these two categories of users (private vs. shared) is required to further understand the necessary regulatory actions, if any, that are needed.

While these findings are important to the regulatory discussion around home sharing platforms, they also provide evidence of the potential for the sharing economy to impact the market structure of local restaurants. As more consumers regard home-sharing as a viable alternative, the presence of visitors in localities without a significant hotel presence will grow. This will impact restaurant demand and could prove crucial to local business owners. Importantly, visitors and locals will likely have different preferences and expectations. Since visitors/tourists are generally more willing to spend money at restaurants, their preferences might impact local restaurant outcomes. As restaurant owners react to these changing demand dynamics, the effect will naturally play a role in determining the type of restaurants that make up the local market structure..

Another important aspect of the changing dynamic of visitors is the source and dissemination of information. Airbnb is pushing its hosts to provide information about the local area through guidebooks and recommendations. While Airbnb's main function is to provide a visitor an accommodation, this results in a potential interaction between host and visitor, where the host can give the visitor information on her favorite local establishments. This means that local restaurants may be served by establishing relationships with popular hosts that can serve as a means of advertising. Even Airbnb has recognized the importance of its role as connector of visitors to restaurants and recently purchased the restaurant reservation platform resy.com.[29] Also, these changing dynamics will have an impact on the role of online reviews (Chevalier and Mayzlin, 2006). Visitors are likely more dependent on reviews when making restaurant selection. Therefore, the importance of reviews to restaurant performance may be magnified when considering that visitors are becoming more distributed across a city.

[29]https://www.theverge.com/2017/9/20/16338668/airbnb-restaurant-reservation-resy

4. CONCLUSION

In this dissertation, we have utilized data from a multitude of different data sources in empirical models to study geographic and demographic heterogeneity driven disparities in digitally transformed markets. We have explored two specific technology-driven disruptions. First, we study the interactions between online P2P lenders and local brick and mortar financial institutions. We explore the geographically dependent strategic reactions of traditional banks to the disruptions imposed by FinTech lenders. Second, we analyze the indirect economic spillovers of home-sharing platforms (specifically, Airbnb) on local establishments offering complimentary services. As the popularity of home-sharing grows, its role in the redistribution of travelers to non-traditional locations in cities will naturally have economic implications on these locations.

In Chapter 2, we find that local competition positively effects the likelihood of strategic reactions from local lending institutions. The impact has economically significant outcome on the debt management of P2P loans. Numerically, our estimates indicate that if an LC borrower moved from an area at the 25^{th} percentile of local competition to an area at the 75th percentile, the probability of prepayment would have increased by approximately 4.49%. To achieve the same increase in the likelihood of prepayment, a borrower would require an increase of at least $75,000 in annual income.

We contribute to the research examining the competition and substitution between online and offline channels (e.g., Forman et al. (2009), Brynjolfsson et al. (2009)). This research stream mainly focuses on retail products, however, our findings indicate that this competition persists when the product category is financial. Financial products, unlike retail products, are often designed to the requirements of individual consumers. Therefore, a local bank's intimate knowledge of the borrower heterogeneity in their area allows them to offer more personalized services. A deeper understanding of the relationship between local market structure and P2P borrowers' debt management may shed light on the factors that are relevant to online and offline channel substitution for financial products. Specifically, we find that local financial market structures with a relatively higher ratio of large banks have a higher prepayment propensity.

In Chapter 3, we study the heterogeneous economic spillover effects of a home sharing platform—Airbnb—on the growth of a complimentary local service—restaurants. We find that, for an average zipcode, if the intensity of Airbnb activity (Airbnb reviews per household) increases by 1%, then the restaurant employment in that zipcode grows by approximately 1.03%. This result is validated across multiple specifications and robustness checks. Our results also indicate that both demographics and market structure have an important role in determining the areas that benefit from the economic spillover of Airbnb. We validate the underlying mechanism behind the main result by evaluating the impact of Airbnb on Yelp visitor reviews – areas with increasing Airbnb activity experience a surge in their share of NYC visitor reviews. This result is further validated by evaluating the impact of a unique Airbnb neighborhood level policy recently implemented in New Orleans. Moreover, we find that the complimentary spillover effect of Airbnb activity on restaurant employment is more pronounced in areas with higher levels of market competition. We also find that restaurants in areas with a relatively high number of Black residents or a relatively high number of Hispanic residents do not benefit from the economic spillover of Airbnb activity.

REFERENCES

REFERENCES

Abhishek, V., Geng, D., Li, B., and Zhou, M. (2018). When the Bank Comes to You: Branch Network and Customer Omni-Channel Banking Behavior. *Information Systems Research*, Forthcoming.

Ambrose, B. W. and Capone, C. A. (1998). Modeling the Conditional Probability of Foreclosure in the Context of Single-Family Mortgage Default Resolutions. *Real Estate Economics*, 26(3):391–429.

Amel, D. F. and Starr-McCluer, M. (2002). Market Definition in Banking: Recent Evidence. *The Antitrust Bulletin*, 47(1):63–89.

Anderson, R. and VanderHoff, J. (1999). Mortgage Default Rates and Borrower Race. *Journal of Real Estate Research*, 18(2):279–289.

Autor, D. H. (2003). Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing. *Journal of Labor Economics*, 21(1):1–42.

Bakos, J. Y. (1997). Reducing buyer search costs: Implications for electronic marketplaces. *Management S cience*, 43(12):1676–1692.

Barron, K., Kung, E., and Proserpio, D. (2018). The Sharing Economy and Housing Affordability. *Working Paper*.

Batt, R., Lee, J. E., and Lakhani, T. (2014). Of human resource practices, turnover, and customer service in the restaurant industry. Technical report, School of International & Public Affairs, Columbia University New York City, NY.

Beck, T., Demirgüç-Kunt, A., and Maksimovic, V. (2004). Bank Competition and Access to Finance: International Evidence. *Journal of Money, Credit and Banking*, 36(3):627–648.

Becker, B. (2007). Geographical Segmentation of US Capital Markets. *Journal of Financial Economics*, 85(1):151–178.

Berger, A. N., Demirgüç-Kunt, A., Levine, R., and Haubrich, J. G. (2004). Bank concentration and competition: An evolution in the making. *Journal of Money, Credit and Banking*, 36(3):433–451.

Berger, A. N., Miller, N. H., Petersen, M. A., Rajan, R. G., and Stein, J. C. (2005). Does Function Follow Organizational Form? Evidence From the Lending Practices of Large and Small Banks. *Journal of Financial Economics*, 76(2):237–269.

Berkovich, E. (2011). Search and Herding Effects in Peer-to-Peer Lending: Evidence from Prosper.com. *Annals of Finance*, 7(3):389–405.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How Much Should We Trust Difference-in-Differences Estimates? *Quarterly Journal of Economics*, 119(1):249–275.

Brynjolfsson, E., Hu, Y. J., and Rahman, M. S. (2009). Battle of the Retail Channels: How Product Selection and Geography Drive Cross-Channel Competition. *Management Science*, 55(11):1755–1765.

Burtch, G., Carnahan, S., and Greenwood, B. N. (2018). Can You Gig It? An Empirical Examination of the Gig Economy and Entrepreneurial Activity. *Management Science*.

Butler, A. W. and Cornaggia, J. (2011). Does Access to External Finance Improve Productivity? Evidence from a Natural Experiment. *Journal of Financial Economics*, 99(1):184–203.

Butler, A. W., Cornaggia, J., and Gurun, U. G. (2016). Do local capital market conditions affect consumers borrowing decisions? *Management Science*, 63(12):4175–4187.

Caliendo, M. and Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1):31–72.

Cetorelli, N. and Strahan, P. E. (2006). Finance as a Barrier to Entry: Bank Competition and Industry Structure in Local U.S. Markets. *Journal of Finance*, 61(1):437–461.

Chevalier, J. A. (1995). Capital structure and product-market competition: Empirical evidence from the supermarket industry. *American Economic Review*, 85(3):415–435.

Chevalier, J. A. and Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*, 43(3):345–354.

Cole, R. A., Goldberg, L. G., and White, L. J. (2004). Cookie Cutter vs. Character: The Micro Structure of Small Business Lending by Large and Small Banks. *Journal of Finance and Quantitative Analysis*, 39(2):227–251.

Coles, P. A., Egesdal, M., Ellen, I. G., Li, X., and Sundararajan, A. (2018). Airbnb Usage Across New York City Neighborhoods: Geographic Patterns and Regulatory Implications. *Working Paper*.

Cramer, J. and Krueger, A. B. (2016). Disruptive Change in the Taxi Business: The Case of Uber. *American Economic Review*, 106(5).

Cyrnak, A. W. and Hannan, T. H. (1999). Is the Cluster Still Valid in Defining Banking Markets? Evidence from a New Data Source. *Antitrust Bulletin*.

Degryse, H. and Ongena, S. (2005). Distance, Lending Relationships, and Competition. *Journal of Finance*, 60(1):231–266.

Dehejia, R. H. and Wahba, S. (2002). Propensity Score-Matching Methods for Nonexperimental Causal Studies. *Review of Economics and Statistics*, 84(1):151–161.

DeYoung, R., Lang, W. W., and Nolle, D. L. (2007). How the Internet affects output and performance at community banks. *Journal of Banking and Finance*, 31(4):1033–1060.

Dick, A. A. (2006). Nationwide branching and its impact on market structure, quality, and bank performance. *Journal of Business*, 79(2):567–592.

Dick, A. A. (2007). Market Size, Service Quality, and Competition in Banking. *Journal of Money*, 39(1):49–81.

Edelman, B., Luca, M., and Svirsky, D. (2017). Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment. *American Economic Journal: Applied Economics*, 9(2).

Einav, L., Farronato, C., and Levin, J. (2016). Peer-to-Peer Markets. *Annual Review of Economics*, 8(1):615–635.

Elliehausen, G. E. and Wolken, J. D. (1992). Banking markets and the use of financial services by households. *Fed Res Bull*.

Elyasiani, E. and Goldberg, L. G. (2004). Relationship lending: a survey of the literature. *Journal of Economics and Business*, 56(4):315–330.

Filippas, A. and Horton, J. (2017). The Tragedy of your Upstairs Neighbors: When is the Home-Sharing Externality Internalized. *Working Paper*.

Forman, C., Ghose, A., and Goldfarb, A. (2009). Competition Between Local and Electronic Markets: How the Benefit of Buying Online Depends on Where You Live. *Management Science*, 55(1):47–57.

Fraisse, H., Hombert, J., and Lé, M. (2018). The competitive effect of a bank megamerger on credit supply. *Journal of Banking & Finance*, 93(2018):151–161.

Freeman, L. (2005). Displacement or Succession? Residential Mobility in Gentrifying Neighborhoods. *Urban Affairs Review*, 40(4):463–491.

Glaeser, E. L., Kim, H., and Luca, M. (2017). Nowcasting the Local Economy: Using Yelp Data to Measure Economic Activity. *NBER Working Paper No. 24010*.

Glaeser, E. L., Kim, H., and Luca, M. (2018). Nowcasting Gentrification: Using Yelp Data to Quantify Neighborhood Change. *Working Paper*.

Gong, J., Greenwood, B. N., and Song, Y. (2018). Uber Might Buy Me a Mercedes Benz: An Empirical Investigation of the Sharing Economy and Durable Goods Purchase. *Working Paper*.

Gutt, D., Herrmann, P., and Rahman, M. S. (2019). Crowd-driven competitive intelligence: Understanding the relationship between local market competition and online rating distributions. *Information Systems Research*.

Hannan, T. H. and Prager, R. A. (2004). The competitive implications of multimarket bank branching. *Journal of Banking and Finance*, 28(8):1889–1914.

Heckman, J. J., Ichimura, H., and Todd, P. E. (1998). Matching As An Econometric Estimator Evaluation. *Th Review of Economic Studies*, 65(2):261–294.

Ho, D. E., Imai, K., King, G., and Stuart, E. A. (2007). Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis*, 15.

Horn, K. and Merante, M. (2017). Is home sharing driving up rents? Evidence from Airbnb in Boston. *Journal of Housing Economics*, 38.

Iacus, S. M., King, G., and Porro, G. (2012). Causal Inference without Balance Checking: Coarsened Exact Matching. *Political Analysis*, 20:1–24.

Imai, K., King, G., and Stuart, E. A. (2008). Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the Royal Statistical Society*, 171(2):481–502.

Iyer, R., Khwaja, A. I., Luttmer, E. F., and Shue, K. (2015). Screening Peers Softly: Inferring the Quality of Small Borrowers. *Management Science*, 62(6):1554–1577.

Jiménez, G., Lopez, J. A., and Saurina, J. (2013). How does competition affect bank risk-taking? *Journal of Financial Stability*, 9(2):185–195.

Keeley, M. C. (1990). Deposit insurance, risk, and market power in banking. *American Economic Review*, 80(5):1183–1200.

Kwast, M. L., Starr-McCluer, M., and Wolken, J. D. (1997). Market Definition and the Analysis of Antitrust in Banking. *The Antitrust Bulletin*, 42(4):973–995.

Li, Z., Yao, X., Wen, Q., and Yang, W. (2016). Prepayment and Default of Consumer Loans in Online Lending. *SSRN Electronic Journal*.

Liangliang Jiang, Ross Levine, C. L. (2017). Does competition affect bank risk? *NBER Working Paper No. 23080*.

Liebersohn, J. (2017). How does competition affect bank lending? quasi-experimental evidence from bank mergers. *Working Paper*.

Lin, M., Prabhala, N. R., and Viswanathan, S. (2013). Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending. *Management Science*, 59(1):17–35.

Lin, M. and Viswanathan, S. (2016). Home Bias in Online Investments: An Empirical Study of an Online Crowdfunding Market. *Management Science*, 62(5):1393–1414.

Marquez, R. (2002). Competition, adverse selection, and information dispersion in the banking industry. *Review of Financial Studies*, 15(3):901–926.

Mazzeo, M., Seim, K., and Varela, M. (2014). The Welfare Consequences of Mergers with Endogenous Product Choice. *Working paper, Northwestern University*.

Morse, A. (2015). Peer-to-Peer Crowdfunding: Information and the Potential for Disruption in Consumer Lending. *Annual Review of Financial Economics*, 7(1):463–482.

Moulton, B. R. (1990). An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. *The Review of Economics and Statistics*, 72(2):334.

Nguyen, H.-L. Q. (2019). Are credit markets still local? evidence from bank branch closings. *American Economic Journal: Applied Economics*, 11(1):1–32.

Pedersen, C. and Delgadillo, L. (2007). Residential Mortgage Default in Low—and High—Minority Census Tracts. *Family and Consumer Sciences Research Journal*, 35(4):374–391.

Petersen, M. A. and Rajan, R. G. (2002). Does Distance Still Matter? The Information Revolution in Small Business Lending. *Journal of Finance*, 57(6):2533–2570.

Pope, D. G. and Sydnor, J. R. (2011). What's in a Picture?: Evidence of Discrimination from Prosper.com. *Journal of Human Resources*, 46(1):53–92.

Quattrone, G., Proserpio, D., Quercia, D., Capra, L., and Musolesi, M. (2016). Who Benefits from the "Sharing Economy" of Airbnb? *International World Wide Web Conference. WWW*.

Rosenbaum, P. R. and Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1):41–55.

Seamans, R. and Zhu, F. (2013). Responses to Entry in Multi-Sided Markets: The Impact of Craigslist on Local Newspapers. *Management Science*, 60(2):476–493.

Sheppard, S., Udell, A., et al. (2016). Do airbnb properties affect house prices. *Williams College Department of Economics Working Papers*, 3.

Tang, H. (2019). Peer-to-peer lenders versus banks: Substitutes or complements? *Review of Financial Studies*, 32(5):1900–1938.

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242.

Wallsten, S. (2015). The Competitive Effects of the Sharing Economy: How is Uber Changing Taxis? *Technology Policy Institute*.

Zervas, G., Proserpio, D., and Byers, J. W. (2017). The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry. *Journal of Marketing Research*, 54(5).

Zhang, J. and Liu, P. (2012). Rational herding in microloan markets. *Management science*, 58(5):892–912.

Zhang, S., Lee, D., Singh, P. V., and Mukhopadhyay, T. (2018). Demand Interactions in Sharing Economies: Evidence from a Natural Experiment Involving Airbnb and Uber/Lyft. *Working Paper*.

VITA

Mohammed Alyakoob

Purdue University
Krannert School of Management
403 W. State Street
West Lafayette, IN, 47906
Phone: + 1 (765) 418-8525
Email: alyakoob@purdue.edu

Education

Ph.D., Purdue University **2019**
Krannert School of Management

M.S., The George Washington University **2012**
School of Business

B.S., Indiana University **2008**
Kelley School of Business

Working Papers

- "Where You Live Matters: The Impact of Local Financial Market Competition in Managing Online Peer-To-Peer Loans", with Mohammad Rahman and Zaiyan Wei.

  Available on SSRN: https://papers.ssrn.com/abstract_id=2985099

- "Shared Prosperity (or Lack Thereof) in the Sharing Economy", with Mohammad Rahman.

  Available on SSRN: https://papers.ssrn.com/abstract_id=3180278

  Media Coverage:

  – **Washington Post**, https://www.washingtonpost.com/business/2018/07/11/airbnb-benefits-local-economies-only-white-neighborhoods-study-finds
  – **Chicago Tribune**, http://www.chicagotribune.com/business/ct-biz-airbnb-economic-impact-neighborhoods-20180711-story.html
  – **Market Watch**, https://www.marketwatch.com/story/airbnb-guests-spend-less-money-in-black-and-hispanic-neighborhoods-2018-07-13
  – Los Angeles Times, The Verge, The Oregonian, and others.

Conference/Workshop Proceedings & Presentations

- Shared Prosperity (or Lack Thereof) in the Sharing Economy, with Mohammad Rahman, *Workshop on Information Systems Economics (WISE)*, San Francisco, CA, 2018.
- Airbnb: Is it a Curse or a Blessing for Restaurant Employment, with Mohammad Rahman, *International Conference on Information Systems (ICIS)*, San Francisco, 2018.
- Shared Prosperity (or Lack Thereof) in the Sharing Economy, with Mohammad Rahman, *Conference on Information Systems and Technology (CIST)*, Phoenix, AZ, 2018.
- Shared Prosperity (or Lack Thereof) in the Sharing Economy, with Mohammad Rahman, *Conference on Digital Experimentation (CODE)*, Boston, 2018.
- Shared Prosperity (or Lack Thereof) in the Sharing Economy, with Mohammad Rahman, *Sixth Annual Platform Research Symposium*, Boston, 2018.
- Shared Prosperity (or Lack Thereof) in the Sharing Economy, with Mohammad Rahman, *Fourteenth Symposium on Statistical Challenges in Electronic Commerce Research (SCECR)*, Rotterdam, Netherlands, 2018.
- Where You Live Matters: The Impact of Local Financial Market Competition in Managing Online Peer-to-Peer Loans, with Mohammad Rahman and Zaiyan Wei, *Conference on Information Systems and Technology (CIST)*, Houston, TX, 2017.
- Where You Live Matters: The Impact of Local Financial Market Competition on a Borrower's Debt Management Strategies, with Mohammad Rahman and Zaiyan Wei, *Third Workshop on Marketplace Innovation*, Stanford University, Stanford, CA, 2017.
- Where You Live Matters: The Impact of Local Financial Market Competition on a Borrower's Debt Management Strategies, with Mohammad Rahman and Zaiyan Wei, *Boulder Summer Conference on Consumer Financial Decision Making*, Boulder, CO, 2017.
- To Prepay or Default? The Impact of Local Financial Market Competition in Online Debt Crowdfunding, with Mohammad Rahman and Zaiyan Wei, *POMS 28th Annual Conference (POMS)*, Seattle, WA, 2017.
- To Prepay or Default? The Role of Local Financial Market Competition in Online Peer-o-Peer Lending, with Mohammad Rahman and Zaiyan Wei, *Workshop on Information Systems Economics (WISE)*, Dublin, Ireland, 2016.
- To Prepay or Default? The Role of Local Financial Market Competition in Online Peer-to-Peer Lending, with Mohammad Rahman and Zaiyan Wei, *INFORMS Annual Meeting*, Nashville, TN, 2016.

Ph.D. Tutorials

- NBER Economics of Digitization Tutorial, Stanford University, March 2018
  - Mentors: Susan Athey (Stanford), Erik Brynjolfsson (MIT), Judith Chevalier (Yale), and Michael Smith (Carnegie Mellon)
- The Structural Modeling and Machine Learning Applications for Research on Technology Workshop (SMART), University of Washington, August 2017
  - Instructors: Stephanie Lee, Yingfei Wang, Hema Yoganarasimhan (University of Washington), Panagiotis Adamopoulos, Vilma Todri (Emory University), Julian Guo (Michigan State University), Yan Huang (University of Michigan), and Junming Yin (University of Arizona)
- Quantitative Marketing and Structural Econometrics Workshop, Washington University in St. Louis, July 2017
  - Instructors: Tat Chan, Steve Ryan, Seethu Seetharaman, Raphael Thomadsen (Washington University in St. Louis), Michaela Draganska (Drexel), Paul Ellickson (University of Rochester), Avi Goldfarb (University of Toronto), Brett Gordon (Northwestern), Carl Mela (Duke), Peter Rossi (UCLA), and Ting Zhu (Purdue)

Teaching Experience

- **Instructor**, Krannert School of Management, Purdue University
  - MGMT 382, Management Information Systems (undergraduate core course), Fall 2016.
    - Teaching Evaluation: 5.0/5.0, Enrollment: 48
    - Awarded *Krannert Certificate for Distinguished Teaching*, the highest recognition at Krannert for excellence in teaching by a doctoral student.
  - MGMT 382, Management Information Systems (undergraduate core course), Summer 2016.
    - Teaching Evaluation: 4.9/5.0, Enrollment: 49
    - Awarded *Krannert Certificate for Distinguished Teaching*
- **Teaching Assistant**, Krannert School of Management, Purdue University
  - MGMT 590, Web Data Analytics (graduate elective course).
    - Fall 2015, Fall 2017, Fall 2018
  - MGMT 682, Management Information Systems (graduate, MBA core course).

    * Fall 2014, Spring 2015, Fall 2015, Spring 2016
  - MGMT 590, Python Programming (graduate elective course).
    * Spring 2018
  - MGMT 590, Design for Instincts (graduate elective course).
    * Fall 2014, Fall 2015
  - MGMT 382, Management Information Systems (undergraduate core course).
    * Summer 2015

Awards & Honors

- Best Presentation Award, Krannert Doctoral Research Symposium (Fall 2018)
- Best Presentation Award, Krannert Doctoral Research Symposium (Fall 2016)
- Financial Award, School of Business, The George Washington University, Finance Department (2010 - 2012).
- Merit Scholarship, Kuwait Ministry for Higher Education Merit Scholarship Program (2004 - 2008)
- Faculty Scholarship, Indiana University (2004 - 2008).

Selected Graduate Coursework

- Economics
  - Econometrics 1 & 2, Microeconometrics, Bayesian Econometrics, Micro-Economic Theory 1 & 2, Advance Game Theory, Labor Economics, Experimental Economics
- Machine Learning & Optimization
  - Data Mining, Stochastic Models in Operations Research, and Discrete Optimization Models and Algorithms.
- Information Systems
  - Doctoral Seminar Series (7 courses)

Industry Experience

- Kuwait Capital Markets Authority, Kuwait City, Kuwait 2012-2014
  - Senior Research Assistant, Market and Economic Research Department.
- Jiblah Holding Company 2008-2010
  - Investment Analyst, Market Research Department.

Professional Services and Affiliations

- Reviewer: Conference on Information Systems and Technology (CIST) (2017, 2018)
- Reviewer: International Conference on Information Systems (ICIS) (2017, 2018)

- Krannert Doctoral Student Association (KDSA) Vice President of Research (2017)
- Member: Association for Information Systems (AIS)
- Member: Institute for Operations Research and the Management Sciences (INFORMS)
- Member: Production and Operations Management Society (POMS)

References

- **Prabuddha De**
  Accenture Professor of Information Technology and Professor of Management
  Krannert School of Management, Purdue University
  403 W. State Street, West Lafayette, IN 47907
  Phone: (765) 494-0699
  pde@purdue.edu
- **Mohammad Rahman**
  Associate Professor of Management
  Krannert School of Management, Purdue University
  403 W. State Street, West Lafayette, IN 47907
  Phone: (765) 494-4464
  mrahman@purdue.edu
- **Justin Tobias**
  Professor and Head, Department of Economics, and Senior Associate Dean
  Krannert School of Management, Purdue University
  403 W. State Street, West Lafayette, IN 47907
  Phone: (765) 494-8570
  jltobias@purdue.edu
