Essays on Entrepreneurship and Local Labor Markets

by

Rahul R. Gupta

Business Administration
Duke University

Date: ______________________

Approved:

____
Aaron K. Chatterji, Advisor

____
Maria Ines de Bivar Black Henriques

____
Alon Brav

____
Wesley Cohen

____
William R. Kerr

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Business Administration
in the Graduate School of Duke University
2020

Abstract

Essays on Entrepreneurship and Local Labor Markets

by

Rahul R. Gupta

Business Administration
Duke University

Date: ________________________
Approved:

Aaron K. Chatterji, Advisor

Maria Ines de Bivar Black Henriques

Alon Brav

Wesley Cohen

William R. Kerr

An abstract of a dissertation submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Business Administration
in the Graduate School of Duke University
2020

Abstract

This dissertation explores the relationship between external shocks local labor markets and entrepreneurship. The first and main essay investigates the effects of a large firm's geographical expansion (anchor firm) on local worker transitions into startup employment through wage effects in industries economically proximate to the anchor firm. Using hand collected data on large firms' site searches matched to administrative Census microdata, I exploit lists of anchor firms' site selection process to employ a difference-in-differences approach to compare workers and employers in winning counties to those in counterfactual counties. Counties are balanced along a number of socio-economic characteristics as well as ex ante industry distribution, firm size distribution, and firm age distribution. The arrival of an anchor firm induces entrepreneurship in industries linked through input-output channels by a magnitude of 120 new establishments that account for over 2,300 jobs. Relative to young firms in counterfactual counties, these new firms grow 12% faster in five-year employment growth and have a 7% lower failure probability. These effects are strongest in the most specialized and knowledge-intensive industries. Attracting an anchor firm to account appears to have limited spillover effects in employment that are mainly driven by reorganization of incumbent firms in input-output industries with occupational similarity of the anchor firm that face rising labor costs.

The second essay provides a blueprint for understanding the dynamics surrounding mass layoffs and business closures. This essay creates a novel data set linking

geocoded Business Registration data to public layoff notifications data. This data can be used to understand how local entrepreneurship can reduce unemployment spells and earnings penalties for low wage displaced workers. Workers eventually employed by startups experience faster post-displacement wage growth than those eventually employed by mature firms. In final essay, I provide motivation for research investigating the spatially heterogeneous effects the advancement of certain industries inhibit entrepreneurship in others. I decompose a Bartik employment measure of demand for a region's labor. The decomposition shows that the recovery from the Great Recession was led by capital-intensive industries (e.g., transportation manufacturing and machinery manufacturing) that are typically inversely associated with local entrepreneurship. Interestingly, the inverse association of these industries and entrepreneurship appears to spillover into other industries. These industries include transportation equipment manufacturing and machinery manufacturing. This set of observations motivates this dissertation's research agenda to understand the cross-industry relationships that drive an area's level of entrepreneurship and labor market dynamism.

Contents

List of Tables

List of Figures

Acknowledgements

Research consistently shows that role models are vital in a person's development. I am lucky to have at least five of them on my committee.

When I first met Ronnie in 2014, I was entering my second year as a Research Associate at Harvard Business School. I had not the slightest sense that I was meeting someone who would be my greatest advocate, mentor, and motivator. Even as early as March 2015, Ronnie's influence on me was profound: He almost literally put me in a suitcase and carried me from Boston to Durham. Now, in more ways than one I feel he has carried me with him throughout these past five years. It has been incredible to engage in so many discussions where we seem to connect on a new level whether personal or in our interests. Wherever our careers take us, I am sure to always feel your presence and influence in it.

Wes has consistently been an invaluable source of knowledge academically and comfort personally. I have learned much more from Wes than cost spreading and industry heterogeneity, though those two ideas certainly rank high. Although my working relationship with Ines has been the shortest in duration, it has been no less influential. Over the past couple years, we talked about my job market paper more than anyone. I have also appreciated the openness with which you have shared your job market and early career experiences. I took Alon's Finance II seminar, and somehow ended up with an advisor I talk mostly about displaced workers. I still think about those paper mill workers who sparked our initial conversations together.

Unbeknownst to me at the time, I truly started my career in June 2013 when I first started working with Bill. The first time I stepped onto Harvard Business School's campus, all I took away was that it would be so cool to have an office steps away from beautiful tennis courts. I did not think I would actually get to spend two years there. There were a number of times that I would look around and wonder if he had made a mistake. I am forever grateful for the opportunity he gave me and his continued service on my committee. My research experiences with Bill equipped me with empirical tools and research insights that have continued to have a profound influence on me. It has been incredible to realize that even in the absence of daily or weekly meetings, my research trajectory is still so deeply influenced by interests formed in those two years prior to entering a Ph.D. program.

There are also two integral members of the Fuqua community to whom I owe a special debt of gratitude. Sonya Wilhelmson and Sharon Bass tackled both the AOM and ASSA/AEA job markets with me. They never missed a deadline or request for a letter of recommendation. They even somehow navigated one of the most complicated and confusing application submission portals imaginable. I hope at some point in my life I develop those intense organizational skills, and patience, to keep a chaotic period so orderly.

Personal relationships both nearby from afar continue to shape me personally and professionally. My officemate, cohortmate, and friend Tatiana Lluent has been been an invaluable confidante who has never fails to make the time to talk through everything from ideas to challenges that arise during a Ph.D.. My roommates, Jonathan Moreno and Viviana Rodriguez (and their dogs Filo and Ofelia) and Alessandro Villa, have been incredible friends throughout these years in Durham. When we all moved in together a couple years ago, we were already friends, but I did not think we would come to be like family as we have. Chelsea Biggs has been the most supportive friend I could ask for. We moved to Boston from Indiana at about the same time and she was a source of inspiration during that first journey. Though in Chicago during my Ph.D. years, I feel we never left. The hours she sank in comforting me when challenges arose to the purest joy I felt from her when things went well are made all the more incredible as she has never stopped being there even after she and her husband had a child and balance ambitious careers.

My parents Ajay and Vibha Gupta and sister Shelley have been an unwavering source of support. Whenever the ups-and-downs of Ph.D. life had me down, they never lost faith in me. I wish I could find the words to express the gratitude I have for having such unconditional love in my life.

With sincerest apologies to those I may have missed and warmest thanks to all, I hope we forge onwards together as I close the first chapter of my career and write the next.[1]

[1]Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 1841. All results have been reviewed to ensure that no confidential information is disclosed. This research uses data from the Census Bureau's Longitudinal Employer Household Dynamics Program, which was partially supported by the following National Science Foundation Grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation. I want to give a special thank you to Bert Grider at the Triangle RDC who has provided invaluable assistance in navigating the Census procedures.

Chapter 1

Introduction

Understanding how agglomerations form, who populates them, and who stands to gain the most are the key questions that motivate my dissertation research. The differential recovery rates across regions in the aftermath of the Great Recession of 2007-2009 sparked a renewed interest in understanding why some labor markets are more robust than others and why firms of similar quality appear to coalesce around each other. These agglomerations of firms and economic activity have long been thought to facilitate innovation, firm productivity, optimal allocation of resources and labor to firms, wage growth, and local amenities (e.g., Jacobs 1961; Glaeser 2011; Moretti 2012). To internalize the benefits of agglomeration economies, governments often seek to institute ambitious policy agendas to create clusters (Chatterji, Glaeser, and Kerr 2014). One such policy includes incentivizing large firms to establish a new large facility in a specific location in the hopes the firm serves as an anchor for follow-on private investment. The magnitudes of such programs are large even at the local level. In 2015, local governments spent approximately \$50 billion in tax incentives and subsidies to attract new facilities of large firms.[1]

A few recent studies have questioned the overall welfare impacts of such targeted

[1] Bartik, Timothy. (2017). "A New Panel Database on Business Incentives for Economic Development Offered by State and Local Governments in the United States." *W.E. Upjohn Institute for Employment Research.*

subsidy programs. For example, Glaeser and Gottlieb (2008), Moretti (2010), and Busso, Gregory, and Kline (2013) show that policies aimed at attracting firms to a location are inefficient and seldom generate net welfare increases through agglomeration spillovers. Economic activity merely shifts from one sector to another and the programs lead to the sub-optimal matching of firms to locations and workers. While Greenstone, Hornbeck, and Moretti (2010) find that for existing manufacturing plants in a region, the arrival of an additional large manufacturing plant increases manufacturing productivity conditional on plant survival, these effects may last only as long as the new manufacturing facility remains operational. Moreover, the subsidies move economic activity away from locations with ideal firm-to-location matches and where agglomeration spillovers would be most widespread.[2] The failure of policy experiments geared towards inducing agglomerations to arise in specific locations is troublesome given that labor market density and clustering of economic activity appears to be closely related to economic mobility, firm productivity, and regional resilience. The characteristics of workers and firms who populate these agglomerations remain less understood and examined.

All three essays in this dissertation seek to understand agglomeration forces driving the location of economic activity and in particular for young firms who are thought to be drivers of economic dynamism and innovation (Decker, Haltiwanger, Jarmin, and Miranda 2014). I address questions of how agglomerations form and who populates and benefits from them by using tools and frameworks from labor and personnel economics and by drawing upon a vast literature in entrepreneurship. In the first essay (Chapter 2), I investigate the effects of a large firm's geographical expansion (anchor firm) on startup formation and employment. Chapter 3 investigates the im-

[2] Kline and Moretti (2012) examine the Tennessee Valley Authority and find the large scale investment in the Tennessee Valley during the Great Depression and the years afterwards increased wages and productivity. However, these effects ended as soon as federal investment ended. Moreover, the welfare gains in the region were offset by welfare losses in regions that did not win federal support.

pact of mass layoffs on entrepreneurship as well as the role of young firms in providing employment opportunities for displaced workers. The final essay in Chapter 4 disentangles an often used tool in regional and labor economics to show the importance in understanding between industry dynamics in entrepreneurship research.

Chapter 2 leverages a vast network of administrative and hand-collected data. Chapter 2 (as well as Chapter 3) uses the universe of 200,000,000 U.S. business establishments since 1990 across all 50 states and D.C. matched to worker-job-level data on workers from 25 states that amount to approximately 700,000,000 worker-job-year observations. I match this microdata to hand collected data on large firms' site searches. I employ a difference-in-differences approach to estimate changes in new venture formation and outcomes by comparing startups in winning counties to those in alternative sites the anchor firm considered in the final stages of its site search. The treated and control groups are similar conditional on a number of socio-economic characteristics. The locations are also similar in industry, firm size, and firm age distributions. The arrival of an anchor firm induces entrepreneurship local supply chain industries by a magnitude of 120 new establishments from startups that account for over 2,300 jobs. Relative to economically proximate young firms in counterfactual counties, these new ventures grow 12% faster in five-year employment growth and have a 7% lower failure probability.

On the surface, these facts appear promising for the case of attracting large anchor firms to a county. However, upon closer inspection, I show that the gains reflect a reshuffling of employment along the firm age distribution and not overall wage or employment growth. Moreover, the effects are plausibly driven by increases in wages and labor costs that force county-industry incumbent firms to shed managerial layers. Workers from layers eliminated by firms found new ventures in industries that are economically proximate to their employer and in which they are most likely to have the most relevant experience and skills.[3]

While the national share of startups that are spinoffs in the anchor and supply chain industries is 15%, over 75% of the new ventures induced by the anchor firm's arrival in winning counties are spinoffs.[4] As expected, the firms that yield the most spinoffs and lose the most upper mid-level wage earners are the ones that experience the fastest wage growth after the arrival of the anchor firm. This process spinoffs induced by incumbent firms streamlining their operations is consistent with prior studies that observe increases in incumbent plant labor productivity (Greenstone et al. 2010) and structured management practices (Bloom, Brynjolfsson, Foster, Jarmin, Patnaik, Saporta-Eksten, and Van Reenen 2019).

Chapter 2 further suggests that policy programs aimed at inducing a large firm to open a large facility in a county may not generate the expected agglomeration gains even if the sheer number of firms in the county (or county-industry) rises. This conclusion is aligned with studies estimating general equilibrium and welfare gains due to targeted subsidies. In terms of employment effects, my result that 2,300 jobs are created specifically in young firms implies a multiplier of 1.3 with respect to the average reported size of the anchor firm's facility. This multiplier is in the realm of estimates with prior literature that examine industry-specific place based policies such as Glaeser and Gottlieb (2008) and Moretti (2010). I also document industry heterogeneity. The effects on wage-induced spinoffs are strongest in high-technology sectors.[5] Finally, I conclude that while both the number of startups formed and

[3] I define economic proximity through the supply chain as tabulated by from the Bureau of Economic Analysis' Use of Commodities matrix. For the main analysis, the supply chain consists of the top 5 industries that supply the anchor firm's industry and the top 5 industries that purchase from the anchor industry. All results are robust to using continuous measures of input and output shares.

[4] Spinoffs are defined as startups whose top earners were employed by an incumbent firm in the same industry-county.

[5] Moretti (2010) similarly suggests that the expected effects of place-based policies will be strongest in high technology tradeable sectors.

the share created by workers with industry and location relevant experience is disproportionately high, large-scale subsidy programs aimed at appropriating agglomeration gains by attracting large firms are likely small relative to county and subsidy size. The channel by which startups form does not appear to be due to conventional agglomeration economies or an increase in market opportunities.[6] Instead, local industry wage effects push workers out of firms facing large increases in labor costs. These workers then transition into entrepreneurship in industries economically proximate to their prior employment.

Instead of examining cases where a firm opens a facility in a county as in Chapter 2, Chapter 3 provides a blueprint for understanding the dynamics surrounding mass layoffs and business closures. In particular, I describe the supply of young firms in an area as key determinants of whether displaced workers reduce the detrimental impacts of layoffs of their lifetime earnings. Further, I narrow the analysis to within cities. Displaced low wage workers are less mobile than their higher wage colleagues (Notowidigdo 2019). One often cited contribution of young firms to the economy is that they facilitate job hopping and accumulation of worker skills. These benefits may be particularly critical for the success of low wage workers who may have fewer opportunities in labor markets post-displacement. It is important to recognize entrepreneurship as either transformational in the Silicon Valley sense or as "main street" small businesses (Schoar 2010). Both classes of young or small firms are important to the economy, but have remarkably different founder and worker characteristics (Hurst and Pugsley 2011).

I continue to leverage Census microdata in Chapter 3. I identify instances where large multiunit firms shutdown one establishment, but retain other operations. Conditional on observable characteristics and wage category, workers in the shutdown establishment are not displaced due to differences in individual productivity (Gib-

[6] Overall supply chain employment remains relatively stable.

bons and Katz 1991). To understand disparate impacts of workers with immediate wage-earning opportunities from those less likely to do so, I focus on a set of multiunit firms who consolidate operations by shutting down one facility, but not all. Workers eventually employed by startups experience 67% faster post-displacement wage growth than those eventually employed by mature firms. Displaced workers who find employment in mature firms exhibit no difference in wage growth from workers of similar relative wages in the firm who exit voluntarily. The effects are strongest for workers displaced in thicker labor markets with smaller than average firm sizes.

An additional noteworthy contribution to the research stream emanating from Chapter 3 is the compilation of public data sources to capture the microgeography of dynamism in local labor markets through entrepreneurship. I scrape information on mass layoff firms and locations through Worker Adjustment and Retraining Notification (WARN) Act postings on state government websites. I geocode these addresses and map the geocoded address list to bulk metadata on business registration filings.[7] Due to current data availability, I focus on Florida whose sunshine laws provide ample information on business starts as well as mass layoffs, though subsequent work will continue to develop a large scale public database. The purpose of this exercise is to show how entrepreneurial formation occurs proximate to sites of mass layoffs, facilitating job hopping for displaced workers. This research stream will also provide a within city perspective on entrepreneurship and labor market dynamism missing from Chapter 2's county-level and Chapter 4's Commuting Zone level analyses. Understanding local details is important for policymakers undertaking the task of formulating policies that assist displaced workers and facilitate business formation all the while considering impacts such policies may have on neighborhood characteristics within a city.

[7] The Business Registration data is compiled similar to Guzman and Stern (2015) and Guzman (2019).

The third and final essay in Chapter 4 investigates how employment levels across geographic regions in the U.S. recovered after the Great Recession. I first document the persistent and disparate impacts of the recession across Commuting Zones in terms of job creation and firm entry. Then using a Bartik (1991) framework to leverage variation across regions according to their industry composition, I uncover several new facts about the response of young firms to local income shocks. I show a new pattern during the recovery from the Great Recession. Regions that experienced the most dramatic downturns do not recover as quickly as expected, and this deficit is largely explained by a diminished role for startups in job creation. I find that common perceptions about local labor market quality and credit availability to startups do not explain the startup deficit post-crisis. The most plausible explanation is that specific industries such as transportation equipment manufacturing and machinery manufacturing drove the recovery and are inversely related to startup formation.

Chapter 4's main contribution is in unpacking the oft-used Bartik (1991) variable. Bartik shift-share designs are increasingly used in entrepreneurship and regional economics literature, but my final dissertation essay points to a number of considerations when employing the empirical method. I use a familiar setup of job creation in startups regressed on a Bartik variable (Adelino, Ma, and Robinson 2017). Bartik variables (and other shift-share designs) are constructed by interacting local industry employment shares with national growth rates (leaving out the focal location) and summing over all industries. Higher values on the Bartik variable theoretically indicate increased demand for local labor inputs and thereby increase wages in the short-run. Higher wages should be associated with more investment in young firms. This relationship holds before the crisis. Interestingly, I find that in the post-crisis period[8] Commuting Zones with favorable Bartik shocks have less entrepreneurial job creation than those with worse Bartik shocks. Because Bartik variables are an aggre-

[8]I refer to the years 2009-2013 as the recovery period.

gation of individual industry elasticities, I decompose the Bartik coefficient into its industry components to calculate Rotemberg (1983) weights.[9] I find that the puzzling change in direction in the variable is driven mainly by shifts in the spatial dispersion of capital-intensive industries after the crisis. In transitioning from pre-crisis to post-crisis, the movement of capital intensive industries into fewer locations led to these industries accounting for a disproportionate share of loading in the Bartik variable in the regressions. These industries are commonly thought to be inversely related to overall entrepreneurship in a county (Chinitz 1961; Glaeser, Kerr, and Kerr 2015). This chapter concludes that identification assumptions must be tested and understood through an examination of industries receiving disproportionate loading.

My findings in Chapter 4 show that unpacking shift-share variables into their components is essential to validate identification as well as understanding micro-level dynamics that are masked by the variable. Chapter 4 coincided with other studies critiquing shift-share variables from an econometric theory standpoint (Borusyak, Hull, and Jaravel 2018; Goldsmith-Pinkham, Sorkin, and Swift 2018). While it is unclear whether parsing the Bartik variable is mechanical or economically meaningful (Goldsmith-Pinkham et al.), the decomposition serves as the basis from which my dissertation and future research stream are constructed. The key insight from the paper is that the advancement of some industries in an area can spell the decline of another. Understanding why these cross-industry relationships exist is essential for modeling and estimating how innovations or disruptions in one sector of the economy have heterogeneous effects across firms in other sectors and drive the location of economic activity.

[9] Goldsmith-Pinkham, Sorkin, and Swift (2018) use a GMM framework to show the industry composition shares serve as the instruments in a Bartik regression. They then provide a method to compute the relative magnitude that each industry share contributes to the identified variation in outcomes.

Chapter 2

Does Goliath Help David? Anchor Firms and Startup Clusters

2.1 Introduction

Given the importance of agglomerations on firm productivity and labor market outcomes, one important question policymakers confront is whether attracting a large firm to a region induces an industrial cluster to form through the reallocation of workers towards these locations and industries related to the large firm (Chatterji, Glaeser, and Kerr, 2014). In the hopes of becoming the next Silicon Valley, local policymakers often view attracting large firms as a means to jump start local economic activity and induce follow-on business formation. If this is the case, then geographic expansions of large firms can serve as "anchors" for new businesses in the county.[1] The arrival of an anchor firm may also induce increase in the cost of local factors, particularly in the short-run. These factor price changes may force some extant firms in the county-industry to reorganize and shed layers which may have implications far different from a story of entrepreneurs capturing entry opportunities unattainable by incumbent firms.

[1] I refer to the geographic expansions to a county by a large firm as an anchor firm. These geographic expansions are represented by the opening of a new facility by the anchor firm. An example of an anchor firm to a county is BMW's first opening of a large facility in Spartanburg, SC in 1992.

Workers may depart incumbent firms to start a new venture for a variety of reasons. One view is that large firms incubate employees who eventually spin-out to form new ventures by equipping them with technical expertise, nontechnical knowledge on regulatory and marketing strategies, and access to financiers (Gompers, Lerner, and Scharfstein, 2005; Chatterji, 2009). Another strand of literature uses the lack of entrepreneurship in historically manufacturing-dependent cities as evidence that large capital-intensive firms deplete resources required by entrepreneurs (Chinitz, 1961; Jacobs, 1970; Glaeser, Kerr, and Kerr, 2015).[2] Reconciling the two views is empirically difficult for two reasons. First, data limitations have restricted analyses to a narrow set of industries. Second, studies discussing entrepreneurship and agglomeration face the key identification challenge posed by the endogeneity between firms' location choice and the location's entrepreneurial climate.[3]

I demonstrate that the arrival of a large anchor firm induces startup formation specifically in its own supply chain. New ventures are more likely to be spinoffs from high quality incumbent supply chain firms.[4] In line with the role of a founder's industry knowledge in startup performance, these new startups grow faster and fail less than startups prior to the anchor firm's arrival. Comprehensive employee-employer matched Census microdata allows me to overcome typical data limitations and describe both the spinoff process and the shock that a large anchor firm's arrival provides to a local supply chain. This restricted-use data covers all businesses and workers in the United States regardless of industry since 1990.[5] Firm and establishment data

[2] Agrawal, Cockburn, Galasso, and Oettl (2014) find that cities with a mix of small firms and large incumbent labs are most innovative by facilitating spinoffs.

[3] The identification challenge also inhibits isolating precise mechanisms describing how collocation with large firms and industry agglomeration affect new venture formation and performance.

[4] I refer to "incumbent" firms as firms that already exist in the county prior to the anchor firm's arrival.

[5] Census business data cover all 50 states and the District of Columbia. Employee-employer data include all workers in 25 states.

come from the Longitudinal Business Database (LBD) which represents the universe of over 200,000,000 establishment-year observations. The LBD allows me to observe outcomes for anchor firms, incumbent firms, and startups. To construct lifetime work histories on startup founders, I match the LBD to the Longitudinal Employer Household Dynamics (LEHD) data that includes 500,000,000 worker-job-year observations.[6]

To address endogeneity, I leverage instances where a large firm conducted a site selection process and revealed its leading alternative sites for a new facility. *Site Selection* magazine is corporate real estate publication that often publishes articles describing site searches of large, generally Fortune 500 firms. [7]The 240 site searches I document describe facilities usually employing over 500 employees and requiring an investment of a quarter billion dollars.[8] Table 1 displays the industry breakdown of the new facilities in the magazine and Table 1b shows basic characteristics of parent firms in the *Site Selection* sample. Observing that winning and runners-up counties are similar on dimensions related to entrepreneurial formation and industrial composition (Table 2.2), I obtain plausibly exogenous variation in startup formation and employ a difference-in-differences design to compare startups and their founders in winning with those in runners-up counties.[9]

This study focuses on two channels through which agglomeration affects startup collocation with large firms. These two mechanisms are proximity to buyers and suppliers and local industry knowledge, and I test their importance using a difference-in-differences identification strategy. The arrival of an anchor firm serves as an unex-

[6]Section 3 describes identification of founders in the LEHD using firm age and wage rank within a firm.

[7]These new openings by large firms in the data are "anchor firms".

[8]Greenstone, Hornbeck, and Moretti (2010) provide an overview of the magazine.

[9]For ease of exposition, I refer to these large firms conducting the site search as "anchor firms". "Winning counties" are the counties that the anchor firms ultimately decide to locate. Sites considered at the end of the site search process but not selected are "finalist" or "runners-up" counties. Greenstone et al. provide additional statistics on county similarity.

pected shock to the demand for labor in firms geographical proximate to a potential customer or supplier of intermediate parts. Prior literature such as Ellison, Glaeser, and Kerr (2010) establish the coagglomeration patterns of input-output industries. Physical proximity can enable firms to better monitor their suppliers or tailor products to their customers. Highlighting the importance of proximity to the supply chain for startups, I find that the effects of the anchor firm's arrival on startup formation are localized to industries upstream and downstream to the anchor firm.[10] The average county-industry pair (defined at the NAICS-4 level) in the anchor firm's supply chain generates 22 more new startup business establishments representing 430 jobs.[11] For the average winning county, this amounts to over 120 new establishments from startup firms and 2,300 jobs.[12] The most pronounced effects on supply chain entry and startup job creation are in knowledge intensive supply chains such as those relating to R&D Services, Semiconductors, and Aircraft Parts Manufacturing, and not in less knowledge intensive industries such as textile manufacturing. Overall, supply chain startups in winning counties grow 12% faster in terms of employment and exit 7% less relative to those in runners-up counties after the anchor firm's arrival.[13]

This paper's most novel contribution is showing that the rise in supply chain entrepreneurship is driven by spinoffs from incumbent supply chain firms.[14] After the anchor firm arrives, the number of employees departing incumbent upstream firms to start new upstream firms increases by almost 50% and for incumbent downstream

[10]The supply chain is defined for each of the 240 site searches and consists of upstream supplier industries, the focal industry, and downstream buyer industries. These designations are defined using BEA Use of Commodities Tables.

[11]For each county, there can be up to 5 upstream and downstream industries.

[12]These totals amount to approximately 20-23% of the total number of supply chain establishments and employees across all age groups.

[13]Growth and failure are tabulated over the first five years of a startup's life. Startups in winning counties perform significantly better post-anchor firm than prior to the anchor firm's arrival as well.

[14]In this paper, a spinoff is defined as employees who leave an incumbent firm and start a new company in the same industry category as their employer. Incumbent firms that generate spinoffs are referred to as "parent" firms.

firms the number of spinning employees doubles. As a share of all new supply chain startups, 25% of upstream startups are spinoffs as are 91% of anchor industry startups and 71% of downstream startups. These magnitudes are particularly striking given that the baseline spinoff share of startups in these industries is 15%.

The reason spinoffs form may not be that some employees perceive new viable entry opportunities. Instead, rising labor costs force firms to reorganize (similar to literature in organizational economics such as Caliendo, Monte, and Rossi-Hansberg, 2015).[15] Firms that reorganize shed layers and these newly pushed out workers form businesses in industries that most suit their skills. In line with this prediction, I find that firms with the fastest wage growth and highest average wage also generate the most spinoffs. A one standard deviation increase in an upstream firm's average pay relative to its industry increases the number of spinoffs by 10%. For anchor industry firms, a one standard deviation increase in average firm pay increases spinoffs by 21%. The effect on downstream firms is 17%.

2.1.1 Related Literature

I contribute to literature that describes agglomeration spillover effects and the geography of supply chains. Describing coagglomeration patterns, Ellison, Glaeser, and Kerr (2010) show that industries linked through intensity of shipments to each other as inputs and outputs tend to collocate. However, less clear is why this trend continues particularly as transportation costs continue to fall. One explanation could be that geographic proximity to buyers and suppliers enables better monitoring of transactions across segments of the supply chain and allows for a greater ability to tailor specialized services and products to downstream firms. These forces are greatest in industries related to R&D (Alcacer, 2006; Alcacer and Delgado, 2016) and I show

[15] Alternatively, prior works in strategy the spinoffs literature suggest that the average wage a firm pays relative to its industry is a measure of parent firm quality (Campbell, Ganco, Franco, and Agarwal, 2011; Carnahan, Agarwal, and Campbell, 2012).

that the mechanism is also relevant for knowledge intensive supply chains. Density of local labor markets is an additional channel discussed as a source of agglomeration benefits in clusters. The anchor firm's arrival reflects an increase in labor market density (Greenstone, Hornbeck, and Moretti, 2010), and in particular labor markets most related to the anchor firm's supply chain. Indeed, wages and within firm pay variance increase substantially in both the anchor industry as well as the most economically connected industries to the anchor industry. The literature in personnel economics (e.g., Lazear and Oyer 2004; Lazear and Shaw 2007) and organizational economics (e.g., Caliendo, Monte, and Rossi-Hansberg 2015) discuss the balance between firm compensation policies and the number of occupations or layers within the firm. However, this literature is not typically connected to the literature in agglomeration. I connect to this latter literature by discussing that the gains to firm growth through collocation in related industries can be identified through the transition of workers who might be pushed out of firms rather than individuals who are necessarily high productivity types. This means that the fast growth rate and increased probability of survival is driven by agglomeration spillovers and not selection of high type individuals into entrepreneurship in dense labor markets (Combes, Duranton, Gobillon, Puga, and Roux, 2012).

This paper also contributes to literature on spinoffs that have shown incumbent firms to be incubators of future entrepreneurs (Gompers, Lerner, and Scharfstein, 2005; Chatterji, 2009). The extant literature points to two primary explanations for spinoffs. One reason employees spin-out is that they recognize market opportunities arise that their employer is unable to capture or the employer is unable to appropriate returns to employees who discovered them (Klepper 2007/10; Klepper and Sleeper, 2005; Franco and Filson, 2006). Another explanation is that some incumbent firms are provide employees with networks and nontechnical expertise about the market and equip them success in knowledge-intensive product markets (Chatterji 2009). However, unlike prior literature I demonstrate that shocks geographically proximate and in the employer's supply chain induce employee spinoffs. I further show that spinoffs may not necessarily be driven by star employees capturing returns bureaucratic incumbent firms cannot. Instead, spinoffs may also arise in agglomerations where workers are pushed out firms that must reorganize to adapt to externally driven shocks to their wage structure. These employees are likely to select into firms or start firms in industries that they are most familiar which is plausibly that of their recent employer.

2.1.2 Organization

The remainder of this paper is as follows: Section 2 discusses possible mechanisms describing the process of spinoffs and agglomeration with a discussion of related literature. The extensive data used in this paper is detailed in Section 3. Section 4 explains the empirical methodology and Section 5 discusses results. I describe robustness of measurements as well as present evidence of subsidy similarity in 6. Finally, I conclude this chapter in Section 7 summarizing my main findings and describing avenues for future research.

2.2 Conceptual Framework on Spinoffs and Agglomeration

2.2.1 Spillovers versus Reorganization

This paper relates closely to two studies examining the spillover effects onto incumbent firms when an anchor firm arrives to a county. Greenstone, Hornbeck, and Moretti's (2010) are the first to use new establishment openings of large firms from Site Selection magazine and provide the basis for this paper's identification strategy. Focusing on manufacturing plants in the 1980's and early 1990's identified in the Census of

Manufactures, they show that productivity of a county's incumbent plants in the same industry as the opening establishment increases. Their model assumes that the arrival of the anchor firm leads to an increase in the number of employers in an industry. They attribute productivity gains to enhanced labor productivity through knowledge spillovers arising from increased thickness in local markets.

Contemporaneous to this paper, Bloom, Brynjolfsson, Foster, Jarmin, Patnaik, Saporta-Eksten, and Van Reenen (2019) focus on manufacturing firms in the Annual Survey of Manufactures matched to their own *Site Selection* magazine sample spanning 2010-2014. Bloom et al. suggest that one component of Greenstone et al.'s knowledge spillovers mechanism is that the anchor firm induces better management practices in the county's incumbent manufacturing firms. These improvements may be reflected in productivity gains in incumbent firms.

While this paper uses its own *Site Selection* sample, it covers a different time horizon (1990-2015) and matches to employee-employer data. In addition, this paper expands beyond manufacturing plants and uses the movement of workers in the LEHD to discuss knowledge and worker ability. Prior works focus on outcomes at the intensive margin to investigate the effects of anchor firms on incumbent firm productivity and management practices. This paper presents a framework for understanding the impact of anchor firms on entrepreneurship while also capturing flows of labor out of incumbent firms. Finally, I discuss a setting where spillover effects are not driven by shocks directly to an incumbent firm's local industry, but instead to the supply chain.

The determinants of wages in the short-run are determined mostly by industry and location factors external to the firm (Lazear and Oyer 2004). However, firms are tasked with responding to their external environment through reorganization. While the data does not allow me to determine whether firms are explicitly reorganizing by adding and dropping layers (Caliendo, Monte, and Rossi-Hansberg 2015), firms exhibiting fastest wage growth after the anchor firm's arrival also shed the most workers who form spinoffs. Though these workers earn more relative to their industry, county, and industry-county pair, they are not necessarily the top wage earners in the firm.[16] This behavior of employee transitions in firms undergoing the most rapid post-anchor arrival wage growth is consistent with literature on firm reorganization (Caliendo et al. 2015) as well as compensation policies for and mobility of mid-level managers (Lazear and Shaw 2007). Incumbent firms' efforts to reorganize in response to rising labor costs arising from the anchor firm's arrival may reflect improvement in management practices. Bloom et al. (2019) find evidence that increased density of economic activity in a county leads to improvements in structured management, perhaps reflecting reorganization and personnel management in surviving incumbent firms after the anchor firm arrives. Reorganization and structured management practices also would be reflected in measured labor productivity improvement in surviving incumbent firms found by Greenstone, Hornbeck, and Moretti (2010).

2.2.2 Agglomeration

Agglomeration economies suggest spinoffs will collocate near their parent firms after a shock to the overall density of firms and labor markets. Dense labor markets facilitate better matching of workers to their employers particularly in industries relying on specialized labor (Duranton and Puga, 2004; Moretti, 2011). These thick labor markets improve sorting of workers into firms and occupations (Helsley and Strange, 1990) and allow firms and workers to appropriate returns to skill (Rotemberg and Saloner, 2000). If the introduction of an anchor firm increases density in local supply chain industries, then the composition of workers in supply chain startups will change. This selection effect of who enters entrepreneurship will right shift the

[16] I find evidence they may be above the firm average, but not in the top 3 or 5 wage earners.

distribution of startup quality and performance.[17]

Observational evidence shows that industries in the same supply chain tend to collocate (Ellison, Glaeser, and Kerr, 2010). One explanation as to why supply chains are spatially concentrated is that proximity to buyers and suppliers improves a firm's ability to monitor suppliers and tailor products to customers. This mechanism is most important in industries that are more knowledge intensive. For example, Alcacer (2006) and Alcacer and Delgado (2016) find spatial clustering tends to be most pronounced for R&D facilities versus sales and production units. Even large and vertically integrated firms rely on external suppliers and contract with other firms to transport and ship goods (Atalay, Hortacsu, and Syverson, 2014). Some large firms strategically open establishments in clusters as a substitute for vertical integration (Helsley and Strange, 2006). For entrepreneurs, proximity to buyers and suppliers may be of particular importance. Glaeser and Kerr (2009) show the importance of thick local supply chain markets in supporting entrepreneurship.

2.2.3 Spinoffs

Another perspective on worker reallocation into young firms departs from traditional agglomeration mechanisms but instead relates to the literature on spinoffs. Spinoffs are empirically defined as instances where employees depart their employer to found a new firm that operates in the same or similar industry. Often, spinoffs are thought to be disruptive to their industry and potentially displace the incumbent employer. Why employees spinoff and which firms facility such transitions from wage-earning employment to firm founder has two main explanations. Some incumbent firms serve as incubators for workers with entrepreneurial ability and other incumbent firms suffer from organizational challenges that do not allow workers to appropriate returns to

[17] Combes, Duranton, Gobillon, Puga, and Roux (2012) attribute half the spatial variation in firm productivity differences to selection of workers to firms.

their inventiveness. I refer to this mechanism as the "bureaucratic channel."

Gompers, Lerner, and Scharfstein (2005) and Chatterji (2009) show that larger firms, often previously venture-backed firms, provide employees with knowledge essential to forming a business. Relevant knowledge for an entrepreneur may be non-technical in nature and include regulatory strategy, marketing, and connections to financiers and investors. In other specialized industries, technical expertise provided to a potential entrepreneur by a high quality employer is also particularly important. Franco and Filson (2006) show that high performing incumbents firms produce the most spin-outs. Interestingly, prior literature has not discussed whether incumbent firms face employee departures due to spinoffs because of shocks outside their industry.

The bureaucratic channel focuses on departures of top performing employees. Contracting frictions between these employees and upper management induce disagreements over which high-value projects to pursue and how to appropriate their returns. Spinoffs may arise because older firms are less able to add new product varieties in response to industry shocks (Franco and Filson, 2006; Klepper and Sleeper, 2005). Spinoffs through disagreements explain a number of industry clusters including the Akron tire inudstry (Buenstorf and Klepper 2009) and Detroit automobiles (Klepper 2007/10). A similar argument suggests that employee departures from incumbent firms into entrepreneurship arise because some employers are unable to properly appropriate returns to their employees' innovations (e.g., Anton and Yao, 1995; Chatterjee and Rossi-Hansberg, 2012).

Closely related to the personnel economics literature in Section 2.2.1 (Lazear and Oyer 2004; Lazear and Shaw 2007) is the literature describing compensation characteristics of firms that lose the most employees to spinoffs. Carnahan, Agarwal, and Campbell (2012) describe a framework where employees are less likely to depart high paying firms with greater wage variance than the competition. They present a theoretical model with empirical evidence that these firms are better at compensating and retaining high performing employees than competitors and are therefore of higher quality. While they may lose fewer employees overall, conditional on departure, high ability employees are more likely to form a new venture in their former employer's industry than join a competitor.

2.3 Data

2.3.1 Anchor Firm Data

Site Selection is a corporate real estate publication with issues dating back to the late 1960s. The magazine variably publishes either quarterly or bi-monthly depending on year. Traditionally, the magazine included regular features titled "Location Reports", "Million Dollar Plants", "Blockbuster Deals" and "Top Deals" in which companies, generally Fortune 500 firms, disclosed information on the selection of sites for new facilities. The articles often, but not always, indicate the projected size of the opening facility, number of employees, and incentives/subsides offered by local economic development bodies. The announcements in the articles are sizeable economic events. They tend to involve openings of plants that are stated to eventually at least five hundred employees with an initial investment that often totals over a quarter of a billion dollars. Because of the size of the events, I refer to these new facilities as "large establishments" and their associated firms as "large firms".

I hand collected information on each announcement reported between 1990 and 2015 using historical microfiche archives, online editions, and searches on the magazine's website. I confirmed events and their locations through LexisNexis and Google searches. I assign company industry (NAICS) designations to the announcements using descriptions of the opening facility in the articles which often describe the product being manufactured or characteristics of the facility and/or company. I refer to these companies as "anchor firms".

Following other studies that have used these articles (e.g., Greenstone, Hornbeck, and Moretti, 2010, and a contemporaneous work by Bloom et al., 2019), I refer to these announcements as “cases" and the openings of these facilities as “plant openings”. For each case, I document the year of the event, name of the firm, the county selected (referred to as *winners*), and other counties the firm considered at the end of its location selection process (which refer to as *finalists* and *runners-up* interchangeably). Most cases have one winner and one runner-up, but in select instances the article will mention multiple alternate sites the firm considered before its final decision. Most of the events involve manufacturing facilities, though R&D labs, telecommunications services, financial institutions, and corporate offices occasionally appear as well. Tablen1 displays the distribution of industries by sector as well as the ten most frequently appearing NAICS-4 industries associated with the facilities openings in the sample.

Table 2.1: Distribution of Announcements by Industry

A.		B.	
Industry Sectors	Cases	10 Most Common Announcement Industries	Cases
11. Agriculture	0	3361. Motor Vehicle Manufacturing	19
21. Mining	1	3364. Aerospace Product and Parts Manufacturing	14
22. Utilities	4	5417. Scientific Research and Development Services	12
23. Construction	0	3254. Pharmaceutical and Medicine Manufacturing	9
31-33. Manufacturing	134	3363. Motor Vehicle Parts Manufacturing	9
42. Wholesale Trade	11	3344. Semiconductor and Other Electronic Component Manufacturing	8
44-45. Retail Trade	15	5171. Wired Telecommunications Carriers	6
48-49. Transportation & Warehousing	12	5222. Nondepository Credit Intermediation	6
51. Information	12	3341. Computer and Peripheral Equipment Manufacturing	5
52. Finance & Insurance	22	5221. Depository Credit Intermediation	5
53. Real Estate	0		
54. Professional, Scientific, and Technical Services	12		
55. Management of Companies	4		
56. Administrative and Support, Waste Management	4		
61. Educational Services	0		
62. Health Care	2		
71. Arts, Entertainment, Recreation	1		
72. Accomodation & Food Services	5		
81. Other Services	1		

Notes: The anchor firm sample includes observations from a broad set of industries. However, the magazine over-represents the manufacturing sector. Panel A displays all private sector NAICS-2 industry sector designations and the number of Site Selection magazine large facilities openings announcement cases associated with each. Panel B shows the 10 most common industries appearing in the sample at the 4-digit NAICS definitions. This table is not produced using restricted Census microdata.

Table 1b shows the large size of firms represented in the magazine.

Figure **??** shows total number of cases each state has had either a winning county

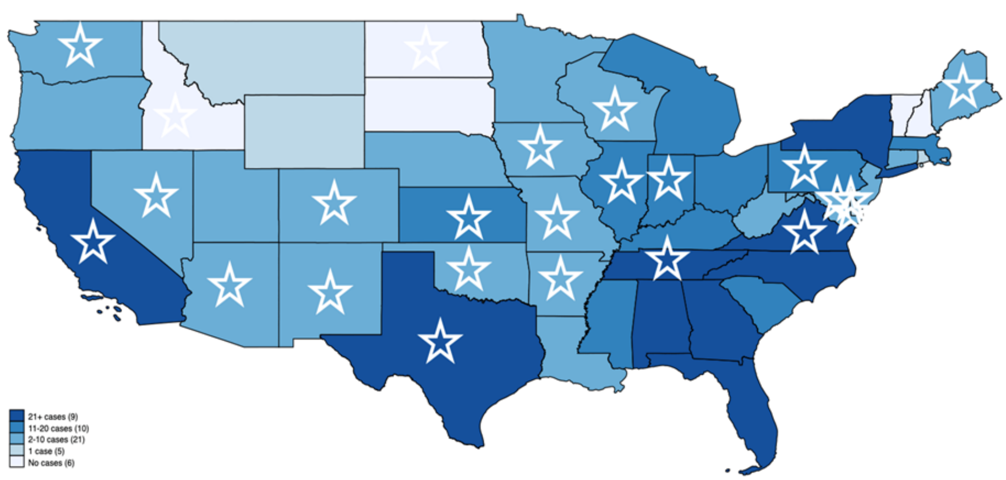
Table 1b: Firm Characteristics

	5th Percentile	Median	95th Percentile
Total Employees	899	37,300	362,063
Total Assets ($ Billions)	0.2898	15.4434	1097.19
Total Revenues ($ Billions)	0.2894	12.9983	1,128.915
Property, Plant, and Equipment ($ Billions)	0.0815	6.0944	828.997
R&D Expense ($ Billions)	0	0.7167	9.571
Share Foreign	0.3316		

Notes: At the overall firm-level, the firms appearing in *Site Selection* magazine typically employ over 37,000 full-time employees and have $15 billion in assets. The median firm conducts nearly a billion dollars of R&D. Most firms represented are heaquartered in the United States, however one-third are foreign firms. Most foreign firms are Japanese. The tabulations are based off the year in which the firm made the announcement for each case. The data are obtained from Compustat and S&P Capital IQ. In cases where Compustat does not provide financial information, the tabulations use the year nearest to the announcement. This table is not produced using restricted Census microdata.

or runner-up county. All but five states are in the magazine sample. The map also shows that the sample over-represents states in the southeast, though states with the largest numbers of cases are distributed across the west coast, midwest, and mid-Atlantic.[18]

FIGURE 2.1: Summary of *Site Selection* locations



Notes: The map shows the number of times each state has a county considered as either a winner or runner-up in a *Site Selection* magazine faciities opening announcement. Data on firm in all 50 states and DC are present in the Longitudinal Business Database. The stars overlayed denoted that the state is providing information on employees and their employes in the Longitudinal Employer Household Dynamics sample: AZ, AR, CA, CO, DC, DE, ID, IL, IN, IA, KS, ME, MD, MO, MT, NV, NM, ND, OK, PA, TN, TX, VA, WA, and WI.

Table 2.2 displays results on comparing similarity between winners and runners-

[18]The stars show states available in the LEHD. However, per disclosure rules the map does not suggest which cases or how many cases are represented for each state in subsequent analyses.

up counties on dimensions commonly thought to be correlated to entrepreneurship. On all dimensions, the set of contest counties (those that are either a winner or runners-up) differ substantially from the average county in the U.S. In my subsequent analysis, I will use Table 2.2 and similar measures pertaining to human capital and employment in the supply chain to support the claim that existing local human capital is an important driver in new venture formation and success in the supply chain.[19]

2.3.2 Longitudinal Business Database

I obtain employment, age, and location information U.S. business establishments from the restricted-use administrative Longitudinal Business Database (LBD) covering the period 1990-2015. The Census Bureau compiles a register of businesses using annual information on U.S. employer establishments from the Social Security Administration's Form SS-4, IRS Master Business File, IRS Form 941 and IRS Form 944. The LBD files are at the annual level, but include time-invariant establishment identifiers allowing researchers to construct a panel of all U.S. business establishments. The LBD covers the universe of all non-farm, non-education/religious employer establishments. Establishments are defined as physical locations of business activity. The industry designations are provided at the 6-digit NAICS level. One particularly attractive feature of the LBD is that industries are assigned at the establishment rather than firm level. Industry assignments reflect the primary economic activity occurring at the physical location.

I use the LBD to measure startup activity. Though the LBD includes firm identifiers, researchers cannot define a startup as the first occurrence of the firm identifier. Transitions from being a single-establishment firm to a multi-unit firm often results in a reassignment of the Census firm identifier as well as the federal Employer Identi-

[19] Namely, I identify that anchor firms target counties that disproportionately have thicker input-output markets. This could indicate higher levels of human capital or resources in these counties that support firms in these industries.

Table 2.2: Similarity of Winning and Runners-up Counties

	Winners	Runners-up	Difference	All counties
Number of counties:	168	205		3,182
Demographics				
Total population	917,390	1,273,125	355,735***	454,197
Total employment	230,910	236,700	-5,790	32,427
Non-white share	29.44%	30.29%	-0.85%	22.10%
Immigrant share	11.19%	12.41%	-1.22%	8.91%
Share older than 65	11.10%	11.35%	-0.25%	12.31%
% High school diploma only	23.87%	23.50%	0.37%	27.93%
% at least Bachelor's degree	32.57%	32.87%	-0.30%	27.28%
Economic				
Prime aged male joblessness	13.96%	14.88%	-0.92%**	14.46%
Self-employment rate	10.28%	10.30%	-0.02%	10.71%
Avg. income	$46,971	$47,610	-$639	$43,105
Mfg. share of employment	14.44%	13.81%	0.63%	11.42%
Single family home value (per sq. ft.)	$88.72	$91.71	-$2.99	$67.77
Community bank deposit share	13.65%	10.72%	2.93%*	8.05%
Announcement industry				
Employment	9,714	9,702	12	474
Employment share of county	3.08%	2.26%	0.81%**	1.38%
Establishment size	72	71	0.91	14

Notes: * 10%, ** 5%, *** 1%. The "All Counties" group differs from runners-up and winner counties at the 1% level for all categories except prime aged male joblessness and self-employment rate. All data reflect values from the mid-point of the sample period in the year 2000. Total population, non-white share, immigrant share, share older than 65, % high school diploma only, % at least bachelor's degree, prime aged male joblessness, self-employment rate and average income are obtained from the Census 2000 5% PUMS sample. Total population includes all individuals living in the county in the 2000 Census. Non-white share, immigrant share, and share $>$ 65 years of age are tabulated as the number of individuals in those categories divided by total population. The remaining PUMS variables are tabulated for prime aged males only. Prime aged males are males aged 25 to 55. Home values per square feet are tabulated using data from Zillow for March 2000. The sample matched to Zillow includes 134 winning counties, 169 runners-up and 1,499 in the 'all counties' sample. Community banks are defined as deposit taking institutions with assets less than $1 billion and either federally chartered and regulated by the Office of Thrift Supervision or state chartered and regulated by the FDIC. Total employment, mfg. share of employment, and all "Case-industry" variables are calculated from County Business Patterns for the year 2000. Employment in County Business Patterns reflect full-time employees on payroll on Tuesday the week of March 12 by county-industry. The tabulations are calculated at the 3-digit NAICS level with total employment reflecting the employment total across all 3-digit NAICS industries in the county. Mfg. share of employment is the percentage of employment between NAICS 310 and 339. The announcement industry variables are defined for the industries of firms in the Site Selection magazine sample. Employment shows the average employment levels in the industry grouped by county, employment share of county is the average employment share of the county employed in the industry, establishment size is the county-level employment over establishment count, and share of industry in county denotes the share of national industry employment in the county.

fication Number (EIN). To ensure firms are not mistakenly classified as startup from mechanical reassignments of firm identifiers, I do the following: Drop establishments that never report positive employment, label a firm as a startup only if all establishments associated with the parent firm appear for the first time, and calculate the age of a firm as the age of the oldest establishment when firm identifier first appears and then allow the firm to age naturally each year regardless of changes in its establishment composition. The key variables of interest tabulated from the Longitudinal

Business Database include county-industry firm entry rates, firm growth rate, and firm survival.

2.3.3 Longitudinal Employer Household Dynamics

The Longitudinal Employer Household Dynamics (LEHD) database allows researchers to construct work histories for all workers employed in a participating state. Observations are at the worker-firm-year-quarter level. Wage and employer information is obtained through state unemployment insurance offices. The LEHD is a complex network of files detailing quarterly wages for workers matched to a business reporting unit. The LEHD also provides geographic information as well as a limited set of demographic variables such as birth country, education, gender, age, and race. A limiting feature of the LEHD in studies over a long time horizon is that states enter the program in different years. Only 11 states participated in the program when it first began in 1992 with many of the current 31 states entering in the late 1990s through 2002. Researchers are granted access to states on a case-by-case basis with many only at the discretion of individual states themselves. Figure **??** shows the 25 states represented in my sample overlaid by the number of county-cases represented in each.

The LEHD is a complex network of files for each state. Figure **??** is the Census Bureau's diagram of the matching required depending on the LEHD subfiles the researcher requested. Part of the complexity of the LEHD is driven by the following: Firms may hold multiple federal EINs with each EIN representing a different tax reporting entity of the firm. Because the LEHD is organized at the state level with data provided through the ES-201 program through state unemployment insurance offices, employers are identified by state tax ids (SEINs). Just as firms have multiple EINs, they may also have multiple SEINs assigned to them by each state in which they operate. Employees are assigned unique personal identifiers which follow the workers throughout the panel and can be matched to their SEIN employer at the tax reporting level (SEINUNIT). Unlike establishment identifiers in other Census products including the LBD, SEINUNITs are not defined by physical locations and are instead tax reporting units within the SEIN. However, the Census does provide a concordance file allowing matching of SEINs to EINs. The difference in establishment definition prevents the perfect mapping of physical establishments in the LBD to a SEINUNIT in the LEHD. I use the combination of EIN-county, and entry date when possible, to obtain approximate employment count, industry, and age of the physical location of worker's job in the LEHD.

I use the Longitudinal Employer Household Dynamics data to construct work histories of individuals who ever start a firm. In the main employee-employer sample, I retain worker-firm-year observations for which average quarterly earnings in quarters worked at the firm is at above a full-time minimum wage salary of $2,678. To identify startups, I use the Longitudinal Business Database to obtain a list of firm-year observations for startup firms. I match these firm-years into the LEHD and the match yields a file of all employees in young firms. Empirically defining an entrepreneur in microdata is a challenge. The challenge arises because founders often are not compensated through income reported on W-2 filings from which LEHD earnings show. Second, Kerr and Kerr (2016) use Mark Zuckerberg's Facebook and Elon Musk's Tesla as examples to note that firms can be considered startups and individuals entrepreneurs long after they initially appear in the microdata. I follow the literature and label founders (also referred to as entrepreneurs in this proposal) to be the top three wage earners in years up to two years after the firm first appears in the LEHD. I then use this list of entrepreneurs to create a panel of their employer and wage histories. This panel allows me to construct variables measuring employee transitions across industries and firms, migrate geographically (in terms of county of employment), and the likelihood of starting a new firm given industry experience.

I use the LEHD to construct four separate, but related samples. Two samples are used to compare winning counties to runners-up counties and the remaining two compare the winning county relative to itself in a pre-period. Sample 1 compares winning and finalist counties at the firm-level. This sample will be used to identify which firms produce the most entrepreneurs. Sample 2 compares contest counties at the entrepreneur-level to capture human capital specific factors in starting a supply chain firm. The last two samples are analogous but focused only on capturing changes in the composition of entrepreneurs pre- and post-anchor firm opening within the winning county itself. I construct additional samples to ensure the winning and runners-up counties do not differ from each other in the outcomes of interest prior to the opening event.[20]

2.4 Empirical Strategy

2.4.1 Defining the Supply Chain

One data limitation is that I do not observe firms' business transactions. This means I cannot directly ascertain whether an entrant is selling to or buying from the anchor firm's new facility. Instead I define the supply chain at the 4-digit NAICS level for each case in the *Site Selection* facility openings sample using the Bureau of Economic Analysis Use of Commodities tables. The supply chain refers to firms in the industry of anchor firm, firms in industries upstream to the anchor firm's industry, and firms in industries downstream to the anchor firm's industry.

The empirical definitions of upstream and downstream come from tabulations of industry-level output and input dependency. An industry's output dependency, $Output_{i \to j}$, is the share of industry i's outputs sold to industry j. I let the anchor

[20] The main specification take the set of firms and individuals in year $t-2$ and estimate entrepreneurial outcomes within the first 7 years of the anchor firm's arrival. The pre-analysis sample takes the set of firms and individuals in year $t-8$ and estimates whether winning and finalist counties differ in entrepreneurial outcomes through year $t-2$.

firm's industry be industry i and label downstream ("buyer") industries for each industry i to be the leading five industry j's corresponding to the largest values of $Output_{i \to j}$. Table 2.4 lists the top 10 announcement industry-buyer industry pairs in the sample. Similarly, $Input_{i \leftarrow j}$ designates the share of industry i's inputs from industry j (Table 2.3). The five industry j's corresponding to the largest values of $Input_{i \leftarrow j}$ are upstream ("supplier") industries.

Table 2.3: Leading upstream industries

Announcement industry	Input provider	Share of inputs received from input producer
3241. Petroleum and Coal Products Manufacturing	2110. Oil and gas extraction	0.89
5250. Funds, Trusts, and Other Financial Vehicles	5239. Other Financial Investment Activities	0.76
3314. Nonferrous Metal (except Aluminum) Production and Processing	3314. Nonferrous Metal (except Aluminum) Production and Processing	0.74
3252. Resin, Synthetic Rubber, and Artificial Synthetic Fibers and Filaments Manufacturing	3251. Basic Chemical Manufacturing	0.69
5242. Agencies, Brokerages, and Other Insurance Related Activities	5242. Agencies, Brokerages, and Other Insurance Related Activities	0.68
3361. Motor Vehicle Manufacturing	3363. Motor Vehicle Parts Manufacturing	0.64
3222. Converted Paper Product Manufacturing	3221. Pulp, Paper, and Paperboard Mills	0.62
3251. Basic Chemical Manufacturing	3251. Basic Chemical Manufacturing	0.61
5241. Insurance Carriers	5418. Advertising, Public Relations, and Related Services	0.61
3116. Animal Slaughtering and Processing	1123. Poultry and egg production	0.60

Notes: Upstream industries are industries that serve as suppliers to the anchor firm's industry ("Announcement Industry"). Data is tabulated using the Bureau of Economic Analysis's "Use of Commodities" tables for 2007 and 2012. The rows in this table should be read as, for example, "52% of inputs used by the computer and peripheral equipment manufacturing industry is purchased from the semiconductor and other electronic components manufacturing industry." The table displays the ten industry pairs with the greatest buyer relationships and only includes buyer industries in the plant openings sample.

Table 2.4: Leading downstream industries

Announcement industry	Purchasing industry	Share of output sold to purchasing industry
3366. Ship and Boat Building	4830. Water Transportation	0.79
3361. Motor Vehicle Manufacturing	3362. Motor Vehicle Body and Trailer Manufacturing	0.75
3252. Resin, Synthetic Rubber, and Artificial Synthetic Fibers and Filaments Manufacturing	3261. Plastics Product Manufacturing	0.66
5239. Other Financial Investment Activities	5250. Funds, Trusts, and Other Financial Vehicles	0.63
3363. Motor Vehicle Parts Manufacturing	3361. Motor Vehicle Manufacturing	0.58
3362. Motor Vehicle Body and Trailer Manufacturing	3361. Motor Vehicle Manufacturing	0.55
3371. Household and Institutional Furniture and Kitchen Cabinet Manufacturing	2334. Building, Developing, and General Contracting	0.55
4231. Motor Vehicle and Motor Vehicle Parts and Supplies Merchant Wholesalers	3361. Motor Vehicle Manufacturing	0.52
3259. Other Chemical Product and Preparation Manufacturing	3231. Printing and Related Support Activities	0.50
5250. Funds, Trusts, and Other Financial Vehicles	5242. Agencies, Brokerages, and Other Insurance Related Activities	0.47

Notes: Downstream industries purchase inputs from the anchor firm's industry ("Announcement Industry"). Data is tabulated using the Bureau of Economic Analysis's "Use of Commodities" tables for 2007 and 2012. The table should be read as "66% of output produced by firms in the resin, synthetic rubber, and atrificial synthetic fibers manufacturing industry is purchased by plastics products manufacturing firms." The table displays the ten industry pairs with the greatest buyer relationship. The focal industries on the left most column are restricted to the large facilities openings industries.

2.4.2 Difference-in-Differences Estimation

I use the introduction of a large facility of a large firm as a shock to both the local return to as well as the opportunity cost becoming an entrepreneur. I use the revealed preferences of large firms in the *Site Selection* announcements as a quasi-experiment to capture sudden opportunities that arise for individuals with specialized human capitals. These opportunities are represented as sudden changes in the proximity of a large potential buyer or supplier for an entrepreneur. The underlining assumption is that conditional on observed factors that lead an anchor firm to open a new facility in one county over another are orthogonal to unobserved characteristics that predict new firm formation and spinoffs by incumbent firms.

Table 2.2 provides descriptive statistics and test for equality between counties on a number of factors commonly thought to be related to entrepreneurship, the winners and runners-up counties are statistically similar. The tabulations are constructed by taking the averages of counties in each category: Winners, runners-up, and all. The set of factors listed are a small subset of the features that drive selection into entrepreneurship and regional startup rates. The table displays figures for the year 2000 since that year is near the center of period over which openings events were announced and for whom data is consistently available across all publicly available demographic and economic data sets. The variables include some of the most important traits including immigration status, education, size of the manufacturing base (Chinitz 1961), home values[21], labor market conditions (Fairlie 2013), presence of community banks, and establishment sizes (Glaeser, Kerr, and Kerr 2015).[22] I also test the overall establishment size distribution across each supply chain component.

[21]However, some studies including Hurst and Lusardi (2004) and Kerr, Kerr, and Nanda (2015) find evidence that wealth constraints and housing collateral are only modest barriers to entry.

[22]The data on demographic traits as well as prime aged male joblessness are taken from the 2000 Census Public Use Microdata Series which overrepresents larger counties. Only nearly 600 of the 3,100 U.S. counties are in the file. However, all other traits are taken from County Business Patterns, Zillow, and the FDIC Summary of Deposits which have more complete coverage.

I find that winners and runners-up counties do not differ from each other in the distribution of establishments by firm size.[23]

Assuming the winner and runners-up counties within each site selection case are similar on dimensions related to entrepreneurship conditional on observable county-specific characteristics, I estimate a difference-in-differences model. The general form of the equation I estimate is:

$$\begin{aligned} y_{fijct} =& \beta_0 + \beta_1 \times \textit{Pre-trend}_{jt} + \beta_2 \times \textit{Post}_{jt} + \beta_3 \times \textit{Win}_{jct} + \\ & \beta_4 \times \textit{Win}_{jct} \times \textit{Post}_{jt} + \beta_5 \times \textit{Win}_{jct} \times \textit{Post-trend}_{jt} + \\ & \alpha_j + \lambda_{it} + \epsilon_{fijt}, \end{aligned} \tag{2.1}$$

where f denotes a firm in industry i, case j, and county c at time t, and y_{fijct} indicates the dependent variable of interest as described in the variables section. β_4 is the difference-in-differences estimator of interest. λ_{it} are industry-year fixed effects to control for secular trends that vary by industry and ω_c capture county fixed effects. The case fixed effects, α_j, ensure the difference-in-differences estimators are identified off comparisons between the winner and runner-up within the specific announcement event. Reported regression coefficients are restricted to the set of startups in winners and finalist counties, but in robustness checks I use the set of all counties to enhance identification of the industry-year fixed effects. I consider the pre-period to be up through five years before the announcement article and the post-period to run through 7 years afterwards.[24]

To capture the impacts of founder background and firm characteristics I take the set of workers in year $t-2$ relative to the facilities opening and compare the number of employee departures for incumbent firms in winning and runners-up counties between

[23] I test whether winners and runners-up counties have the same number of establishments by size groupings.

[24] The case fixed effects are dummy variables for the case number associated with the firm-county observation interacted with a dummy variable if the observation is between five years prior to and seven years post-announcement.

t and $t+7$. I use the LEHD to observe employee transitions across firms and identify startup founders. Because the state-level LEHD uses a slightly different definition of a business than IRS-based LBD, I define incumbent firms at the SEIN-county level. I first take the set of all employees in each SEIN-county in year $t-2$ and use as the outcome variable the number of employees who depart the firm in the post period to start a firm in the supply chain. In founder-based regressions, my outcome variable is instead defined at the worker-level and I use a linear probability model to estimate the likelihood an individual starts a firm in the supply chain conditional on departing an incumbent firm to become an entrepreneur.

The baseline regression follows a simple difference-in-differences for the number of employee departures:

$$\begin{aligned} y_{fijct} =& \beta_0 + \beta_1 \times Post_{jt} + \beta_2 \times Win_{jct} + \\ & \beta_4 \times Win_{jct} \times Post_{jt} + \beta_5 \times X_{fj} + \alpha_j + \lambda_{it} + \epsilon_{fijt}, \end{aligned} \tag{2.2}$$

Here, the firm f is at the SEIN-county level and X_{fj} is a vector of firm controls including total employment and age. To capture the change in likelihood that an entrepreneur selects into the supply chain, I define the outcome as y_{pijct} which takes on the value of 1 if the entrepreneur selects into a supply chain industry. These entrepreneur-level regressions control for race, sex, education, age, and age squared.

To estimate the variation across the supply chain by experience within specific components of the supply chain, I add industry interactions to Equation 2:

$$\begin{aligned} y_{fijct} = & \beta_0 + \beta_1 \times Win_{fjc} + \beta_2 \times Upstream_{fijc} + \\ & \beta_3 \times Anchor_{fijc} + \beta_4 \times Downstream_{fijc} + \beta_5 \times Upstream_{fijc} \times Win_{fjc} + \\ & \beta_6 \times Anchor_{fijc} \times Win_{fjc} + \beta_7 \times Downstream_{fijc} \times Win_{fjc} + \\ & \beta_8 \times X_{fjt} + \alpha_j + \lambda_{it} + \epsilon_{fijt}, \end{aligned} \tag{2.3}$$

I make the necessary substitutions of y_{pijct} in the outcome and X_{fjt} as a control for entrepreneur-level regression.[25] The coefficients of interest are β_5, β_6, and β_7.

To estimate the role of founder characteristics how firm and founder characteristics affect spinoffs as well as selection of entrepreneurs into supply chain industries, I run analogous regressions to Equation 3 by adding interactions with C_fjc, the characteristic of interest:

$$
\begin{aligned}
y_{fijct} = \beta_0 + \beta_1 \times Win_{fjc} + \beta_2 \times Win_{fjc} \times C_{fjc} + & \\
\beta_3 \times Upstream_{fijc} + \beta_4 \times Anchor_{fijc} + \beta_5 \times Downstream_{fijc} + & \\
\beta_6 \times Upstream_{fijc} \times Win_{fjc} + \beta_7 \times Anchor_{fijc} \times Win_{fjc} + & \\
\beta_8 \times Downstream_{fijc} \times Win_{fjc} + \beta_9 \times Upstream_{fijc} \times Win_{fjc} \times C_{fjc} + & \\
\beta_{10} \times Anchor_{fijc} \times Win_{fjc} \times C_{fjc} + \beta_{11} \times Downstream_{fijc} \times Win_{fjc} \times C_{fjc} + & \\
\beta_{12} \times X_{fj}\alpha_j + \lambda_{it} + \epsilon_{fijt}, &
\end{aligned}
\tag{2.4}
$$

where β_2, β_9, β_{10}, and β_{11} are the coefficients of interest. The excluded group from the industry designations are the "all other industries" category. This means that coefficients are interpreted as relative to a supply chain founder from an industry unrelated to the focal firm's defined supply chain. I interpret the β_3's to suggest the relative propensity of an entrepreneurial individual selecting into the announcement firm facility's supply chain.

To lend credibility that the opening of the anchor firm drives variation in the post period, I repeat all specification using only the pre-period. I compare outcomes for businesses and individuals in winners and runners-up counties between 8 and 2 years prior to the facilities openings. Overall, I find no evidence that the counties differ in the number of spinoffs incumbent firms yield or the likelihood of entrepreneurs to select into the industries of the anchor firm that eventually arrives. Online Appendix

[25] Even though these regressions fix the firm (or eventual entrepreneur) in year $t-2$, the treatment of the facilities opening occurs anytime between 1995 and 2008 which is why the secular industry trend λ_{it} remains as a fully identified fixed effect.

B outlines the robustness checks on spinoff and entrepreneur-level specifications.[26]

2.5 Results

2.5.1 Economic Proximity and Internal Employment Growth

I first document changes in firm entry in the supply chain after the arrival of an anchor firm at the county-industry level.[27] I measure the entry rate of new businesses in a county as the share of establishments in the county-industry attributable to new startup firms. Following the literature, I consider startups to be firms aged less than three years.[28]

The sample consists of all firms in the contest county for each announcement event's associated supply chain. The supply chain is defined independently for each individual announcement case. Therefore, the industries included in case 1 are different than those in case 2. The number of new establishments of firms aged less than three and employment in these establishments are collapsed at the county-announcement event-industry level. This implies that for each contest county within a case there are annual observations for up to 5 supplier industries, 1 anchor industry, and 5 buyer industries. Industries are defined at the 4-digit NAICS level. The coefficients in Tables 5-8 represent the average effect the anchor firm has on entrepreneurship at the county-industry pair level. Column 1 of Table 2.5 shows an increase of about 23 new establishments in each industry-county pair of all supply chain industries. This represents an increase in county-industry entry rate of approximately 0.75%, or a 10% increase in the entry rate from the county-industry average. Column 2 shows that

[26] An additional check is to measure the impact of effects by size of the arriving anchor firm.

[27] The null results showing no change in startup entry and startup employment have not yet been disclosed.

[28] I use firms aged less than three in part due to noise in microdata related to spurious entry in the initial year that are often populated by 0's for employment, missing payroll, or a 0 for either employment or payroll but not both.

the impact on employment amounts to 429 new jobs added in county-industry pairs associated with the supply chain. In total, this amounts to over 120 total startup establishments and 2,300 jobs in startups across all unique supply chain NAICS-4 industries represented in an average case.

Table 2.5: Supply Chain county-industry entry and startup growth

	(1)	(2)	(3)	(4)
	Establishments of new firms	Employment from new firms	5-year growth rate	5-year failure probability
Constant	7.389***	3.301***	-0.3378***	0.166***
	(0.114)	(0.123)	(0.0160)	(0.024)
Win	-0.8781	23.62	-0.0410	0.01
	(10.24)	(125.2)	(0.0346)	(0.76)
Post	-6.607*	-489.1***	-0.0190	0.01
	(3.446)	(120.3)	(0.0182)	(1.60)
Win * Post	22.05*	429.6**	0.1225*	-0.0680*
	(11.98)	(185.1)	(0.0663)	(0.0340)
Win * Post-trend	-1.538*	-28.82**	-0.0087*	0.005*
	(0.892)	(13.98)	(0.0046)	(0.003)
Pre-trend	-0.573**	-34.93***	-0.0022	0.00
	(0.236)	(10.83)	(0.0020)	(1.50)
Industry x Year FE	Y	Y	Y	Y
Case FE	Y	Y	Y	Y
Effect in t+5	12.82*	256.7**		
Adj. R-Sq.	0.7552	0.6032	0.0988	0.1330
N	91500	91500	215000	215000

Notes: * 0.1 ** 0.05 *** 0.01. The anchor firm's arrival induces the formation of supply chain startups that add 430 new jobs to the county. These startups grow 12% faster in employment over the first five years and fail 7% less. This table displays output from a difference-in-differences regression estimating the impact of a large facilities opening on the outcome indicated for all industries in the supply chain. The unit of observation is at the county-industry-year level. Industries are defined at the 4-digit NAICS level. Industries are classified as being in the same industry as the oepning announcement, as a supplier firm (upstream), or buyer firm (downstream). Supplier and buyer are defined as the top 5 non-zero industry of goods & services flows by the Bureau of Economic Analysis Use of Commodities tables. The outcome variable in Column 1 is the number of new startups at the county-industry level. Column 2 shows the total employment attributable to these new businesses. Column 3 shows the 5-year growth rate of firms after their first year of appearing in the Longitudinal Business Database. The growth rate is calculated using the Davis, Haltiwanger, and Schuh (1996) firm growth calculation aggregated by establishment-level microdata. The measure is symmetric on the interval [-2, 2], where -2 indicates firm exit and 2 indicating firm entry into the industry. Column 4 displays the impact of a facilties opening on the likelihood the startup fails within five years. Column 4 uses a binary variable of 1 indicating the startup fails within five years as the outcome. "Effect in t+5" shows the number of new establishments from new firms (Column 1) and new job from these establishments (Column 2) in year t+5. These values are not shown for Columns 3 and 4, though the results are qualitatively similar to the difference-in-differences estimator. All regressions are weighted by total county-industry employment. Census rounding rules limit the number of significant digits reported to 4. The number of decimal places reported in each column use the minimum number after Census rounding. Standard errors are clustered by county.

Tables 2.6, 2.7, and 2.8 show that the aggregate magnitude is largely driven by upstream entry followed by downstream entry. Entry does not occur in the anchor firm's industry. First restricting to county-industry pairs defined as upstream within each announcement event in Table 2.6, Column 1's estimate of the difference-in-difference estimator shows a treatment effect amounting to an increase in 23 new startup establishments associated with 453 (Column 2) jobs at the county-industry level. This increase amounts to about 15% of total upstream employment. Though currently undisclosed, much of this upstream entry is driven by startup formation in the first three years of the anchor firm's arrival.

Table 2.6: Upstream county-industry entry and startup growth

	(1)	(2)	(3)	(4)
	Establishments of new firms	Employment from new firms	5-year growth rate	5-year failure probability
Constant	76.34***	1322***	-0.3482***	0.167***
	(4.55)	(91)	(0.0217)	(0.017)
Win	-6.72	-190.4*	-0.0678	0.03
	(10.82)	(101.6)	(0.0535)	(1.12)
Post	-10.36	-530.4***	-0.0588*	0.0425**
	(6.89)	(197.0)	(0.0317)	(0.0220)
Win * Post	23.50***	453.1*	0.1888*	-0.104*
	(8.06)	(254.1)	(0.1094)	(0.063)
Win * Post-trend	-1.836***	-27.10*	-0.0135*	0.0075*
	(0.682)	(16.16)	(0.0076)	(0.0045)
Pre-trend	-1.322**	-47.20***	-0.0086**	0.0048**
	(0.611)	(15.24)	(0.0039)	(0.0023)
Industry x Year FE	Y	Y	Y	Y
Case FE	Y	Y	Y	Y
Effect in t+5	12.49***	290.5*		
Adj. R-Sq.	0.7259	0.7717	0.1312	0.1660
N	32000	32000	88500	88500

Notes: * 0.1 ** 0.05 *** 0.01. After the anchor firm's arrival, the number of upstream startup establishments increases by 23 and represent 450 new jobs. These startups grow 19% faster and fail 10% less. This table displays output from a difference-in-differences regression estimating the impact of a large facilities opening on the outcome indicated for all upstream industries in the supply chain. The unit of observation is at the county-industry-year level. Industries are defined at the 4-digit NAICS level. The outcome variable in Column 1 is the number of new startups at the county-industry level. Column 2 shows the total employment attributable to these new businesses. Column 3 shows the 5-year growth rate of firms after their first year of appearing in the Longitudinal Business Database. The growth rate is calculated using the Davis, Haltiwanger, and Schuh (1996) firm growth calculation aggregated by establishment-level microdata. The measure is symmetric on the interval [-2, 2], where -2 indicates firm exit and 2 indicating firm entry into the industry. Column 4 displays the impact of a facilties opening on the likelihood the startup fails within five years. Column 4 uses a binary variable of 1 indicating the startup fails within five years as the outcome. "Effect in t+5" shows the number of new establishments from new firms (Column 1) and new job from these establishments (Column 2) in year t+5. These values are not shown for Columns 3 and 4, though the results are qualitatively similar to the difference-in-differences estimator. All regressions are weighted by total county-industry employment. Census rounding rules limit the number of significant digits reported to 4. The number of decimal places reported in each column use the minimum number after Census rounding. Standard errors are clustered by county.

Table 2.7 shows no increased startup activity in the anchor firm's own industry.

Table 2.8 restricts to county-industries downstream to the anchor firm. A comparison of Columns 1 and 2 shows that while the number of downstream startups do not increase, those that do enter at much larger scale than before with over 260 new

Table 2.7: Anchor firm county-industry entry and startup growth

	(1)	(2)	(3)	(4)
	Establishments of new firms	Employment from new firms	5-year growth rate	5-year failure probability
Constant	25.41***	641.7***	-0.4532***	0.186***
	(1.32)	(102.2)	(0.0793)	(0.071)
Win	-8.911***	-290.5**	0.0571	-0.0197
	(2.927)	(141.2)	(0.0633)	(0.0660)
Post	0.345	-23.5	0.1365	-0.0123
	(1.597)	(187.2)	(0.1458)	(0.0290)
Win * Post	6.134	16.4	-0.2488	0.0136
	(4.993)	(360.9)	(0.1612)	(0.1900)
Win * Post-trend	-0.2924	3.76	0.0119	-0.0015
	(0.3594)	(28.56)	(0.0123)	(0.2500)
Pre-trend	0.0803	104.7	0.0181	-0.0008
	(0.1791)	(69.3)	(0.0165)	(0.1700)
Industry x Year FE	Y	Y	Y	Y
Case FE	Y	Y	Y	Y
Effect in t+5	4.380	38.98		
Adj. R-Sq.	0.9035	0.4986	0.1141	0.1540
N	12000	12000	10500	10500

Notes: * 0.1 ** 0.05 *** 0.01. Startup formation and performance in the anchor firm's industry are unaffected by the facilities opening event. This table displays output from a difference-in-differences regression estimating the impact of a large facilities opening on the outcome indicated for the anchor firm's industry. The unit of observation is at the county-industry-year level. Industries are defined at the 4-digit NAICS level. The outcome variable in Column 1 is the number of new startups at the county-industry level. Column 2 shows the total employment attributable to these new businesses. Column 3 shows the 5-year growth rate of firms after their first year of appearing in the Longitudinal Business Database. The growth rate is calculated using the Davis, Haltiwanger, and Schuh (1996) firm growth calculation aggregated by establishment-level microdata. The measure is symmetric on the interval [-2, 2], where -2 indicates firm exit and 2 indicating firm entry into the industry. Column 4 displays the impact of a facilities opening on the likelihood the startup fails within five years. Column 4 uses a binary variable of 1 indicating the startup fails within five years as the outcome. "Effect in t+5" shows the number of new establishments from new firms (Column 1) and new job from these establishments (Column 2) in year t+5. These values are not shown for Columns 3 and 4, though the results are qualitatively similar to the difference-in-differences estimator. All regressions are weighted by total county-industry employment. Census rounding rules limit the number of significant digits reported to 4. The number of decimal places reported in each column use the minimum number after Census rounding. Standard errors are clustered by county.

jobs in new startup establishments. The effects last at least 5 years after the anchor firm arrives. Tables 6 and 8 show that in the fifth year after the anchor firm arrives, 291 jobs in upstream startups and 175 jobs in downstream startups are still being created.

One possible effect of the anchor firm's arrival is that startup entry occurs *en masse* due to expectations on the ease of finding suppliers and customers or perceived signalling about the quality of the local industry. If the arriving anchor firm leads to more entry driven by unfounded beliefs on the local market, these entrants will be of low quality and fail quickly. To measure startup growth, I use the conventional definition for calculating firm growth from establishment-level microdata

Table 2.8: Downstream county-industry entry and startup growth

	(1)	(2)	(3)	(4)
	Establishments of new firms	Employment from new firms	5-year growth rate	5-year failure probability
Constant	7.896***	739.8***	-0.3182***	0.159***
	(0.180)	(53.8)	(0.0220)	(0.062)
Win	-6.083	-56.41	-0.0286	0.0035
	(11.03)	(85.76)	(0.0416)	(0.2002)
Post	-8.969	-42.0	-0.0072	-0.0040
	(6.188)	(123.8)	(0.0240)	(0.4401)
Win * Post	25.79	260.6**	0.1511**	-0.0593*
	(18.63)	(121.1)	(0.0741)	(0.0362)
Win * Post-trend	-1.829	-14.30*	-0.0092*	0.004
	(1.145)	(7.27)	(0.0052)	(1.900)
Pre-trend	-0.9872**	-4.63	0.0006	-0.0004
	(0.4386)	(12.18)	(0.0026)	(0.3400)
Industry x Year FE	Y	Y	Y	Y
Case FE	Y	Y	Y	Y
Effect in t+5	14.82	174.8**		
Adj. R-Sq.	0.7966	0.3737	0.0953	0.1300
N	28000	28000	115000	115000

Notes: * 0.1 ** 0.05 *** 0.01. After the anchor firm's arrival, downstream startups generate 260 more jobs, grow 15% faster, and fail 6% less. This table displays output from a difference-in-differences regression estimating the impact of a large facilities opening on the outcome indicated for all downstream industries in the supply chain. The unit of observation is at the county-industry-year level. Industries are defined at teh 4-digit NAICS level. The outcome variable in Column 1 is the number of new startups at the county-industry level. Column 2 shows the total employment attributable to these new businesses. Column 3 shows the 5-year growth rate of firms after their first year of appearing in the Longitudinal Business Database. The growth rate is calculated using the Davis, Haltiwanger, and Schuh (1996) firm growth calculation aggregated by establishment-level microdata. The measure is symmetric on the interval [-2, 2], where -2 indicates firm exit and 2 indicating firm entry into the industry. Column 4 displays the impact of a facilties opening on the likelihood the startup fails within five years. Column 4 uses a binary variable of 1 indicating the startup fails within five years as the outcome. "Effect in t+5" shows the number of new establishments from new firms (Column 1) and new job from these establishments (Column 2) in year t+5. These values are not shown for Columns 3 and 4, though the results are qualitatively similar to the difference-in-differences estimator. All regressions are weighted by total county-industry employment. Census rounding rules limit the number of significant digits reported to 4. The number of decimal places reported in each column use the minimum number after Census rounding. Standard errors are clustered by county.

(Davis, Haltiwanger, and Schuh 1996):

$$g_{ft} = \sum_{k \in f} \gamma_{kt} \times \frac{\text{emp}_{k,t+5} - \text{emp}_{kt}}{0.5 * (\text{emp}_{kt} + \text{emp}_{kt+5})} \tag{2.5}$$

where γ_{kt} is establishment k's share of firm f's total average employment between year t and $t + 5$. My focus is on understanding how much firms grow specifically in the supply chain. Therefore, I measure firm growth with respect to the total employment in the supply chain position. For example, Table 2.6 uses a startup's growth in all upstream industries as the outcome.[29] This outcome variable is used to estimate

[29] A startup could operate in industry A and industry B, both of which are supplier industries. I

Equation 1 and is reported in Column 3 of Tables 5-8. Table5 demonstrates that the effect of the anchor firm's arrival in winning counties on startup's supply chain growth is 12.25%. Table 2.6 shows that the impact on upstream startup growth is 18.88%. Table 2.7 continues to suggest the lack of an effect within the anchor firm's industry itself. Table 2.8 shows the treatment effect of a 15.11% increase in employment growth.

I also test whether the survival probabilities differ as well. To accurately capture startup exits, I follow existing common practice and only consider a startup an exit if all establishments underneath a firm identifiers in year t exit by year $t + 5$. Formally, in Equation I use the following outcome:

$$Pr(\text{Firm exit}_{ft} = 1), \tag{2.6}$$

if all establishments exit by year $t + 5$. Column 4 of Tables 5-8 show the magnitude of the treatment effect of an anchor firm on failure probabilities. Table 2.5 shows a decrease of 6.8% across all supply chain industries. Table 2.6 estimates a decline of 10.4% for upstream industries. Tables 7 and 8 and show no change in failure and a decline of 5.9% in anchor and downstream industries, respectively.

For all regressions, I also test for parallel trends. I use a dynamic difference-in-difference where the win dummy variable is interacting for a time dummy for all times of the case period. I find no significant results on any of the time dummies. Please refer to Online Appendix B for details on these additional checks.[30] In addition, I show that aggregate employment and establishment counts across the full firm age distribution in the county and supply chain do not increase.[31]

combine firm-county employment for these two industries and calculate the firm's upstream growth. To identify industry-by-year fixed effects, I use the NAICS-4 of greatest employment (i.e., greatest employment between A and B).

[30]Forthcoming when robustness check disclosure is passed.

[31]These additional results are under disclosure review as well. However, the aggregate results may also be identified using County Business Patterns. Effects of the anchor firm on total employment

In summary, Tables 5-8 show firm entry, new venture growth, and survival improves in upstream and downstream industries. This set of results shows: 1) Aggregate employment and establishment counts across all firm ages are unchanged using the set of all industries as well as limiting to supply chain industries for each announcement case. The arrival of the anchor firm appears to only shift the composition of industry employment from old to young in supply chain industries. 2) Anchor firms induce entry and employment gains in startups in upstream and downstream firms. 3) Anchor firms lead to faster growth and survival of supply chain startups. Similar to works discussing the importance of local factor markets in supporting entrepreneurship or clustering of firms (e.g., Glaeser and Kerr 2009), these results suggest potential increases in the thickness of local supplier-buyer networks support entry and startup performance.

2.5.2 Spinoffs and Employee Departures

All analyses in Section 4.2 are at the level of incumbent firms and their employees who spin-out in winning and runners-up counties. I find descriptive evidence that anchor firms typically focus their final location candidates to counties that have greater employment specialization in industries upstream and downstream to them (Table 2.2 and Online Appendix B). I use this descriptive tendency as motivation for probing at whether the treatment effect found in Section 5.1 is specific to founders of high human capital as measured by industry experience and wages. If high quality firms incubate entrepreneurs (e.g., Gompers et al., 2005, and Chatterji, 2009), then proxies of quality such as wages should relate positively to the number of entrepreneurs they produce. If industry experience is essential then not only will high quality firms generate more entrepreneurs, the entrepreneurs will select into their employer's industry. Finally,

and establishment counts are near 0 across all industries, supply chain industries only, upstream industries only, anchor industries only, and downstream industries only.

I investigate heterogeneity within incumbent firms to test whether firms lose more individuals to spinoffs or individuals of higher ability or ranking within the firm. Namely, I focus on industry experience and ability using earnings as a proxy. I conduct the analysis from two perspectives. In each characteristic of interest I first investigate which firms generate the greatest number of spinoffs as measured by the number of individuals who leave the firm to start their own in the post-opening period. The second set of regressions under each heading estimate a linear probability model of industry selection of entrepreneurs conditional on the individual eventually becoming an entrepreneur.[32]

Spinoffs in the Supply Chain

I first ask whether firm entry is driven by spinoffs or a general transition of workers to entrepreneurship regardless of industry background. Operationally I define a spinoff as firms founded by employees in the same segment (upstream, anchor firm industry, or downstream) of the supply chain as their most recent employer. Following the sample construction outlined in Section 4.2, Table 2.9 estimates Equation 3 and displays the key coefficients of interest on the triple interaction between the winning county indicator, runners-up county dummy variable, and the supply chain category variable. The baseline category is grouped as all industries outside the supply chain which means the coefficients are interpreted as the number of employee departures into entrepreneurship relative to firms outside the supply chain.

Table 2.9 uses the total number of employee departures into any industry of entrepreneurship. Looking within a column, the rows indicate the relative number of employee departures into entrepreneurship of that category. Column 1 shows that

[32]The idea behind estimating industry selection only on the set of eventual entrepreneurs is that I aim to measure changes in the composition of startups after the anchor firm arrives. However, this does assume that individuals decide to select into entrepreneurship and then pinpoint the industry they enter instead of first estimating the latent distribution of would-be entrepreneurs who stay employed.

Table 2.9: Number of Employee Departures to Entrepreneurship by Firm Industry

	(1)	(2)	(3)	(4)
	# of employees founding a firm	# of employees founding upstream firm	# of employees founding anchor industry firm	# of employees founding downstream firm
Win * Upstream	-0.0914	0.1682***	0.0305***	0.0271**
	(0.0873)	(0.0258)	(0.0115)	(0.0119)
Win * Anchor Industry	-0.0981	0.4103***	0.4960***	0.1365*
	(0.3831)	(0.0814)	(0.0793)	(0.0784)
Win * Downstream	0.0403	0.0745***	0.0167	0.3968***
	(0.0910)	(0.0197)	(0.0177)	(0.0432)
Main Effects	Y	Y	Y	Y
Firm Controls	Y	Y	Y	Y
Industry x Year FE	Y	Y	Y	Y
Case FE	Y	Y	Y	Y
Adj. R-Sq.	0.6367	0.1149	0.0596	0.0838
N	1094000	1094000	1094000	1094000

Notes: * 0.1 ** 0.05 *** 0.01. Winning county incumbent firms do not lose more employees to entrepreneurship after the anchor firm's arrival, but the industries their employees select into shifts in favor of supply chain industries. This table shows the variation in a facilities opening on employee departures to entrepreneurship by firm-industry of employment. The outcome variables are the number of employees in a firm who leave within 7 years of the opening to found a new firm. Firm founders are defined as a top 3 wage earner in a firm younger than three years. All regressions regress the outcome variable on a dummy variable indicating whether the firm is in a winning or runners-up county, industry classification in the supply chain, and interaction between winning status and each industry group. The excluded industry category is "all other industries". Industry groupings are determined for each case: Firm-industries are coded differently in case 1 versus case 2 depending on the anchor firm's industry in each case. Column 1's outcome variable is the total number of employee departures to entrepreneurship regardless of industry (including non-supply chain), Column 2 shows the number departing to form an upstream firm to the anchor firm, 3 for anchor firm's industry, and 4 for downstream industry startups. Regressions are weighted by firm size and standard errors are clustered at the SEIN-county.

firms in winning counties after the anchor firm's arrival do not lose more employees to entrepreneurship. Column 2 shows that relative to firms outside the supply chain, upstream incumbents in winning counties lose .17 more employees after the arrival of the anchor firm. Relative to the baseline this coefficient represents an increase of over 30% in employee departures. Put differently, this coefficient implies the anchor firm's arrival leads to 1.7 new spinoffs for every ten upstream firms in the county. Moving through Columns 2, 3, 4, Table 2.9 shows that firms disproportionately lose employees to entrepreneurship in their own segment of the supply chain. Upstream and downstream firms lose employees disproportionately to upstream and downstream entrepreneurship and anchor industry firms generate entrepreneurs across the supply chain. The introduction of an anchor firm to the county also induces employee departures into supply chain entrepreneurship by employees in incumbent anchor industries.[33]

Table 2.10 shows selection into supply chain entrepreneurship conditional on an individual eventually becoming an entrepreneur. All coefficients represent the change in likelihood relative to the counterfactual counties that an entrepreneur selects into the specified industry (column headings) given their industry of employment prior to the anchor firm's arrival. Moving downwards diagonally across columns, the table shows entrepreneurs become significantly more likely to start a new venture in the supply chain position of their prior employer. Entrepreneurs from upstream (anchor firm) industries are 8.7% (4.4%) more likely to start an upstream (anchor industry) firm. The strongest effect is found in entrepreneurs from downstream firms who become 17.2% more likely to have their startup in a downstream industry. Table 2.10 also shows having experience in an anchor industry only has a significant and positive impact of forming an anchor industry startup and corresponds to a strongly negative (-17.3%) impact on founding a downstream venture.

Taken together Tables 9 and 10 demonstrate that the anchor firm's arrival alters the industry composition of entrepreneurship. While firms do not lose more employees to entrepreneurship than before, those that they do lose are more likely to form spinoffs and operate in a similar industry. The change in composition of founder background is driven by the opening of a large facility in the supply chain of the county's incumbent firms. Table 2.9 also shows that anchor industry incumbent firms in treated counties generate more entrepreneurs across supply chain industries, but Table 2.10 suggests the likelihood of starting a firm in the supply chain is unaffected for upstream startups and reduced for downstream startups. These two facts may suggest that the number of employee departures into supply chain entrepreneurship by anchor firm industries is driven by departures by a concentrated set of firms such that

[33]The coefficients allow for computing the implied share of new startups by a county's employees that are spinoffs: 25% of upstream, 91% of anchor industry, and 71% of downstream.

Table 2.10: Likelihood of Entry Industry by Industry of Prior Employment

	(1)	(2)	(3)
	Pr(Upstream startup)	Pr(Anchor industry startup)	Pr(Downstream startup)
Win * Upstream	0.0871***	0.0083*	-0.0054
	(0.0144)	(0.0046)	(0.0054)
Win * Anchor Industry	-0.0126	0.0444**	-0.1730***
	(0.0142)	(0.0179)	(0.0227)
Win * Downstream	-0.0115	-0.0286***	0.1720***
	(0.0071)	(0.0070)	(0.0188)
Main Effects	Y	Y	Y
Worker Controls	Y	Y	Y
Industry x Year FE	Y	Y	Y
Case FE	Y	Y	Y
Adj. R-Sq.	0.1053	0.0589	0.1485
N	2124000	2124000	2124000

Notes: * 0.1 ** 0.05 *** 0.01. The anchor firm's arrival increases the likelihood that an entrepreneur in the winning county selects into supply chain entrepreneurship. This table shows the variation in a facilities opening on likelihood an individual departs to entrepreneurship by firm-industry of employment. The sample takes individuals who ever start a firm in the post-opening period. Each of the three regressions estimate whether the entrepreneur started a firm in the specified industry (columns) depending on industry of employment in the year prior to the announcement event (rows). The excluded industry category is "all other industries". Controls include age, age-squared, sex, race, education, and SEIN FEs. Regressions are unweighted and standard errors are clustered at the SEIN-county.

across all entrepreneurs (Table 2.10) the likelihood of an entrepreneur coming from an anchor industry is largely unaffected by being in a treated county.[34] This finding potentially complements both Greenstone et al. (2010) and Bloom et al. (2019) who discuss productivity and management practices spillovers onto incumbent firms in the anchor firm's industry.

One concern with estimating Equation 3 is pre-existing differences in spinoffs and entrepreneurial selection into specific industry prior to the anchor firm's arrival.[35] To account for this, I run the same analyses using incumbent firms in year $t-8$ and check spinoffs over the subsequent years up until the anchor firm's arrival and find no differences between the counties. In Section 4.3 I further compare winning counties against themselves from a long pre-period to verify that the change in composition

[34]Subsequent analysis in Section 5.2.2 will show the characteristics of anchor industry firms who lose employees to entrepreneurship in the supply chain.

[35]This concern arises because the comparisons are made between winning and runners-up county between year $t-2$ and year $t+7$ without an explicit control from pre-existing differences.

of winning counties specifically is driving the findings.

Firm Compensation and Wage Compression

Though the literature is personnel economics generally does not focus specifically on the relationship between compensation and the characteristics of firms employees may transition to, factors such as average firm wage and pay compression may be closely related to employee departures. Outside of economics, the management literature studies use firm wages relative to the industry as well as pay dispersion as measurements of employee performance and firm quality. Examples include Campbell, Ganco, Franco, and Agarwal (2011) and Carnahan, Campbell, and Agarwal (2012) who discuss that high paying firms retain top earners, but those who do leave select into entrepreneurship. These top earners typically have the most industry expertise and found successful ventures (e.g., Franco and Filson 2000).

Operating under the assumption that a firm's wages relative to the industry are reflective of firm quality, I estimate Equation 4 using mean wages in the firm as the firm characteristic of interest. Firm mean wages are tabulated as the average quarterly earnings of workers retained in the sample (Section 4) by SEIN-county. The variable is standardized to mean 0 and unit standard deviation. The industry-by-year fixed effect implies the coefficients on the interactions are estimated within each incumbent firm's industry and show how compensation differentials relative to the industry influence spinoffs. Pay dispersion in the firm is tabulated as the variance in average worker wages in the firm. This variable is similarly transformed to mean 0 and unit standard deviation.

The main wage interaction with the winning county dummy variable in Column 1 of Table 2.11 shows that firms with higher than industry average wages generate more entrepreneurs. A one standard deviation increase in relative wages increases the number of entrepreneurs generated by the average firm by .23, or of approximately a 10 percentage point increase above the baseline. Looking across all columns, the top row in Table 2.11 shows that firms with better pay generate more entrepreneurs selecting into upstream or anchor industry startups. The triple interactions in the remaining rows show whether the elasticity between wages and employee departures to entrepreneurship varies by supply chain industry. The excluded industry group in the specification is "all other industries". Isolating an individual row and comparing coefficients across columns shows whether firms in that industry grouping yield entrepreneurs disproportionately in certain industries. The rows of Table 2.11 show higher paying firms generate more entrepreneurs and particularly into their own supply chain category. Upstream (downstream) firms have significantly more entrepreneurs who form new ventures in upstream (downstream) industries than in anchor firm industries or downstream (upstream) industries. Consistent with the suggestive evidence in Table 2.9 and Table 2.10 that the number of employee departures from anchor industry incumbent is concentrated in a limited set of firms, Table 2.11 shows that the firms generating entrepreneurs tend to be higher paying firms.

Using the set of all eventual entrepreneurs in a county, I show that the probability of founding a supply chain firm is increasing in individual ability and industry experience. Table 2.12 estimates Equation 4 as a linear probability model. The main interaction in the first row shows that across all supply chain groupings, the probability a founder starts a firm in the supply chain in treated counties increases relative to founding a firm in all other industries. The triple interactions then show the heterogeneous effect of income and being in a winning county by supply chain industry. Row 2 in Table 2.11 Column 1 shows that a 1 standard deviation increase in wages of an entrepreneur from an upstream industry increases the relative likelihood of starting a firm in an upstream industry by 3%. The difference between this coefficient and the other coefficients related to the row is also statistically different, suggesting higher performing upstream entrepreneurs disproportionately select into

Table 2.11: Number of Employee Departures by Firm Pay

	(1)	(2)	(3)	(4)
	# of employees founding a firm	# of employees founding upstream firm	# of employees founding anchor industry firm	# of employees founding downstream firm
Win * Mean Wage	0.2324***	0.0100***	0.0016**	0.0014
	(0.0208)	(0.0010)	(0.0007)	(0.0009)
Win * Upstream * Mean Wage	-0.3181***	0.0975***	0.0448***	0.0267*
	(0.0813)	(0.0306)	(0.0134)	(0.0137)
Win * Anchor Industry * Mean Wage	-0.1029	0.2729***	0.2096***	0.1768**
	(0.2528)	(0.0800)	(0.0786)	(0.0740)
Win * Downstream * Mean Wage	-0.0503	0.0119	0.0409***	0.1651***
	(0.0555)	(0.0137)	(0.0114)	(0.0351)
Main Effects	Y	Y	Y	Y
Firm Controls	Y	Y	Y	Y
Industry x Year FE	Y	Y	Y	Y
Case FE	Y	Y	Y	Y
Adj. R-Sq.	0.6369	0.1186	0.0648	0.0851
N	1094000	1094000	1094000	1094000

Notes: * 0.1 ** 0.05 *** 0.01. High paying firms generate the most entrepreneurs across all industries (row 1), though the impact of wage is heterogeneous across industries (rows 2-4). The strongest effects are on spinoffs (e.g., Column 2 Row 2, Column 3 Row 3, and Column 4 Row 4). This table shows the impact firm wages have on employee departures by industry. The coefficients are interpreted as the relative elasticity of employee departures to firm wages by industry. Wages are normalized to have mean 0 and unit standard deviation. Firm wages are defined as average log quarterly earnings for all employees employed in the SEIN-county. Interactions are conducted after standardization to restore main effects. Regressions are weighted by total SEIN-fips employment. Standard errors are clustered at the SEIN-county.

upstream entrepreneurship. This pattern generalizes as well for entrepreneurs from anchor industries. However, wages do not appear to have a significant impact on the founding industry of entrepreneurs from downstream employers.

Firms with greater variance in compensation relative to their industry are thought to be better at rewarding top performers (e.g., Campbell et al., 2011). These top performers are more likely to stay in firms that appropriate rents from their skill, but conditional on leaving are also more likely to select into entrepreneurship. Table 2.13 corroborates this hypothesis and demonstrates that incumbent firms with greater pay dispersion generate more entrepreneurs even outside the supply chain. The evidence on wages aligns with the industry-specific analysis in Section 5.2.1 to support the human capital based hypothesis where the selection of high ability individuals with relevant industry experience in high quality firms into entrepreneurship drives improved startup performance after a local shock to supplier-buyer markets.

The complication related to pre-existing differences driving the results from Sec-

Table 2.12: Likelihood Entrepreneur Selects into Industry by Wage

	(1)	(2)	(3)
	Pr(Found upstream firm)	Pr(Found anchor industry firm)	Pr(Found downstream firm)
Win * Wage	0.0047***	0.0017***	0.0022***
	(0.0003)	(0.0002)	(0.0003)
Win * Upstream * Wage	0.0305***	0.0115***	0.0071*
	(0.0087)	(0.0036)	(0.0037)
Win * Anchor Industry * Wage	0.0342***	0.0619***	0.0443***
	(0.0095)	(0.0108)	(0.0126)
Win * Downstream * Wage	-0.0097**	0.0043	0.0111
	(0.0043)	(0.0041)	(0.0097)
Main Effects	Y	Y	Y
Worker Controls	Y	Y	Y
Industry x Year FE	Y	Y	Y
Case FE	Y	Y	Y
Adj. R-Sq.	0.1075	0.0626	0.1519
N	2124000	2124000	2124000

Notes: * 0.1 ** 0.05 *** 0.01. Entrepreneurs in winning counties tend to be higher wage employees relative to their firm of prior employment. This table shows the impact firm wages have on employee departures by industry. The coefficients are interpreted as the relative elasticity of supply chain entrepreneurship to individual wage by industry. Wages are normalized to have mean 0 and unit standard deviation. Firm wages are defined as average log quarterly earnings the eventual entrepreneur's most recent SEIN-county of employment prior to the anchor firm's arrival. Interactions are conducted after standardization to restore main effects. Controls include age, age-squared, sex, race, education, and SEIN FEs. Regressions are unweighted. Standard errors are clustered at the SEIN-county.

Table 2.13: Number of Employee Departures by Firm Pay Dispersion

	(1)	(2)	(3)	(4)
	# of employees founding a firm	# of employees founding upstream firm	# of employees founding anchor industry firm	# of employees founding downstream firm
Win * Wage Dispersion	0.4239***	0.0121***	0.0026***	0.0010
	(0.0360)	(0.0009)	(0.0005)	(0.0011)
Win * Upstream * Wage Dispersion	-0.3974***	0.1124***	0.0390***	0.0282**
	(0.0760)	(0.0282)	(0.0136)	(0.0137)
Win * Anchor Industry * Wage Dispersion	0.5206**	0.4151***	0.4034***	0.2764***
	(0.2309)	(0.0983)	(0.0961)	(0.0927)
Win * Downstream * Wage Dispersion	-0.1233*	0.0159	0.0298***	0.1947***
	(0.0694)	(0.0133)	(0.0112)	(0.0360)
Main Effects	Y	Y	Y	Y
Firm Controls	Y	Y	Y	Y
Industry x Year FE	Y	Y	Y	Y
Case FE	Y	Y	Y	Y
Adj. R-Sq.	0.6375	0.1237	0.0716	0.0868
N	1094000	1094000	1094000	1094000

Notes: * 0.1 ** 0.05 *** 0.01. Firms with greater pay dispersion lose more employees to spinoffs. This table shows the impact firm pay dispersion has on employee departures by industry. The coefficients are interpreted as the relative elasticity of employee departures to firm wage dispersion by industry. Firm wage dispersion is defined as the variance of log quarterly earnings for all employees employed in the SEIN-fips. Variance is normalized to have mean 0 and unit standard deviation. Interactions are conducted after standardization to restore main effects. Regressions are weighted by total SEIN-fips employment. Standard errors are clustered at the SEIN-county.

tion 5.2.1 is also relevant for this section's estimation of Equation 4. To account for this, I compare the profile of spinoffs and entrepreneurs across eventual supply chain industries 8 through 2 years prior to the anchor firm's arrival to the county. I do not find evidence to suggest my results are driven by pre-existing difference in spinoffs between winning and runners-up counties.

Employer's Age and Size

Alternative mechanisms may relate to why aggregate employment does not increase, but instead reallocates to young firms. One such mechanism is that employees depart bureaucratic firms that are unable to take advantage of new opportunities that arise from the anchor firm's opening. A number of studies such as Klepper (2007), Buenstorf and Klepper (2009), and Klepper and Thompson (2010) describe a process where employees who are unable to undertake projects within bureaucratic firms spin-out and pursue projects on their own. Canonical examples of spinoffs through employee disagreements include the origins of the Detroit automobile cluster and the Akron tire industry (Klepper, 2007; Buenstorf and Klepper, 2009). On the contrary, entrepreneurs may take on projects complementary to their former employer (e.g., Oettl and Agrawal, 2008). This hypothesis could also be reflected by a business stealing effect. If spinoff employees are also former employees of old large incumbent firms then this could serve as evidence that bureaucratic firms are unable to respond to local supply chain shocks. To test the bureaucracy hypothesis I using size and age as proxies for bureaucracy. I find only inconclusive results, though age and size are crude measures of a firm's bureaucracy.[36]

Large and bureaucratic firms may be less able to undertake new opportunities that arise. Their employees may then spinoff to pursue these projects. This process of the clustering of entrepreneurship has been used to explain the origins of the Detroit

[36]Testing for business stealing is out of the scope of this paper.

automobile cluster and the Akron tire industry (Klepper 2007/10). The microdata do not allow for directly testing for firm bureaucracy specifically or through anecdotal evidence, but do allow for testing the relationship between firm characteristics that may correlate strongly with an incumbent firm's inability to shift and adjust to sudden opportunities.

Tables 2.14 and 2.15 test firm age and firm size. The main interaction shown in Table 2.14 shows that older firms in winning counties generate more entrepreneurs who start firms upstream and anchor industries, though this effect is not particularly economically significant. The triple interactions show the effect does not vary by industry grouping and in some cases older firms spawn more entrepreneurs. Table 2.15 tests the potential firm size dimension using the linear probability framework and similarly demonstrates mixed evidence on the role of an entrepreneur's prior employer's size and selection into supply chain entrepreneurship.

Table 2.14: Number of Employee Departures by Firm Age

	(1)	(2)	(3)	(4)
	# of employees founding a firm	# of employees founding upstream firm	# of employees founding anchor industry firm	# of employees founding downstream firm
Win * Firm Age	0.0056**	0.0009***	0.0003***	0.0002*
	(0.0028)	(0.0001)	(0.0001)	(0.0001)
Win * Upstream * Firm Age	-0.0196	-0.0004	-0.0029**	-0.0037***
	(0.0120)	(0.0032)	(0.0013)	(0.0014)
Win * Anchor Industry * Firm Age	-0.0336	-0.0126	-0.0080	-0.0144*
	(0.0397)	(0.0081)	(0.0081)	(0.0079)
Win * Downstream * Firm Age	0.0119	-0.0024	-0.0055**	0.0022
	(0.0126)	(0.0028)	(0.0022)	(0.0056)
Main Effects	Y	Y	Y	Y
Industry x Year FE	Y	Y	Y	Y
Case FE	Y	Y	Y	Y
Adj. R-Sq.	0.6368	0.1152	0.0604	0.0839
N	1094000	1094000	1094000	1094000

Notes: * 0.1 ** 0.05 *** 0.01. It is unclear whether older firms that may be more bureaucratic lose more employees to spinoffs. This table shows the impact firm age has on employee departures by industry. Firm age is calculated in the LBD as the age of the oldest establishment when the firmid first appears. The firmid then ages naturally. The LBD age measure is then mapped to the LEHD using the firmid-county-EIN-SEIN concordance in the ECF. Interactions are conducted after standardization to restore main effects. Regressions are unweighted. Standard errors are clustered at the SEIN-county.

The mixed evidence related to incumbent firm age and size does not necessarily negate an employee disagreements with employer explanation for the increase

Table 2.15: Likelihood of Departure by Firm Size

	(1)	(2)	(3)
	Pr(Found upstream firm)	Pr(Found anchor industry firm)	Pr(Found downstream firm)
Win * Firm Size	0.0014**	0.0005**	0.0008
	(0.0006)	(0.0002)	(0.0005)
Win * Upstream * Firm Size	-0.0136	0.2039***	0.2384***
	(0.1275)	(0.0309)	(0.0339)
Win * Anchor Industry * Firm Size	0.0434	-0.1837***	0.4800***
	(0.1116)	(0.0522)	(0.1239)
Win * Downstream * Firm Size	0.1063	0.0795**	-0.5916***
	(0.0827)	(0.0339)	(0.1126)
Main Effects	Y	Y	Y
Worker Controls	Y	Y	Y
Industry x Year FE	Y	Y	Y
Case FE	Y	Y	Y
Adj. R-Sq.	0.1078	0.0612	0.1530
N	2124000	2124000	2124000

Notes: * 0.1 ** 0.05 *** 0.01. The evidence on the impact of firm size on the likelihood an individual departs into supply chain entrepreneurship is unclear. This table shows the impact of firm size on employee departures by industry. Firm size is calculated in the LEHD sample. Size reflects the number of employees who worked in the firm during the year with an average wage above minimum wage in all quarters worked. Interactions are conducted after standardization to restore main effects. Notes: * 0.1 ** 0.05 *** 0.01. Entrepreneurs in winning counties tend to be higher wage employees relative to their firm of prior employment. This table shows the impact firm wages have on employee departures by industry. The coefficients are interpreted as the relative elasticity of supply chain entrepreneurship to individual wage by industry. Wages are normalized to have mean 0 and unit standard deviation. Interactions are conducted after standardization to restore main effects. Controls include age, age-squared, sex, race, education, and SEIN FEs. Regressions are unweighted. Standard errors are clustered at the SEIN-county.

in entrepreneurship through spinoffs. Without any contextual evidence, the generic measures of age and size are crude and may mask underlining heterogeneity even within those characteristics. The results in Sections 5.2.1, 5.2.2, and 5.2.3, however, remain consistent with a view supporting the general role of proximity to buyer and supplier markets for entrepreneurial formation and performance alongside individuals with location and industry specific human capital being most able to capture gains to local industry shocks. Future research should probe at understanding whether incumbent firms face a business stealing effect from employee departures or if departing entrepreneurs form businesses in the same supply chain category, but complementary or non-competing product markets.

2.5.3 Within-Firm Wage and Occupation Effects

I examine distributional impacts by taking the set of winning counties and measuring the treatment effect of the anchor firm's arrival within the set of winning counties. Specifically, I compare the characteristics of spinoffs that form between years $t-8$ and $t-2$ with those that form in the defined post anchor firm period. Whereas prior analyses focused on the direct effects of the anchor firm's arrival on employee departures from incumbent firms and the success of those departing, this section turns towards policy implications of anchor firms arriving to a county. The key features are that the introduction of an anchor firm does not increase aggregate employment in supply chain industries, but simply employment in young firms. The young firms that form in a counting become much more likely to be founded by supply chain entrepreneurs from top paying firms.

Table 2.16 shows that incumbent firms in winning counties lose more employees to entrepreneurship after the anchor firm arrives. Consistent with Section 5.2.1, entrepreneurs become more likely to come from supply chain related industries that the set of all other industries. Within the supply chain, the composition of firms who generate entrepreneurs in the supply chain also shifts towards incumbent supply chain industries. Upstream incumbents generate .18 more spinning out employees than prior to the opening, an effect that represents about a 30% increase above the baseline spinoff rate or two spinoffs for every 10 business units in the county. Downstream incumbents similarly generate 4 new spinoffs for every 10 extant business units. Table 2.16 shows that the comparisons between winning and runners-up counties on the dimension of industry background are driven by changes in spinoff composition within winning counties. Though currently suppressed, entrepreneur-level regressions show qualitatively similar and strong effects of founder experience on selection into supply chain entrepreneurship.

Table 2.16: Change in Departures by Parent Firm Industry

	(1)	(2)	(3)	(4)
	# of employees founding a firm	# of employees founding upstream firm	# of employees founding anchor industry firm	# of employees founding downstream firm
Post * Upstream	-0.0389	0.1797***	0.0447***	0.0371***
	(0.0858)	(0.0263)	(0.0122)	(0.0124)
Post * Anchor Industry	1.294***	0.4705***	0.5257***	0.2959***
	(0.2601)	(0.1082)	(0.1004)	(0.0998)
Post * Downstream	0.3380***	0.1035***	0.0880***	0.3918***
	(0.1094)	(0.0281)	(0.0235)	(0.0550)
Main Effects	Y	Y	Y	Y
Firm Controls	Y	Y	Y	Y
Industry x Year FE	Y	Y	Y	Y
Case FE	Y	Y	Y	Y
Adj. R-Sq.	0.5483	0.1611	0.0793	0.0932
N	332000	332000	332000	332000

Notes: * 0.1 ** 0.05 *** 0.01. The anchor firm's arrival induces spinoffs across the supply chain. Incumbent supply chain firms lose employees to entrepreneurship in the supply chain, but the strongest effects are on spinoff entrepreneurship. For example, upstream firms lose more employees to upstream entrepreneurship (row 1). This table displays coefficients of interest from estimating Equation 3 restricted to the set of winning counties and replacing the Win dummy variable with a $Post$ dummy variable. This table tests whether the industries that yield spinoffs after the anchor firm arrives changes in winning counties. The sample includes only those employees whose average quarterly earnings across quarters worked in the firm exceed $2,678 which represents earnings for a minimum wage employee working 40 hours for 13 weeks. Wages are then inflated to 2015 dollar amounts. The regressions include all main effects and interactions shown in Equation 3, but the table displays only key coefficients of interest. Firm controls include controlling for firm size and firm age. Industry by year fixed effects control for secular trends in startup formation that vary by industry. Case fixed effects ensure that comparisons are made between firms within each anchor firm event. The case identifier is interacted with a dummy variables taking the value of 1 if the year is between 5 years prior to the anchor firm event through 7 years after. This facilitates identification of the industry year fixed effects while ensuring the time periods where loading on the difference-in-differences estimator is bounded.

A number of recent studies have examined the causes and effects of increased sorting of high ability workers into top paying firms (e.g., Barth et al., 2006; Card et al., 2018; Song et al. 2019). This study considers the role top paying firms have on incubating employees who become primed to seize opportunities that suddenly arise perhaps through observing shocks that arise in their local market, but outside their direct industry. Though currently undisclosed, loading on the difference-in-differences estimator is largely driven by the earliest entry cohorts–employees inside the county with substantial industry experience prior to the anchor firm's arrival. These individuals are in fact top earners in high quality firms which may further suggest that sorting does not only improve labor outcomes within firm-worker matches, but also through a channel whereby top performing employees are more able to seize an outside option of entrepreneurship rather than wage-earning employment.

Table 2.17: Change in Departures by Pay of Parent Firms

	(1)	(2)	(3)	(4)
	# of employees founding a firm	# of employees founding upstream firm	# of employees founding anchor industry firm	# of employees founding downstream firm
Post * Mean Wage	0.2645***	0.0092***	0.0014*	0.0021**
	(0.0130)	(0.0012)	(0.0008)	(0.0008)
Post * Upstream * Mean Wage	-0.0801	0.1127***	0.0785***	0.0678***
	(0.0605)	(0.0288)	(0.0171)	(0.0173)
Post * Anchor Industry * Mean Wage	-0.0432	0.2072**	0.1866**	0.1842**
	(0.2265)	(0.0826)	(0.0785)	(0.0731)
Post * Downstream * Mean Wage	-0.0634	0.0304	0.0509***	0.1699***
	(0.0797)	(0.0198)	(0.0154)	(0.0378)
Main Effects	Y	Y	Y	Y
Firm Controls	Y	Y	Y	Y
Industry x Year FE	Y	Y	Y	Y
Case FE	Y	Y	Y	Y
Adj. R-Sq.	0.5492	0.1640	0.0867	0.0988
N	332000	332000	332000	332000

Notes: * 0.1 ** 0.05 *** 0.01. Firms that pay employees more than their industry competitors generate more entrepreneurs after the anchor firm's arrival to the county. The magnitude of the effects varies by industry with anchor industry incumbents and downstream incumbents losing the most. This table shows coefficients of interest from estimating Equation 4 restricted to the set of winning counties and replacing the Win dummy variable with a $Post$ dummy variable. This specification tests whether the elasticity of spinoffs to firm mean wages varies by industry. Mean wages is measured as the average worker's average quarterly earnings in a SEIN. The sample includes only those employees whose average quarterly earnings across quarters worked in the firm exceed $2,678 which represents earnings for a minimum wage employee working 40 hours for 13 weeks. Wages are then inflated to 2015 dollar amounts. The mean firm wage is transformed to have mean 0 and unit standard deviation. The regressions include all main effects and interactions shown in Equation 4, but the table displays only key coefficients of interest. Firm controls include controlling for firm size and firm age. Industry by year fixed effects control for secular trends in startup formation that vary by industry and imply that firm wages are relative to other firms in their own industry. Case fixed effects ensure that comparisons are made between firms within each anchor firm event. The case identifier is interacted with a dummy variables taking the value of 1 if the year is between 5 years prior to the anchor firm event through 7 years after. This facilitates identification of the industry year fixed effects while ensuring the time periods where loading on the difference-in-differences estimator is bounded.

I show that the composition of supply chain entrepreneurs changes when an anchor firm arrives by the growth rate of wages in the firms economically proximate industries to the anchor firm. These employees are typically above the mean wage-earners in their firm, are not considered to be the top wage earners in their firms. Table 2.17 shows that high wage growth firms increase the number of employee departures into entrepreneurship. Table 2.18 shows that supply chain industries also strongly correlate with occupational similarity. The table shows that the correlation across all industries represents a shift in the earnings profile of entrepreneurs in the county. Table 2.18 follows Table 2.13 in demonstrating that firms in the supply chain that are considered to be better at compensating top performers within the firm do not necessarily yield more entrepreneurs overall but do generate more entrepreneurs who select into the supply chain. Entrepreneur-level results show that the main effect on the $Post \times Wage$ variable is positive and strongly significant on regressions estimating the likelihood of selecting into a supply chain industry. The main effect does not vary by industry of a founder's origin.

Table 2.18: Change in Departures by Pay Dispersion of Parent Firms

	(1)	(2)	(3)	(4)
	# of employees founding a firm	# of employees founding upstream firm	# of employees founding anchor industry firm	# of employees founding downstream firm
Post * Wage Dispersion	0.5397***	0.0131***	0.0019***	0.0029***
	(0.0178)	(0.0011)	(0.0007)	(0.0008)
Post * Upstream * Wage Dispersion	-0.0834	0.1446***	0.0722***	0.0666***
	(0.0622)	(0.0274)	(0.0170)	(0.0171)
Post * Anchor Industry * Wage Dispersion	0.6352***	0.3992***	0.3881***	0.3445***
	(0.2327)	(0.1247)	(0.1169)	(0.1132)
Post * Downstream * Wage Dispersion	-0.1179	0.0384*	0.0588***	0.2348***
	(0.0738)	(0.0208)	(0.0161)	(0.0370)
Main Effects	Y	Y	Y	Y
Firm Controls	Y	Y	Y	Y
Industry x Year FE	Y	Y	Y	Y
Case FE	Y	Y	Y	Y
Adj. R-Sq.	0.5530	0.1695	0.0966	0.1063
N	332000	332000	332000	332000

Notes: * 0.1 ** 0.05 *** 0.01. The top row shows that incumbent firms with greater pay dispersion lose more employees to entrepreneurship across a broad set of industries. The main effect shows that a one standard deviation increase in wage dispersion leads to approximately a 30 percentage point increase in the number of employee departures. This magnitude is similar for changes in the number of employees who spinoff from upstream, anchor industry, and downstream firms. This table shows coefficients of interest from estimating Equation 4 restricted to the set of winning counties and replacing the Win dummy variable with a $Post$ dummy variable. Wage dispersion is measured as the variance in SEIN's average worker quarterly wages. The sample includes only those employees whose average quarterly earnings across quarters worked in the firm exceed $2,678 which represents earnings for a minimum wage employee working 40 hours for 13 weeks. Wages in this cleansed sample are inflated to 2015 dollar amounts. The variance in wages is transformed to have mean 0 and unit standard deviation. The regressions include all main effects and interactions shown in Equation 4, but the table displays only key coefficients of interest. Firm controls include controlling for firm size and firm age. Industry by year fixed effects control for secular trends in startup formation that vary by industry. Case fixed effects ensure that comparisons are made between firms within each anchor firm event. The case identifier is interacted with a dummy variables taking the value of 1 if the year is between 5 years prior to the anchor firm event through 7 years after. This facilitates identification of the industry year fixed effects while ensuring the time periods where loading on the difference-in-differences estimator is bounded.

This paper posits that this result corresponds to layer shedding in the firm (Caliendo, Monte, and Rossi-Hansberg 2015).[37] Strikingly, the results continue to show that the effects are specific to supply chain firms who disproportionately contribute to the number of employees spinning out to found other supply chain firms. Tables 2.17 and 2.18 fix the log wage pre-anchor arrival and display log wages after the anchor firm arrives. This means that the coefficients represent a log wage growth relative to pre-arrival. The first row of Table 2.17 shows that the combination of wage growth

[37] Greenstone, Hornbeck, and Moretti 2010 show that wages increase in industries treated by the anchor firm.

firms generating more entrepreneurs as well as overall industry employment remaining constant suggests a reallocation of workers across firms. If workers are displaced from their jobs, it may be the case that they simply transition into former occupations in industries that disproportionately hire those same occupations.[38] Lazear and Oyer (2004) and Lazear and Shaw (2007) discuss that wage pressures on firms is largely driven by external local industry factors. Combined with Caliendo et al., given that compensation is a key component of employee retention across occupational sectors, growing firms that shed mid-management layers may also be those that generate the most employees who select into firms that complement their own skills.

I use the Bureau of Labor Statistics (BLS) Occupational Employment Statistics (OES) to provide descriptive evidence that supply chain industries are also occupationally similar industries. The OES creates occupational codes that are constant across industries and provides the number of employees that are accounted for in each occupation. I tabulate the percentage of industry employees in each occupation and compare the occupational share of industry across anchor and supply chain industry. Table 2.19 uses the industry pairs in Table 2.3 and 2.5 to show the pairwise correlations between occupational shares of those industries.

The table shows that supply chain linked industries display stronger occupational similarity than all other industry pairs. The most intensely related anchor firms and paired industries have a pairwise occupational share correlation of 0.2894. This correlation is significantly different from the correlation between the full set of industries (0.0598). As displayed, many of the paired supply chain industries are within the same 3-digit NAICS sector which may mechanically yield stronger correlations by virtue of being already classified as similar industries. The correlation between industries of the same NAICS-3 sector is 0.4891 which either implies a mechanical

[38]Firms facing the greatest exogenous shock to wages are also in the most high-tech industries where the supply of labor may be relatively inelastic by occupation.

Table 2.19: Occupational Similarity for Top Supply Chain Pairs

Announcement industry	I/O industry	Occupational correlation
3222. Converted Paper Product Manufacturing	3221. Pulp, Paper, and Paperboard Mills	0.9396
3361. Motor Vehicle Manufacturing	3362. Motor Vehicle Body and Trailer Manufacturing	0.9228
3361. Motor Vehicle Manufacturing	3363. Motor Vehicle Parts Manufacturing	0.9054
5250. Funds, Trusts, and Other Financial Vehicles	5239. Other Financial Investment Activities	0.8008
3252. Resin, Synthetic Rubber, and Artificial Synthetic Fibers and Filaments Manufacturing	3251. Basic Chemical Manufacturing	0.6363
4830. Water Transportation	3366. Ship and Boat Building	0.5119
3252. Resin, Synthetic Rubber, and Artificial Synthetic Fibers and Filaments Manufacturing	3261. Plastics Product Manufacturing	0.4158
3241. Petroleum and Coal Products Manufacturing	2110. Oil and gas extraction	0.3469
3116. Animal Slaughtering and Processing	1123. Poultry and egg production	0.1993

Notes: The table shows similarity of occupational composition in the leading anchor industry supply chain pairs in terms of percentage of all input/output to/from the paired industry. Occupational composition is tabulated using the Bureau of Labor Statistics' Occupational Employment Statistics (OES). The OES provide occupational codes and the share of industry employment in each code. This table displays between industry correlations of the percentage of industry workers employed in an occupation. The table shows that the share of an industry's shipments to/from an industry correlate strongly with the occupationsal mix in those industries. Overall, the correlation of occupational shares between anchor firm industries and their supply chain industries is 0.2894. By comparison, the correlation of occupational mix in anchor industries and the full set of industries is 0.0598. Predictably industries within the same sector have similar occupational mixes as well. Within the same 3-digit NAICS designations, the correlation is 0.4891. Within 3-digit NAICS and within the supply chain the correlation is 0.6673. The table excludes shipments from an industry to itself.

relationship or shows that even within the set of neighboring industries, supply chain industries also share a common occupational component. Taking the set of 3-digit NAICS industries, the occupational composition between the anchor firm and supply chain industries is a significantly stronger correlation of 0.6673. These correlations suggest that the proper interpretation of the upstream and downstream designations combined with the wage effects suggest that labor is shared between *economically proximate* industries in the anchor industries.

Table 2.20 shows the most correlated industry supply chain pairs in the anchor firm sample. The table combined with Table 2.1 shows that the most represented industries on which most loading occurs in the difference-in-differences estimator is by anchor industry pairings most occupationally similar.

These findings together reveal that the arrival of an anchor firm does not result in an aggregate rise in entrepreneurship in a county through a discovery of new opportunities. Firm entry is specific to economically proximate, as measured by input-output intensity and occupational similarity, and by workers who seem to transition between those industries when pushed out of incumbent firms. Entry within the supply chain is not balanced across individual characteristics such as worker skills or experience, but may be driven by differences in firm-specific effects as evidenced by

Table 2.20: Leading Supply Chain Pairs by Correlation

Announcement industry	I/O industry	Occupational correlation
3342. Communications Equipment Manufacturing	3345. Navigational, Measuring, Electromedical, and Control Instruments Manufacturing	0.9716
3369. Other Transportation Equipment Manufacturing	3362. Motor Vehicle Body and Trailer Manufacturing	0.9639
3119. Roasted Nuts and Peanut Butter Manufacturing	3114. Fruit and Vegetable Preserving and Specialty Food Manufacturing	0.9467
3115. Dairy Product Manufacturing	3114. Fruit and Vegetable Preserving and Specialty Food Manufacturing	0.9400
3221. Pulp Mills	3222. Converted Paper Product Manufacturing	0.9400
3222. Converted Paper Product Manufacturing	3221. Pulp Mills	0.9400
3369. Other Transportation Equipment Manufacturing	3363. Motor Vehicle Parts Manufacturing	0.9348
3115. Dairy Product Manufacturing	3119. Roasted Nuts and Peanut Butter Manufacturing	0.9284
3115. Dairy Product Manufacturing	3112. Grain and Oilseed Milling	0.9283
3361. Leather and Hide Tanning and Finishing	3362. Motor Vehicle Body and Trailer Manufacturing	0.9228

Notes: See Table 2.19. The table displays the Top 10 correlations between the anchor industry and any one of its supply chain pairs.

wages and personnel decisions firms make after the anchor firm arrives. Additional work, however, may provide further evidence on the differences between effects on entrepreneurial opportunity and a layer shedding effect. This paper provides evidence that extant firms face rising factor prices through labor market channels. However, this paper does not test interactions between firm size and firm age, nor do I explicitly examine the extent to which departing founders employ lower wage employees from their prior employer.

2.6 Robustness

2.6.1 Placebo Tests

A more detailed explanation of placebo checks and results on time periods where loading on the difference-in-differences estimators occur are detailed in Online Appendix B.[39] The main checks testing the validity of the empirical strategy related to Section 5.1 include directly testing for parallel trends using a dynamic difference-in-differences estimation and conducting placebo tests.

The parallel trends test uses time interactions for each year between five years prior to the event and seven years post event. In these regressions, there is no evidence to reject the null hypothesis that the coefficients $win \times time$ interactions for periods

[39] Forthcoming

prior to the announcements are 0. The dynamic panel also allows for computing the specific time periods in the post-event interval where most of the loading occurs on the difference-in-differences estimators. I find that the first three years of entry cohorts and founders drive the impacts in new venture creation and performance. This also tends to be the period during which top earners in local incumbent firms spinout from their employers.

Placebo tests ensure that the coefficients of interest on the difference-in-differences estimators and interactions are driven by the treatment of the specific case's winning county and its counterfactual county. The first placebo test randomizes the treated county within each set of case counties.[40] A second set of placebo test randomizes the treatment industry within each case's contest counties.[41] The lack of significance on coefficients of interest support that the differences are driven by the anchor firm's arrival in the winning county.

The analysis on spinoffs from incumbent firms and entrepreneurs compares subsequent outcomes of spinoffs and entrepreneurs in the winning and runners-up counties in year $t-2$ over the first 7 years of the anchor firm's arrival to the county. The spinoff process may be inherently different between winning and runners-up counties and the interactions of the variables of interest with the Win dummy variable may reflect these innate differences. Using work histories and spinoffs from years prior to the anchor firm's arrival to a county, I do not find evidence there are inherent differences between the counties *ex ante* in terms of selection into spinoffs or entrepreneurship.

[40] Suppose Case 1 involves counties A and B. The placebo tests randomly selects either A or B as the winning county. This randomization is repeated over all cases.

[41] For each case's counties, the placebo test randomizes the industry the "anchor firm" opens, thereby creating a fictitious opening event.

2.6.2 Subsidies

My identification strategy assumes that profit maximizing firms select locations that minimize operating costs associated with the its production function. Therefore, the anchor firm's shortlisted sites is assumed to reflect locations that deliver equivalent operating costs with respect to its production function. However, subsidies may allow a firm to consider a location with lower profitability if the opportunity cost is compensated by the locale using tax incentives and other support mechanisms.

The prevalence of subsidies is non-trivial. During 2015, state and local governments spent an estimated $45 billion in incentive programs targeting specific companies (Bartik, 2019). Using GoodJobsFirst and information in *Site Selection* magazine, I collect data on subsidy programs relevant to the anchor firm location choices in my sample.[42] However, obtaining information on bids the counterfactual counties submitted is difficult because these locales do not disburse funds and avoid disclosure to media and government spending watchdog organizations. In cases where magazine articles or the company's own public statements describe the bidding process for counterfactual counties, I record the subsidy types and amounts.

I find that subsidies are almost universal in my sample, but both winning and runners-up counties offer similar packages with the firm's ultimate location choice driven by idiosyncratic preferences. Subsidies involved range from less than $100,000 in public support (e.g., Kinko's in Dallas County, Texas) to over $1 billion (e.g., Nissan in Madison County, Mississippi). The size of the incentive programs increase in the capital intensity of the project. Aircraft and auto manufacturing facilities command the largest subsidies, though counterfactual counties offer competitive packages. Hyundai received nearly $120,000,000 from state and local governments in Alabama to facilitate its location choice, but received an offer of $123,000,000 from Kentucky.

[42] Collecting complete subsidy information is a work in progress. I currently have information on a random sample of 40 cases.

Hyundai selected the Alabama site because Kentucky requested a project design four months before it was ready to and the C.E.O. "in an unusually strongly worded comparison of the two contenders... mentioned average temperatures" as being more favorable in the Alabama site. In research and development facilities, incentives appear to play a less important role relative to the presence of universities, existing corporate labs, and a skill-specific labor pool. Companies opening these types of facilities even refuse incentive packages. Novartis' decision to locate its main research office in Cambridge, MA over San Diego, CA was notably absent any subsidies and the firm "reportedly turned down... offers of assistance" from state officials.

Other location-specific factors unobserved to the econometrician that simultaneously affect the entrepreneurial environment may also drive the anchor firms' site search. Firms may consider sites where they project to have greater influence over future laws and regulations they face. Political economy considerations aside, firms may also anticipate follow-on activity by large established suppliers or customers who will collocate. If this is true, then my results on entrepreneurship may be responses to follow-on anchor firms and not the initial firm I attribute my results to. I document notable cases where large manufactures attracted its own large suppliers and customers to a location as well (Mobis in Alabama). However, my results for entrepreneurship are driven by the earliest years after the anchor firm locates to a county whereas increases in the number of large firms and establishments does not occur until the last years of a case's time horizon.

2.6.3 Identifying Spinoffs in the LEHD

In the final LEHD sample, I identify an entrepreneur as a top 3 wage earner in a new *SEIN* associated with a *firmid* aged less than three years of age.[43] Firm

[43]Wages are defined by the average quarterly wage associated with a Personal Identification Key (PIK) in non-zero wage quarters associated with a SEIN-SEINUNIT. The sample excludes jobs for which the average quarterly wage is less than full-time minimum wage ($5.15 * 40 hours/week * 13

characteristics on SEINs are obtained by merging establishment-level data in the Longitudinal Business Database summed to the federal EIN-county level since this is the narrowest observational unit common to the LBD and LEHD. Matching on EIN-county and importing age based on the *firmid*'s age, I obtain a list of matched SEIN-counties. I take SEIN-counties associated with SEINs that have never appeared before in the data as new firms.

Spinoffs are formed when a new SEIN operates in the same 4-digit NAICS code as the founder's most recent employer. Parent firm and spinoff size are tabulated using the Employer Characteristics File in the LEHD with robustness checks using the number of individuals in the cleansed LEHD sample. New firms are defined as firms aged less than three years as in the Longitudinal Business Database cleansing procedure. Because a firm in year $t-1$ could appear as a "startup" in year $t+2$ in terms of age, I require a founder to operate in a different SEIN post-announcement than in the pre-period.

2.6.4 Alternative Measures of Supply Chains

The current measure of supply chains does not include all value enhancing activities associated with a focal industry. In robustness checks, I use NAICS-4 industries with the same NAICS-3 codes, but different NAICS-4 designations to signify "adjacent industries". Often, these industries are complementary to the focal NAICS-4 industry. My results are qualitatively similar, if not stronger, to including adjacent industries as the set of supply chain industries. This alternative measure similarly does not capture all value enhancing industries. For example, video games (software) would not be included as part of the supply chain of a video game console (manufacture) firm. However, constructing a complete set of complementary activities will require a more contextualized and refined industry analysis for future research.

weeks). After cleansing, wages are adjusted to 2015 dollar amounts.

An anchor firm's industry may disproportionately source inputs or ship outputs to a startup's industry, but the startup's dependence may be relatively small. To account for this, I include interaction terms in Equation 1 for the share of a startup firm's industry inputs (output) from (to) the anchor firm industry. I do not find evidence of heterogeneity within the set of upstream and downstream firms by dependence of the startup industry. In other robustness checks, I include all industries without defining a subset of upstream and downstream industries and a continuous measure of input/output shares to and from industries as an interaction with the treatment variable. I find results qualitatively similar to the main findings on the importance of input-output linkages in startup formation and spinoffs.

Glaeser and Kerr (2009) provide alternative continuous measure that describe the availability of inputs and the local market size for output. The input measure sums the difference between the share of an industry i input from industry k ($Input_{i \leftarrow k}$) and the share of the county's total employment in industry k.

$$Input_{ict} = -\sum_{k=1,\ldots,K} \left| Input_{i \leftarrow k} - \frac{E_{kct}}{E_{ct}} \right|, \tag{2.7}$$

Larger values (closer to 0) correspond to the county's employment across industries being in proportion to the share of total inputs used by a focal industry. The output measure is constructed to capture the ability of an industry to sell locally.

$$Output_{ict} = \Big[\sum_{k=1,\ldots,K} Output_{i \rightarrow k} \frac{E_{kct}}{E_{kt}} \Big] \Big[\sum_{k=1,\ldots,K} Output_{\cdot \rightarrow k} \frac{E_{kct}}{E_{kt}} \Big]^{-1} \tag{2.8}$$

As before, $Output_{i \rightarrow k}$ is the share of i's output to k. This term is multiplied by the share of industry k's employment in county c ($\frac{E_{kct}}{E_{kt}}$). The first bracketed term effectively interacts output intensity to an industry the degree of localization of that industry.[44] This interaction is scaled by the second bracketed term that measures the total size of the output market. I construct these measures for the year prior to the anchor firm's announcement and tabulate changes in the measure post-anchor firm arrival. These county-industry level measures are similar between winners and runners-up counties prior the anchor firm's announcement.

2.6.5 Loading on Time Periods and Events

A growing body of literature critiques event-study difference-in-differences designs. Borusyak and Jaravel (2017) and Goodman-Bacon (2018) discuss estimation issues when subjects are treated at different times. The variation in timing may lead to irregular and inconsistent weighting of events in the difference-in-differences estimator. I conduct preliminary checks on disproportionate loading by time periods and events. I find that loading on the difference-in-differences estimator is driven timing within each event and the industry of the event rather than the years of the individual events or counties themselves.

I test entry cohort effects to capture the persistence of the anchor firm's arrival shock over time. I construct entry cohorts as firms aged 0 in year $t-k$ or $t+k \ \forall \ k$ relative to the opening announcement. I compare startups in winners and runners-up county in terms of cumulative employment growth for each year through 5 years after the startup cohort enters. I estimate the following modified empirical response function:

$$\log\left(\frac{emp_{fjct}}{emp_{fjc0}}\right) = \alpha_0 + \theta_t \times Win_{jc} + \gamma_j + \gamma_c + \epsilon_{fijct} \tag{2.9}$$

for each $\theta_t \in (t, t+5]$. The strongest effects on employment growth in winning counties are found over the first three years of the startup's formations. This structure

[44] The degree of localization is somewhat reflective of how tradeable the sector's goods are (Ellison and Glaeser 1997).

also allows me to compare employment growth rates for startups formed in each year since $t-5$. I refer to each of startups formed in a year t as an entry cohort. The findings suggest the entry cohorts in the first four years after the anchor firm's arrival announcement drive the startup activity documented in the main findings of the paper. I interpret this result to lend credibility to the importance of localized knowledge by specialists already in the county. After several years elapse, general equilibrium effects may dominate and moderate subsequent entrepreneurial activity. I leave these longer-term adjustments in the county to future research.

In other checks, I group events by timing (e.g., announcements between 1990 and 1994, 1995-1999, 2000-2004, etc.) as well as industry groupings (e.g., Bureau of Labor Statistics defined high-technology categories). The strength of startup formation, entry size, and employment growth effects are qualitatively similar across timing groups, but not industry groupings. Events associated with the high-tech sector appear to have the strongest effects in a manner consistent with literature on local employment multipliers Moretti (2010).

2.7 Conclusion

I use the geographic expansion of a large firm through openings of large establishments in new counties as a shock to geographic proximity of potential buyers and suppliers to local firms and potential entrepreneurs. This paper shows that the process of spinoffs can be driven by shocks outside an incumbent firm's industry. In particular, a shock to the firm's local supply chain either motivates employees to depart their employer or forces extant firms to streamline and shed layers in the face of rising labor costs. Firms in the anchor firm's industry as well as industries that purchase from or supply to the anchor's industry face the sharpest rise in wages. Incumbent firms already in a county-industry pair of the anchor's supply chain lose employees to entrepreneurship either because these employees perceive a new opportunity due to the anchor firm's proximity or because incumbent firms reorganize. These new entrepreneurs disproportionately found new ventures in the same industry as their former employer, suggesting a role of local industry experience. Incumbent firms in industries outside the anchor firm's supply chain and who do not face rising labor costs do not experience an increase in employee departures to entrepreneurship. The overall startup entry rate does not increase in the county, but instead the anchor firm's arrival alters the industry composition of entrepreneurship.

My findings are consistent with theories of startup performance in clusters driven by both proximity to buyers and suppliers as well as specialized human capital. These forces dictate which employees are primed to seize opportunities that arise by a sizable shock through the arrival of a firm upstream or downstream to their employer. Founders with supply chain experience disproportionately enter entrepreneurship in the supply chain and these effects are magnified for higher wage employees from high paying firms. These large local industry shocks, however, do not induce a mass of entry by a broad set of individuals with diverse experience or through migration. Instead, individuals from firms most affected by rises in local labor costs are most likely to depart the firm and found their own venture. Whether these individuals are explicitly capturing an opportunity incumbents are unable to capture or are being pushed out of incumbent firms and into smaller, young firms of their own is less clear. However, the rise in supply chain wages and stability in employment support predictions from a layer shedding view of incumbent firms. Because these departing workers have local industry knowledge, they are more likely to work in firms that best utilize that knowledge. Therefore, it is not surprising that my effects are strongest in industries that require the most specialized knowledge and where labor supply may be relatively more inelastic.

This paper connects the literature describing the clustering of input-output industries to literature on management practices and productivity spillovers from anchor firms. Ellison, Glaeser, and Kerr (2010) show that coagglomeration patterns are particularly strong for industries that purchase from and supply to each other. While I show a clustering of new firms about an anchor firm are linked through input-output industries, startup formation most likely occurs because rising wages and labor costs in industries most directly affected by the anchor firm's arrival are forced to streamline operations and potentially shed extra layers of management. These workers then use their industry-specific skill to start firms in the same industry as their employers or industries that complement their accumulated skills. This process is particularly strong in more knowledge intensive industries such as RD services and high-tech manufacturing where local labor supply is relatively more inelastic relative to less knowledge intensive industries such as textile manufacturing.

The determinants of wages in the short-run are determined mostly by industry and location factors external to the firm (Lazear and Oyer 2004). In responding to a sudden and unexpected rise in labor costs, firms may respond to their external environment by reorganizing. While the data does not allow me to determine whether firms are explicitly reorganizing by adding and dropping layers (Caliendo, Monte, and Rossi-Hansberg 2015) or employees willingly depart, firms exhibiting fastest wage growth after the anchor firm's arrival also shed the most workers who form spinoffs. These workers come from higher paying firms, but are not disproportionately the top earners within their firms. These employee transitions in response to shocks to the prevailing wage may help explain findings in prior literature exploiting a similar empirical framework. For example, Bloom et al. (2019) show that management practice scores increase for extant firms when the anchor firm arrives. They attribute their finding to managers learning from the arriving firm. This paper suggests a specific channel through which managers improve: The anchor firm forces incumbents to reorganize and shed less productive layers. These changes may be reflected in structured management practice scores. Similarly, better management practices through reorganization and layer shedding also would be reflected in measured labor productivity improvement in surviving incumbent firms found by Greenstone, Hornbeck, and Moretti (2010).

This paper's finding that the anchor firm's arrival induces entry and employment in segments of the supply chain, but not across wide industry groups is consistent with prior literature assessing the impacts of subsidies and place-based policies. For example, Moretti (2010) summarizes research on local economic multipliers and concludes that dynamic effects may offset employment gains in one industry by reducing employment in another and that job gain multipliers are largest in knowledge intensive and high-technology sectors. In the ballpark of Moretti's (2010) estimate of targeted place-based policies, a back of the envelope calculation places the local multiplier at 1.3 jobs created for every anchor firm job. While I discuss agglomeration channels, like Glaeser and Gottlieb (2008/9) I cannot prescribe a policy solution or optimal subsidy strategy without establishing a parametric form to agglomeration spillovers. However, the aggregate effects appear relatively small given subsidy amounts that often exceed $100 million.

This study must be qualified as it abstracts away from two key components related to the shock in the difference-in-differences and direct testing of the knowledge mechanism. The empirical strategy and interpretation assumes the only shock occurring in the county is the arrival of the anchor firm. However, anchor firms may locate in a county with the expectation of the arrival of other large suppliers. These suppliers may be large established firms who also expand into the county for the first time following the anchor firm. The startup response I observe may then be linked to these follow-on entrants. I also do not match the location of universities and R&D labs to counties where anchor firms arrive. Future research will explore both dimensions using the Longitudinal Business Database to identify follow-on expansions into the county and the Business R&D and Innovation Survey (BRDIS) for information of the location of research facilities. More work must also be conducted that considers the endogenous matching of workers and firms to locations with the highest productivity. Unobserved selection of firms that incubate entrepreneurs or matching of anchor firms to winning locations with an expected entrepreneurial climate or improving economic conditions may also explain at least some of the results. While the balance tests and various checks on difference-in-differences assumptions may assuage these concerns, the interpretation of the results and agglomeration channels may point to future discussions on labor pooling and variation in worker-firm matches across locations.

A research stream emanating from this paper relate to understanding how labor market effects vary with the industry of the anchor firm. The effects on incumbent wages should vary by how specialized labor is in those industries. Incorporating measures of knowledge intensity of the industry, innovation intensity of the firm, and exploiting information on the location of corporate labs in the BRDIS may provide additional evidence to highlight the labor-specific channels suggested in this paper. Though the evidence presented in this paper is consistent with views on firm reorganization, additional work should be done to specifically identify whether firms are eliminating layers from their organizations. This paper also stops short of discussing welfare implications. The jobs per subsidy dollar seems unreasonably high given the localized nature of the treatment effect, but it is unclear whether incumbent firm reorganization and the movement of labor towards younger firms provides other welfare gains. Understanding the longer-term dynamics of these startups and production in these counties may provide a more complete picture on the welfare implications of these anchor firm arrivals.

Chapter 3

City Entrepreneurship and Transitory Employment: Evidence From Mass Layoffs

3.1 Introduction

Local entrepreneurship is seen as a key driver of regional growth and economic mobility. Whether owed to the difficulty in the measurement of entrepreneurial activity or the variation in the precise definition of entrepreneurship, urban and labor economists perceive entrepreneurship, and especially local entrepreneurship, in a drastically different light. Depending on perspective, entrepreneurship can be understood both as an occupation of choice as well as one of necessity. The role of entrepreneurship in the local economy is often either thought of as an indicator of creativity and economic choice in a neighborhood or city or as a specific vehicle to support an individual's labor market outcomes through employment opportunities or experimentation and learning. This paper seeks to understand whether local entrepreneurship is an important channel enabling workers to rebound following a negative shock to their formal employment opportunities.

This paper discusses the extent to which entrepreneurship facilitates the transition of displaced workers into employment when fewer formal labor market opportunities exist. In essence, I discuss whether entrepreneurship contributes to, or is the product of, a city's resilience. In the first part of this paper I use administrative employee-employer matched Census microdata to identify job transitions of displaced workers. By exploiting mass layoffs where workers within an establishment[1] are not displaced due to observable differences in individual productivity (Gibbons and Katz 1991), I estimate the impact of a location's entrepreneurial where a shutdown occurs on wages and employment in the local area. To understand disparate impacts of workers with immediate wage-earning opportunities from those less likely to do so, I focus on a set of multiunit firms who consolidate operations by shutting down one facility, but not all. I show that these firms reallocate many employees from shutdown establishments to existing establishments elsewhere. In particular, these retained employees tend to be workers above the median of the closing facility.

The set of workers permanently laid-off by the firm are either hired by existing incumbent firms or by local startups. Conditional on becoming displaced, workers employed by startups experience faster post-displacement wage growth than those eventually employed by mature firms. I compare these wage estimates to wage effects from voluntary[2] workplace transitions. Displaced workers who find employment in mature firms exhibit no difference in wage growth from workers of similar relative wages in the firm who exit voluntarily. By contrast, displaced workers who find employment in startups experience faster wage growth. The effects are strongest for eventual startup workers displaced in thicker labor markets with smaller than average firm sizes.

The geographically local nature of worker transitions for lower wage employees mo-

[1]Throughout this dissertation, establishments are physical locations of a firm's economic activity. Firms with operations in multiple physical locations have more than one establishment and are defined as multiunit.

[2]Employee departures that are not accompanied with at least 50 workers exiting the establishment during the same year.

tivates examining the microgeography of shutdowns and entrepreneurship. I describe a new avenue for understanding regional labor markets and entrepreneurship using publicly available data sets on business registration filings to identify entrepreneurial formation and web scraped data on mass layoffs. Offices of Secretary of State at the state-level are often tasked with managing filings for new businesses, though sole proprietors are not required to file with the state in which they are domicile. Some states provide bulk download files of all registrations that include company name, address, date incorporated, active or inactive status and the date this status was determined. Prior studies such as Guzman and Stern (2015) and Guzman (2019) show the utility of compiling such business filings to understand the spatial dynamics of entrepreneurship and entrepreneurial quality. I extend their work by geocoding addresses of filings to map the location of entrepreneurial entry in cities. I then combine the business registration data with a novel hand-collected data set from federally mandated layoff notifications through the Worker Adjustment and Retraining Notification (WARN) Act. The Act requires firms above a size threshold to notify workers and submit a letter to the state's department of labor when a major establishment shutdown or layoff is to occur. These notices can be found on most state websites in pdf files with either the text of the actual letters companies produced for workers and government officials or a summary of information. The most common pieces of information included in the public document across states are company name, address, filing date, layoff date, number of affected employees, and whether the layoff is permanent or temporary. As with the Business Registration data, I geocode the closure addresses.

The geocoding of both the Business Registration and WARN Act data allow me to conduct a rich spatial analysis of the location of shutdowns and subsequent entrepreneurial formation. The geocoding also allows me to construct maps of specific cities to pinpoint more precisely the nature of entrepreneurship and regional resilience in the aftermath of a negative shock to employment opportunities. Due to the intensive nature of both scraping the various data sources as well as paywalls many states enforce for their Business Registration data, I focus on Florida whose sunshine laws make more easily available data on both the entrepreneurship and closure dimensions. Within Florida, I focus on the largest and most dense labor market area of Miami. The focus on an individual city provides the unique ability to capture neighborhood and city-level dynamics.[3]

3.2 Background

3.2.1 Local Entrepreneurship and City Growth

Entrepreneurial formation can be understood through the demand side forces whereby entrepreneurs are required to capture an unmet market opportunity or as an outcome of the supply of entrepreneurs (Chinitz 1961; Glaeser, Kerr, and Ponzetto 2010). The latter channel describes the spatial distribution of entrepreneurs as largely driven by disparities in the location of entrepreneurial talent and perhaps the local ecosystems that support entrepreneurs (Jacobs 1961; Glaeser 2000). One consideration, however, is that recent literature tends to focus on growth or transformational entrepreneurship, which is strikingly different from the most common forms of entrepreneurship (Schoar 2010; Hurst and Pugsley 2011).

Accounting for only a sliver of aggregate firm entry and employment in the economy, limiting the definition to Silicon Valley-style high-technology sector startups as the outcome of interest does not capture the labor market benefits of a vibrant local entrepreneurial (Decker, Haltiwanger, Jarmin, and Miranda 2014; Adelino, Ma, and Robinson 2017). From a sociological standpoint, the earliest urban scholars such as Jane Jacobs described the role of main street entrepreneurs in revitalizing neighborhoods, a description of urban entrepreneurship that shaped subsequent works in

[3] I am extending the data set to include a broader set of cities.

economics (Jacobs 1961; Glaeser, Kallal, Scheinkman, and Shleifer 1992). Understanding the benefits to workers and cities from entrepreneurship requires a more encompassing of the entrepreneurial activities residents engage in. This study uses both administrative microdata as well as complete Business Registration data to identify "main street" firms.

3.2.2 Entrepreneurship by Distressed Workers

Motivation for entering self-employment or starting a firm can arise due to lack of wage-earning opportunities. Exploiting regional variation in post-Great Recession job losses, Fairlie (2013) provides ample empirical evidence showing the positive relationship between financial distress and self-employment. This paper relates to transitions to entrepreneurship driven by displacement from potentially distressed firms. Babina (2015) finds entrepreneurs who found companies after departing a distressed firm are typically higher wage earners in the firm and start firms that grow faster than firms created by founders from non-distressed employers. In effect, the paper finds that entrepreneurs are instrumental in reallocating resources to new firms more able to exploit investment opportunities away from failing incumbent firms.

A long literature links individual characteristics to the propensity to become self-employment or start a new venture (e.g., Evans and Leighton 1990; Doms, Lewis, and Robb 2010; Azoulay, Jones, Kim, and Miranda 2020). These studies find that middle aged workers and workers who have had higher variance in earnings and more job transitions are more likely to enter entrepreneurship at some point in their lives. Another thread of literature discusses lasting effects of displacement on a worker's future labor market outcomes (e.g., Gibbons and Katz 1991; Jacobson, LaLonde, and Sullivan 1993; Kletzer 1998). Finding that workers' familiarity with firm-specific knowledge is key component of earnings, these studies generally find persistent post-displacement earnings losses when workers transition to new jobs. Corroborating this evidence is that earnings losses are less for workers who are able to find work in the same industry as their defunct former employer. However, beyond the new employer's industry the characteristics of firms who eventually hire these workers and how they affect worker and city-level outcomes is less understood.

Linking local entrepreneurship to displacement is important for two reasons. Young firms and small businesses may provide transitory employment opportunities for workers whose accumulated skills are no longer utilized. Displaced workers may also serve as a source of lower wage labor for liquidity constrained young firms. From a local policy perspective, low wage displaced workers are also far less spatially mobile and unable to capture spatial arbitrage opportunities (Notowidigdo 2019). Finding cost-effective avenues that enable these workers to find gainful employment is critical for welfare within local jurisdictions. Second, if entrepreneurship is assumed to be a key component of a city's resilience to an economic shock, then new ventures should serve as a vehicle that attenuates the impacts of displacement. With these considerations in mind, this descriptive paper will establish that displaced workers who transition into entrepreneurship are indeed lower wage workers from closing establishments. I then show that while these workers experience slower earnings growth post-displacement than their colleagues who find employment in older incumbent firms, they fare better than low wage workers who voluntarily[4] leave a firm. Of course, opportunities to become employed by a young firm depends on the availability of young firms in the city prior to displacement. I use Business Registration and public notices on layoffs to identify the extent to which new firms arise proximate to large shutdowns and the correlation with overall job growth across cohorts of young firms.

[4] Or fired from their job. This distinction is impossible to identify in the microdata.

3.3 Data and Empirical Summary

3.3.1 Aggregate Entrepreneurial Job Creation from Quarterly Workforce Indicators

The Census Bureau provides publicly available Quarter Workforce Indicators (QWI) data that allow researchers to understand business dynamics along a number of worker and firm characteristics including location, industry, and age. With the source data coming from the LEHD, states enter the sample in different years. However, the focus of this paper is on post-recession years after 2010 at which point all 50 states and the District of Columbia have complete data.

Location variables are available at the national, state, Core-Based Statistical Area (CBSA), and county-level. For consistency with the focus on within city-level dynamics, I use the CBSA level of analysis.[5] CBSAs represented in the sample include both metropolitan and micropolitan statistical areas allowing for understanding differences between larger and smaller labor market areas. QWI provides employment totals by firm age category. The categories are 0-1 years, 2-3, 4-5, 6-10, and 11+. This age structure allows tracking employment dynamics by firm entry cohorts at two year increments. For example, all firms represented in the 0-1 age category in 2010 comprise the 2-3 grouping in 2012 and 4-5 in 2014. The broader age bins for older firms do not allow for tracking an entry cohort beyond the initial five years. One important limitation of the public data is that when creating an age-based data extract at geographic levels finer than the state, the Census Bureau only provides public data at the sector-level (i.e., 2-digit NAICS).

Data suppression is a considerable issue limiting analysis using the public version of the QWI. When too few firms are represented in a cell, the Census Bureau suppresses employment totals for that city-sector-age combination. The Census Bureau

[5] I refer to CBSAs as cities.

does explicitly state a cell has 0 employment when possible.[6] For each city-sector-age triplet, I require all age bins to have non-missing data. I retain city-sector combinations that do not have any gaps in years that the pair is in the data. Therefore, I may retain only a subset of city-sectors. Finally, for each city I only retain the set of years for which the remaining set of city-sectors have no gaps in years. This cleansing procedure provides a balanced panel for each city.

Table 3.1: Data Availability in QWI by Sector

NAICS Sector	CBSA Count
11. Agriculture, Forestry, Fishing and Hunting	183
21. Mining, Quarrying, and Oil and Gas Extraction	44
22. Utilities	4
23. Construction	661
31-33. Manufacturing	340
42. Wholesale Trade	339
44-45. Retail Trade	734
48-49. Transportation and Warehousing	323
51. Information	134
52. Finance and Insurance	319
53. Real Estate and Rental and Leasing	349
54. Professional, Scientific and Technical Services	527
55. Management of Companies and Enterprises	46
56. Administration and Support and Waste Management/Remediation Services	434
61. Educational Services	135
62. Health Care and Social Assistance	622
71. Arts, Entertainment, and Recreation	181
72. Accommodation and Food Services	660
81. Other Services	609
92. Public Administration	N/A

Notes: The table displays the number of CBSAs represented in each NAICS sector in the public Quarterly Workforce Indicators (QWI) data.

The final cleansed sample retains a large portion of locations, sectors, and location-sector combinations. The sample includes CBSAs from all 50 states and the District of Columbia as well as 834 unique cities of the total possible 952 CBSAs. There are 6,644 unique CBSA-sector pairs and on average CBSAs have 8 sectors (of 19 possible)

[6] This implies that missing data should not be assumed to be 0's.

represented. Table 3.3 shows the number of CBSAs represented in each sector. For more detail on QWI, refer to Chapter 4.

3.3.2 Closures and Entrepreneurship in Administrative Census Microdata

I identify large establishment closings in the Longitudinal Business Database (LBD) and capture the movement of workers within and across firms using the Longitudinal Employer-Household Dynamics (LEHD) data come from the U.S. Census Bureau. The LBD sample used in this chapter is an establishment-level panel covering the universe of approximately 200,000,000 U.S. non-farm private businesses between 1990 and 2015. Physical establishments retain a unique time-invariant identifier throughout the panel. These establishment identifiers are linked to a federal Employer Identification Number (EIN) as well as a Census created firm identifier (firmid). The EINs and firmids allow researchers to match establishments into parent firms. In a given year, establishments map into a unique EIN and firmid. EINs and firmids may have multiple establishments beneath them indicating multiunit firm status. Establishment-year observations are assigned to a unique industry (4-digit NAICS in this setting). Due to differences in the definition of an establishment between the LBD and LEHD, I construct a synthetic definition of an establishment as EIN-county-industry triplets as this represents the narrowest physical workplace location between the LBD and LEHD. I consider a firm to an EIN and I define a large shutdown as instances where all establishments beneath an EIN-county-industry cease to exist in the LBD. Operationally, a shutdown in this study is defined as instances where a firm (EIN) closes an entire industry of operation in a county and may represent the closure of multiple physical establishments. I focus on EIN-county-industry shutdowns involving more than 50 employees.[7]

Table 3.2: Employment & Payroll Relative to Final Year

	(1) Employment	(2) Payroll per Worker
t-5	0.2663	3.596**
	(3.760)	(1.451)
t-4	1.648	2.530**
	(3.324)	(129.2)
t-3	2.844	3.232*
	2.983	(1.779)
t-2	-4.491	1.390
	(7.792)	(1.153)
t-1	Base group	
Establishment FE	Y	Y
N	19,500	19,500
Adj. R-sq.	0.9363	0.4397

Notes: *** 1% ** 5% * 10%. This table shows changes in employment and wages in closing establishments in the Longitudinal Business Database relative to the final year of positive employment (t-1). Column 1 shows that firms do not gradually reduce employment in closing establishments. None of the coefficients in Column 1 are statistically different from each other. Column 2 shows that firms do reduce pay over the long horizons as the difference between the coefficient on t-2 and t-5 is significant (p-value 0.0943). However, year-to-year changes (e.g., t-5 to t-4) do not differ significantly. Closures are defined as EIN-county-industry combinations that have all establishments exit in the subsequent year. This table shows employment (Column 1) and payroll per worker (Column 2) in the years leading up to the final year of positive employment. The sample is restricted to EIN-county-naics pairs that employ at least 50 full-time workers. The sample is a balanced panel of 3,900 unique EIN-county-NAICS closures. Standard errors are clustered by EIN-county-NAICS.

Table 3.2 shows changes in employment and wages ahead of the closure. Firms do not appear to gradually downsize a closing facility prior to the mass layoff. This feature of large scale closures will be useful to identify differences between workers who select into entrepreneurship and those who do not post-displacement.

Matching shutdown EIN-county-industry combinations into the LEHD allows me to capture whether workers who were employed by the EIN-county-industry remain

[7] This matching strategy between the LBD and LEHD is similar to Tate and Yang's (2015) study of wage differentials between sexes following a mass layoff.

in the same EIN (but different county and/or industry) or transition into another EIN suggesting employment in a different firm after the shutdown. The LEHD consists of wage records for workers in the United States provided to the Census Bureau through state unemployment insurance offices. Through the Employee History File (EHF), workers are identified through personal identification keys (PIKs) and matched to a state-level tax reporting entity (SEIN) and unit (SEINUNIT). Each state has its own tax identifier for the same firm. Therefore, firms may have multiple SEINs both across states and within states if they have multiple tax reporting entities and places of business. The EHF provides the establishment of employment as well as quarterly wages in the establishment for all workers. Business information is merged into the EHF using the Employer Characteristics File (ECF) at the SEIN-SEINUNIT level. The ECF includes such information as NAICS classification, total quarterly employees on payroll, total quarterly payroll, and geocoordinates and address. Matched EHF-ECF records are known as worker-job observations. The ECF also provides federal tax identifier Employer Identification Numbers (EINs). A firm may be organized to have multiple tax reporting EINs that consist of multiple physical establishments. Each SEIN-SEINUNIT pair is associated with one EIN, but as in the Longitudinal Business Database (LBD) where EINs correspond to multiple physical establishments of a firm, the EIN may also include multiple SEIN-SEINUNITs.

I cleanse the employee-employer matched sample to include only those worker-job-year observations for which the average wage worked in non-zero earnings quarters during the calendar year is greater than minimum wage of $2,678 ($5.15/hour x 40 hour work week x 13 weeks). This procedure retains approximately 400,000,000 of over 700,000,000 total worker-job-year observations between 1990 and the first quarter of 2012. Final matched worker sample allows for construction of a worker's relative wages compared to coworkers as well as tracking movements and job transitions. Beyond wages and industry of employment at the worker's establishment, the Individual Characteristics File (ICF) provides limited demographic traits on workers (PIKs) including sex at birth, date of birth, place of birth/immigrant status, race, and education level.

One limitation of LEHD analysis is incomplete geographic coverage. States are not required to provide the Census Bureau with employee wage and employer information, and currently 31 states including the District of Columbia participate in the data sharing program. Researchers are granted access to the LEHD's infrastructure files on a case-by-case basis. This project has access to 25 states: Arkansas, Arizona, California, Colorado, DC, Delaware, Idaho, Illinois, Indiana, Iowa, Kansas, Maine, Maryland, Missouri, Montana, New Mexico, Nevada, North Dakota, Oklahoma, Pennsylvania, Tennessee, Texas, Virginia, Washington, and Wisconsin. Workers who drop out of the sample may do so because they are unemployed, have left the labor force, or are employed in an out of sample state. The second limitation of the LEHD is the lack of occupational codes, allowing only for inferring a worker's position in the firm through wage and the reporting unit's industry.

I match the Longitudinal Business Database identified closures into the LEHD using EIN-county-industry triplets. The focus of this chapter is on understanding the dynamics of within firm reallocation and outside options for displaced workers. Therefore, I focus on the subset of shutdowns where the firm (EIN) closes an industry-county pair but does not exit completely. Further, I retain workers in each event for whom I can construct a balanced panel that includes observations for the three years preceding the closure as well as the three years post-closure. The final sample includes approximately 4,000 shutdown events.

My analysis compares outcomes for workers displaced from the same closing facility. I compare workers within similar wage levels of the firm from which they are displaced. I estimate the effect working for a young firm has on post-displacement earnings. I find that relative to displaced workers who find employment in mature firms, these workers exhibit faster wage recovery when compared to another set of workers who voluntarily depart firms.

The table shows that workers from the same closure and within firm income quintile who obtain subsequent employment in a local young firm earn 67% over the next three years than their displaced coworkers who find employment elsewhere. I control for three year log changes in wages for the 3 preceding years to the closure event to capture changes from their pre-existing wage trajectories. I also test for differences in wages between the workers in the four groupings (startup, mature, within, reallocation) and do not find evidence that their wages differ significantly in the pre-period. This table provides motivational evidence for understanding the relative importance of local entrepreneurship in job searches, particularly for displaced or low wage workers.

3.3.3 The Worker Adjustment and Retraining Notification Act

A second contribution of this chapter is to provide researchers in the urban economics and the economics of entrepreneurship avenues for exploiting public data sources to understand clusters of entrepreneurship and local labor markets. I use public disclosure requirements of mass layoffs to identify specific cases of establishment shutdowns. The Worker Adjustment and Retraining Notification (WARN) Act of 1988 was intended to provide greater transparency to employees surrounding closures and layoffs. The WARN Act requires firms to provide workers and state officials at least 60 days advance notice of a closure or mass layoff. The WARN Act was intended to increase the amount of time workers have to prepare for a layoff and potentially find new employment opportunities. Though few recent studies have exploited information in the WARN Act, early studies such as Addison and Blackburn (1994ab) find the Act did not affect firms' notification behavior to their workers. The Act has, however, led to the systematic development of public databases to store records of establishment

Table 3.3: Post-displacement Wage Growth by Employer Type

	(1) ln(Wage in t+3 / Wage in t)
Young firm	0.6723***
	(0.0600)
Mature firm	-0.2474***
	(0.0434)
Within firm-county	Not disclosed
	Not disclosed
Relocation	Not disclosed
	Not disclosed
Pre-growth	Not disclosed
	Not disclosed
Event FE	Y
Year FE	Y
Wage quintile FE	Y
Demographic controls	Y
Observations	49,000
Adj. R-sq.	0.0988

Notes: *** 1% ** 5% * 10%.This table is based off the Longitudinal Business Database constructed closure sample matched to the Employee History File of the LEHD. The sample restricts to workers who have observations in all three pre-displacement years as well as all three post-displacement years to construct a balanced panel. The table shows earnings growth relative to the worker's final wage in the closed establishment. The cateogries of subsequent work are: Employment in a firm aged $<$ 3 years, in firm $>$ 3 years, different establishment within the firm-county, or in the same firm but different county. These later two coefficients are not disclosed due to potential disclosure concerns. The pre-growth variable controls for worker's wage growth in the three years preceding closure. based off the displaced workers' subsequent job. The Event FE forces coefficients to be estimated between individuals of the same closure event, the year FE controls for secular wage trends, the wage quintile forces comparisons to be made between workers in the same quintile of their employer's wage distribution. Demographic controls include age, sex at birth, country of birth, and education level. Standard errors are clustered by closure establishment.

shutdowns.

The Act generally applies to closures of facilities operated by large firms. The Act requires firms with over 100 full-time employees who layoff at least 50 workers at a single physical location, an employment loss of 500 or more employees, or a loss of 50-499 employees if these workers constitute at least a third of the firm's overall workforce.[8] Exceptions to the disclosure require include layoffs/closures induced by natural disasters, circumstances argued to be "unforeseeable", or situations where the company demonstrates "good faith effort" with reasonable expectation to acquire capital to avoid triggering a mass layoff.

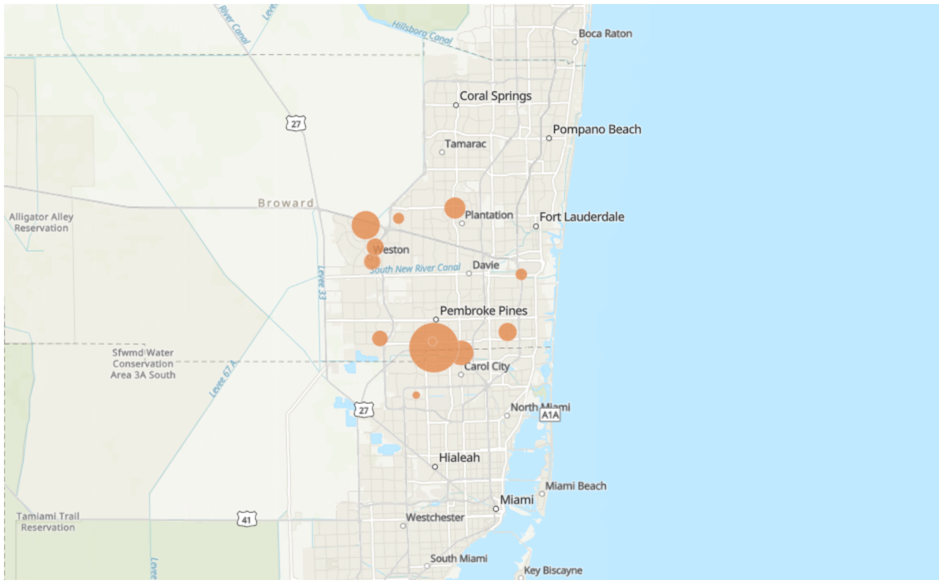
State governments post online the letters they receive from firms describing an upcoming layoff and some states maintain a public database with historical information on the shutdown/layoff company, physical address or city of displacement, estimated number of affected workers, an indication of whether the layoff is temporary or permanent, and on occasion a brief description of the establishment's primary business operation. However, coverage and data availability is imperfect. Though the WARN Act took effect in 1988, only Oregon's database provides information going back to 1989 with the next most complete coverage being Alabama which commences in 1998. States incrementally begin providing WARN filings online throughout the 2000s with all states except Nevada, New Hampshire, North Dakota, Vermont, and Wyoming participating by 2017.

Figure 3.1 is a map of WARN Act notices in Miami, Florida during the years 2014 and 2015.[9] The map shows the concentration of filings around larger cities. The bubble sizes indicate the relative size of the layoff. The largest bubble corresponds

[8]A job loss is defined as cases where the layoff is expected to last longer than 6 months or a reduction in an employee's hours worked by at least 50% over any 6 month period.

[9]Florida's WARN notices and business openings data are particularly deep. While the final dataset will expand all possible states for the widest coverage, this analysis relies on Florida. Because Miami is the largest and most dense labor market area in Florida, I use it as my example throughout this essay.

FIGURE 3.1: Geography of Mass Layoff Events in Miami, FL



to Univita of Florida's mass layoff of 596 employees in 2015.

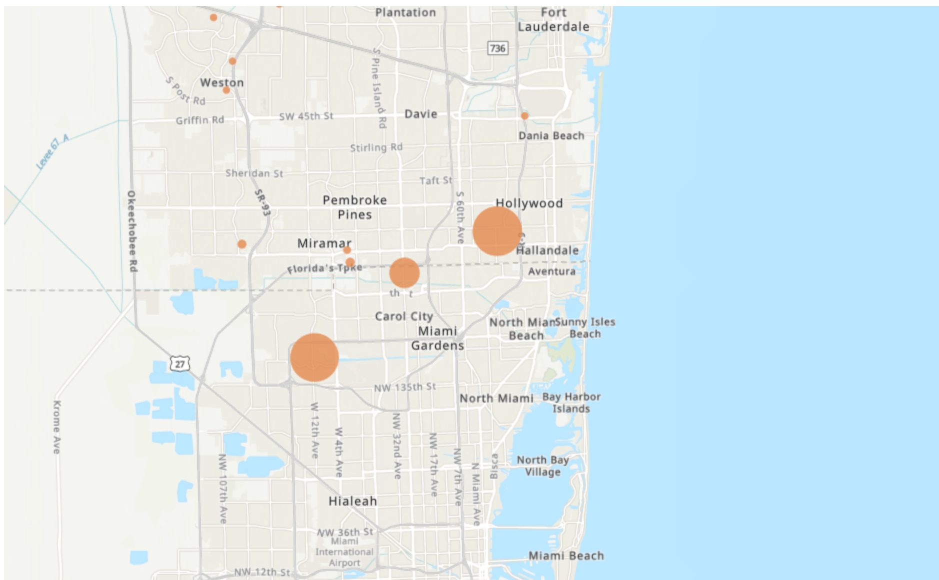
3.3.4 Business Registration Files

Most firms, aside from sole proprietors, register their businesses with the state's Secretary of State's office. Most states allow the public to search for business names and obtain information of the firm's location, date of filing, and active or inactive status. States such as Connecticut, Florida, New York, Virginia, and Washington as well as select cities such as Chicao allow researchers to download bulk data of all active and inactive business registrations. Most other states allow researchers to purchase the data for a nominal fee.

Guzman and Stern (2015) which focuses on Massachusetts and Guzman's (2019) registration data for 26 states are the first to compile and exploit these state records to explore geographic clustering of entrepreneurial formation. Florida provides free and detailed Business Registration information that goes back to at least 1998. Therefore, the Business Registration analysis in this chapter explores Florida's microgeography.

The Florida data provides precise business location address, date of filing, date of approval, and date of last filing/active status for each company name. These data allow me to infer longitude and latitude coordinates for new business filings using address lists matched to ArcGIS mapping data.

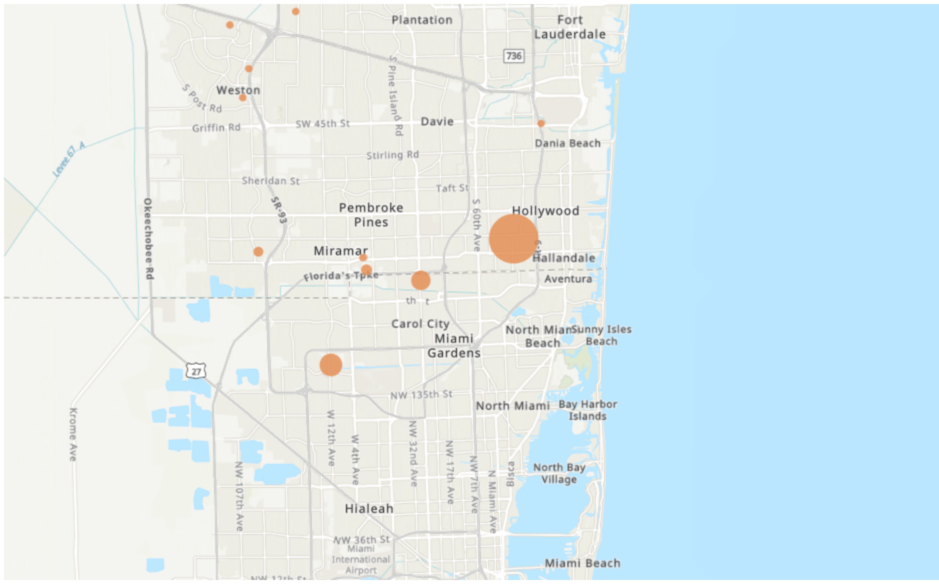
FIGURE 3.2: Registrations within 10km of closures



The combination of the ArcGIS coordinate matched WARN Act notices and the Business Registration filings allows me to conduct a microgeographic analysis of business formation after a large firm closure. The following maps demonstrate the variation in business formation within varying radii of the focal closure event. Figure 3.3 maps the same WARN notices as Figure 3.1 but instead uses the number of registrations within 10 kilometers to determine bubble sizes. The largest bubbles correspond to healthcare provider Outpatient Services, Inc. (1,143 new businesses) and military apparel manufacturer Short Bark Industries, Inc. (1,171 new businesses).

Figure ?? reduces the radius to 5 kilometers. The map shows that Univita's closure occurs in an area of denser economic and entrepreneurial activity.

FIGURE 3.3: Registrations within 5km of closures



These figures do not establish an association between the location of layoffs and the likelihood of business formation in an area. However, the maps provide a unique look at economic geography, location choice, and local labor markets that are missed by studies relying exclusively on Census microdata. The maps do suggest that attention to neighborhood-level and more precise location details is instrumental in understanding why some areas even within cities are more dynamic than others. Alongside prior literature that establishes low wage displaced workers find work locally as well as the documentation of positive earnings effects of working for a young firm post-displacement, finer geographic detail can add far more to our understanding local economic dynamism, business formation, and how entrepreneurship facilitates employment for displaced workers.

3.4 Conclusion

This essay provides an outline for understanding the microgeography of layoffs and entrepreneurship. Because entrepreneurship is thought to be an important aspect of economic dynamism. One possible contribution of young firms is that they provide employment opportunities for displaced workers. In this essay, the entrepreneurs I refer to are less likely to be the "transformative" entrepreneurs, but instead are "main street" style businesses.

Understanding the microgeography of where these new businesses form is important. A broad literature in urban planning discusses the importance of small businesses in city neighborhoods as a source of economic vitality and creativity for local area residents. It may be the case that areas of the city where starting a business is more feasible are also the places that can absorb the most displaced workers when a mass layoff occurs.

Data on both layoffs and business formation are difficult to obtain. I combine business registration data with hand collected WARN notice information to construct a database that allows researchers to understand the microgeography of layoffs/closures and business starts. By combining these novel data with other sources on local economic conditions and demographics, entrepreneurship and urban scholars can potentially uncover additional channels through which entrepreneurship facilitates labor mobility.

Chapter 4

Regional Resilience, Startups, and the Great Recession

4.1 Introduction

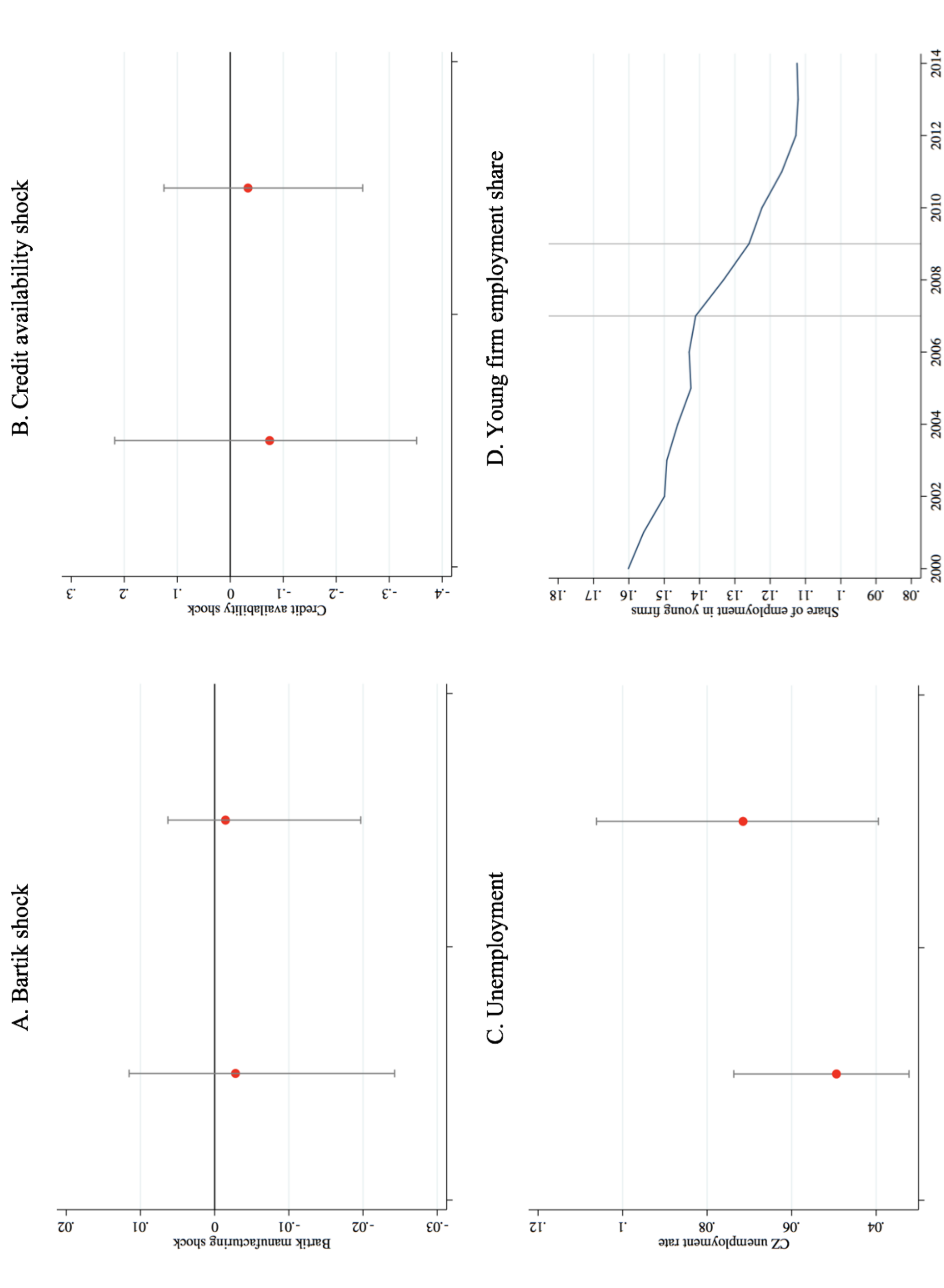
The Great Recession saw a sharp rise in unemployment, a sluggish recovery in jobs, and a continued decline in the share of employment in new businesses. The impacts of the Great Recession were not homogeneous across local areas (Commuting Zones) of the United States (Figure 4.2). Rather, places such as Las Vegas, NV or Phoenix, AZ, saw employment growth during the 2007-2009 crisis 10% below Austin, TX. Perhaps more striking and concerning is the fact that the disparity between local areas severely affected (10th percentile of employment growth between 2007 and 2009) by the recession and those least affected (90th percentile) continued to grow at least six years after the crisis ended. The lack of regional convergence over five years after the crisis ended is a unique feature of this particular recovery (Yagan, 2017). Yagan (2016, 2017) discuss the lasting negative impact of being located in a severely affected Great Recession location on job prospects, earnings, and likelihood of remaining in the labor force. While these worker-level impacts of the crisis are beginning to be studied, less attention has been given to understanding how firm characteristics, specifically firm age and industry, contributed to regional resilience in job creation before the crisis and during the recovery.

Prior works, such as Adelino, Ma, and Robinson (2016), have explored the relative contributions to employment growth by firms of different ages. They find startups account for over 90% of non-tradeable sector job creation. Moreover, startups are thought to facilitate the optimal allocation of resources in the economy, enable job hopping while improving labor market matches, and contribute to dynamism and aggregate productivity growth (Decker, Haltiwanger, Jarmin and Miranda, 2014, 2017; Dent, Karahan, Pugsley and Sahin, 2016; Alon, Berger, Dent and Pugsley, 2017). While the importance of startups in the economy is clear, they are declining in employment share of the economy (Figure 4.1, Panel D) and clustered among regions. In regions suffering the largest employment losses during the Great Recession, the decline in startup rates may have limited labor flows and dampened their recovery.

To motivate our study, we first document several striking facts about the Great Recession. Using an empirical response function that satisfies an autoregressive growth process (Blanchard and Katz, 1992), we show that unlike in previous recoveries, Commuting Zones differentially impacted by the recession continued to diverge up to five years after the crisis ended. By 2014, the most severely affected Commuting Zones grew 17% less in employment and 11% less in net establishment entry than the least affected Commuting Zones. Put differently, the disparity between the local areas nearly doubled in cumulative employment growth (from -10% in 2009 to -17% in 2014) and almost quadrupled in net entry (-3% in 2009 and -11% in 2014). We demonstrate that a startup deficit in the bottom half of local area Great Recession performance explains this dramatic divergence.

Our paper focuses on job creation resulting from short-term local industry shocks before and after the Great Recession. We investigate whether regions and firms responded differently to exogenous shocks during the recovery than they did before the crisis. We implement a framework developed by Adelino, Ma, and Robinson (2016) in

Figure 4.1: Summary of Great Recession Shocks



Notes: The lines in Panels A, B, and C show the pre-crisis (2000-2007; left most line of each graph) and post-crisis (2011-2014; right most line of each graph) distributions of the Bartik manufacturing shock variable, credit availability shock, and unemployment rates. The end points of each line indicate the 10th and 90th percentiles. The points are the median of the distribution. Panel D displays the share of US employment in firms aged 0 to 5 years with grey lines indicating the 07-09 recession. The Bartik shock uses data from County Business Patterns, credit shock from FFIEC disclosure reports, unemployment rates from the BLS employment files, and young firm employment share from Quarterly Workforce Indicators.

which a Bartik (1991) tradeable sector shock generates variation in a region's income. The Bartik shock is a shift share approach to exogenously shift a local area's labor demand. The oft-used shock variable interacts the share of a Commuting Zone's employment in a manufacturing sub-industry with the national growth rate of the industry outside the Commuting Zone.[1] Summing over all manufacturing industries in a Commuting Zone, the Bartik shock can be thought of as predicted employment growth in the industrial base. This shock, in part, captures changes in local income arising from productivity growth or product demand changes in the tradeable sector. We connect these local industry shocks to non-tradeable sector job creation along the firm age distribution. To do this, we utilize the Census Bureau's Quarterly Workforce Indicators to measure job creation in young and old firms. Similar to Adelino et al., we find that better Bartik shocks between 2000 and 2007 generate more employment largely driven by new establishments. During the recovery, however, we strikingly find that local areas receiving better shocks to their industrial base tended to have far less job creation in new businesses.[2]

We show that common explanations for startup formation such as local labor market and credit conditions do not explain the reversal in startup job creation's relationship to industrial shocks. On the other hand, a decomposition of the shock's individual industry components reveals that the effect of shocks to local industry on startup job creation is dominated by a select set of industries. Most notably, the transportation equipment manufacturing industry, which includes automobile, aerospace and railroad stock manufacturing, accounts for over 30% of the Bartik

[1]Select recent examples of the Bartik variable's usage include Acemoglu and Restrepo (2017), Diamond (2016), Charles, Hurst and Notowidigdo (2013), Notowidigdo (2011) and Serrato and Zidar (2016). Others use the Bartik setup to construct local shocks on different dimensions such as Autor, Dorn and Hanson's (2013) measures of international trade competition.

[2]In our firm age and QWI analysis, we define periods at 2-year intervals. This means the recovery is 2009-2011, 2010-2012, 2011-2013 and 2012-2014. We label the periods using the last year of the interval: 2011, 2012, 2013 and 2014.

shock coefficient after the crisis. The covariance between the size of the industry and job creation in startups was negative in both the pre-crisis and recovery periods, but now that the industry is growing implies that the industry is a substantial factor in driving the relationship between local industrial base shocks and startup job creation negative.

While we do not theorize why the inverse relationship between employment in transportation equipment manufacturing and startup job creation exists, prior studies including Chinitz (1961) and Glaeser, Kerr, and Kerr (2015) discuss features of heavy manufacturing and capital intensive industries that are associated with far less regional entrepreneurship and employment growth. The two works suggest that large establishment sizes correspond to industries that are more capital intensive. These industries effectively dry up a region's entrepreneurial capital–both human capital and financial–as well as limiting the existing of supply-chain networks. Their discussion of the Pittsburgh steel industry (Chinitz, 1961) or the historical location of mines (Glaeser et al., 2015) is consistent with the industrial characteristics of transportation equipment manufacturing in terms of capital intensity and large establishment sizes. In the Bartik framework, the finding suggests that as the transportation equipment manufacturing industry grows in the US economy, startup employment in the non-tradeable sector is likely to remain most depressed in regions with a larger transportation manufacturing sector. We do not make any welfare claims regarding this tendency.[3] We believe unpacking the Bartik shock is a useful and novel contribution of our paper.

Our study relates directly to Adelino, Ma, and Robinson (2016), but framed as an investigation on the dynamics of entrepreneurship during the recovery. Using their study as a motivation and our observations of regional divergence after the Great

[3]It could be that employment in transportation equipment and consumption from existing establishments is optimal relative to employment in new restaurants and retail.

Recession, we contribute to a body of literature describing the lasting impacts of recessions on labor markets. Uniquely, we focus on regional entrepreneurship. In a macroeconomic framework, Barlevy (2002) describes the "sullying effect" recessions have using a job quality and job search model. Similarly describing declines in job quality in the wage-sector, Fairlie (2013) empirically discusses the rise in self-employment during the 2007-2009 crisis in the context of diminished wage-earning opportunities. He finds that workers in cities with larger rises in unemployment are more likely to transition into self-employment. Using startup job creation instead of self-employment, we do not find a similar relationship in the interaction of local labor market conditions and the Bartik shock.

Our paper contributes to several lines of literature. We discuss the role of firms and startups in the recovery. We relate to Adelino et al. in investigating the relationship between local shocks and entrepreneurship. Second, we build on regional economics scholarship that seeks to understand the long-term, spatial effects of the 2007-2009 crisis and the slow recovery. We take a different stance from the worker-level studies in recent literature and turn our focus to businesses. Our hope is that we can gradually understand how firms and entrepreneurs respond and adjust to economic shocks. Lastly, we provide a useful examination of the Bartik shock instrument used widely in literature across many disciplines.

4.2 Data

We primarily utilize two publicly available Census Bureau datasets including County Business Patterns and Quarterly Workforce Indicators in addition to Quarterly Census on Employment and Wages and Federal Financial Institutions Examiners Council data. The Census Bureau makes publicly available files detailing industry employment by establishment size and industry at the county-level through the County Business

Patterns (CBP) data series. Industries are defined up to the 6-digit NAICS level. Each state file contains county-industry observations. The pairings are provided for 2-, 3-, 4-, 5-, and 6- digit industry codes. The Longitudinal Business Database underlies the CBP data and employment counts are include only those employees on the establishment's payroll on the Tuesday of the week of March 12.

While a rich setting of information, the main concern with public Census data is data suppression to protect confidentiality. This issue is greatly exacerbated in smaller counties and narrowly defined industries. Data suppression can be severe. Summing employment in 4-digit industries across counties in a state may account for far less the state's total employment figure. At the county-level, often only 60% of county total is accounted for in the disaggregations. We impute missing data using the following regression for each year using 2-digit NAICS establishment size data.[4] Because the regression is estimated on the 2-digit NAICS observations, employment and establishment information is generally complete. The idea behind the imputation is that industry-county employment in each year is a function of the size of establishments in the industry-county:

$$emp_{ic} = \beta_0 + \sum_n \beta_{nic} est_{nic} + \epsilon_{ic},$$

where i indexes industry, c denotes county, and n are the establishment size bins (1-4 employees, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, 1,000-1,499, 1,500-2,499, 2,500-4,999, and 5000+). Each year regression yields a R-squared between .97 and .99. The coefficients on the specification are used to obtain $\widehat{emp_{ic}}$ and we use the imputed employment value to replace suppressed data. We check the estimations by again summing the predicted employment across industry-counties and industry-state pairings and find that 97-99% of state and county total employment is accounted for. Summing 4-digit industries at the county-level across the United States similarly

[4]This procedure is similar to the that described by Duranton, Morrow, and Turner (2014).

yields close to perfect approximations of national industry employment. Finally, we aggregate county data to the Commuting Zone (CZ) level using the Department of Agriculture's 2000 definitions.

The main data series of our paper is the Quarterly Workforce Indicators (QWI) and we rely on the procedure outlined by Adelino, Ma, and Robinson (2016) to cleanse and construct our sample. QWI is based off the Census Bureau's Longitudinal Employer-Household Dynamics Database (LEHD) and provides employment totals in two-digit NAICS sectors by county for five age categories: 0-1 years, 2-3, 4-5, 6-10, and 11+. By taking differences at two year intervals, we can construct a measure of job creation along the firm age distribution. This data structure allows us to track cohorts of firms by formation year at two-year increments by differencing total employment in, for example, age bin 4-5 at time t with employment in age bin 2-3 at time $t-2$. We collapse the 6-10 and 11+ age categories to 6+. For the 6+ category we tabulate job creation by taking total employment in 6+ at time t minus employment in 4-5 and 6+ at time $t-2$. [5] Job creation is scaled by Commuting Zone total employment from the Bureau of Labor Statistics Quarterly Census of Employment and Wages in 2000 to facilitate comparisons across years and age categories. Though we will refer to job creation in a firm size bin at time t, these figures represents employment changes between $t-2$ and t.

The primary limitation of QWI is that states enter the data at different years of our analysis. When we commence our sample in 2000, 41 states participate in the LEHD program. By 2005 the number is 50 including the District of Columbia and Massachusetts is last to join in 2010. We ensure that job creation is not artificially inflated by the inclusion of new counties into the CZ as time progresses by using

[5] We check cohort growth from 6-10 to 11+ at 4 year increments which excludes firms age bucket 6 in t-4, but we interpret our checks to suggest that net job creation in firms aged 6+ is driven by firms aged at least 11 years.

only those counties included in the sample the first time a CZ appears.[6] QWI does not allow us to capture an increasingly important component of labor markets. Katz and Krueger (2016) discuss that nearly all net job growth over the last decade is attributable to alternative work arrangements. Neither these work arrangements nor the self-employment captured in the Current Population Survey, as in Fairlie (2013), are in our analysis.

Similar to County Business Patterns, data suppression can be an issue. County-industry-firm age-year bins can contain missing employment values. We exclude any county-industry-year observations with at least 1 missing value. We must keep a stable county-industry sample to avoid artificially inflating or deflating employment totals. We further ensure that county-industry pairings are stable throughout the sample period. To facilitate comparability with the Adelino et al. results, our pre-crisis sample only requires CZ information to be complete between 2000 and 2007. Our post-crisis results use only county-industry pairs that never drop from the sample once they first appear starting in 2000. This ensures that the post-crisis sample includes only those CZs that are in the pre-crisis sample and CZs that appear for the first time in subsequent years. We believe our pre- and post-crisis sample are most comparable to each other using this procedure. The pre-crisis sample includes 554 CZs at its peak and the 2000-2014 sample requiring stable county-industries across the full period includes 507 CZs. However, our results are robust to constructing pre- and post-crisis samples separately in which the post-crisis sample includes 532 CZs.

Several additional datasets are used in our analysis. We obtain control variables for CZ size, share of high school educated workers, and total CZ wages from the Bureau of Labor Statistics' Quarterly Census of Employment and Wages (QCEW),

[6]Suppose in 2005 CZ A has counties 1,2, and 3. However, county 1 appeared in 2000 and 2 and 3 did not appear until 2005. CZ A will be defined as county 1. This is generally not understating CZ sizes as most CZs are constructed within state. For the major multistate CZs such as New York City-Jersey City-White Plains or Chicago-Naperville-Kenosha, states enter the data in the same year.

the 2000 decennial Census, and the IRS county-level Statement of Income (SOI) files, respectively. Our procedure for obtaining controls mimics that of Adelino et al. (2016). Other pieces of the analysis utilize the 2014 National Establishment Time Series Manufacturing database (NETS) and the Federal Financial Institutions Examiners Council (FFIEC) disclosure reports, both of which we discuss in their relevant sections.

4.3 Evidence of Economic Scarring from the Great Recession

We motivate our study by first documenting the extent to which regions recovered from the crisis. In terms of employment growth and establishment entry and exit, the effects of the Great Recession persisted well into the recovery. Regions most severely affected by the Great Recession remain laggards in employment growth and establishment entry.

For documenting these initial facts, we show the growth trajectory of Commuting Zones before and after the crisis using an empirical response function. The idea behind the response function is to show the cumulative impact of the Great Recession's employment declines after the crisis. We first show cumulative growth of CZs before the crisis (from 2003 to 2007), then show the impact of the crisis by plotting cumulative growth from 2007 to 2014. The specification follows that of Dix-Carneiro and Kovak (2017):[7]

$$log\left(\frac{y_{c,t}}{y_{c,2007}}\right) = \alpha_0 + \theta_t g_c^{GR} + \gamma_t log\left(\frac{y_{c,2007}}{y_{c,2003}}\right) + \epsilon_{c,t}, \tag{4.1}$$

using data from the Bureau of Labor Statistics Quarterly Census of Employment and Wages. We estimate Equation 1 separately for each $t \in [2009, 2014]$. $y_{c,t}$ is our

[7]Dix-Carneiro and Kovak (2017) investigate differential regional impacts of trade liberalization in Brazil, considering trade liberalization policy changes in the early 1990s as a one-time shock.

outcome of interest (total employment and establishment counts) in CZ c at time t. g_c^{GR} represents employment growth in a Commuting Zone between 2007 and 2009 transformed to reflect the difference between the 10th and 90th percentile in Great Recession growth rates. g_c^{GR} is the same for a Commuting Zone in each year regression. The coefficients θ_t's therefore show the cumulative effect of the Great Recession at time t on total CZ employment and establishment counts.[8] Equation 1 satisfies the autoregressive process of local labor market outcomes by controlling for the pre-crisis trend (Blanchard and Katz, 1992). The specification is similar to a first difference so we implicitly control for CZ-specific characteristics. By transforming g_c^{GR} to the 10-90th percentile difference, we show how the most severely affected regions rebounded relative to those least affected.[9][10] Negative coefficients suggest larger declines the outcome variable in regions suffering most severely during the Great Recession. We do not use this specification to make a causal claim about the Great Recession, but present evidence suggestive that locations most adversely affected by the crisis remain depressed relative to those not as severely impacted (Figure 4.2).

We measure the divergence between Commuting Zones in terms of two outcomes: Cumulative employment growth and cumulative net entry. Cumulative net entry measures the growth in establishment counts. Figure 4.3 plots each θ_t from the year regressions in Equation 1 separately for each outcome.

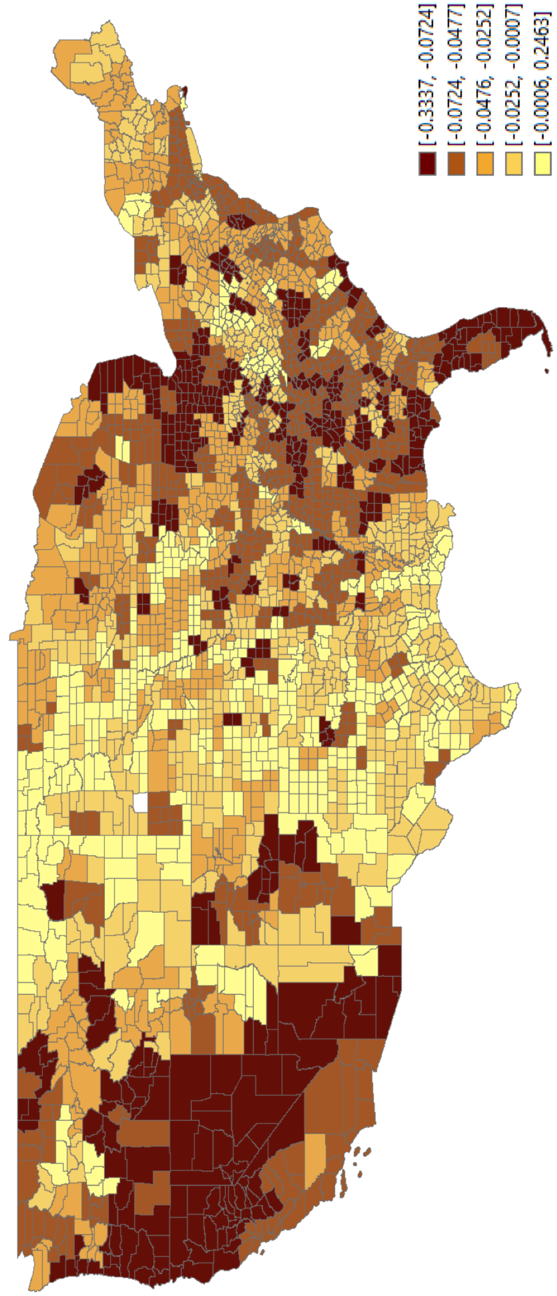
The point at 2009 in the cumulative employment growth regression is the actual difference in 2007-2009 employment growth between the 10th and 90th percentiles. The circle point at 2014 shows that since 2007, the hardest hit regions saw cumulative

[8] In the total employment specification, because the dependent variable is the actual growth from 2007 to t and the explanatory is the scaled 10-90 growth difference, θ_{2009} is the actual difference between the 10th and 90th percentiles of Great Recession employment growth. This means negative coefficients indicate the extent to which the 10th percentile CZs have lagged the 90th percentile.

[9] Conceptually, the coefficients are similar to representing the difference between Las Vegas, NV and Austin, TX which are near the 10th and 90th percentiles, respectively.

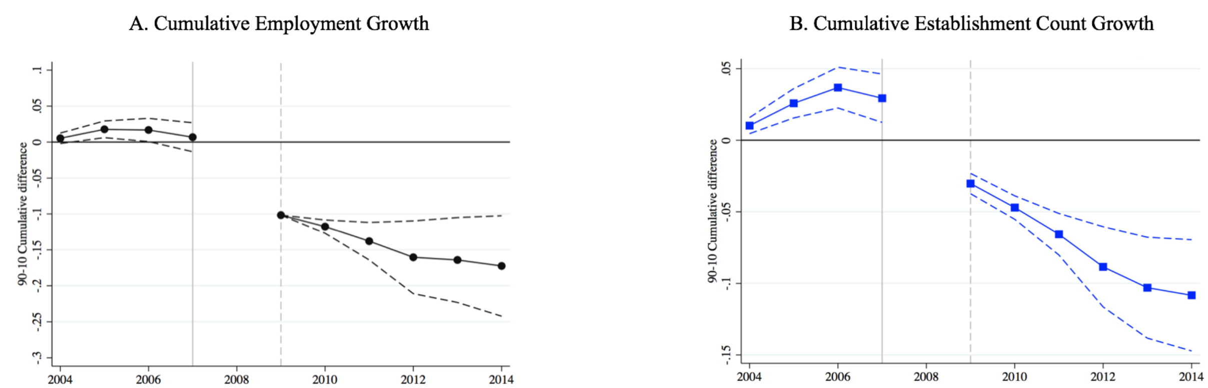
[10] Results are robust to using other points on the distribution such as the 75th and 25th percentile differences.

Figure 4.2: Great Recession Employment Change



Notes: The map displays quintiles of employment growth between 2007 and 2009. Areas shaded darkest saw the most severe employment declines, whereas lighter colors indicate Commuting Zones with better growth during the crisis. Data is taken from County Business Patterns and counties are matched to Commuting Zones using the Department of Agriculture's 2000 definitions.

FIGURE 4.3: Impulse Response Function



Notes: The graphs plot the coefficient on actual Great Recession employment changes from Equation 1. The Great Recession employment change is the log growth in employment between 2007 and 2009. The variable is scaled such that the coefficient reflects the difference in the outcome between the CZs at the 90th and 10th percentiles of Great Recession employment changes. This means that the 2009 point in Panel A is the actual difference in 2007-2009 employment growth between the 90th and 10th percentile CZs. The y-axis is the cumulative difference in the dependent up until time point t of between the 10th and 90th percentile Great Recession shock. The pre-2007 points plot the cumulative difference from 2003 to time t. The post-crisis points reflect cumulative growth since 2007. Dashed lines show 95% confidence bounds. Standard errors are robust.

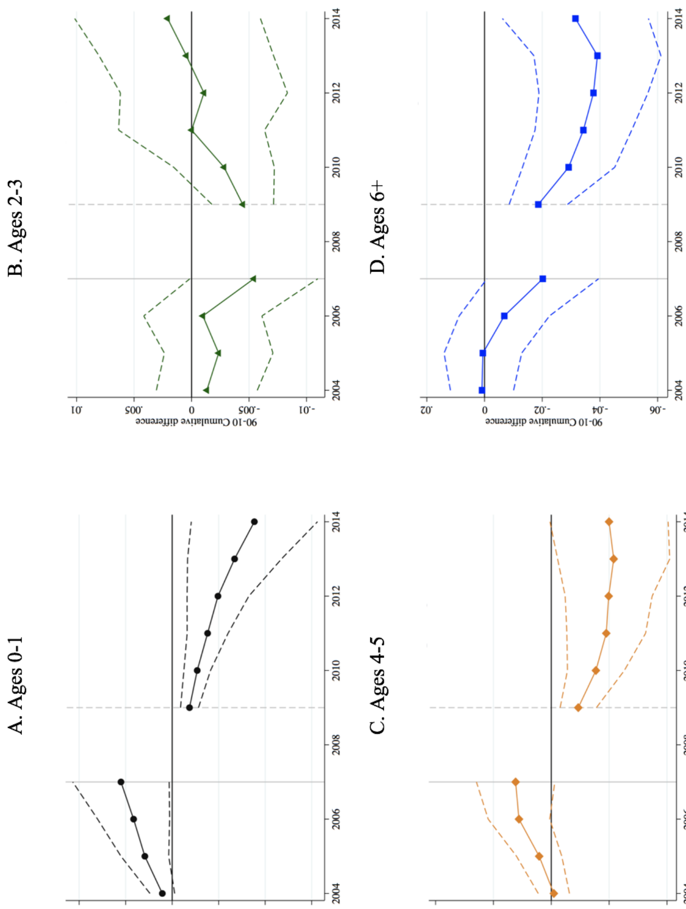
employment growth 17.2% less than the best Great Recession CZs. The functions show that the differences between regions grew substantially even after the crisis ended. To first understand how young firms and startups respond after a crisis. We document patterns of net entry by depicting cumulative establishment growth. Net entry rates remain depressed in the severely affected regions. By 2014, the CZs in the first decile of Great Recession employment growth had establishment count growth 10.8% lower than that of the least affected local areas. This discrepancy may be driven by differences in entry rates, exit rates, or some combination of the two. Strikingly, in both employment and net entry the most severely and least affected Commuting Zones continued to diverge from each other years after the crisis ended.

To disentangle the divergence in employment growth and show that the decline in net entry rates is driven by the startup deficit, we estimate Equation 1 using job creation in different age categories as the outcome variable. Figure 4.4 is interpreted in the same manner as Figure 4.3 and shows cumulative job creation differences in startups and mature firms.[11] Between 2007 and 2014, the the most severely affected Commuting Zones saw approximately 4% less job creation in startup firms with much of the decline occurring during the recovery years, a period in which the disparity in startup job creation tripled. Net job creation in mature firms, by contrast, seems to be relatively level over the recovery. Taken together, the set of four coefficient plots in Figure 4.4 demonstrate that much of the disparity in employment growth is driven by the youngest firms.

The descriptive evidence presented here suggests that many regions did not rebound during the recovery. This finding is in line with other long time-horizon studies that document the persistence of regional growth differences (e.g., Kline and Moretti, 2014). We attempt to understand why this is the case by investigating the way re-

[11]The figure plots job creation in the non-tradeable sector corresponding to NAICS codes 44-45 and 72 referring to retail and food and beverage and accommodation establishments. We explain the logic behind using the non-tradeable sector in Section 3.

Figure 4.4: Impulse Response Function



Notes: The graphs plot the coefficient of actual Great Recession employment changes from Equation 1 using cumulative job creation up to time t in an age category as the dependent variable. The Great Recession employment change is the log growth in employment between 2007 and 2009. The variable is scaled such that the coefficient reflects the difference in the outcome variable between the CZs at the 90th and 10th percentiles of Great Recession employment changes.The y-axis is the cumulative difference in the dependent up until time point t of between the 10th and 90th percentile Great Recession shock. The grey lines represent the Great Recession period. The pre-crisis points plot the cumulative difference from 2003 to time t. The post-crisis period points represented by the circles reflect the cumulative difference since 2007. 95% confidence bounds are shown by the dashed lines. Standard errors are robust.

gions respond to short-term local industry shocks. These shocks will have hit regions on a shorter-term basis and impacted decisions of entrepreneurs to enter the market or older firms to expand or contract. Regions receive additional, smaller shocks throughout the sample period and understanding how firms responded to these local shocks lends insight into why regions did not rebound after the crisis. Our empirical discussion explores how firms of different ages translate short-term local shocks into job creation before and after the crisis.

4.4 Discussion of the Bartik Shock Framework

To understand the relationship between job creation and firm age, we implement an empirical strategy centered around the Bartik (1991) shock. We consider how growth in manufacturing industries affects the demand for labor–and therefore income. A region's exposure to national changes in labor demand motivate a Bartik (1991) shock measure to capture local shocks. We treat region c as a small, open economy with K workers and measure how shocks to the nation affect region c's job creation along the firm age distribution.

Regions produce both traded goods and a non-traded good. The traded goods sector produces differentiated products whose demand comes from both region c and all other regions $-c$. The non-traded goods sector produces a homogeneous good for region c's consumption only. We abstract away from commuting across regions for consumption of the local good. The labor-market outcome of interest for region c is the change in non-traded industry, denoted by N, employment in firms of age a, $\Delta L^a_{N,c}$.

Worker k is employed in either the traded goods sector or the non-traded goods sector and earns income $I_{k,j}$ and uses this income to consume non-traded goods, traded goods produced locally in c, and traded goods produced in $-c$. Firms pay workers wages from consumers' expenditures on their products. Let $j \in \{N, T\}$, where N (T) denotes the non-traded (traded) sector. Region c's total income is thus,

$$W_c = \underbrace{\sum_{k=1}^{K} I_{k,N}}_{W_{c,N}} + \underbrace{\sum_{k=1}^{K} I_{k,T}}_{W_{c,T}}.$$

Income from the non-traded goods sector, $W_{c,N}$, is determined endogenously in region c, whereas the traded goods sector's income, $W_{c,T}$, has an exogenous component coming from $-c$'s demand for c's traded goods. The basis of our empirical approach is to exploit changes in income arising from external changes in demand for region c's tradeable goods. Changes in demand for c's tradeable goods translate into shocks in the demand for local labor.

To capture these external effects of industry-wide changes on local labor demand, we employ a Bartik shock (Bartik, 1991). The Bartik shock is a shift share approach to exogenously shift labor demand. In using the Bartik shock, we are non-trivially assuming that industry composition in the traded goods sector is fixed in the short-run. Because local labor supply is fixed in the short-run, changes in either industry productivity or national demand for tradeable products affects the marginal product of labor and shifts local labor demand.[12] These shifts in local labor demand therefore change local traded goods sector wages. Since workers use these wages to consume local goods, changes in wages in the tradeable goods sector translate into demand shifts in the non-tradeable sector goods market. At the end of the period, firms in the non-tradeable either enter, exit, expand, or contract to adjust to demand and we define $\Delta L^a_{N,c}$ as our variable of interest.

The Bartik variable allows us to capture external shocks to the local economy without conflating changes in the value of local amenities, region-specific productivity

[12]We have described a framework around the premise of shifts in the demand for goods. However, in practice a Bartik shock to employment does not disentangle productivity shocks from demand shocks. The effects on local labor demand, however, are unchanged with the nature of the shock.

shocks, or market power by local employers.[13] Empirically, we calculate the Bartik shock using data from County Business Patterns. Along the lines of Adelino, Ma, and Robinson (2016), we define the Bartik shock to region c at time t as,

$$B_{c,t} = \sum_{j=1}^{J} \omega_{c,j,(t-\tau)} \times g_{-c,j,t}, \tag{4.2}$$

The first component of the Bartik shock is the industry weight, $\omega_{c,j,(t-\tau)}$, which captures the region's exposure to these national-level demand or productivity shocks. $\omega_{c,j,(t-\tau)}$ denotes the share of total employment across *all* sectors in region c attributable to non-traded goods industry j. The second Bartik component is industry j employment growth outside CZ c, $g_{-c,j,t}$, from $t - \tau$ to t. The industry growth variable reflects either changes in national tradeable industry productivity or shifts in national demand for the industry's goods. If we do not exclude region c from the tabulation, then we cannot separate industry and national-level shocks from CZ-specific productivity and demand. Industry j comes from the set of 86 4-digit NAICS manufacturing sub-sectors.

4.5 Explanation of Empirical Approach and Results

We began by discussing how regions differentially impacted by the Great Recession significantly diverged in economic outcomes post-crisis. Our hypothesis is that regions must be responding to short-term local shocks differently after the crisis than they were before the crisis. For example, the year-to-year changes in Figure 4.3 and Figure 4.4 reflect a region's response to a short-term change in income. Therefore, we use the Bartik shock from Equation 2 as the explanatory variable of interest in regressions for job creation along the firm age distribution.

[13]For example, using actual or predicted variation in wages would reflect any number of these factors. Diamond (2016) provides a discussion.

Since the non-traded goods sector is consumed only by individuals in its CZ, we are able to estimate the direct effect an exogenous change in labor demand has on local job growth. The Bartik shock captures exogenously generated income from the tradeable sector to a Commuting Zone. Because the Bartik shock describes exogenous changes in the tradeable (manufacturing) sector, we measure job creation in the non-tradeable sector.

4.5.1 Main Empirical Specification

So far we have documented that regions hit hardest by the Great Recession did not appear to rebound during the recovery. In fact, many regions continued to fall behind local areas less affected by the crisis. Moreover, Figure 4.4 shows that divergence between these Commuting Zones began with the crisis and manifested itself primarily through a reduction in startup employment. In the periods before the crisis and during the recovery, local areas receive shocks to labor demand stemming from the composition of their industrial bases. The ability of regions to translate these exogenous shocks to their local economy into new jobs dictates the extent to which regions recover from crises. For example, we expect that as local areas come out of the crisis, they begin receive comparatively favorable draws from the shock distribution which affects their ability to regain jobs and converge back to their pre-crisis levels.

To investigate regional responsiveness to local labor demand shocks, we implement an empirical approach similar to that of Adelino et al. (2016). We use shocks to the tradeable sector to measure job creation in the non-tradeable sector. We define the tradeable sector as manufacturing industries (NAICS 3100-3399) and the non-tradeable sector as retail, restaurants, and accommodations (NAICS 44-45). We capture local demand shifts by directly introducing the Bartik shock into a reduced form regression of employment growth along the lines of Charles, Hurst and Notowidigdo (2013), Diamond (2016) and Notowidigdo (2011). Our headline fixed effects model to estimate is,

$$\Delta L^{a}_{N,c,t} = \alpha_0 + \beta^{a} \times B_{c,t} + \gamma^{a} \times X_c + \delta_t + \epsilon_{c,t}, \tag{4.3}$$

setting $\tau = 2$ to see local responses to short-term shocks. X_c includes the CZ-specific controls of log CZ total employment in 2000 from QCEW, share of persons aged 25+ that are high school educated taken from the 2000 Decennial Census, and total wages in 2000 from the IRS Statement of Income. δ_t are year fixed effects. Job creation, $\Delta L^{a}_{N,c,t}$ is scaled to total non-tradeable sector employment in 2000 to facilitate comparisons in job creation across regions and over time. The job creation variable captures net job creation between $t-2$ and t. By construction, all jobs in the 0-1 age group enter the model as net job creation. All variables are at the county-level and we aggregate to the CZ-level. We consider 2011 as the first year in the recovery sample because our job creation at time t reflects job creation over the previous two years. 2009-2011 is therefore the first interval outside of the 2007-2009 recession.

In interpreting the coefficients, care must be given since in practice our Bartik manufacturing shocks tend to be negative. Figure 4.1 Panel A shows approximately three-quarters of the shock distribution is below 0. The location of the distribution reflects the secular decline in manufacturing over the past several decades. Unless we extrapolate to out-of-sample positive shock values, we are unable to say that the effects we see are from Commuting Zones receiving positive shocks. We instead conservatively interpret our results as answering if CZ with comparatively better draws from the shock distribution have better/worse outcomes compared to those with worse draws. Nevertheless, we believe we are still able to understand how regions collectively respond to an economic crisis. For example, suppose a region was in the bottom percentiles of performance during the Great Recession. Even if a CZ receives negative shocks post-crisis, so long as its shocks are less severe in magnitude, the

CZ should still make up ground to the CZs that performed better during the Great Recession.

Our empirical analysis will show that the CZs appear to be responding differently to shocks after the crisis than they did before the crisis began. Rather than translating better shocks into relatively more non-tradeable startup and non-tradeable aggregate job creation, these shocks corresponded to *less* job creation after the crisis. We further find that this tendency is largely driven by the Commuting Zones suffering the most severe employment losses between 2007 and 2009.

4.5.2 Bartik Manufacturing Shocks and Retail Employment

Our main tests to understand the reasons regions continue to diverge after the crisis utilize the local labor demand shock variable. By comparing pre- and post-crisis responsiveness to income shocks, we demonstrate the link between an exogenous shock to the local industrial base and non-tradeable job creation fundamentally changed during the recovery. We further show that local areas below the median of Great Recession employment growth drive the shift in the relationship.

Table 4.1: OLS Regresions of job creation in non-tradable sector on manufacturing Bartik

* 5%, ** 1%, *** 0.1%	Non-overlapping sample: 01, 03, 05, 07					Non-overlapping sample: 11, 13				
	Aggregate	0-1	2-3	4-5	6+	Aggregate	0-1	2-3	4-5	6+
Mfg Bartik	0.337**	0.207***	0.002	-0.035	0.163	-1.138***	-0.531***	-0.010	0.038	-0.635**
	(3.11)	(3.02)	(0.09)	(-1.47)	(1.75)	(-4.80)	(-5.07)	(-0.17)	(0.93)	(-2.88)
log(Total employment)	-0.002	0.005	-0.001	-0.000	-0.005	0.007	0.007	-0.001	-0.002*	0.003
	(-0.32)	(1.22)	(-0.94)	(-0.35)	(-0.98)	(1.13)	(1.80)	(-0.90)	(-2.48)	(0.73)
% High school educated	-0.008	0.071***	-0.011**	-0.003	-0.065***	-0.021	0.048***	-0.018***	-0.001	-0.050**
	(-0.48)	(4.85)	(-3.29)	(-1.07)	(-5.60)	(-1.03)	(3.59)	(-4.34)	(-0.27)	(-3.20)
log(Total CZ wages)	-0.005	-0.001	0.001	-0.000	0.005	0.002	-0.001	0.001	0.002*	0.001
	(0.84)	(-0.17)	(0.50)	(-0.20)	(1.00)	(0.34)	(-0.41)	(0.68)	(2.19)	(0.22)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	2056	2056	2056	2056	2056	1008	1008	1008	1008	1008
R-sq	0.081	0.137	0.019	0.047	0.080	0.143	0.183	0.046	0.041	0.057

Notes: The table shows regressions of net job creation in NAICS 44-45 and 72 on the local manufacturing Bartik shock calculated at 2-year intervals. The shock interacts the share of total CZ employment across all manufacturing subsectors at the 4-digit NAICS level with the national growth of the subsector outside of the focal CZ. Net job creation is scaled by total CZ employment in 2000 to facilitate consistent comparisons of job creation over time adjusted for CZ size. Regressions are run separately for each age category (columns). The regressions use a non-overlapping sample in pre-recession years 2001, 2003, 2005, and 2007 and similarly for post-recession years. Control variables include log total CZ employment in 2000 from the QCEW, share of 25+ year olds with at least a high school education from the 2000 decennial Census, and log total CZ wages in the 2000 IRS Statement of Income. Parentheses report t-statistics. Standard errors are clustered at the level of 553 CZs in the pre-crisis sample and 531 in the post-period. Regressions are weighted by log 2000 total CZ employment.

We first estimate Equation 3 separately for the pre- and post-crisis sample periods in Table 4.1. We weight our regressions by log total CZ employment from QCEW in 2000 to upweight areas of larger size. In our regressions, we use two-year non-overlapping growth rates and consider the recovery periods to be 2009 to 2011 and 2011 to 2013 in one set of regressions and 2010 to 2012 and 2012 to 2014 in another. In the pre-crisis sample, we find a strong and positive coefficient from better income shock draws in startup and aggregate job creation. No other age groups display significant values. This suggests that before the crisis, CZs with comparatively better draws saw more job creation, largely from startups, than those with worse draws. A one percent increase in the Bartik shock value corresponds to a .21% increase in new establishment job creation. Given that non-tradeable job creation rates are generally around 5% before the crisis this represents a non-trivial effect on new establishment job creation. Our results further suggest almost 60% of aggregate net job creation in a CZ from an increase in the local labor demand shock is accounted for by new establishments.

In the post-crisis sample we see a complete reversal of the startup and aggregate employment coefficients. Using the 2011/2013 non-overlapping sample, the coefficient on aggregate job creation falls to -1.32% and the coefficient on the startup age bin drops to $-.62\%$. This suggests that during the recovery, CZs drawing better shocks actually saw comparatively lower levels of startup employment and aggregate job creation. Given that non-tradeable startup job creation rates are on average around 4% in the post-crisis period, this represents a substantial magnitude. Though we find negative effects in the mature firm age category in 2011 and 2013, this effect is not robust to recovery sample years and goes away in the 2012/2014 sample.

Table 4.2: OLS Regressions of post-crisis job creations by severity of Great Recession employment losses

* 5%, ** 1%, *** 0.1%	Severe=0					Severe=1				
	Aggregate	0-1	2-3	4-5	6+	Aggregate	0-1	2-3	4-5	6+
Mfg Bartik	-0.731*	-0.387*	0.034	0.093	-0.471	-1.505***	-0.678***	-0.017	-0.001	-0.809**
	(-2.15)	(-2.00)	(0.30)	(1.56)	(-1.47)	(-4.76)	(-5.59)	(-0.25)	(-0.03)	(-2.75)
log(Total employment)	0.008	0.010	-0.003	-0.003	0.003	0.006	0.006	-0.001	-0.002*	0.003
	(0.47)	(1.02)	(-0.75)	(-1.35)	(0.26)	(1.08)	(1.51)	(-0.47)	(-2.32)	(0.69)
% High school educated	-0.064*	0.017	-0.015**	-0.004	-0.061**	0.027	0.087***	-0.022**	0.002	-0.040
	(-2.31)	(0.95)	(-2.67)	(-1.09)	(-2.81)	(0.92)	(4.67)	(-3.27)	(0.47)	(-1.83)
log(Total CZ wages)	0.000	-0.004	0.002	0.002	0.000	0.003	-0.001	0.000	0.002	0.002
	(0.01)	(-0.45)	(0.68)	(1.27)	(-0.01)	(0.61)	(0.22)	(0.20)	(1.91)	(0.58)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	503	503	503	503	503	505	505	505	505	505
R-sq	0.103	0.120	0.040	0.040	0.038	0.202	0.282	0.059	0.048	0.087

Notes: See Table 4.1. Severe=0 includes only thoe CZs above the 50th percentile in growth rates from 2007 to 2009. Severe=1 includes only those CZs below the 50th percentile of growth rates during the 2007-2009 recession. The sample includes non-overlapping years 2011 and 2013. Standard errors are clustered by the 264 CZs in the Severe=0 group and 263 in Severe=1. Regressions are weighted by log total CZ employment in the year 2000.

The change in CZ response to a comparatively favorable Bartik shock may be in line with the observation that CZs facing severe employment declines during the recession continued to diverge from the least affected areas. We conduct two tests of this hypothesis. First, we split the sample based on the CZ being above or below the median employment growth from 2007 to 2009. Table 4.2 show that while both groups see the negative effect on the Bartik coefficient, the severe group is much stronger. Table 4.3 summarizes the findings in our main group of interest–startup establishments.

The split sample suggests that the CZs facing the harshest Great Recession performance are also those that responded negatively to (better) local shocks after the crisis. Less job creation in startup establishments account for anywhere from 45 to 74% of the net job losses arising from a one percent increase in the shock. In Table 4.3 we summarize our results for the startup establishment age bin using the 2011/2013 fixed effects regression as the post-crisis period. Here, it is easy to observe the large extent to which the results are driven by the CZs facing the most severe Great Recession employment losses.

The split sample regressions implicitly make the additional assumption that the control variables differentially affect the two recession groups. To relax this assumption, we more rigorously assess the heterogeneous effect Bartik shocks have on CZs. We interact an indicator variable designating whether the CZ is below the median Great Recession growth with the Bartik manufacturing shock. The interaction terms show the local industry shock effect by recession group. Formally we estimate,

$$\begin{aligned}\Delta L^a_{N,c,t} = \alpha_0 + \mathbb{1}\left[Below\ median\right]_c + \beta^a \times \mathbb{1}\left[Below\ median\right]_c \times B_{c,t} + \\ \gamma^a \times \mathbb{1}\left[Above\ median\right]_c \times B_{c,t} X_c + \epsilon_{c,t} \quad (4.4)\end{aligned}$$

Table 4.3 shows the estimates for Equation 4. The “Below median” (“Above median”) variable refers to indicators for CZs below (above) the median employment

Table 4.3: OLS regressions splitting sample by severity of Great Recession

* 5%, ** 1%, *** 0.1%	All CZs		Severe=0		Severe=1	
	Pre-crisis	Post-crisis	Pre-crisis	Post-crisis	Pre-crisis	Post-crisis
	0-1	0-1	0-1	0-1	0-1	0-1
Mfg Bartik	0.207**	-0.531**	0.141	-0.387*	0.238*	-0.678***
	(3.02)	(-5.07)	(1.39)	(-2.00)	(2.50)	(-5.59)
log(Total employment)	0.005	0.007	0.012	0.010	0.003	0.006
	(1.22)	(1.80)	(1.21)	(1.02)	(0.85)	(1.51)
% High school educated	0.071***	0.048***	0.050*	0.017	0.110***	0.087***
	(4.85)	(3.59)	(2.51)	(0.95)	(5.58)	(4.67)
log(Total CZ wages)	-0.001	-0.001	-0.007	-0.004	-0.000	-0.001
	(-0.17)	(-0.41)	(-0.75)	(-0.45)	(-0.14)	(0.22)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	2056	1008	1028	503	1028	505
R-sq	0.137	0.183	0.103	0.120	0.197	0.282

Notes: Severe=1 indicates regions above the median in terms of Great Recession employment *declines*. The pre-crisis sample includes odd years between 2001 and 2007. The post-crisis period includes non-overlapping year 2011 and 2013. The table shows that CZs least affected by the Great Recession also did not experience a run-up in startup employment pre-crisis (Column 3). The most affected CZs underwent substantial employment growth in startups pre-crisis (Column 5) and saw a substantial drop in gross job creation in startups after the crisis. The difference between the Bartik coefficient in Columns 4 and 6 are significant at the 10% level. The F-stat on the difference-in-differences Bartik change between least and most affected CZs is also statistically significant.

growth between 2007 and 2009. We find robust evidence that places facing the more severe downturn during the recession see a strong negative response to a percentage increase in the Bartik shock in age 0-1 employment after the crisis. The results on the interactions between the Above median indicator and the manufacturing Bartik are robust to post-crisis sampling period, suggesting the inverse relationship between local labor demand shocks and startup employment rates are driven by the severely affected CZs. Before the crisis, startups in both groups responded favorably to better draws from the Bartik distribution. Interestingly, the CZs that would eventually experience below median Great Recession employment growth saw the strongest response to increases in shocks to the local industrial base. Our results are robust to defining the groups as being above or below the 10th percentile.

We augment our sample with data from the National Establishment Time Series to see to which of these two groups entering establishments belong. Our NETS sample includes only manufacturing establishments so our results in Table 4.4 are not a one-to-one matching to the retail and restaurants job creation figures, but the analysis is once again suggestive of a shift in responses to pre- and post-crisis local labor demand shocks.

NETS is a panel version of Dun & Bradstreet identified establishments with employment, PayDex Scores, and a number of additional firm- and establishment-level variables added by Walls & Associates. Establishments are defined as unique business lines at a location. Following prior literature (e.g., Neumark, Wall, and Zhang, 2011), we retain records with imputations in fewer than 50% of establishment-year observations. We further cleanse establishments that never exceed 3 employees to reduce the risk of establishments that enter, but do not engage in any meaningful business activity. We identify new establishments as Dunsnumbers that appear in the sample for the first time. NETS also provides us with a headquarters identifier that is the same across all establishments under the same firm. We consider an establishment entrant to be a new firm if its HQ identifier appears for the first time. An entering establishment is an existing firm expansion if the Dunsnumber is new, but the HQ identifier is already in existence. We also measure establishment exits as Dunsnumbers that appeared in year t but not $t+1$. Just as in our job creation variables, we sum the counts of establishment entry and exit and scale by establishment counts in a CZ in the year 2000 to facilitate consistent interpretation across categories.

We continue to estimate Equation 3, but substitute in for the dependent variable CZ exit rates, entry rate from new firms, and entry rates from existing firms. Table 4.4 shows that while a one percent increase in the local labor demand shock increased entry rates of establishments from new firms and of existing firms, the effect on existing firms was 3 times as strong. As in our QWI sample presented earlier, we continue to find evidence that increases in the Bartik shocks are associated with declines in entrant activity post-crisis. The decline in entry, however, is driven largely by fewer existing firm expansions.

In sum, the QWI and NETS results seem to complement each other in that a one

Table 4.4: OLS regressions of establishment entry and exit

* 5%, ** 1%, *** 0.1%	Pre-crisis			Post-crisis		
	Exit rate	Entry from new firms	Entry from existing firms	Exit rate	Entry from new firms	Entry from existing firms
Mfg Bartik	0.081* (1.95)	0.041** (2.01)	0.126** (2.52)	-0.171 (-1.53)	0.015 (0.81)	-0.271*** (-4.64)
log(Total employment)	-0.054 (-0.68)	-0.072 (-0.56)	-0.091 (-0.65)	-0.033 (-0.16)	-0.007 (-0.26)	-0.049 (-0.43)
% High school educated	-0.011** (-2.33)	-0.002 (-0.50)	-0.024*** (-3.09)	-0.011 (-1.17)	0.002 (0.79)	-0.009*** (-2.87)
log(Total CZ wages)	0.301*** (4.00)	0.056 (0.47)	0.214 (1.61)	0.167 (0.87)	-0.006 (-0.23)	0.0571 (0.54)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	837	837	837	396	396	396
R-sq	0.702	0.558	0.778	0.251	0.609	0.752

percent increase in the shock correlate to an increase in entrant activity before the crisis and less after the crisis. We link these findings to the idea of economic scarring. The reversal in response to Bartik shocks is far more pronounced in regions below the median in Great Recession employment growth. The implication is that even if these regions received Bartik shocks better than those less affected by the crisis, they would not return to their pre-crisis levels. Indeed, it appears the Great Recession had a significant lasting impact on startup activity that persisted through the recovery.

While we have confirmed existing literature in the pre-crisis sample and documented several additional findings, we have not discussed why the pre-crisis and recovery periods look so different. The mechanism underlining the changing nature of a region's reaction to a shock is less understood, though we explore several possibilities.

4.6 Potential Causes of the Change in Startup Response to Bartik Shocks

The finding that increases in the Bartik shock correlated to less job creation in non-tradeables and fewer entrants is striking. However, answering why this reversal in the

Bartik effect takes place is a far more dificult challenge. While we cannot definitively assert a mechanism, we do believe we can eliminate some seemingly obvious culprits and discuss the mechanics of the Bartik shock that may be at work.

4.6.1 Labor Market Slack

One possibility is that our results reflect the fact that places receiving worse shocks have more slack in their local labor markets. Fairlie (2013) notes that cities most severely impacted by the Great Recession actually saw more startup businesses. He finds this because when wage employment opportunities disappear, people enter self-employment.

To address this hypothesis, we compare Commuting Zones with unemployment rates below and above the median. Prior studies suggest that places with higher unemployment rates see more entry into entrepreneurship (e.g., Fairlie, 2013). The literature generally posits that the reason is when formal wage-employment opportunities decline and labor market conditions deteriorate, individuals enter self-employment as an alternative. If this is the case, then we might expect CZs with higher unemployment rates to have more employment in new establishments and the negative coefficient on the Bartik shock to be driven by CZs with lower unemployment rates.

Using employment data from the Bureau of Labor Statistics, we obtain county-level unemployment and employment levels. We calculate unemployment rates in year t as the average unemployment rate between year t and $t-1$. Next, we create an indicator variable for high unemployment and another for low unemployment, defined as whether the CZ-year observation is above or below the median unemployment rate. In the pre-crisis sample we use the median rate between 2000 and 2007 and in the post-crisis sample we use the median rate between 2011 and 2014. Figure 4.1 Panel C shows the distribution of CZ unemployment rates. We estimate Equation 4, but using the unemployment level indicator.

Table 4.5: OLS regressions showing heterogeneous effects by labor market quality

* 5%, ** 1%, *** 0.1%	Pre-crisis					Post-crisis				
	Aggregate	0-1	2-3	4-5	6+	Aggregate	0-1	2-3	4-5	6+
High unemployment * Mfg Bartik	0.285*	0.194**	0.011	-0.037	0.116	-0.852***	-0.400***	-0.040	0.016	-0.428
	(2.39)	(2.67)	(0.40)	(-1.31)	(1.14)	(-3.31)	(-3.79)	(-0.63)	(0.36)	(-1.77)
Low unemployment * Mfg Bartik	0.376***	0.250**	-0.029	-0.039	0.195	-1.599***	-0.740***	0.041	0.074	-0.974***
	(2.59)	(2.92)	(-0.84)	(-1.28)	(1.47)	(-5.80)	(-5.41)	(0.53)	(1.58)	(-3.80)
Main Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	2056	2056	2056	2056	2056	1008	1008	1008	1008	1008
R-sq	0.084	0.137	0.021	0.048	0.084	0.151	0.189	0.047	0.043	0.061

Notes: See Table 4.3. High unemployment is an indicator variable coded as 1 if the CZ is in the top third of unemployment rates. Low unemployment is similarly defined as the lowest third. All regressions include main effects of interaction terms as well the full vector of control variables as in all other OLS tables. A Wald test shows that the difference in the high unemployment interaction and low unemployment interaction in Column 2 of the post-crisis panel is significant. The difference shows CZs with higher unemployment rates post-crisis generate more jobs in startups relative to places with lower levels of unemployment after the crisis. There was no statistical significance in the difference of coefficients in Column 2 of the pre-crisis panel.

Table 4.5 shows that in both the pre- and post-crisis samples we do not see heterogeneous effects of the Bartik shock by unemployment group. Pre-crisis, the difference in the Bartik shock effects between the high and low unemployment groups is not statistically significant in either the aggregate (0.302 and 0.386, respectively) or 0-1 bin (0.198 or 0.241, respectively). Moreover, the point estimates are close to the overall point estimate from Table 4.1 (.352 for aggregate and .207 in 0-1). In the post-crisis regression for years 2011 and 2013, we see that the negative effect of Bartik is indeed stronger in the low unemployment group than the high unemployment group. However, in both cases the effect is still negative, meaning that even though a CZ may have slack in its labor markets, it still sees a negative Bartik effect. The negative coefficient in the low unemployment interaction goes away in the 2012 and 2014 group further suggesting that local labor market conditions are not driving the negative coefficient in the post-crisis Bartik shock.

4.6.2 Local Credit Availability

Another possible explanation for the reversal in the sign on the Bartik shock coefficient after the crisis is that increases in Bartik manufacturing shocks were mitigated by less credit availability. This hypothesis asserts that banks either became constrained or more risk-averse and did not lend as much to new firms and small businesses, thereby contributing to the decline in startup activity. To determine extent to which banking affects firm growth along the firm age distribution, we follow Greenstone, Mas, and Nguyen's (2014) and construct a Great Recession credit availability measure for each CZ.

Following Greenstone et al.'s procedure, we obtain county-level small business loan data from the Federal Financial Institutions Examiners Council Community Reinvestment Act disclosure files. One data limitation we face is that lending information by firm age is unattainable and instead must be proxied by small business lending which will include mature firms as well. The difficulty in using loan originations as a measurement of credit availability is that it reflects not only the supply of credit but also demand for credit. Loan originations can fall simply because businesses do not need credit for expansions, for example, during a downturn. To attempt to separate these effects, we employ a Bartik-style shifter for small business loans by bank branch j in each CZ. The identification strategy exploits the fact that each bank j can have branches in multiple CZs c and a CZ can have multiple banks. In the first step regress,

$$log(\frac{l_c^t}{l_c^{t-2}}) = d_c^t + s_j^t + e_{c,j}^t, \tag{4.5}$$

where the LHS represents the growth in total small business lending in CZ c over a two year period $t-2$ to t, d_c^t CZ fixed effects captures local demand factors (health of the local economy), and the remaining s_j^t bank fixed effects captures credit supply net of local demand. We run the regression over each year between 2000 and 2014. We obtain estimates of $s_j \forall t$. We use these estimates to construct a measurement of local credit availability in CZ c at time t as,

$$p_c^t = \sum_j \omega_{c,j}^t \times s_{c,j}, \tag{4.6}$$

where $\omega_{c,j}$ is the share of bank j's small business lending of the total in CZ c and $s_{c,j}$ is transformed to have bank asset-weighted mean 0. Figure 4.1 Panel B shows the distribution of these credit availability shocks in the pre- and post-crisis sampling periods.

Table 4.6: OLS regressions showing heterogeneous effects by credit availability

* 5%, ** 1%, *** 0.1%	Pre-crisis					Post-crisis				
	Aggregate	0-1	2-3	4-5	6+	Aggregate	0-1	2-3	4-5	6+
High credit availability * Mfg Bartik	0.204	0.241*	0.031	-0.045	-0.022	-0.636	0.060	0.061	-0.013	-0.743**
	(1.33)	(2.42)	(0.80)	(-1.35)	(-0.17)	(-1.95)	(0.37)	(0.75)	(-0.24)	(-2.62)
Low credit availability * Mfg Bartik	0.405***	0.191**	-0.012	-0.030	0.256*	-1.302***	-0.747***	-0.042	0.057	-0.570*
	(3.55)	(2.75)	(-0.46)	(-1.08)	(2.47)	(-4.99)	(-6.21)	(-0.72)	(1.33)	(-2.42)
Main Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	2056	2056	2056	2056	2056	1008	1008	1008	1008	1008
R-sq	0.082	0.137	0.020	0.020	0.082	0.159	0.206	0.053	0.043	0.067

Notes: See Table 4.4. High credit availability is an indicator variable coded as 1 if the CZ is in the top third of the credit measure. Low credit availability is similarly defined as the lowest third. The credit availability measure is constructed similar to Greestone, Mas, and Nguyen (2014) where the credit measure is a weighted average of bank fixed effects in a regression of bank small business lending in the CZ on CZ and bank fixed effects. All regressions include main effects of interaction terms as well the full vector of control variables as in all other OLS tables.

We substitute this bank shock variable in for the Bartik shock variable in Equation 3 and obtain estimates of β over the pre- and post-crisis period. We are unable to find a significant effect either pre- or post-crisis from the Greenstone et al. credit shock. In Table 4.6, the variable "Low credit shock" is an indicator taking the value of 1 if the credit shock the CZ faces over a two-year interval is in the bottom half of the credit shock distribution. The table shows that in the 2000-2007 period, better credit shocks accentuated the positive Bartik effect, suggesting better credit conditions facilitate job creation. However, the effect is not driven by startups. The difference in the Bartik effect between high and low credit shocked areas is due to the positive effect credit availability has for mature businesses aged at least six years. It is likely that these older establishments are those that account for much of the small business lending activity.

Taking the pre-crisis results may suggest that better credit shocks perhaps either attenuate the negative Bartik effect or even turn the effect positive. However, we find evidence suggesting the opposite is the case. The negative Bartik effect seems to only coincide with better credit shocks. We interpret this finding as simply suggesting that the effect of the Bartik variable swamps any impact credit conditions may be having, though it is curious that regions with less credit availability see no negative effect on employment. Greenstone et al. similarly find no relationship between credit lending shocks and county-level employment outcomes.[14]

4.6.3 The Mechanics of the Bartik Shock Variable

The Bartik shock's construction may itself drive the results. We decompose the measure to understand the variation the instrument captures. We run a procedure to decompose the Bartik shock into its industry components. We believe this procedure

[14]Davis and Haltiwanger, 2017, similarly use the Greenstone et al. measure and find no evidence local credit conditions explain the decline in startup growth.

will allow us to understand if certain industries are driving the results or if the shock generates mechanical coefficients.

We regress job creation in age bin 0-1 on the $\omega_{c,j,(t-\tau)} \times g_{-c,j,t}$ component of the Bartik shock separately in each industry-year. We obtain coefficients on each of these industry components. We are able to identify which industries switch signs from positive to negative, and obtain a rough measure of the weight an industry receives in the overall Bartik shock. Suppose we conduct cross-sectional univariate regressions of non-tradeable startup job creation on the Bartik manufacturing shock, excluding the control variables, in each year between 2000 and 2014. Further, let $x_j = \omega_{c,j,(t-\tau)} \times g_{-c,j,t}$ where $j \in \{1, 2\}$. Also, let ΔL^a be job creation in age bin a. Then we can write the coefficient on the Bartik shock as,

$$
\begin{aligned}
\beta^a &= \frac{Cov(\Delta L_N^a, B_c)}{Var(B_c)} \\
&= \frac{Cov(\Delta L_N^a, \sum_j x_j)}{Var(\sum_j x_j)} \\
&= \frac{1}{Var(\sum_j x_j)} \sum_j Cov(\Delta L_N^a, x_j) \\
&= \sum_j \frac{Var(x_j)}{Var(\sum_j x_j)} \frac{Cov(\Delta L_N^a, x_j)}{Var(x_j)} \\
&= \sum_j \frac{Var(x_j)}{Var(\sum_j x_j)} b_j \\
&= \frac{\sum_j Var(x_j)}{Var(\sum_j x_j)} \left(\sum_j \frac{Var(x_j)}{Var(\sum_j x_j)} b_j\right)
\end{aligned}
$$

where Cov and Var are covariance and variance, respectively. This formulation shows that the Bartik shock coefficient in our regressions is a weighted sum of the coefficients on individual industry-level shocks (b_1 and b_2 are univariate OLS coefficients). The weights are the shares of Bartik variance across CZs explained by the industry.

Table 4.7: Decomposition of Bartik and Industry Loading

	Weight in Bartik		Univariate Coeff.		Employment Growth	
	Pre	Post	Pre	Post	Pre	Post
311. Food	1.13%	3.17%	-2.45	-4.72	-0.27	0.00
312. Beverage & Tobacco Products	0.51%	0.98%	0.64	-11.69	0.03	0.10
313. Textile Mills	20.37%	1.57%	0.99	26.64	-0.63	-0.08
314. Textile Product Mills	4.30%	1.47%	0.67	3.97	-0.19	-0.04
315. Apparel	26.50%	0.99%	2.14	-22.64	-0.83	-0.10
316. Leather and Allied Products	0.81%	0.22%	2.05	-17.28	-0.33	-0.14
321. Wood Product	8.52%	6.47%	-6.58	-3.34	-0.06	0.03
322. Paper	2.47%	1.23%	3.97	2.66	-0.26	-0.03
323. Printing & Related Support	0.73%	1.96%	17.55	21.20	-0.21	-0.06
324. Petroleum & Coal Products	0.91%	0.14%	-1.41	13.83	0.13	0.01
325. Chemical	0.30%	0.35%	-3.88	-26.28	-0.07	0.02
326. Plastics & Rubber Products	1.36%	1.75%	4.75	25.22	-0.18	0.05
327. Nonmetallic Mineral Products	1.43%	1.21%	-6.97	-6.74	-0.03	0.03
331. Primary Metals	2.56%	2.92%	3.14	14.05	-0.28	0.04
332. Fabricated Metal Products	3.39%	5.46%	-10.06	-6.80	0.12	0.06
333. Machinery	5.54%	14.20%	3.05	-7.55	-0.18	0.09
334. Computer & Electronic Products	7.85%	4.37%	-2.06	-4.55	-0.37	-0.05
335. Electrical Equipment, Appliances & Components	3.06%	1.50%	20.15	-62.89	-0.33	-0.01
336. Transportation Equipment	5.74%	40.15%	-6.56	-0.66	-0.12	0.15
337. Furniture & Related	1.91%	7.62%	-3.54	-5.87	-0.16	0.02
339. Miscellaneous	0.60%	2.29%	13.13	11.32	-0.03	-0.05

Notes: The above shows the results from a Bartik variable decomposition. The coefficient on the Bartik variable is a weighted average elasticity of individual regressions on each sub-industry. The weights are based on the dispersion of local industrial growth across CZs and its interaction with each CZ's share of employment in the focal industry. Summing across all CZs yields the focal industry's overall weighting in the Bartik coefficient. Columns 2 and 3 show the change in industry loading between the pre- and post-periods. Columns 3 and 4 display the coefficient from a univariate regression of startup job creation on each CZs "shift-share". These CZ-level shift-shares are calculated as the growth of the industry outside the CZ and the weight of the industry in the CZ. The final two columns are the simple log employment growth of the national industry grouping. Regressions in the prior OLS tables use NAICS-4, but for simplicity the table is aggregated to the NAICS-3 level. The table shows the substantial changes in each industry's weighting in the overall Bartik variable. and the elasticity of startup entry with respect to the industry's growth.

For ease of interpretation and to avoid noise driven by small sample sizes, we use 3-digit NAICS manufacturing industries in our Bartik decomposition. Table 4.7 displays the average industry coefficient value between 2000 and 2007 and then separately for 2011 and 2014. The table similarly provides the average weight of the industry in the Bartik shock.[15] These weights dictate the share of the overall Bartik

[15] i.e., $\frac{Var(x_1)}{Var(x_1)+Var(x_2)}$

coefficient from Equation 3 that is attributable to the industry. There are 21 NAICS-3 industries which means that if the weights were evenly distributed, each industry should receive a weight of 0.048. Using this procedure and estimating the b_j's in each period, Table 4.7 shows the sign change in the Bartik coefficient in the age 0-1 regressions is driven largely by NAICS 336 Transportation Equipment Manufacturing. Between 2000 and 2007, the transportation equipment manufacturing industry receives an average weight of 0.057 with a coefficient that is on average positive. In the post-crisis period this weighting rises to an average of 0.401 with a negative coefficient. Years 2000 to 2007 are more heavily weighted towards apparel and textiles which both see a substantial decline in weights from over 20% to around 1%. We can also compute a Bartik-style index by summing the x_j's over all industries. This gives us a Bartik variable of 0.39 from 2000-2007 and -2.16 in the 2011-2014 period, following a similar pattern to that of our headline regression.

One additional complication remains. In computing the x_j's, the individual industry coefficients and Bartik weights contain three components: Local industry weight, national industry growth outside the CZ and the interaction between the two. To disentangle the former two components, suppose we interact the CZ-industry weights with a common industry growth rate across all CZs using the national industry growth rate, g_j, such that $x_j = g_j \times \omega_{c,j,t-\tau}$.[16]

We reproduce the output from Table 4.7, but use this new definition of the x_j's in Table 4.8. Considering the industry growth component to be common across all CZs means each coefficient's sign depends on whether the industry is declining or growing in aggregate and on the covariance between the size of the industry in a CZ and startup job creation. Transportation equipment manufacturing underwent log growth of -0.12 between 2000 and 2007 and $+0.15$ between 2011 and 2014. This implies that while the overall coefficient on the industry flipped from positive to neg-

[16] We focus on decomposing the separable components of the Bartik instrument.

Table 4.8: Weight Changes Due to Within CZ Variation in Pre- and Post Industry Mix

	Weight in Bartik		Univariate Coeff.	
	Pre	Post	Pre	Post
311. Food	5.69%	7.65%	-5.84	9.68
312. Beverage & Tobacco Products	0.36%	1.57%	4.53	-12.03
313. Textile Mills	14.87%	2.25%	3.75	1.24
314. Textile Product Mills	14.87%	2.25%	-7.79	5.41
315. Apparel	20.94%	1.84%	4.91	4.52
316. Leather and Allied Products	0.71%	1.26%	16.56	3.34
321. Wood Product	15.40%	7.43%	-3.62	28.56
322. Paper	1.86%	2.08%	-4.97	-51.09
323. Printing & Related Support	0.55%	1.45%	11.80	124.74
324. Petroleum & Coal Products	0.65^	0.72%	-114.06	-4.61
325. Chemical	0.70%	0.28%	-5.22	-32.11
326. Plastics & Rubber Products	1.95%	1.52%	-30.42	-42.05
327. Nonmetallic Mineral Products	1.50%	1.61%	62.89	-12.18
331. Primary Metals	2.53%	3.30%	-1.45	-16.63
332. Fabricated Metal Products	2.63%	4.83%	-24.06	-31.68
333. Machinery	4.47%	12.19%	-10.61	-1.54
334. Computer & Electronic Products	7.45%	1.82%	-1.03	-0.54
335. Electrical Equipment, Appliances & Components	2.57%	0.19%	-14.07	56.32
336. Transportation Equipment	6.58%	31.71%	3.95	-5.59
337. Furniture & Related	3.63%	6.54%	1.94	9.83
339. Miscellaneous	1.37%	1.75%	386.87	28.55

Notes: See Table 4.7. This table repeats the same decomposition, but fixes each industry's growth rate to the national rate. In otherwords, the table does not employ the "leave-one-out" calculation. This table shows the change in industry loading attributable to changes in local industry concentration.

ative, the covariance between the industry size in a CZ and startup job creation in non-tradeables is negative in both pre-crisis and recovery periods. This suggests that part of the decline in non-tradeable job creation rates form the Bartik shock may largely be the result of an industry with a negative relationship with non-tradeable startups growing disproportionately faster. Indeed, regions at the 10th percentile of employment growth, which we showed did not recover in Section 2, have over 20% of their manufacturing and 2% of total employment in transportation equipment manufacturing. Regions at the 90th percentile, by contrast, have just 8% of manufacturing and 1% of total employment in the industry.

The industry's weights in the Bartik are dependent on both the dispersion of an industry's weight across CZs and the growth rate of the industry. We find that the correlation between industry weights over time is stable. For example, the share of CZ employment in transportation equipment manufacturing in 2000 have a correlation of 0.73 with those of 2014. The average and median absolute changes of the industry's employment share of a CZ is 1% and 0.3%, respectively. These facts are suggestive that rise in the weight of the transportation equipment manufacturing industry is also driven by increases by the industry's growth. Summing x_j over all j industries (taking the sum of $g_j \times \omega_{c,j,t-\tau}$ over all j's), gives a value of 4.01 before the crisis and 0.31 after. This seems to suggest that changes in industry mix in CZs between 2000-2007 and 2011-2014 cannot capture the decline in the overall Bartik instrument.

In our final Bartik instrument decomposition (Table 4.9, we turn to holding the employment shares, $\omega_c, j, t-\tau$, the same across CZs and only allow $g_{-c,j,t}$ to be CZ-specific. These industry shares are set to the national share of employment in industry j at $t-\tau$. Letting $x_j = \omega_{j,t-\tau} \times g_{-c,j,t}$ and comparing the pre- and post-crisis coefficients and weights, we are attempting to uncover the portion of the overall change in the Bartik instrument attributable to changes in CZ-industry growth differences. The sign on the coefficients will reflect the relationship between non-tradeable startup job creation and industry growth outside the CZ. The weights show the dispersion of these growth rates. For example, in the automobile industry, suppose Detroit auto begins to grow strongly after the crisis, but in Durham, NC it is in decline. As in the normal Bartik case, the growth rate attributed to the Detroit CZ in the calculation is Durham's industry growth whereas Durham's growth component in the calculation will be based off Detroit's rebound. This industry will likely have a relatively high weight because of the dispersion in regional industry growth rates. Table 4.9

Table 4.9: Weight Changes Due to Regional Growth Differences

	Weight in Bartik		Univariate Coeff.	
	Pre	Post	Pre	Post
311. Food	4.21%	12.01%	82.42	130.47
312. Beverage & Tobacco Products	0.74%	0.52%	103.17	-1274.61
313. Textile Mills	1.02%	0.19%	76.97	280.91
314. Textile Product Mills	0.76%	0.32%	222.78	-246.57
315. Apparel	7.98%	1.11%	93.97	-943.82
316. Leather and Allied Products	0.12%	0.11%	-103.34	130.80
321. Wood Product	0.68%	0.62%	83.06	-116.69
322. Paper	0.63%	0.61%	257.32	-43.80
323. Printing & Related Support	0.75%	0.70%	320.47	-58.00
324. Petroleum & Coal Products	0.30%	0.26%	508.12	132.65
325. Chemical	8.07%	6.68%	-33.59	6.49
326. Plastics & Rubber Products	1.82%	1.97%	109.88	201.21
327. Nonmetallic Mineral Products	0.76%	0.55%	-441.84	160.99
331. Primary Metals	1.16%	1.17%	132.68	-8.13
332. Fabricated Metal Products	4.11%	6.23%	113.62	5.83
333. Machinery	4.20%	6.72%	67.41	-31.32
334. Computer & Electronic Products	28.48%	20.68%	14.57	-11.82
335. Electrical Equipment, Appliances & Components	1.57%	1.25%	-187.26	-457.71
336. Transportation Equipment	29.69%	35.07%	21.83	27.25
337. Furniture & Related	1.33%	0.92%	-241.83	-556.80
339. Miscellaneous	1.62%	2.31%	135.69	-143.83

Notes: See Table 4.7. This table repeats the same decomposition, but fixes each CZ's industry employment share to the overal national employment share of the industry. In otherwords, the table is less influenced by the ex ante location of industries. The table represents loading driven by regional variation in industry growth.

shows quite vividly the large increase in magnitudes of the coefficients. The table also indicates that the transportation equipment manufacturing industry carries a large weight in both the pre- and post-crisis periods (29.69% and 35.07%), suggesting the regional variation in transportation equipment manufacturing growth is consistently large relative to the other manufacturing industries. Its coefficient is positive in both periods, though these coefficients are more difficult to interpret. Summing the x_j's over all j industries yields a change from 31.85 in the 2000-2007 period to -40.40

in the post-crisis period. This further suggests that shifts in the sign of the Bartik variable are driven by industries declining from 2000-2007 and then rebounding after the crisis, or vice versa. Industries with the largest x_j values in the post-crisis period are food manufacturing, beverage and tobacco manufacturing, and apparel manufacturing. Summing x_j over these three industries yields a value of -32.80. The large contribution of these three industries to this particular decomposition suggests that regional variation in their industry growth rates is substantial. Comparing Tables 9 and 10 suggest changes in national industry growth and the substantial regional variation in industry growth drive the switch in signs on our headline Bartik instrument.

Decomposing the Bartik variable into its industry components is a potentially useful form of analysis to understand the industry-determinants relating the outcome variable to the Bartik shock. These decomposition outputs show that the industry mix in the Bartik shock changed dramatically between the two periods. By unpacking the two separable components of the Bartik instrument, we find that the shock can be driven by a select few industries. When analyzing data of long time-horizons, changes in the sign of the Bartik need not reflect a fundamental change in the relationship of industries with the outcome variable, rather the changes in the sign of coefficient also reflect economy-wide and CZ-industry growth rates.

While our decomposition results suggest the large contribution of the transportation manufacturing industry to the changing sign of the Bartik instrument is largely the result of its substantial regional variation in growth rates and that the industry went from declining to growing, the inverse relationship between transportation manufacturing and non-tradeable sector startup employment is a puzzle. Why do regions with employment in this now growing industry actually see fewer new restaurants and retailers? Plausible factors could be those described by Chinitz (1961) or Glaeser, Kerr, and Kerr (2015) where the presence of large establishments and heavy manufacturing industries can dampen entrepreneurship and employment growth. Prior literature suggests regions with industries dominated by big companies and higher fixed costs have less entrepreneurship because they lack an entrepreneurial culture, face significant capital constraints as small businesses are more likely to obtain financing from financial institutions if they are in cities with small businesses, large firms obtain intermediate goods internally so regions with large firms have less developed supply chains. We do not have an answer to which mechanism is at work in the transportation equipment manufacturing industry, but we observe similarities between the this industry and the heavy manufacturing/mining industries discussed by Chinitz (1961) and Glaeser et al. Transportation equipment manufacturing includes manufactures of automobiles, aerospace, ships, and their related parts, all of which are dominated by larger firms and are characterized by captial intensity and large fixed costs.

4.7 Conclusion

From the relative stagnation in Appalachia and the Rust Belt to the fast paced technology hubs of Silicon Valley, Boston, and the Research Triangle, the spatial variation in economic activity across the United States is striking. Prior literature established substantial differences in regional economic performance stemming from human capital and concentration of well-educated workers (Berry and Glaeser, 2005; Moretti, 2012), agglomerations (Carlino and Kerr, 2014), and the abundance of entrepreneurs (Glaeser, Kerr and Ponzetto, 2010; Chatterji, Glaeser and Kerr, 2014). With most academic research focused on the contributions of a few growth oriented startups on aggregate productivity, much of the finer details about regional resilience and dynamism are missed. Most regions across the country lack such technology clusters and a quarter of the workforce is employed in the non-tradeable industries of restaurants and retail. The severity and spatially disparate effects of the Great Recession make it a unique laboratory to study the relationship between firm age and the resilience of regions.

In this paper, we use Adelino, Ma, and Robinson (2016) as a baseline for understanding the recovery from the Great Recession. We show that regions with differential exposure to the Great Recession varied strongly in their recovery from the crisis. In describing this tendency, we contribute to a long literature discussing spatial variation in economic outcomes. Moreover, we have highlighted a troubling tendency for regions to continue to diverge in economic performance.

The second phase of our analysis sought to explain the divergence of regions through their response to short-term local labor demand shocks. We employ a Bartik framework and find that while job creation and startup employment responded favorably to increases in the Bartik shock, the post-crisis period saw the opposite. The split sample shows that this relationship is largely driven by CZs in the bottom half of Great Recession employment growth. We find evidence of a heterogeneous impact of the Bartik shock variable dividing CZs into Great Recession severity groups. The negative relationship between the Bartik shock and startup job creation is stronger and more robust in regions below the median Great Recession employment growth. Using the NETS sample of manufacturing establishments, we find that much of the positive local labor demand shock on startup establishment employment growth pre-crisis came from new establishments of existing firms and not the creation of new firms. Similarly, the decline in the recovery may be driven by less entry from these same types of establishments. We note that our NETS analysis is not a perfect mapping to our headline results since our main results consider only the non-tradeable sector.

We then attempt to pinpoint causal mechanisms, but find that common explanations for differences in regional entrepreneurship do not seem to explain the striking post-crisis dynamics. We test hypotheses on labor market slack and local credit shocks. The labor market slack hypothesis suggests that places with poorer labor market conditions may see more entrepreneurship due to a lack of formal wage-employment opportunities. This implies that the shock to the local industrial base should be positive on startup job creation in places with higher unemployment rates. We do not find such a relationship.

The second mechanism we test is local credit availability. We use Greenstone, Mas and Nguyen (2015) to construct local credit supply shocks. Again, we find no evidence suggesting credit availability explains the post-crisis dynamics. Finally, we discuss the role of individual industries in contributing to local labor demand shocks. We decompose the Bartik tradeable sector shock into its industry components as it pertains to regressions of startup job creation on the shock. The Bartik variable overweights industries that are more spatially concentrated. In the post-crisis period, the transportation equipment manufacturing industry accounts for over 40% of the Bartik shock coefficient in startups. This industry is inversely correlated with non-tradeable startup employment, but is now growing and could therefore be dampening job creation in new restaurants and retail stores. We do not theorize why this manufacturing industry is negatively correlated with non-tradeable sector job creation, but prior works, notably Chinitz (1961) and Glaeser, Kerr and Kerr (2015), offer insight into the relationship between capital intensive industries and regional entrepreneurship. We believe using their frameworks to discuss the relationship between growth in industries similar in capital intensity to transportation equipment manufacturing and the decline startup job creation is well-deserving of further research.

While our results show that a percentage increase in the Bartik shock variable has a negative effect on job creation, particularly from a dampening of job creation in startups, we remain conservative in our conclusions. We recognize limitations in the Bartik shock in terms of interpretation and its mechanics. For example, housing markets are an important determinant in small business creation (Mian and Sufi, 2014; Adelino, Schoar, and Severino, 2015). Regions may not be exposed to manufacturing shocks, but they may be exposed to national income shocks through their local housing markets which are not captured by industry shocks. Second, though we use publicly available county loan data in a Greenstone, Mas, and Nguyen (2014) style credit shock, we do not directly discuss the role credit frictions and banking play in entrepreneurship though a body of literature suggests they play an important role in entrepreneurship (e.g.: Black and Strahan, 2002; Barlevy, 2003; Kerr and Nanda, 2009, 2010; Osotimehin and Pappada, 2016). Other lines of literature document that contracting frictions dampen the recovery process after a recession (Caballero and Hammour, 1996) or that for firms, the relationship of their primary lender with distressed firms such as Lehman Brothers weakened their employment growth (Chodorow-Reich, 2014).

While we cannot determine whether startups cause regions to be more reslient or resilient regions have more startups, we have implemented a framework developed by Adelino, Ma, and Robinson (2016) to begin to probe at the question why regions have diverged so strongly after the Great Recession. Clearly, firm characteristics such as age and industry are substantial pieces of the story. Our framework is not without limitations. For example, the Bartik shock relies on strong assumptions on the exogeneity of industry employment. The shock is also limited to the manuacturing sector and we therefore are unable to capture fluctuations in local income arising from housing markets, credit frictions, or other confounding regional and industry factors. Our framework is further limited to startups and job creation in the retail and restaurant sectors of the economy. These industries are not the growth industries that scholars discuss in the context of aggregate productivity growth or innovation. Future scholarship must enhance our understanding of the relationship between industry advancement and spillovers into regional entrepreneurship.

Chapter 5

Conclusion

The social and economic differences across regions of the United States remain striking. Spatially differential impacts of increased trade competition (Autor, Dorn, and Hanson 2013), automation (Acemoglu and Restrepo 2017), the Great Recession (Yagan 2019), and political polarization (Autor, Dorn, Hanson, and Majlesi 2019) have further entrenched an economic landscape with a few dense metropolitan areas as engines for growth and prosperity with a remaining large share of the country lagging behind (Glaeser 2011; Moretti 2012; Hsieh and Moretti 2015). All three dissertation essays describe just one aspect of economic dynamism–entrepreneurship–that is thought to be an integral part of robust labor markets, innovation, and job creation. Understanding how and why high-growth and productive economic activity clusters geographically is important for both policymakers focused on employment and wages as well as the broader business community seeking to understand the relationship between their location and personnel decisions and employee mobility.

Chapter 2 discusses whether local governments competing for an anchor firm can stimulate cluster formation geographically proximate to the anchor firm. I find that while startup entry and job creation increases, it only does so in the direct supply chain industries of the anchor firm. I contribute to a literature that shows place-based policies that target job growth in specific industries by providing incentives to large firms to expand operations to a new location do not necessarily create the jobs required to recover subsidy expenditures. These subsidies can be large. Firms such as Nissan received subsidies amounting to over $1.2 billion in their site search that ended in Canton, MS (GoodJobsFirst). A back of the envelope calculation finds the local jobs multiplier to be approximately 1.3 in supply chain industries which is in line with prior literature on place-based policies (Glaeser and Gottlieb 2008; Moretti 2010). Glaeser and Gottlieb (2008) and Moretti (2010) predict job gains in one sector are often offset by losses in other sectors, and the subsidy dollars are not welfare enhancing.[1] Moreover, subsidy programs and policies aimed at inducing an anchor firm's arrival lead to suboptimal matching of firms to low productivity places.

The second contribution from Chapter 2 connects the literature describing the clustering of input-output industries to literature on management practices and productivity spillovers from anchor firms. Ellison, Glaeser, and Kerr (2010) show that coagglomeration patterns are particularly strong for industries that purchase from and supply to each other. While I show a clustering of new firms about an anchor firm are linked through input-output industries, startup formation most likely occurs because rising wages and labor costs in industries most directly affected by the anchor firm's arrival are forced to streamline operations and potentially shed extra layers of management. These workers then use their industry-specific skill to start firms in the same industry as their employers or industries that complement their accumulated skills. This process is particularly strong in more knowledge intensive industries such as RD services and high-tech manufacturing where local labor supply is relatively more inelastic relative to less knowledge intensive industries such as textile manufacturing.

Chapter 2 presents a picture of startup formation centered about personnel and

[1]Slattery (2018) suggests that any welfare gains through agglomeration that may occur would be fully internalized by the anchor firm.

organizational decisions incumbents make in response to rising labor costs. In line with the literature in personnel economics, the determinants of wages in the short-run are determined mostly by industry and location factors external to the firm (Lazear and Oyer 2004). However, firms are tasked with responding to their external environment through reorganization. While the data does not allow me to determine whether firms are explicitly reorganizing by adding and dropping layers (Caliendo, Monte, and Rossi-Hansberg 2015), firms exhibiting fastest wage growth after the anchor firm's arrival also shed the most workers who form spinoffs. Though these workers earn more relative to their industry, county, and industry-county pair, they are not necessarily the top wage earners in the firm.[2] This behavior of employee transitions in firms undergoing the most rapid post-anchor arrival wage growth is consistent with literature on firm reorganization (Caliendo et al. 2015) as well as compensation policies for and mobility of mid-level managers (Lazear and Shaw 2007). Incumbent firms' efforts to reorganize in response to rising labor costs arising from the anchor firm's arrival may reflect improvement in management practices. Bloom et al. (2019) find evidence that increased density of economic activity in a county leads to advancements in structured management, perhaps reflecting reorganization and personnel management in surviving incumbent firms after the anchor firm arrives. Reorganization and structured management practices also would be reflected in measured labor productivity increases in surviving incumbent firms found by Greenstone, Hornbeck, and Moretti (2010).

Chapter 3 transitions from examining the arrival of a large firm on local labor markets and entrepreneurship to instead focus on entrepreneurship as a means for workers to mitigate displacement. This essay provides a blueprint for a burgeoning research stream that will exploit microgeography of large establishment closures and business registration filings. By focusing on the location of economic activity within

[2] I find evidence they may be above the firm average, but not in the top 3 or 5 wage earners.

a city, we can better understand how entrepreneurship contributes to local economic dynamism. Because lower wage displaced workers tend to be less economically mobile (Notowidigdo 2019), the availability of job opportunities locally is important to reduce unemployment spells and long-term wage losses. I find that conditional on wage level, workers who select into younger firms experience the least wage losses compared to other displaced workers. These workers also tend to be at the lower end of the wage distribution with the highest wage earners being workers in shuttering establishments who relocate to other establishments within the same firm.

In terms of entrepreneurial formation, the availability of fixed capital and cheaper labor inputs may present an opportunity for *de novo* entrants in areas where shutdowns occur, depending on other characteristics of the local area (e.g., education level and housing markets). In this essay, I discuss the pertinent literature related to this research stream. Moreover, I assemble a novel dataset consisting of Business Registration filings geocoded by street address matched to geocoded Worker Adjustment and Retraining Notification Act notices. This data assembly will allow researchers to explore entrepreneurship and regional dynamism using publicly available sources independent of Census microdata.

The final essay in Chapter 4 examines an often used tool in entrepreneurship and regional studies. The Bartik (1991) framework has been used by a number of prior literature and most relevantly by Adelino, Ma, and Robinson (2017). In this framework, local industrial composition is interacted with the national growth rate to create a summary statistics of the demand for local labor. Before the Great Recession, Commuting Zones receiving better draws from the distribution of Bartik shocks also generated the greatest rates of job creation in startups. However, this association inverts post-crisis.

Recognizing the Bartik variable as a weighted average elasticity of its individual industry components, I compute the share of the Bartik variable attributable to each industry. Similar to the procedure contemporaneously outlined by Goldsmith-Pinkham, Sorkin, and Swift (2018) who show the calculation of Rotemberg weights for Bartik components, I find that weights in the Bartik variable adjust substantially for nearly every industry after the crisis. The changes in the weights suggests industry migration and variation in local industry growth rates between the pre-recession and recovery periods. The weight changes also cast doubt on the validity of identification assumptions of the Bartik framework in this regional entrepreneurship context. However, calculating the weights and observing which industries changed dramatically motivates research seeking to understand why growth of some industries imply stagnation of others. This motivation is the basis for this dissertation.

Throughout the three essays, I explore various dimensions of local industry composition and entrepreneurship. I exploit microdata to explore how shocks to local labor markets affect the movement of labor and transitions into entrepreneurship. I find that industry heterogeneity in where these shocks occur is instrumental in understanding whether and where the shock induces or inhibits entrepreneurship and thereby affects labor market outcomes as well as the growth of a particular region. I further show that selection into young firms can be driven by reorganization and personnel decisions of incumbent firms. With respect to local dynamism, I construct a new data set, describe literature, and outline a research stream that shows entrepreneurship to be a channel to insulate displaced workers from longer-term unemployment spells and earnings losses. Finally, I motivate a research agenda focused on understanding how shifts in the growth of industries nationally may affect the location and spatial distribution of economic activity.

Bibliography

Acemoglu, Daron, and Pascual Restreop. 2017. "Robots and Jobs: Evidence from US Labor Markets.".

Addison, John T., and McKinley L. Blackburn. 1994*a*. "The Worker Adjustment and Retraining Notification Act." *Journal of Economic Perspectives* 8(1): 181–190.

Addison, John T., and McKinley L. Blackburn. 1994*b*. "The Worker Adjustment and Retraining Notification Act: Effects on Notice Provision." *Industrial and Labor Relations Review* 47(4): 650–662.

Adelino, Manuel, Antoinette Schoar, and Felipe Severino. 2015. "House Prices, Collateral, and Self-Employment." *Journal of Financial Economics* 117: 288–306.

Adelino, Manuel, Song Ma, and David Robinson. 2017*a*. "Firm Age, Investment Opportunities, and Job Creation." *Journal of Finance* 72(3): 999–1038.

Adelino, Manuel, Song Ma, and David Robinson. 2017*b*. "Firm Age, Investment Opportunities, and Job Creation." *Journal of Finance* 72(3): 999–1038.

Agrawal, Ajay, Iain Cockburn, Alberto Galasso, and Alexander Oettl. 2014. "Why Are Some Regions More Innovative Than Others? The Role of Small Firms in the Presence of Large Labs." *Journal of Urban Economics* 81: 149–165.

Alcacer, Juan, and Mercedes Delgado. 2016. "Spatial organization of firms and location choices through the value chain." *Management Science* 62(11): 3213–3234.

Alcacer, Juan, and Wilbur Chung. 2014. "Location Strategies for Agglomeration Economies." *Strategic Management Journal* 35(12): 1749.

Alon, Titan, David Berger, Robert Dent, and Benjamin Pugsley. 2017. "Older and Slower: The Startup Deficit's Lasting Effects on Aggregate Productivity." *Journal of Monetary Economics* 93: 68–85.

Anton, Jim, and Dennis Yao. 1995. "Start-ups, spin-offs, and internal projects." *Journal of Law, Economics and Organization* 11: 362–378.

Atalay, Enghin, Ali Hortacsu, and Chad Syverson. 2014. "Vertical integration and input flows." *American Economic Review* 104(4): 1120–1148.

Audia, Pino G., Olav Sorenson, and Jerald Hage. 2001. "Tradeoffs in the Organization of Production: Multiunit Firms, Geographic Dispersion and Organizational Learning." *Advances in Strategic Management* 18: 75–105.

Autor, David, David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review* 103(6): 2121–2168.

Autor, David, David Dorn, Gordon Hanson, and Kaveh Majlesi. 2016. "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure." *NBER Working Paper* 222637.

Azoulay, Pierre, Benjamin F. Jones, J. Daniel Kim, and Javier Miranda. 2020. "Age and High-Growth Entrepreneurship." *American Economic Review* 2(1): 65–82.

Babina, Tania. 2015. "Destructive Creation at Work: How Financial Distress Spurs Entrepreneurship." *SSRN 2692841* .

Barlevy, Gadi. 2002. "The Sullying Effect of Recessions." *Review of Economic Studies* 69(1): 65–96.

Barth, Erling, Alex Bryson, James Davis, and Richard Freeman. 2016. "It's where you work: Increases in the dispersion of earnings across establishments and individuals in the United States." *Journal of Labor Economics* 34(2): 67–97.

Bartik, Timothy. 1991. *Who Benefits From State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Berry, Christopher, and Edward L. Glaeser. 2005. "The Divergence of Human Capital Levels Across Cities." *Papers in Regional Science* 84(3): 407–444.

Bhide, Amar. 2003. *The Origin and Evolution of New Businesses.* Oxford University Press.

Black, Sandra, and Philip E. Strahan. 2002. "Entrepreneurship and Bank Credit Availability." *Journal of Finance* 57(6): 2807–2833.

Blanchard, Oliver, and Lawrence Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity* 1: 1–75.

Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen. 2019. "What drives differences in management practices?" *American Economic Review* 109(5): 1648–1683.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2018. "Quasi-Experimental Shift-Share Research Designs." *NBER Working Paper* w24997.

Bresnahan, Timothy, Alfonso Gambardella, and AnnaLee Saxenian. 2001. "'Old Economy' Inputs for 'New Economy' Outcomes: Cluster Formation in the New Silicon Valleys." *Industrial and Corporate Change* 10(4): 835–860.

Busso, Matias, Jesse Gregory, and Patrick Kline. 2013. "Assessing the Incidence and Efficiency of a Prominent Place Based Policy." *American Economic Review* 103(2): 897–947.

Caballero, Ricardo J., and Mohammad L. Hammour. 1996. "On the Timing and Efficiency of Creative Destruction." *Quarterly Journal of Economics* 111(3): 805–852.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg. 2015. "The Anatomy of French Production Hierarchies." *Journal of Political Economy* 123(4): 809–852.

Campbell, Benjamin, Marco Ganco, April Franco, and Rajshree Agarwal. 2012. "Who leaves, where to, and why worry? Employee mobility, entrepreneurship and effects on source firm performance." *Strategic Management Journal* 33(1): 65–87.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and labor market inequality: evidence and some theory." *Journal of Labor Economics* 36(1): 13–70.

Carlino, Gerald, and William R. Kerr. 2015. "Agglomeration and Innovation." *Handbook of Regional and Urban Economics* 5: 349–404.

Carnahan, Seth, and Deepak Somaya. 2013. "Alumni effects and relational advantage: The impact on outsourcing when a buyer hires employees from a supplier's competitors." *Academy of Management Journal* 56(6): 1578–1600.

Carnahan, Seth, Rajshree Agarwal, and Benjamin Campbell. 2012. "Heterogeneity in turnover: The effect of relative compensation dispersion of firms on the mobility and entrepreneurship of extreme performers." *Strategic Management Journal* 33: 1411–1430.

Charles, Kerwin, Erik Hurst, and Matthew Notowidigdo. 2013. "Manufacturing Decline, Housing Booms, and Non-Employment." *Chicago Booth Research Paper* pp. 13–57.

Chatterjee, Satyajit, and Esteban Rossi-Hansberg. 2012. "Spinoffs and the market for ideas." *International Economic Review* 53(1): 53–93.

Chatterji, Aaron K. 2009. "Spawned with a silver spoon? Entrepreneurial performance and innovation in the medical device industry." *Strategic Management Journal* 30(2): 185–206.

Chatterji, Aaron K., Edward L. Glaeser, and William Kerr. 2014*a*. "Clusters of entrepreneurship and innovation." *Innovation Policy and the Economy* 14(1): 129–166.

Chatterji, Aaron K., Edward L. Glaeser, and William R. Kerr. 2014*b*. "Clusters of Entrepreneurship and Innovation." *Innovation Policy and the Economy* 14(1): 129–166.

Chauvin, Jasmina. 2017. "When Distance Shrinks: The Effects of Competitor Proximity on Firm Survival.".

Chinitz, Benjamin. 1961. "Contrasts in Agglomeration: New York and Pittsburgh." *American Economic Review* pp. 279–789.

Chodorow-Reich, Gabriel. 2013. "The Employment Effects of Credit Market Disruptions: Firm-Level Evidence from the 2008-9 Financial Crisis." *Quarterly Journal of Economics* 129(1): 1–59.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sebastien Roux. 2012. "The productivity advantages of large cities: Distinguishing agglomeration from firm selection." *Econometrica* 80(6): 2543–2594.

Couch, Kenneth A., and Dana W. Placzek. 2010. "Earnings Losses of Displaced Workers Revisited." *American Economic Review* 100(1): 572–589.

Davis, Steven J., and John Haltiwanger. 2017. "Dynamism Diminished: The Role of Housing Markets and Credit Conditions." *NBER Working Paper* w25466.

Davis, Steven J., John Haltiwanger, and Scott Schuh. 1998. *Job creation and destruction.* Vol. 1 1 ed. MIT Press.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2014*a*. "The Role of Entrepreneurship in US Job Creation and Economic Dynamism." *Journal of Economic Perspectives* 28(3): 3–24.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2014*b*. "The Role of Entrepreneurship in US Job Creation and Economic Dynamism." *Journal of Economic Perspectives* 28(3): 3–24.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2016*a*. "Declining Business Dynamism: What We Know and the Way Forward." *American Economic Review: Papers & Proceedings* 106(5): 203–207.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2016*b*. "Where Has All the Skewness Gone? The Decline in High-Growth (Young) Firms in the U.S." *European Economic Review* 86: 4–23.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review* 106(3): 479–524.

Dix-Carneiro, Rafael, and Bryan Kovak. 2017. "Trade Liberalization and Regional Dynamics." *American Economic Review* 107(10): 2908–2946.

Doms, Mark, Ethan Lewis, and Alicia Robb. 2010. "Local Labor Force Education, New Business Characteristics, and Firm Performance." *Journal of Urban Economics* 67(1): 61–77.

Duranton, Gilles, and Diego Puga. 2001. "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products." *American Economic Review* 91(5): 1454–1477.

Duranton, Gilles, and Diego Puga. 2004. "Micro-foundations of urban agglomeration economies." *Handbook of Regional and Urban Economics* 4: 2063–2117.

Duranton, Gilles, Peter M. Morrow, and Matthew A. Turner. 2014. "Roads and Trade: Evidence from the US." *Review of Economic Studies* 81(2): 681–724.

Ellison, Glenn, and Edward L. Glaeser. 1997. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy* 105(5): 889–927.

Ellison, Glenn, Edward L. Glaeser, and William Kerr. 2010. "What causes industry agglomeration? Evidence from coagglomeration patterns." *American Economic Review* 100: 1195–1213.

Evans, David S., and Linda S. Leighton. 1990. "Some Empirical Aspects of Entrepreneurship." *The Economics of Small Firms* 11: 79–99.

Fairlie, Robert W. 2013*a*. "Entrepreneurship, Economic Conditions, and the Great Recession." *Journal of Economics & Management Strategy* 22(2): 207–231.

Fairlie, Robert W. 2013*b*. "Entrepreneurship, Economic Conditions, and the Great Recession." *Journal of Economics & Management Strategy* 22(2): 207–231.

Franco, April, and Darren Filson. 2006. "Spin-outs: knowledge diffusion through employee mobility." *RAND Journal of Economics* 37(4): 841–860.

Gaubert, Cecile. 2018. "Firm Sorting and Agglomeration." *American Economic Review* 108(11): 3117–3153.

Gibbons, Robert, and Lawrence Katz. 1991. "Layoffs and Lemons." *Journal of Labor Economics* 9(4): 351–380.

Glaeser, Edward. 2000. "Urban and Regional Growth." *The Handbook of Economic Geography* pp. 83–98.

Glaeser, Edward L. 2011. *Triumph of the City.* New York: Penguin Press.

Glaeser, Edward L., and Joshua D. Gottlieb. 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States." *Journal of Economic Literature* 47(4): 983–1028.

Glaeser, Edward L., and Joushua D. Gottlieb. 2008. "The Economics of Place-Making Policies." *Brookings Papers on Economic Activity* pp. 155–253.

Glaeser, Edward L., and William Kerr. 2009. "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?" *Journal of Economics & Management Strategy* 18(3): 623–663.

Glaeser, Edward L., Giacomo A.M. Ponzetto, and Kristina Tobio. 2014. "Cities, Skills and Regional Change." *Regional Studies* 48(1): 7–43.

Glaeser, Edward L., Hedi D. Kallal, Jose Scheinkman, and Andrei Shleifer. 1992. "Growth in Cities." *Journal of Political Economy* 100(6): 1126–1152.

Glaeser, Edward L., Sari Pekkala Kerr, and William Kerr. 2015. "Entrepreneurship and urban growth: An empirical assessment with historical mines." *Review of Economics and Statistics* 97(2): 498–520.

Glaeser, Edward L., William R. Kerr, and Giacomo A.M. Ponzetto. 2010*a*. "Clusters of Entrepreneurship." *Journal of Urban Economics* 67(1): 150–168.

Glaeser, Edward L., William R. Kerr, and Giacomo Ponzetto. 2010*b*. "Clusters of Entrepreneurship." *Journal of Urban Economics* 67(1): 150–168.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2018. "Bartik Instruments: What, When, Why, and How." *NBER Working Paper* w24408.

Gompers, Paul, Josh Lerner, and David Scharfstein. 2005. "Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999." *Journal of Finance* 60(2): 577–614.

Greenstone, Michael, Alexander Mas, and Hoai-Luu Nguyen. 2020. "Do Credit Market Shocks Affect the Real Economy? Quasi-experimental Evidence from the Great Recession and "Normal" Economic Times." *American Economic Journal: Economic Policy* 12(1): 200–225.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2010. "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings." *Journal of Political Economy* 18(3): 536–598.

Guzman, Jorge. 2019. "Go West Young Firm: Agglomeration and Embeddedness in Startup Migrations to Silicon Valley." *Columbia Business School Research Paper* 18-49.

Guzman, Jorge, and Scott Stern. 2015. "Nowcasting and Placecasting Entrepreneurial Quality and Performance." *NBER wp 20954* .

Helsley, Robert L., and William C. Strange. 2006. "Agglomeration, opportunism, and the organization of production." *Journal of Urban Economics* 62(1): 55–75.

Helsley, Robert W., and William C. Strange. 1990. "Matching and agglomeration economies in a system of cities." *Regional Science and Urban Economics* 20(2): 189–212.

Hsieh, Chang-Tai, and Enrico Moretti. 2015. "Why Do Cities Matter? Local Growth and Aggregate Growth." *CEPR Discussion Papers* .

Hurst, Erik, and Annamaria Lusardi. 2004*a*. "Liquidity constraints, household wealth, and entrepreneurship." *Journal of Political Economy* 112(2): 319–347.

Hurst, Erik, and Annamaria Lusardi. 2004*b*. "Liquidity Constraints, Household Wealth, and Entrepreneurship." *Journal of Political Economy* 112(2): 319–347.

Hurst, Erik, and Benjamin Pugsley. 2011. "What Do Small Businesses Do?" *Brookings Papers on Economic Activity* 2011(2): 73–143.

Jacobs, Jane. 1961. *The Death and Life of Great American Cities.* New York: Random House.

Jacobs, Jane. 1970. *The Economy of Cities.* New York: Vintage.

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan. 1993. "Earnings Losses of Displaced Workers." *American Economic Review* 83(4): 685–709.

Katz, Lawrence, and Alan B. Krueger. 2016. "The Rise and Nature of Alternative Work Arrangements in the United States, 1995–2015." *Industrial and Labor Relations Review* 72(2): 382–416.

Kerr, Sari Pekkala, William Kerr, and Ramana Nanda. 2015. "House money and entrepreneurship." *NBER Working Paper* No. w21458.

Kerr, William R., and Ramana Nanda. 2009. "Democratizing Entry: Banking Deregulations, Financing Constraints, and Entrepreneurship." *Journal of Financial Economics* 94(1): 124–149.

Kerr, William R., and Ramana Nanda. 2010. "Banking Deregulations, Financing Constraints, and Firm Entry Size." *Journal of the European Economic Association* 8(2-3): 582–593.

Klepper, Steven. 2007. "Disagreements, spinoffs, and the evolution of Detroit as the capital of the U.S. automobile industry." *Management Science* 53(4): 616–631.

Klepper, Steven. 2010. "The origin and growth of industry clusters: The making of Silicon Valley and Detroit." *Journal of Urban Economics* 67(1): 15–32.

Klepper, Steven, and Guido Buenstorf. 2009. "Heritage and agglomeration: The Akron tyre cluster revisted." *The Economic Journal* 119(537): 705–733.

Klepper, Steven, and Peter Thompson. 2010. "Disagreements and intra-industry spinoffs." *International Journal of Industrial Organization* 28(5): 526–538.

Klepper, Steven, and Susan Sleeper. 2005. "Entry by spin-offs." *Management Science* 112: 1291–1306.

Kletzer, Lori. 1998. "Job Displacement." *Journal of Economic Perspectives* 12(1): 115–136.

Kline, Patrick, and Enrico Moretti. 2014. "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority." *Quarterly Journal of Economics* 129(1): 275–331.

Lazear, Edward, and Kathryn Shaw. 2007. "Personnel Economics: The Economist's View of Human Resources." *Journal of Economic Perspectives* 21(4): 91–114.

Lazear, Edward, and Paul Oyer. 2004. "The Structure of Wages and Internal Mobility." *American Economic Review* 94(2): 212–216.

Lazear, Edward P. 2004. "Balanced skills and entrepreneurship." *American Economic Review* 94(2): 208–211.

Lyne, Jack. 2001. "Hyundai's $1B Plant Alabama Bound After 11th-Hour Bargaining." *Site Selection* .

Marshall, Alfred. 1890. *Principles of Economics*. McMaster University Archive for the History of Economic Thought.

Mian, Atif, and Amir Sufi. 2014. "What Explains the 2007–2009 Drop in Employment?" *Econometrica* 82(6): 2197–2223.

Moretti, Enrico. 2010. "Local Multipliers." *American Economic Review: Papers & Proceedings* 100(2): 373–377.

Moretti, Enrico. 2011. "Local Labor Markets." *Handbook of Labor Economics* 4(B): 1237–1313.

Moretti, Enrico. 2012. *The New Geography of Jobs*. Houghton Mifflin Harcourt.

Neumark, David, Brandon Wall, and Junfu Zhang. 2011. "Do Small Businesses Create More Jobs? New Evidence for the United States from the National Establishment Time Series." *Review of Economics and Statistics* 93(1): 16–29.

Notowidigdo, Matthew. 2019. "The Incidence of Local Labor Demand Shocks.".

Oettl, Alexander, and Ajay Agrawal. 2008. "International labor mobility and knowledge flow externalities." *Journal of International Business Studies* 39(8): 1242–1260.

Osotimehin, Sophie, and Francesco Pappada. 2016. "Credit Frictions and The Cleansing Effect of Recessions." *Economic Journal* 127(602): 1153–1187.

Overman, Henry, and Diego Puga. 2010. "Labor Pooling as a Source of Agglomeration: An Empirical Investigation." In *Agglomeration Economics.* University of Chicago Press pp. 133–150.

Rosen, Sherwin. 1982. "Authority, Control, and the Distribution of Earnings." *The Bell Journal of Economics* 13(2): 311–323.

Rotemberg, Julio J. 1983. "Instrumental Variable Estimation of Misspecified Models." *Working Paper MIT Sloan* pp. 1508–15083.

Rotemberg, Julio J., and Garth Saloner. 2000. "Competition and human capital accumulation: A theory of interregional specialization and trade." *Regional Science and Urban Economics* 30(4): 373–404.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2020. *IPUMS USA: Version 10.0 [dataset].* Minneapolis, MN: IPUMS.

Saxenian, AnnaLee. 1994. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128.* Cambridge, MA: Harvard University Press.

Schoar, Antoinette. 2010. "The Divide between Subsistence and Transformational Entrepreneurship." *Innovation Policy and the Economy* 10: 57–81.

Serrato, Juan Carlos Suarez, and Owen Zidar. 2016. "Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms." *American Economic Review* 106(9): 2582–2624.

Song, Jae, David Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter. 2019. "Firming up inequality." *Quarterly Journal of Economics* 134(1): 1–50.

Sorenson, Olav. 2017. "Regional Ecologies of Entrepreneurship." *Journal of Economic Geography* 17: 959–974.

Soreson, Olav, and Pino G. Audia. 2000. "The social structure of entrepreneurial activity: Geographic concentration of footwear production in the United States, 1940–1989." *American Journal of Sociology* 106(2): 424–462.

Suarez-Serrato, Juan Carlos, and Owen Zidar. 2016. "Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms." *American Economic Review* 106(9): 2582–2624.

Tate, Geoffrey, and Liu Yang. 2015. "Female Leadership and Gender Equity: Evidence from Plant Closure." *Journal of Financial Economics* 117(1): 77–97.

Walls, Don. 2014. "National Establishment Time-Series (NETS) Database.".

Yagan, Danny. 2016. "Is the Great Recession Really Over? Longitudinal Evidence of Enduring Employment Impacts.".

Yagan, Danny. 2019. "Employment Hysteresis from the Great Recession." *Journal of Political Economy* 127(5): 2505–2558.

Biography

Rahul Raj Gupta is a Ph.D. candidate in Duke University's Fuqua School of Business. His expected degree conferral date is May 2020. Rahul will be an Assistant Professor of Strategy, Economics, and Public Policy at Georgetown University's McDonough School of Business starting in July 2020.

Rahul graduated from Homestead High School in Fort Wayne, Indiana in June 2008. He then attended Indiana University–Bloomington and graduated in May 2012 with a Bachelor of Science in Finance and a Bachelor of Arts with majors in Economics and Mathematics. After undergraduate studies, Rahul completed a Master's of Arts in Economics at Boston University in 2013. Prior to his enrollment in Duke's Ph.D. program, Rahul was a full-time Research Associate at Harvard Business School for William R. Kerr and Joan Farre-Mensa between June 2013 and June 2015.