Does Goliath Help David? Anchor Firms and Startup Clusters

Rahul R. Gupta*
Georgetown University
September 2024

Abstract

Does attracting a "Goliath" firm to a location anchor the formation of a local industrial cluster? I study whether geographic expansions of large, typically multinational firms induce startup entry. I construct a hand-collected data set of "Million Dollar Plants" (MDPs) where a firm announces a winning and a next-best site for a new large facility. Matching these site search contests to restricted-use U.S. Census data shows that winning an MDP leads to economically limited spillover effects. The earliest startup cohorts in winning county supply chain industries are founded by local industry workers and experience 12% faster employment growth relative to those in counterfactual counties, though the effects diminish quickly. Winning counties see supply chain employment that is .71% more concentrated in startup firms than in runners-up counties. These findings highlight the need for scholarship on startup and entry dynamics for the successful implementation of place-based policies aimed at creating industrial clusters.

JEL Codes: H25, R12, R38, L21
Keywords: Entrepreneurship, Firm Growth, Agglomeration, Million Dollar Plants, Place-Based Business Subsidies

*rahul.gupta@georgetown.edu. I am grateful to my advisers Inês Black, Alon Brav, Ronnie Chatterji, Wes Cohen, and Bill Kerr. Any views expressed are those of the author and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 1841. (CBDRB-FY19-R8124, CBDRB-FY19-R8165, CBDRB-FY20-R8265, CBDRB-FY24-P1841-R11529). 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.

1 Introduction

In 2017, Foxconn announced a landmark deal that promised to bring over 13,000 jobs to Mount Pleasant, Wisconsin, a small community with a population just under 30,000. Foxconn's decision came at the end of a long site search for a LCD screen manufacturing plant that pit the Wisconsin community against towns in Illinois, Indiana, Ohio, Pennsylvania, and Texas. Then Governor Scott Walker touted the state's $3 billion subsidy offer to Foxconn as "the single largest economic development project in the history of Wisconsin" and the governor's office pointed to an estimated 22,000 additional jobs to be created in Foxconn's local supply chain (Marley and Stein, 2017). Nearly six years after the deal was struck, Foxconn is yet to produce a single LCD screen and its Wisconsin employment never exceeded 1,000 employees (Chávez, 2022).

Though these place-based firm subsidies are typically administered at the state, county, and city-levels of government, more recently politics in the United States is trending toward national industrial policy. For example, in January 2022 Intel announced its plans to construct two new semiconductor fabrications plants (fabs) in New Albany, Ohio. The decision came after a $2.1 billion subsidy offer from the governor and a promise to change Ohio law to extend job creation tax credits from 15 years to 30 years. However, by mid-summer Intel announced the Ohio plants would be scaled down unless the U.S. Congress passed the CHIPS Act that called for $80 billion in place-based incentives for business R&D spending, funding for new R&D hubs, and over $50 billion in subsidies for new construction of domestic semiconductor chip plants. After the bill passed in August 2022, Intel announced it would augment the local subsidies it procured with the CHIPS monies to invest in the two new Ohio fabs (Williams, 2022). The passage of the CHIPS Act signified a marked change in U.S. economic policy with the widespread adoption of industrial policy that explicitly sought to use typically large, multinational firms, or Goliath firms, to seed and anchor future business investment in specific communities (Juhász, Lane, and Rodrik, 2023; Chatterji, Glaeser, and Kerr, 2014).

This paper quantifies the effect of attracting large "Goliath" firms through the arrival of their large facilities ("Million Dollar Plants") on local industry agglomerations by focusing on follow-on firm formation in the supply chain. The analysis uses both administrative Census microdata and hand collected data. The hand collected data consists of nearly 240 site searches between 1990 and

2015 where a Goliath firm detailed its process for narrowing down to a final location for a new large facility. I then analyze the impact of winning a Million Dollar Plant on new firm formation and the workers that these new firms employ.

The theoretical underpinning for the analysis comes the vast literature relating labor market size to firm productivity, employment, and wages through agglomeration. Prior literature has described the strong positive relationship between the size of a labor market and outcomes for both firms and workers who populate those labor markets. Considering the number of unique firms employing workers in a location and industry as the summary statistic for labor market size, the arrival of a large firm to an area effectively increases its labor market size for that particular industry. This increase in labor market size potentially enhances job hopping and job search for workers, facilitates chance encounters between people to share novel ideas that turn into inventions or new businesses, or reduces distance needed to transport goods and services between firms and their suppliers and customers.

The main empirical challenge in identifying the impact of a large firm on local outcomes is that forward looking firms sort into locations where they expect growth, subsidies, and firm-specific supply chains to be strongest. To address this endogeneity I leverage instances where a large firm conducted a site selection process and revealed its leading alternative sites for a new facility under the assumption that revealed next-best locations are on a similar growth path to the winning location. Similar to Greenstone, Hornbeck, and Moretti (2010) whose sample pre-dates this analysis, I scrape *Site Selection* magazine articles that detail site searches for generally large Fortune 500 firms that usually employ over 500 employees, require an investment of at least a quarter billion dollars, and mention a runner-up location. I match this hand-collected *Site Selection* magazine data into comprehensive employee-employer matched Census microdata. The business-level microdata covers the universe of 200,000,000 establishment-year observations between 1990 and 2015 while the employee-matched data includes approximately one billion worker-job-year observations in 25 states (Figure 4). Table 2 displays the industry breakdown of the new facilities in the magazine and Table 1 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 (Figure 5), I obtain plausibly exogenous variation in startup formation and employ a difference-in-differences design to compare firm formation and workers in winning

with those in runners-up counties. One complication is that anchor firms may select locations both explicitly on the likelihood of follow-on startup entry in economically proximate industries as well as on factors that co-determine entrepreneurship. The first validation checks I run test for parallel trends in wage growth, startup entry rates, startup exit rates, and employee departures from incumbent firms into new firms to assuage concerns that winning counties are "trending better" than counterfactual counties leading up to the announcement. A second set of checks randomize the treated industries and counties within each case to show the identified effects are particular to the industries defined ex ante as economically proximate. Additional tests ensure for similarity of several distributions between winners and counterfactual counties in each case: firm age, firm size, and firm quality.

I find that the effects of the anchor firm's arrival on startup formation are localized to industries in the supply chain.[1] 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.[2] For the average winning county, this amounts to over 120 new establishments from startup firms and 2,300 jobs.[3] 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.[4] After the anchor firm arrives, the number of employees departing an employer to start a new venture in the employer's input or output industries increases by almost 50%. As a share of all new supply chain startups, the coefficients imply that 25% of upstream startups are spinoffs as are 91% of anchor industry startups and 71% of downstream startups.

Related Literature. I primarily contribute to literature that links location specific productivity and industrial conditions to entrepreneurial formation. Guzman (2024) describes the causal impact location has on performance by comparing entrepreneurial ventures that relocate to Silicon Valley to counterfactual startups that do not, suggesting location-specific agglomerations play a role in startup success. Building on the notion of location-specific agglomerations, this paper generalizes across a broad range of location-industry shocks to examine the impact of cluster size

[1]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.

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

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

[4]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.

on entrepreneurial outcomes. More specifically, I exploit shocks to local labor market size that are induced by the marginal increase in the number of incumbent firms as represented by the opening of a "Million Dollar Plant". I also contribute to literature that describes the role of both large and small firms in explaining regional performance similar to Agrawal, Cockburn, Galasso, and Oettl (2014), though I focus on the impact of large firms on young firms. Counter to traditional Marshallian agglomeration arguments for cluster formation, Klepper (2010), Klepper (2007), and Klepper and Sleeper (2005) model local entrepreneurship and cluster formation as an idiosyncratic process driven by incumbent firm-worker (mis-)match. This paper makes a novel contribution to our understanding of entrepreneurship and agglomeration by building on Steve Klepper's seminar works on spin-offs by showing the process of spin-offs can be induced by shocks in a parent firm's supply chain.

The second stream of literature I contribute to is that examining the spatial organization of firms and policy efforts influence cluster formation. I extend the literature documenting the drivers of cluster (county-industry pairs) growth and success (Delgado, Porter, and Stern, 2014, 2010) by highlighting the role of ex ante local industry experience in promoting cluster growth. Because of potential productivity, wage, and innovation benefits to local economies by cluster formation, another line of literature seeks to identify new areas of cluster formation (Guzman and Stern, 2015) or policy efforts to stimulate clusters (Chatterji, Glaeser, and Kerr, 2014). I extend this work by testing whether one tool for seeding a cluster, attracting a new establishment of a typically large, multinational firm to a location leads to a new agglomeration. Similar to Alcacer and Delgado (2016) and Ellison, Glaeser, and Kerr (2010) I identify that interfirm linkages through the supply chain are critical in understanding agglomeration patterns.

Finally, this paper builds on the literature describing Million Dollar Plants (MDP) both as an identification tool to understand geographic agglomerations and as a critical economic phenomenon in its own right. Greenstone, Hornbeck, and Moretti (2010) and Greenstone and Moretti (2003) are the initial works that developed the MDP identification strategy and show that employment, wages, and incumbent plant productivity rise in counties winning a MDP relative to those in runners-up counties. Seeking to understand the nature of these potential agglomeration spillovers, Bloom et al. (2019) finds that incumbent plants see significant improvements in management practices after the arrival of a MDP to their county. Other studies mostly in the domain of public economics seek to

understand MDPs as a place-based business subsidy policy and evaluate whether subsidy sizes are socially optimal (e.g., Slattery and Zidar, 2020; Slattery, 2018) or as a growing form of industrial policy Juhász, Lane, and Rodrik (2023). I connect this literature to scholarship on entrepreneurship by showing the effects of these policies on local firms and business formation.

Organization The remainder of this paper is as follows: Section 2 highlights the relationship between labor market size and entrepreneurship and then details the mechanics of place-based business subsidies. The extensive data used to arrive at the estimation sample is detailed in Section 3. Section 4 explains the empirical methodology and Section 5 discusses results. Finally, Section 6 discusses the implications of the main findings.

2 Background and Setting

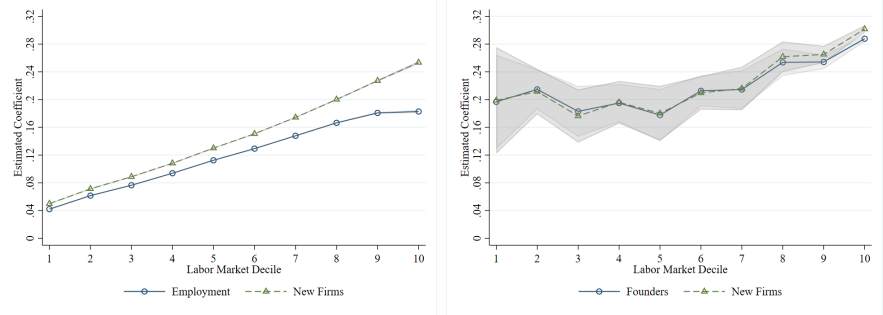
2.1 Agglomeration Economies and Business Formation

Agglomeration broadly refers to the spatial concentration of firms. The number of active employer firms in a place describes the size, or thickness, of its labor market. An increase in the thickness of a local labor market is linked to increased productivity growth and innovation in firms and higher wages and employment rates for workers (see Carlino and Kerr, 2015; Combes et al., 2012; Greenstone, Hornbeck, and Moretti, 2010; Feldman, 1999). Anecdotally, Saxenian (1994) describes dense labor markets as facilitating entrepreneurship by increasing the likelihood of chance encounters with financiers and industry experts using the growth of Silicon Valley as an example. With the benefits of geographic density of economic activity in mind, policymakers incentivize agglomerations to form by attracting the marginal employers to their locales (e.g., Juhász, Lane, and Rodrik, 2023; Chatterji, Glaeser, and Kerr, 2014).

I first show that labor market size correlates to increased local entrepreneurial activity by regressing several measures of entrepreneurship on labor market size. I use the LBD (see Section 3.2) to measure labor market size at the county-industry, or "cluster", level. For each county c and industry i pair, I tally the number of unique firms, $N_{c,i}$, with positive employment at the start of each decade, 1990, 2000, and 2010. For each national industry i, I rank the $N_{c,i}$'s to obtain a decile rank for county-industry with one representing the county-industry pairs with the fewest employers

Figure 1: Labor Market Size and Local Firm Formation

(a) Share of New Firms (b) Share of Local Firm Founders



Notes: Panel A plots the coefficients of county-industry level startup share of local employment (blue circles) and firms (green triangles) on dummies for 10 deciles of initial labor market size. Panel B plots the coefficients of startup-level share (blue circles) of founders who were previously employed by an incumbent firm in the same county-industry. The green triangles are coefficients from regressing a dichotomous variable that takes the value of one if a majority of the startup's founders were previously employed in the same county-industry. Labor market size is calculated by ranking counties by the number of unique employer firms in each industry at the start of each decade (1990, 2000, and 2010) so that each county-industry pair may be assigned a different decile within a single county. Regressions include Year x NAICS and Year x County fixed effects. Standard errors are clustered at the county-level. 90% confidence intervals are in grey (intervals in Panel A are tightly bound about the plotted line and may not appear clearly). $N = \{3,615,000, 1,416,000\}$ in Panels A and B, respectively, where counts have been rounded to accord by U.S. Census Bureau disclosure rules. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 1841. (CBDRB-FY24-P1841-R11529).

and ten being the thickest decile. The decile ranks represent labor market size, $LM_{c,i,t}$. I then regress measures of subsequent local entrepreneurship on the initial decade decile ranks and fixed effects to control for secular industry and county trends.[5] Formally, I estimate the following

$$y_{c,i,t} = \beta_0 + \sum_{d=1}^{10} \beta_d \times \mathbf{1}[LM_{c,i,t} = d] + \gamma_{i,t} + \eta_{c,t} + u_{c,i,t}, \tag{1}$$

and plot the estimated β_d's in Figure 1.

Panel A measures the share of annual cluster activity that is accounted for by startups based on a cluster's *initial* size at the start of each decade. The blue circles show the share of county-industry employment in year t in new firms, $\frac{emp^{new}_{c,i,t}}{emp_{c,i,t}}$. The green triangles represent the estimated share of firm counts in year t that are startups. Panel A shows that county-industry pairs in the bottom decile in number of employers at the start of the decade have just 4.2% of their employment in startup firms whereas the largest clusters in terms of unique employers have 18.3% of employment in startups. The plot also shows a striking monotonically increasing relationship though the difference between the ninth and tenth deciles are not economically significant (18.1% versus 18.3%). The green triangles in Panel A show a similar relationship when plotting the share of firms in a cluster

[5]County fixed effects are identified because counties are assigned ranks for each industry rather than a single rank for the whole county.

that are startups, $\frac{N^{new}_{c,i,t}}{N_{c,i,t}}$. Approximately 5% of firms in clusters that start a decade with the fewest employers are startups versus 25.4% for the largest initial size. Taken together, the two regressions in Panel A show that even when fixing clusters to their *initial* size and accounting for industry and county trends, initial density correlates to greater subsequent entrepreneurial activity.

Next, I test the Saxenian (1994) thesis that thicker labor markets induce entrepreneurship by facilitating local job hopping within clusters. Panel B matches the LBD to the LEHD (see Section 3.3) to identify whether startup founders have prior work experience in the cluster. If founders in thicker labor markets are more likely to have prior local industry experience, then entrepreneurial formation in clusters is more likely driven by positive selection into entrepreneurship than other explanations. The first exercise in Panel B takes the share of startup firm founders whose job prior to the startup firm was in the same cluster (blue circles).[6] In clusters with the least dense labor markets, I estimate that 19.7% of founders have prior work experience in the cluster. In the most dense clusters, the share is 28.8%. The second regression in Panel B is a linear probability model that uses as the outcome a dichotomous variable that designates whether a majority of the startup firm previously worked in the same cluster. I estimate that approximately 20% of startups in the smallest initial clusters have founding teams that have a majority of founders who are local-industry workers whereas 30.2% of startups in the largest initial clusters have founding teams with prior same cluster experience. Unlike Panel A, the coefficients in Panel B do not display the same monotonic relationship for all deciles with deciles three through five having lower shares than the first decile. However, for larger initial deciles we observe that startup firms in larger clusters benefit by having more founders with prior industry experience.

Figure 1 shows that clusters with thicker initial labor markets see more subsequent entrepreneurial activity driven by founders with local industry experience. This descriptive fact motivates this paper's empirical framework that answers whether increasing labor market size forms industry agglomerations through entrepreneurship. In the subsequent sections I describe my methodology for testing the effect of a marginal increase in labor market size on entrepreneurship and cluster formation. I specifically focus on regional efforts to attract a new large employer ("Goliath firms") and investigate whether their "Million Dollar Plants" stimulate local agglomerations by serving as

[6]Though the LEHD does not contain occupational information, I follow convention by classifying top earners over the first three years in startup firms identified in the LBD as founders. For larger startups with more than 40 employees, I use the top 10% of earners to approximate the "founding team".

an "anchor" for startup firms.

2.2 Attracting Goliath Firms

The arrival of a large firm's ("Goliath") new business establishment to an area signifies a marginal increase in the number of local employers in the establishment's industry. Firms consider factors that include workforce talent, local cost structure, ease of moving and accessing goods and services, proximity to suppliers, proximity to customers, their own spatial structure and strategic positioning, and public infrastructure. The largest and most capital intensive projects such as aerospace manufacturing plants or a major pharmaceutical R&D lab often involve relocating labor and physical assets, hiring thousands of employees, and incurring billions in new capital expenditures. In these cases, firms often hire site search consultants to aid in the search process.

The site search consulting industry is not new. In 1929, Leonard Yassen arranged the earliest known place-based business subsidy in helping textile manufacturer Reliance Manufacturing locate to the severely Great Depression-struck town of Columbia, Mississippi with an $85,000 in funds from the city (Cobb, 2019). Columbia Mayor Hugh Lawson White went on to assume governorship of Mississippi touting his local jobs program that generated six years of growth at a time when other small towns in the state suffered annual job losses. As governor, White created the Balance Agriculture with Industry Program (or Mississippi Industrial Act) to make the process used to lure Reliance to Columbia a standardized process across the state using the now oft-used tax-exempt Industrial Revenue Bonds to provide financing for corporate subsidy deals. Having found early success in arranging a business subsidy, Leonard Yassen went on to found Fantus Factory Locating Service in 1934, credited as the first site search consulting firm (LeRoy, 2005; Due, 1961).

Over the next several decades, Fantus Factory Locating Service ushered through more than 4,000 subsidy deals for firms opening new facilities. The firm was eventually acquired by Deloitte, and since the 1970s, all of the Big 4 consulting firms have site search services that help firms select locations suitable for their production needs and target state and local development officials for attracting the most lucrative subsidy deals.[7] Consulting firms often receive payouts based on the size of subsidy deal awarded to the client. Estimates suggest these deals are large in size and have approached $70 billion in value (Bartik, 2017). These public expenditures are justified if the

[7]The Big 4 refer to Deloitte, KPMG, PricewaterhouseCoopers, and Ernst & Young.

returns to agglomeration spillovers created by the "Goliath" firm accrue to the local population and businesses (Slattery and Zidar, 2020; Slattery, 2018).

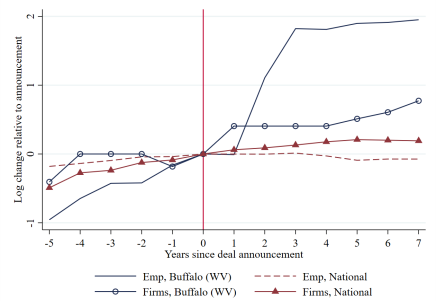
Just as in the Foxconn announcement, when promoting deals to attract new establishment from large, typically multinational Goliath firms, policymakers often make job creation claims that include the impact indirect jobs. Indirect jobs are jobs created not by the opening of the target firm's new facility, but jobs related to construction, the plant's supply chain, or any other services up- or downstream to the opening firm. I focus on supply chain jobs that relate to a clear Marshallian channel to understand whether a focal target firm, or "anchor firm", serves as the seed for an industrial agglomeration (also referred to as a "cluster").

Figure 2 shows four examples of firms described in a corporate real estate magazine, *Site Selection*, that involved a major subsidy deal by local economic development officials. Panel A shows Toyota's decision to locate to Buffalo, West Virginia on the back of a nearly $400 million dollar subsidy for its automobile assembly plant in 1997. The plot shows the log change in supply chain firm count and employment relative to when the deal was announced. I also plot the log change in firm count and employment in those supply chain industries at the national level. Panel A would seem to suggest attracting Toyota to Buffalo, West Virginia induced firm entry and employment growth in excess of the national average and the city's pre-Hyundai trend. Eli Lilly's $425 million subsidy to establish a new drug manufacturing facility in Prince William County, VA shows a similar trajectory (Panel B), though the facility was ultimately completed by Covance Inc. Panel C shows railroad manufacturer Union Tank Car's muted effect on Alexandria, LA's rail industry despite its $65 million deal to locate there in 2007. Finally, Panel D shows Kia's greenfield entry into the United States with the establishment of its first U.S. autoassembly plant in West Point, GA with a $400 million dollar deal.

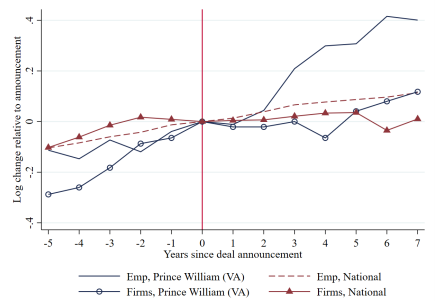
Figure 2 highlights one of the central challenges in assessing the impact of a targeted firm's arrival to a location. Without a valid counterfactual, isolating the impact of the firm, subsidy program, and location-specific trends is impossible to identify. The goal of this paper is collect data on sites that were considered by the firm and its consultants in the final stage of the site selection process. This enables estimating an empirical specification that controls for location-specific pre-trends, and compare post-announcement changes in the winning location to the county.

Figure 2: Employment Growth and Firm Counts in Supply Chain

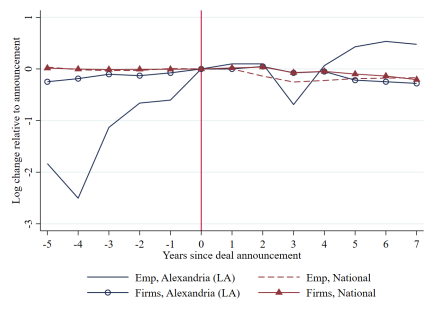
(a) Toyota in Buffalo, WV



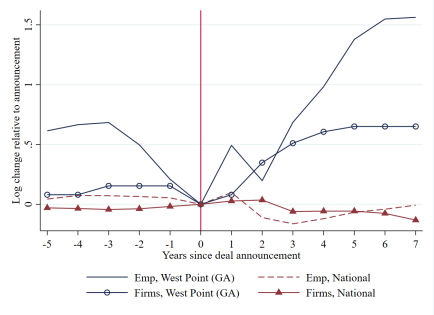
(b) Eli Lilly in Prince William, VA



(c) Union Tank Car in Alexandria, LA



(d) Kia in West Point, GA



Notes: These examples do not necessarily represent any Census matched samples used in subsequent analysis and are constructed using data external to restricted Census data. Figure 2 shows growth in supply chain industries of four firms announcing the construction of new establishments. The plot uses the NAICS code associated with the establishment in the National Establishment Time Series data. Using the facility's industry code, I identify the supply chain as the five leading input-output transaction industries in the 2007 BEA Use of Commodities matrix and plot log employment growth. Employment data comes from County Business Patterns with imputed missing values (Eckert et al., 2021). Log changes in the supply chain firm counts are calculated from NETS. Changes in firm counts are due to *de novo* firm formation, firm exit, or growth of incumbent firms. Source: *Site Selection* magazine.

3 Data

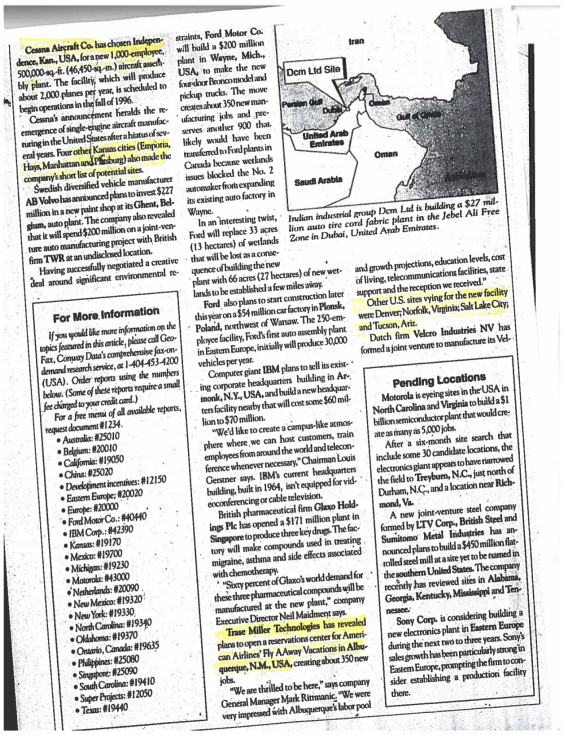
3.1 *Site Selection* Magazine

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 sporadically mention incentives or subsides offered by local economic development bodies. The announcements in the articles are sizeable economic events. They tend to involve

openings of plants that promise, when reported, an average of 1,500 full-time employees and an initial capital investment of $500 million. Figure 3 is an example of a microfilm of an April 1995 *Site Selection* magazine article.

Figure 3: Archival Data



Cessna Aircraft Co. has chosen Independence, Kan., USA, for a new 1,000-employee, 500,000-sq.-ft. (46,450-sq.-m.) aircraft assembly plant. The facility, which will produce about 2,000 planes per year, is scheduled to begin operations in the fall of 1996.

Cessna's announcement heralds the re-emergence of single-engine aircraft manufacturing in the United States after a hiatus of several years. Four other Kansas cities (Emporia, Hays, Manhattan and Pittsburg) also made the company's short list of potential sites.

Swedish diversified vehicle manufacturer AB Volvo has announced plans to invest $227 million in a new paint shop at its Ghent, Belgium, auto plant. The company also revealed that it will spend $200 million on a joint-venture auto manufacturing project with British firm TWR at an undisclosed location.

Having successfully negotiated a creative deal around significant environmental restraints, Ford Motor Co. will build a $200 million plant in Wayne, Mich., USA, to make the new four-door Bronco model and pickup trucks. The move creates about 350 new manufacturing jobs and preserves another 900 that, likely would have been transferred to Ford plants in Canada because wetlands issues blocked the No. 2 automaker from expanding its existing auto factory in Wayne.

In an interesting twist, Ford will replace 33 acres (13 hectares) of wetlands that will be lost as a consequence of building the new plant with 66 acres (27 hectares) of new wetlands to be established a few miles away.

Ford also plans to start construction later this year on a $4 million car factory in Plonsk, Poland, northwest of Warsaw. The 250-employee facility, Ford's first auto assembly plant in Eastern Europe, initially will produce 30,000 vehicles per year.

Computer giant IBM plans to sell its existing corporate headquarters building in Armonk, N.Y., USA, and build a new headquarters facility nearby that will cost some $60 million to $70 million.

"We'd like to create a campus-like atmosphere where we can host customers, train employees from around the world and teleconference whenever necessary," Chairman Louis Gerstner says. IBM's current headquarters building, built in 1964, isn't equipped for videoconferencing or cable television.

British pharmaceutical firm Glaxo Holdings Plc has opened a $171 million plant in Singapore to produce three key drugs. The factory will make compounds used in treating migraine, asthma and side effects associated with chemotherapy.

"Sixty percent of Glaxo's world demand for these three pharmaceutical compounds will be manufactured at the new plant," company Executive Director Neil Maidment says.

Trase Miller Technologies has revealed plans to open a reservations center for American Airlines' Fly AAway Vacations in Albuquerque, N.M., USA, creating about 350 new jobs.

"We are thrilled to be here," says company General Manager Mark Rittmanic. "We were very impressed with Albuquerque's labor pool and growth projections, education levels, cost of living, telecommunications facilities, state support and the reception we received."

Other U.S. sites vying for the new facility were Denver; Norfolk, Virginia; Salt Lake City; and Tucson, Ariz.

Dutch firm Velcro Industries NV has formed a joint venture to manufacture its Vel-

For More Information

If you would like more information on the topics featured in this article, please call GeoFax, Company Data's comprehensive fax-on-demand research service, at 1-404-953-4200 (USA). Order reports using the numbers below. (Some of these reports require a small fee charged to your credit card.)

For a free menu of all available reports, request document #1234.

- Australia: #25010
- Belgium: #20010
- California: #19050
- China: #25020
- Development incentives: #12150
- Eastern Europe: #20020
- Europe: #20000
- Ford Motor Co.: #40440
- IBM Corp.: #42390
- Kansas: #19170
- Mexico: #19700
- Michigan: #19230
- Materials: #43000
- Netherlands: #20090
- New Mexico: #19320
- New York: #19330
- North Carolina: #19340
- Oklahoma: #19370
- Ontario, Canada: #19635
- Philippines: #25080
- Singapore: #25090
- South Carolina: #19410
- Super Projects: #12050
- Texas: #19440

Indian industrial group Dem Ltd is building a $27 million auto tire cord fabric plant in the Jebel Ali Free Zone in Dubai, United Arab Emirates.

Pending Locations

Motorola is eyeing sites in the USA in North Carolina and Virginia to build a $1 billion semiconductor plant that would create as many as 5,000 jobs.

After a six-month site search that include some 30 candidate locations, the electronics giant appears to have narrowed the field to Treyburn, N.C., just north of Durham, N.C., and a location near Richmond, Va.

A new joint-venture steel company formed by LTV Corp., British Steel and Sumitomo Metal Industries has announced plans to build a $450 million flat-rolled steel mill at a site yet to be named in the southern United States. The company recently has reviewed sites in Alabama, Georgia, Kentucky, Mississippi and Tennessee.

Sony Corp. is considering building a new electronics plant in Eastern Europe during the next two to three years. Sony's sales growth has been particularly strong in Eastern Europe, prompting the firm to consider establishing a production facility there.

Notes: Figure 3 is a microfilm capture of the April 1995 "Global Overview" section of *Site Selection* magazine. The highlighted portions describe two site search events for Cessna Aircraft and Trase Miller Technologies. The article includes several other site searches, though they are either foreign expansions outside of the U.S. or do not include both a winning and finalist site and are therefore excluded from the data collection.

The two examples highlighted in Figure 3 state:

> Cessna Aircraft Co. has chosen Independence, Kan., USA, for a new 1,0000 employee 500,000 sq.ft. (46,450 sq.m.) aircraft assembly plant. The facility which will produce about 2,000 planes per year is scheduled to begin operations in the fall of 1996. Cessna's announcement heralds the re-emergence of single engine aircraft manufacturing in the United States after a hiatus of several years. Four other Kansas cities (Emporia, Hays, Manhattan, and Pittsburg) also made the company's short list of potential sites.
>
> Trase Miller Technologies has revealed plans to open a reservations center for American Airlines' Fly AAway Vacations in Albuquerque, N.M., USA, creating about 350 new jobs. "We are thrilled to be here," says company General Manager Mark Rittmanic. "We were very impressed with Albuquerque's labor pool and growth projections, education levels, cost of living, telecommunications facilities, state support and the reception we received." Other U.S. sites vying for the new facility were Denver; Norfolk, Virginia; Salt Lake City; and Tucson, Ariz.

Using archival microfiche of magazine editions, physical magazine copies, and online editions of the

magazine, 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.

Following other studies that have used these articles (e.g., Bloom et al., 2019; Greenstone, Hornbeck, and Moretti, 2010; Greenstone and Moretti, 2003), I refer to these announcements as "cases" or "events" and the openings of these facilities and business establishments as "plant openings", "large facilities", or "Million Dollar Plant" (MDP). For each article mention of a MDP site search process, I document the year of the announcement, corporate entity, the *winner* county, and the *finalists* (*runners-up*) considered at the end of its location selection process. All cases have one winner and approximately 75% of contests conclude with a single runner-up county. However, some cases have multiple with one case concluding with eight runners-up. This skewness leads a mean runner-up count per case of 1.49 counties. I include all case runners-up as counterfactuals for the ultimate winner in the empirical analysis.

Table 1 shows a summary of firm characteristics represented in the magazine sample. Firms reporting in *Site Selection* magazine are typically large firms with the median firm employing over 37,000 employees. The median firm has assets worth over $15 billion and conducts approximately $700 million in R&D spending annually.

Table 1: 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: This table is produced using *Site Selection* magazine and Compustat. The firms and data do not reflect inclusion or exclusion in Census microdata. 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 domicile 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.

Most of the events involve manufacturing facilities, though R&D labs, telecommunications ser-

vices, financial institutions, and corporate offices occasionally appear as well. Table 2 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.[8]

Table 2: Distribution of Announcements by Sector

A. Industry Sectors	Cases	B. 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. Accommodation & Food Services	5		
81. Other Services	1		

Notes: The anchor firm sample includes observations from a broad set of industries, though 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. The industry designations come from descriptions of the facility announcements in the magazine as well as Google searches on the facilities. This data is not produced using Census microdata. Firms in this sample may or may not be in the matched Census sample.

I consider economically proximate firms to be those in industries that purchase inputs from and sell outputs to the anchor firm's industry. I refer to these industries as the anchor firm's "supply chain." 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 (BEA) 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 MDP's industry.

The empirical definitions of upstream and downstream come from tabulations of industry-level output and input dependency. An industry's input dependency, $Input_{i \leftarrow j}$ designates the share of industry i's inputs from industry j. In this formulation, the Million Dollar Plant's industry is i and its input providing industries are j. Similarly, $Output_{i \rightarrow j}$, is the share of MDP industry i's outputs sold to industry j. Table 3 lists the top ten industry pairs that are upstream (Panel A) and downstream (Panel B) in the MDP sample along with their corresponding input and output shares.

In tabulation based on Census microdata, I categorize the MDP's top five input providing

[8]Tables 2 and 1 do not describe the Census microdata sample. These tabulations are based solely on matches between firms mentioned in *Site Selection* magazine that are matched to Compustat.

Table 3: Leading Industry Pairs

A. Leading Upstream Pairs to MDP Industry

Announcement Industry	Paired Industry	Input Share
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 Re-lated 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

B. Leading Downstream Pairs to MDP Industry

Announcement Industry	Paired Industry	Output Share
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: Upstream industries supply the Million Dollar Plant industry with inputs. Panel A shows the ten leading industry pairs by $Input_{i \leftarrow j}$. Downstream industries purchase inputs from the Million Dollar Plant's industry. Panel B displays the ten leading industry pairs by $Output_{i \rightarrow j}$. These input and output shares are tabulated from the Bureau of Economic Analysis's "Use of Commodities" tables for 2007 and 2012.

industry as upstream and the top five output purchasing industries as downstream. In checks using public data, I describe the MDP's supply chain as the five most "input-output" linked industries.

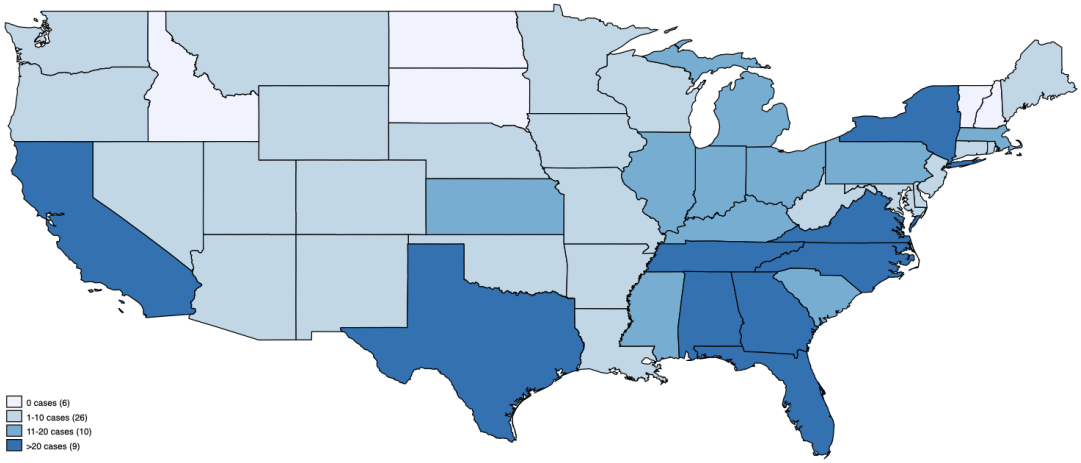
Figure 4 describes *Site Selection* magazine's geographic detail. The map shows total number of cases the magazine references for each state as having either a winning or runner-up county. All but six states are found in the hand-collected magazine sample. The map shows that the magazine over-represents states in the southeast, though states with the largest numbers of cases are distributed across the West Coast, Midwest, and Mid-Atlantic.[9]

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 es-

[9]Figure 4 describes the data coming from *Site Selection*, but not the internal Census data that the magazine sample is matched to or the resulting matched dataset.

Figure 4: Geography of *Site Selection* Announcements



tablishments from the Social Security Administration's Form SS-4, IRS Master Business File, IRS Form 941 and IRS Form 944.[10] 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. This facilitates treating MDP announcements as a shock to the labor market size of a particular industry cluster (county-industry).

The analysis uses the LBD to measure startup activity. Tracking startup formation in the LBD is imperfect because while the LBD includes firm identifiers, mechanical changes in Employer Identification Number (EIN) for tax purposes result in changes in Census firm identifiers. For example, a firm that transitions from being a single-establishment firm to a multi-unit firm results in a new EIN and Census firm identifier. To track cohorts of startup firms, I do the following: Drop establishments that never report positive employment, label a firm as a startup only if all establishments associated with a new Census firm identifier 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 Decker et al.

[10] Jarmin and Miranda (2002) provide technical documentation for the Longitudinal Business Database and its construction.

(2015). The key variables of interest tabulated from the Longitudinal Business Database include county-industry firm entry rates, firm growth rate, and firm survival.

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.[11] 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.

The LEHD is a complex network of files for each state. 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 depending on their tax structure, they may also have multiple SEINs assigned to them by each state in which they operate. Wage-earners in the U.S., however, are assigned unique personal identifiers which follow the workers throughout the panel across states 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. Similar to other Census Bureau working papers such as Tate and Yang (2013) and Babina (2017), 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.

[11] Vilhuber and McKinney (2014) provide an extensive documentation of the LEHD infrastructure files and source data.

We 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. Following prior literature, I consider founders to be the top three wage earners in years up to two years after the firm first appears in the LEHD. I create a panel of work histories to track industry and location experience along with wage growth for both founders and employees in startup firms. This panel allows us to construct outcome variables that measure employee transitions to entrepreneurship based on industry and location experience.

4 Empirical Strategy

4.1 Difference-in-Differences

The entry and exit of firms is a constant feature of the U.S. economy. The dynamism of a region's startup and business activity, labor pool, and productivity is linked to the composition of firms, workers, and institutions that cluster there (Glaeser, Kerr, and Kerr, 2015; Fujita, Krugman, and Venables, 2001; Chinitz, 1961). Identifying the causal effect of the arrival of a Million Dollar Plant to a location is complicated by the fact that firms are forward-looking and do not randomly select locations for their new facilities. Firms strategically locate into places where their expected profit is greatest. Unobserved shocks to the MDP firm's future profitability correlate across firms and potential entrepreneurs. This inhibits making causal claims about the impact of a new large facility on local firms and entrepreneurs based on compiling case studies as in Figure 2. Therefore, I compare winning counties to next-best locations for a facility type (see Section 3.1). This revealed preferences approach to constructing treatment and counterfactual groups allows me to identify changes to a local supply chain that are driven by the arrival of a Million Dollar Plant.

The *Site Selection* magazine sample covers 230 events where a Million Dollar Plant was announced along with at least one next-best location. In total, the sample includes 167 unique winning counties and 206 runners-up counties. These 230 events and their associated facility types

and contest counties are matched into both public and restricted-use Census data which reduces the dimensionality of the microdata and enables identifying outcomes in winning and counterfactual counties. One concern with this quasi-experimental design is that cases occur throughout the full panel spanning 1990 to 2015. Recent literature critiquing two-way fixed effects and staggered difference-in-differences designs describe bias introduced by the heterogeneity in treatment timing. Variation in when counties are treated could lead to disproportionate loading in the standard two-way fixed effects estimator (Goodman-Bacon, 2021; Sun and Abraham, 2021; Callaway and Sant'Anna, 2021; Borusyak, Jaravel, and Spiess, 2022). This research design is not sensitive to the concern because the empirical specification is identified *within plant opening cases* and not across cases.

The identification strategy assumes parallel trends in our outcomes of interest between winning and runners-up counties conditional on time and county controls. More precisely, the main estimating equation is,

$$y_{cijt} = \gamma_0 + \beta_W \mathbf{1}(Win_{cij}) + \sum_{k=-5}^{7} \tau_{jk} \mathbf{1}[t=k] + \sum_{k=-5}^{7} \beta_k \mathbf{1}(Win_{cij}) \times \mathbf{1}[t=k] + \eta_c + \alpha_j + \lambda_{it} + \epsilon_{cijt}, \quad (2)$$

where c indicates county, i denotes industry, j specifies the plant opening case, and t is time. The specification enables observing time dynamics surrounding a Million Dollar Plant event. $\mathbf{1}(Win_{cij})$ takes the value of 1 if county c in event j wins the MDP. The τ_{jk} capture the case-time main effects common to all counties in case j. The coefficients of interest are the β_k's. t is normalized about the opening announcement within each case to retain years in $[t-5, t+7]$. This time window facilitates observing any changes in the winning county's cluster outcomes while also preventing an individual cluster from being treated multiple times within the window.[12] α_j are the critical MDP event fixed effects that force identification of the β_k's to be within case. η_c captures county fixed effects and λ_{it} controls for secular industry trends. λ_{it} industry trends are identified by the actual calendar year and not the case-time. This allows the estimation to adjust for national industry trends during the sample period. The omitted time period is $t-1$ to account for anticipation effects ahead of the MDP announcement and to adjust for timing differences between when an article appears in *Site Selection* magazine and when the decision concluded. Across all specifications, standard errors are

[12] Out of 230 cases, only one county-industry pair wins a Million Dollar Plant more than once.

always clustered at the county level to adjust for shocks that correlate across all industries and firms within a county.

After showing results at the cluster-level using publicly available data to estimate Equation 2, the analysis proceeds to startup and founder analysis in the microdata. Restrictions to cell counts for disclosure require estimating a version of the Greenstone, Hornbeck, and Moretti (2010) specification that reduces the time interactions in Equation 2 to county pre- and post-trend variables. This analysis also focuses on cohorts of startup firms rather than county-industry pairs. I use the sample of startup firms aged 0 or 1 in each year $\in [t-5, t+7]$ and show changes in growth and survival rates after the introduction of a Million Dollar Plant. The main estimating equation in the microdata is:

$$\begin{aligned} y_{fijct} =& \gamma_0 + \beta_1 \times \textit{Pre-trend}_{cijt} + \beta_2 \times \textit{Post}_{jt} + \beta_3 \times \textit{Win}_{fjc} + \\ & \beta_4 \times \textit{Win}_{fjc} \times \textit{Post}_{jt} + \beta_5 \times \textit{Win}_{fjc} \times \textit{Post-trend}_{cijt} + \\ & \alpha_j + \lambda_{it} + \epsilon_{fijct}, \end{aligned} \tag{3}$$

where f denotes a firm. y_{fijct} are the startup-level outcomes of interest: 5-year growth rate and 5-year failure probability. Following the intuition of Equation 2, Equation 3 includes the α_j case fixed effects to force identification of the model's parameters to be from comparing the winning county to its next-best county *within* each case. β_4 measures the shift in the mean outcome after the MDP announcement relative to counterfactual counties and β_5 summarizes change in the slope of the outcome during the post-period in winning counties.

4.2 Balance and Parallel Trends

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.

Our first test on this identifying assumption compares similarity of winner and runners-up counties on a variety of observed demographic and economic characteristics. I do this using a regression approach where I test whether variables commonly thought to predict local firm formation also predict whether a county ultimately wins the Million Dollar Plant. Specifically, using the final

MDP data set described in Section 3.1 I estimate the following regression:

$$\mathbf{1}[Win]_{j,c} = \alpha + X_{j,c}\beta + \alpha_j + \alpha_c + u_{j,c} \tag{4}$$

The outcome variable is a dichotomous variable taking the value one if the county wins the MDP event and is regressed on a vector of covariates, $X_{j,c}$, that include variables that correlate to local economic conditions conducive to both startup formation as well as a potential megafirm's decision to open a MDP there. I use case fixed effects, α_j, to force β coefficients to be identified within rather than across cases. Because some counties may be considered in multiple site search contests, I also include county fixed effects α_c. Figure 5 plots the β coefficients which show whether the corresponding county characteristic predicts the county's win status for a case.

$X_{j,c}$ variables include immigration status, education, size of the manufacturing base (Chinitz, 1961), joblessness (Fairlie, 2013), community banking, and pre-existing supply chain presence (Glaeser, Kerr, and Kerr, 2015). I run Equation 4 on two sets of variables. The first set (Panel A and Panel B) use growth variables to test pre-announcement trends between year $t-5$ and year $t-1$. The second set of (Panel C and Panel D) are variables tabulated in year $t-1$. Panel B and Panel D also include two variables related to local banking conditions.

Figure 5 shows that within each case, the vector of observable traits do not have predictive power in which county wins the Million Dollar Plant. I confirm this using both growth variables measuring for changes in the county's economic conditions in the years leading up to the announcement decision and static variables at the time of the announcement.

5 Results

5.1 Cluster Formation

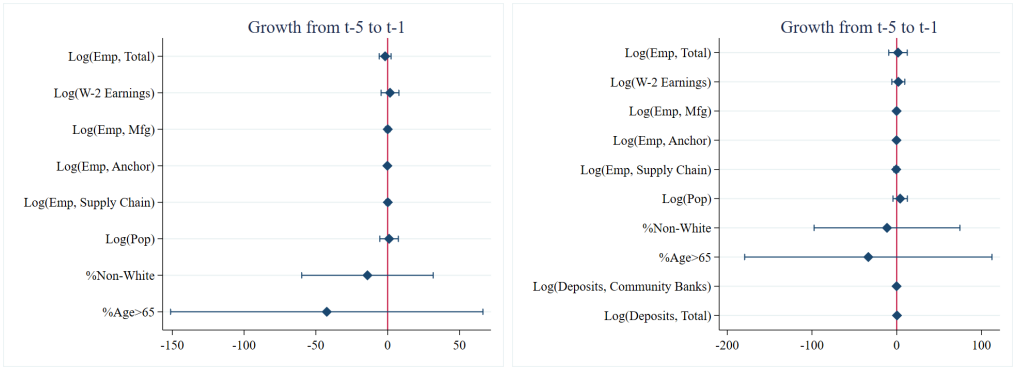
5.1.1 Industry Growth

The first portion of the analysis serves to validate the use the Million Dollar Plant openings as a plausibly exogenous shock to an industrial cluster and county-industry employment. Because this analysis is aggregated to the county-industry level and excludes entry or firm dynamics, I use publicly available County Business Patterns data with imputations for missing data (Eckert et al.,

Figure 5: Similarity of Winning & Runners-Up Counties

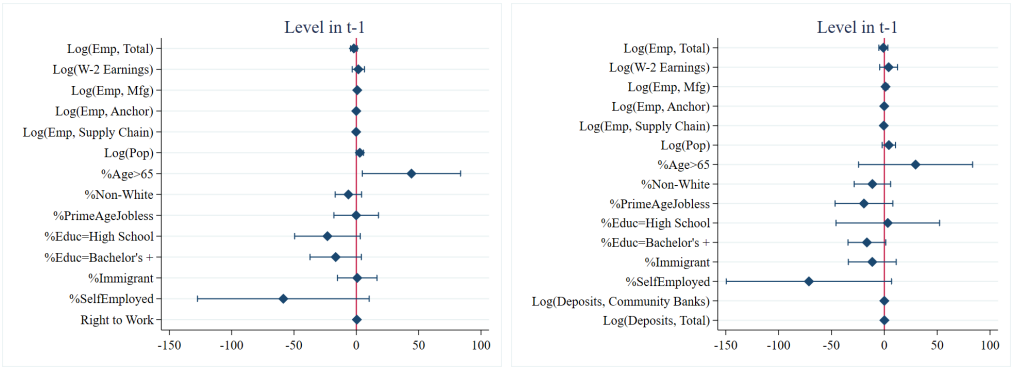
(a) Excluding Bank Sample

(b) Including Bank Sample



(c) Excluding Bank Sample

(d) Including Bank Sample



Notes: Figure 5 plots β coefficients and confidence intervals when regressing case-county win status on a vector of county characteristics and fixed effects for case and county as shown in Equation 4. Population, %Non-White, and %Age¿65 are obtained from the National Cancer Institute's SEER population characteristics registry. Employment values (Emp, Total; Emp, Mfg; Emp, Anchor; Emp, Supply Chain) are tabulated from County Business Patterns using Eckert et al. (2021) imputations at the 4-digit NAICS level. %PrimeAgeJobless, %Educ=High School, %Educ=Bachelor's+, %Immigrant, and %SelfEmployed are obtained from the 5% PUMS (Ruggles et al., 2020) sample from Census 1990 or 2000 depending whichever decennial Census immediately preceded the MDP announcement. %PrimeAgeJobless is the percentage individuals between 25 and 55 and identifying as male in the Census not working or working fewer than 40 hours per week. Self-employment is the numbed of individuals working at least 40 hours as self-employed divided the total number of full-time adult workers. Average income is tabulated at the household-level using the IRS Statistics of Income files for year preceding the opening announcement. 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. Right to Work is an indicator variable taking the value one if the county is in a state with either a statute or constitutional amendment defined as right to work in year $t-1$ relative to the contest.

2021).

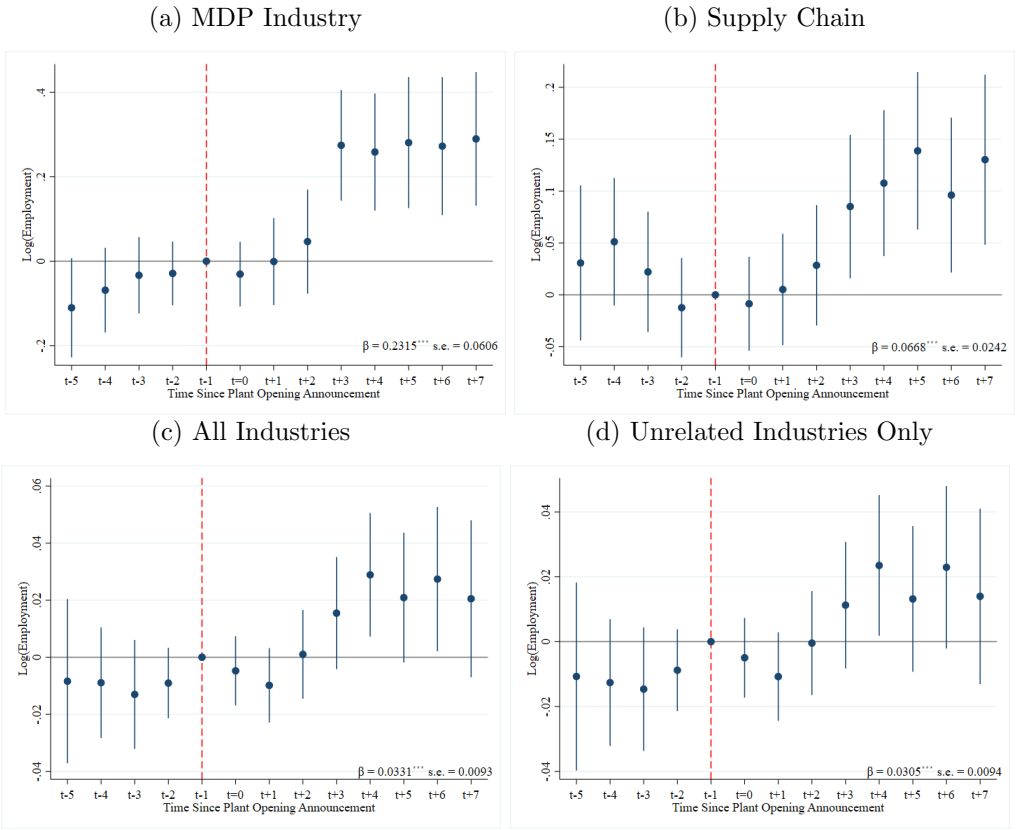
The public data enables displaying estimates of the full dynamic specification from Equation 2. In this first exercise, I regress log employment in a county-industry pair (cluster). Because the excluded time period is $t-1$ the coefficients are interpreted as employment growth relative to the counterfactual county in the year prior to the *Site Selection* announcement. Figure 6 presents the evolution of cluster employment growth following the announcement of a Million Dollar Plant. To facilitate interpretation of the dynamic difference-in-differences, I also present estimates from a simple difference-in-differences estimation to summarize the average effect of the MDP. More specifically, in the bottom right corner of each panel in Figure 6 is the estimated β from the following regression:

$$emp_{cijt} = \alpha_0 + \beta_W \mathbf{1}(Win_{cij}) + \beta \mathbf{1}(Win_{cij}) \times \mathbf{1}[t \geq 0] + \eta_c + \lambda_{jt} + \gamma_{it} + \epsilon_{cijt}. \tag{5}$$

The important detail in Equation 5 is that the λ_{jt} forces the DiD estimator to be identified by comparing the winning county cluster to the finalist county cluster in the same case-time to prevent comparisons between different times, cases, or case-times from biasing estimates (De Chaisemartin and d'Haultfoeuille, 2020).

Figure 6 shows county-industry (cluster) growth after the arrival of the Million Dollar Plant. As expected, the MDP leads to an economically and statistically significant increase in employment in its own industry (Panel A). It is also noteworthy that employment gains commence three years after the announcement year, though prior literature documents that agglomeration spillover effects on measured incumbent plant productivity and management practices accrue immediately after announcement (Bloom et al., 2019; Greenstone, Hornbeck, and Moretti, 2010; Greenstone and Moretti, 2003). Panel B expands analysis to industry clusters that are part of the MDP industry supply chain and similarly shows supply chain industries see a significant increase in employment starting three years after the announcement. Despite that these Million Dollar Plant openings are significant economic events that might spur job creation across a wide range of industries, we see that its overall impact on industry growth is limited. Panel C shows only slight increases when expanding the sample to include all industries in the county and Panel D estimates employment changes in industries that are unrelated to the MDP industry.

Figure 6: Cluster Growth After Million Dollar Plant Arrival



Notes: Figure 6 plots β_k time coefficients and confidence intervals estimated from Equation 2. The panels also present the β estimate from Equation 5 so summarize the average effect of the MDP on winning county industries relative to finalist counties. Employment values are obtained from County Business Patterns using Eckert et al. (2021) imputations at the 3-digit NAICS level. I use the 3-digit NAICS to facilitate matching with the Bureau of Economic Analysis Use of Commodities matrix to construct supply chains for each Million Dollar Plant. Standard errors are clustered at the county-level to capture correlation in employment changes across industries in the same county.

Though Panel C shows a broad increase in winning county industry growth relative to counterfactual counties, growth disproportionately driven by the Million Dollar Plant industry (Panel A) and its linked supply chain industries (Panel B). Focusing on the set of industries that are neither the MDP industry nor its supply chain, undergo relatively muted effects suggesting the shock of the MDP's arrival is specific to its supply chain and economically proximate industries. Placebo tests validate that the treated industries are supply chain industries and that the MDP does not lead to widespread growth in industries and firms unrelated to the MDP (Figure A.4). In the subsequent sections, I show firm and founder dynamics in these treated industries.

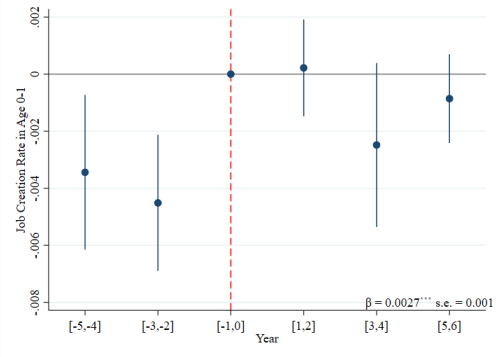
5.1.2 Startup Clusters in the Supply Chain

Million Dollar Plants lead to local growth of industries related to the large plant. These are jobs effects that are typically cited in proposals to attract these large establishments to communities. However, to understand who populates clusters induced by these place-based business subsidies to attract MDPs, I examine outcomes related to business formation and entrepreneurial success. These outcomes related to new firms are important for understanding two key considerations in understanding the implications for industrial policies that lure large firms to specific locations: i) Concern over supporting megafirms at the expense of entrepreneurial ventures that could risk amplifying market power of select firms and ii) Understanding implications for firms and workers already in these locations.

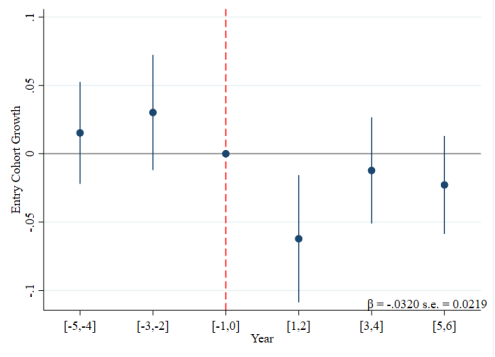
I utilize publicly available Quarterly Workforce Indicators (QWI) and the Census Longitudinal Business Database (LBD). The QWI allows researchers to examine growth in employee counts by firms born in two year intervals up through age of 5. The publicly available data is at the county-sector-age level where firm ages are grouped by age 0-1, 2-3, 4-5, 6-10, and 11+. In year t the firm age bin 0-1 represents employees in firms born in year t and $t-1$. In year $t+2$, these firms graduate to the age 2-3 firm age bin. In year $t+4$, this cohort has progressed into firm age bin 4-5. By comparing jobs in firm age 4-5 in year $t+4$ to jobs in the 0-1 bin in year t, I calculate startup cohort growth rates. This structure facilitates estimating the impact of the MDP's arrival on total job creation in new firms and whether the MDP leads to faster growing startup cohorts. I estimate a version of Equation 2 using the startup job creation rate and 5-year cohort job growth rate as the outcomes.

Figure 7: Startup Cohort Analysis

(a) Age 0-1 Job Creation



(b) Cohort Growth



Notes: This figure presents estiamtes of Equation 2 with the TWFE estimator from Equation 5. The TWFE is provided for ease of interpreting mean shifts. The outcome in panel A is the share of sector employment in firms aged 0 or 1. The outcome variable in panel B is the 5-year growth rate of the startup cohort. The data are tabulated using publicly available Quarterly Workforce Indicators. QWI provides employment totals by firm age at the county-sector (2-digit NAICS) level. The age bins facilitate tracking employment growth in entry cohorts at two year intervals up through the age of 5. Models include county and sector by year fixed effects. Standard errors are clustered at the county-level to adjust for auto-correlation across sectors within a county.

Figure 7 shows the dynamic estimates from Equation 2 as well as the Two-Way Fixed Effects estimator from Equation 5. I measure the intensity of startup job creation as the number of jobs in firms aged 0-1 in county c sector s, emp^{0-1}_{csjt}, divided by county-sector employment in the year $t-1$ relative to the MDP announcement as the outcome, $emp_{csj,t-1}$. Panel A shows that relative to counterfactual counties, winning counties had much lower job creation intensities in startup firms. However, post-announcement these counties experience a significant increase in startup jobs. The aggregate nature of the QWI data prevents creating supply chain linkages using subsector industry codes. Nevertheless, the MDPs arrival leads to more jobs in firms aged 0-1 even after controlling for sector-specific effects. However, this increase could represent entrepreneurs who move to the county chasing opportunities behind the MDP or uptake in workforce development programs, better supplier networks (Glaeser and Kerr, 2009), and other policy-related changes that lead to increased entry after the MDP. However, these new startup firms may be of low quality and not lead to longer-term employment in the county. Panel B employs five-year startup cohort growth as the outcome of interest. Each estimate in Panel B represents growth of an entry cohort relative to the entry cohort in the counterfactual county. The dynamic estimates show a slight *decline* in cohort growth after the MDP's arrival, though the TWFE estimator is noisily estimated.

While the public data shows that MDP's lead entrepreneurs to start new businesses (Panel A)

Table 4: Supply Chain Startups

	(1)	(2)	(3)	(4)
	Startup Entry Rate	Startup Employment Share	Establishments of New Firms	Employment from New Firms
Constant	7.389***	3.301***	81.00***	1212***
	(0.1135)	(0.1234)	(5.974)	(85.17)
Win	0.0001	0.0540	-0.8781	23.62
	(0.1084)	(0.1116)	(10.24)	(125.2)
Post	-0.5587***	-0.7012**	-6.607*	-489.1***
	(0.1999)	(0.2723)	(3.446)	(120.3)
Win * Post	0.8692***	0.7101**	22.05*	429.6**
	(0.2549)	(0.2790)	(11.98)	(185.1)
Win*Post-trend	-0.0540***	-0.0473***	-1.538*	-28.82**
	(0.0162)	(0.0176)	(0.892)	(13.98)
Pre-trend	-0.0597***	-0.0509**	-0.573**	-34.93***
	(0.0147)	(0.0223)	(0.236)	(10.83)
Industry x Year FE	Y	Y	Y	Y
Case FE	Y	Y	Y	Y
Effect in t+5	0.5388***	0.4203**	12.82*	256.7**
Adj. R-Sq.	0.6535	0.3240	0.7552	0.6032
N	91500	91500	91500	91500

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. 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 opening 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 entry rate in county c industry i. The entry rate is calculated as the number of establishments associated with firms age 0 in a year divided by the total number of the cluster's business establishments one year prior. Column 2 shows the share of total industry employment in new establishments of firms aged three years or younger. Columns 3 and 4 are in levels to highlight the number of new businesses and jobs created in new businesses after the MDP's arrival. "Effect in t+5" shows the MDP's impact on startup formation and job creation in year t+5 after the announcement. Standard errors are clustered by county to capture shocks correlated across industries within a county. All regressions are weighted by total county-industry employment. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8124, CBDRB-FY19-P1841-R8165). Values are rounded to comply with Census disclosure rules.

but that have potentially worse outcomes (Panel B) when adjusted for cohort effects, Figure 7 does not facilitate observing startup-level outcomes or linking of industries to each MDP's supply chain. Turning to the administrative data, I estimate Equation 3 using various measures of entrepreneurial entry and startup jobs as the outcomes of interest by 4-digit NAICS industry. The sample includes firms in industries in the supply chain of the Million Dollar Plant. Column 1 of Table 4 shows the effect of the MDP's impact on the entry rate of new businesses in a cluster. Specifically, the outcome is the number of new establishments in firms aged less than 3-years divided by the total number of business establishments in the county-industry cluster in the prior year. For ease of interpretation, values in columns 1 and 2 are multiplied by 100. The standard $Win \times Post$ estimator suggests increase in new business establishments of startup firms amounting to approximately 0.87%. The dynamic estimates using public data suggested effects of the MDP vary by time. The $Win \times Post - Trend$ captures whether the slope of outcomes in the winning county differ after

the arrival of the MDP. We see that the MDP's effect on entry rates diminishes each year after the announcement. Using these coefficients, I estimate that five years after the MDP's announcement, clusters in winning counties have a 0.54% higher entry rate. Column 2 measures the share of county-industry employment in these new establishments of startup firms. I find that the MDP leads to 0.71% greater startup employment shares with the size of the effect decreasing by 0.05% each year after the announcement. In columns 3 and 4, I show levels rather than rates (column 1) or log growth (column 2) to provide context for the absolute size of the MDP's effects. Column 3 shows that 22 new business establishments in each supply chain industry after the arrival of the MDP that employ approximately 430 employees (column 4). Much of these gains are concentrated in the earliest years post-MDP announcement, though 5-years post-announcement young firms still create 12.8 establishments in a year that employ 257 people. For a given case these magnitudes sum to over 120 total startup establishments and 2,300 startup jobs across all the MDP's supply chain industries. Publicly available County Business Patterns data places these total effect sizes as a 10 percentage point increase in the size of a local supply chain following the arrival of a MDP that is directly attributable to startups.

The next set of estimates mimic Panel B in Figure 7, but in the LBD. 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 are predicted to be of lower quality and therefore grow less and fail faster. To measure startup growth at the firm-level, I employ the Davis, Haltiwanger, and Schuh (1998) growth measure (DHS) for firm f in year t with a set of establishments k:

$$g_{ft} = \sum_{k \in f} \gamma_{kt} \times \frac{\text{emp}_{k,t+5} - \text{emp}_{kt}}{0.5 * (\text{emp}_{kt} + \text{emp}_{kt+5})} \tag{6}$$

where γ_{kt} is establishment k's share of firm f's total average employment between year t and $t+5$. This measure allows us to capture organic growth of young firms that is not driven by mergers and acquisitions. The measure also facilitates capturing exit since it is bounded by -2 reflecting shutdown and 2 representing entry. I estimate Equation 3 using the and report impact of the MDP's announcement on the DHS growth measure and report the estimates in Column 1 of

Table 5. Startup firms in supply chain industries grow 12.25% faster after the MDP announcement, consistent with findings thus far this effect diminishes over time (-.87% per annum). In column 2, I ask whether these startup firms in winning counties have better survival rates. Specifically, I regress a dichotomous variable that takes the value of 1 if all establishments of the startup firm exit within five years of the firm's birth and report estimates in column 2. I find that supply chain startups have failure rates that are 6.8% lower in winning counties relative to counterfactual counties following the MDP, though the effect diminishes each year. To investigate heterogeneity within the supply chain, I split the results by upstream and downstream in Tables A.1 and A.3. The tables show the impact on upstream startups–startups that provide inputs to the MDP–are stronger than on startups in industries that use inputs from the MDP's industry. As an additional check, I show that my overall effects in Table 5 are not driven by policies designed to induce growth in the MDP's industry specifically. Table A.2 shows null effects on startups in the MDP's own industry.

Table 5: Startup Growth and Survival

	(1)	(2)
	5-year growth rate	5-year failure rate
Constant	-0.3378***	0.166***
	(0.0160)	(24.20)
Win	-0.0410	0.0125
	(0.0346)	(0.76)
Post	-0.0190	0.0113
	(0.0182)	(1.60)
Win * Post	0.1225*	-0.0680*
	(0.0663)	(-2.55)
Win*Post-trend	-0.0087*	0.00487*
	(0.0046)	(2.47)
Pre-trend	-0.0022	0.00117
	(0.0020)	(1.50)
Industry x Year FE	Y	Y
Case FE	Y	Y
Adj. R-Sq.	0.0988	0.133
N	215000	215000

Notes: * 0.1 ** 0.05 *** 0.01. See Table 4. Supply chain startups grow 12% faster in employment over the first five years and fail 7% less. Column 1 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 (1998) 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 2 displays the impact of a facility's opening on the likelihood the startup fails within five years. Standard errors are clustered by county to capture geographic correlation of shocks affecting all startups in a county. All regressions are weighted by total county-industry employment. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8124).

5.1.3 Robustness

The main effects of the MDP's impact on cluster formation are robust to a variety of alternative specifications and falsification tests. These additional tests highlight that winning a Million Dollar Plant has a sizable impact on local supply chain industries and that spillover effects onto other industries is *limited*.

Weighting: The regressions analysis presented thus far do not weight county-industry clusters by size. However, we may expect that clusters that are *ex ante* larger undergo smaller effects from the MDP because the MDP only accounts for a small share of local employment. Figure A.1 presents estimates of Equation 2 when weighting county-industry observations by log employment in year $t-1$. We see that the employment growth dynamics are broadly similar to the main results presented in Figure 6.

Growth Calculation: Figure 6 presents estimates for the impact of the MDP on log employment in a county-industry by year. While this setup facilitates interpreting time coefficients as representing employment growth relative to year $t-1$, it drops county-industry pairs with zero employment. This potentially attenuates estimates by ignoring cases where the MDP was the initial source of employment in those clusters. The measure in Equation 6 from Davis, Haltiwanger, and Schuh (1998) bounds growth rates from -2 to 2 where -2 represents employment collapsing entirely in a cluster and 2 the initial employment in a new cluster. In Figure A.2, I present results using this alternative growth calculation. The main results are similar to our main estimates, though effects in non-supply chain industries are weaker.

Multiple Treatments: Counties appear in multiple Million Dollar Plant contests as both winners and runners-up. If a county wins multiple events, then our estimates are confounded by multiple treatments and bias estimates upwards. To ensure results are not biased by counties receiving multiple Million Dollar Plants or from counties that previously won an MDP but are now runners-up, I drop counties after their first win and re-estimate Equation 2. Figure A.3 shows robustness to dropping these counties with results on the MDP and supply chain industries potentially stronger and spillover effects in unrelated industries weaker.

Placebo Tests: This paper uses the arrival of a Million Dollar Plant as a shock to a specific local industry by affects its labor market size. Shocks to this focal industry affect other industries that are linked through input-output channels. However, an alternative explanation to the results is that Million Dollar Plants entail large policy changes and subsidies at the local level that lead to growth across sectors or that as the county develops, all treated industries see an increase in growth and entrepreneurship. To test this alternative hypothesis, I randomize the treated industry within each MDP event and construct a "placebo" supply chain and re-estimate the Equation 5 as if the winning county received a plant in the placebo industry. I repeat this exercise 1,000 times and plot the kernel density of the DiD estimator in Figure A.4. The peak of the kernel density for the placebo MDP (red solid lines) and supply chain (blue dashed lines) are different from the actual estimator. The placebo supply chain is indeed closer to the actual than the placebo MDP industry because of overlap in supply chain industries, though the actual estimator is still in the right tail of the kernel density. Overall, the placebo tests confirm that the effect of the MDP is localized to the specific MDP industry of treatment and its most related industries.

Case Sensitivity: Another concern is that results are driven by single canonical examples of Million Dollar Plants such as the arrival of BMW to Spartanburg, SC in 1992 (Greenstone, Hornbeck, and Moretti, 2010). To test for this, I re-estimate the DiD model excluding one case each time. Figure A.5 plots the kernel densities of the coefficients obtained in this procedure. The peaks of the kernel densities are all tightly centered about the actual coefficient. This test supports that the MDP results are broader than any single large MDP event.

Industry Heterogeneity: Turning to startup cohort results, effects from Million Dollar Plants may vary by industry. The typically larger and capital intensive manufacturing MDPs may lead to larger direct employment effects but with muted, or even, negative impacts on startups. Prior literature documents that cities dominated by large manufacturing establishments typically see depressed entrepreneurship in subsequent years (Glaeser, Kerr, and Kerr, 2015). By contrast, we might expect large R&D intensive MDPs facilitating startup formation (Agrawal et al., 2014). Figure A.6 presents a subset of cases in Figure 7 that relate to manufacturing establishments or R&D facilities. Figure A.6 shows broadly similar trends to the main results. We observe an

increase in job creation in new firms, but depressed growth rates in these new startups. These results also contrast with the microdata which enables us to go beyond the sector-level and explore heterogeneity in related 4-digit NAICS industries, confirming that the impact of MDPs are highly localized and confined to industries economically proximate.

5.2 Founder Characteristics

In this section, I explore whether the success of startup cohorts, particularly those shortly after the MDP announcement, is driven by founders with local industry experience transitioning into entrepreneurship at higher rates than in counterfactual counties. If high quality firms incubate entrepreneurs (Chatterji, 2009; Gompers, Lerner, and Scharfstein, 2005), then proxies of quality such as wages will positively correlate to the number of entrepreneurs they produce. If local industry experience is essential for successful entry post-MDP then not only will high quality firms generate more entrepreneurs, the entrepreneurs will select into their employer's industry.

Following the convention of Klepper entrepreneurship through spin-outs, I use the terms spinoffs and spin-outs to refer to workers become a top 3 wage earner in startups (i.e. "founders") in the same industry and county as their former employer. To explore the variation across incumbent firms by their positioning in the MDP's supply chain, I estimate the following:

$$\begin{aligned} y_{fijct} = \beta_0 + \beta_1 \times Win_{fjc} + \beta_2 \times Upstream_{fijc} + \\ \beta_3 \times Anchor_{fijc} + \beta_4 \times Downstream_{fijc} + \beta_5 \times Upstream_{fijc} \times Win_{fjc} + \\ \beta_6 \times Anchor_{fijc} \times Win_{fjc} + \beta_7 \times Downstream_{fijc} \times Win_{fjc} + \\ \beta_8 \times X_{fjt} + \alpha_j + \lambda_{it} + \epsilon_{fijt}, \end{aligned} \tag{7}$$

where y_{fijct} is the number of employee departures from firm f into entrepreneurship and X_{fjt} is a vector of firm-level controls such as firm age and size. Following the sample construction outlined in Section 3.3, these regressions compare incumbent firms in winning and runners-up counties in year $t-2$ relative to the MDP announcement. The outcome of interest is the number of employees who depart these incumbent firms to found a new firm by the year $t+5$. The coefficients of interest are β_5, β_6, and β_7 which provide estimates of whether incumbent firms in the MDP's supply chain lose disproportionately more employees to entrepreneurship than firms outside of the

supply chain.[13] Table 6 presents estimates of the main coefficients of interest from Equation 7. The column headings correspond to the number of entrepreneurs departing incumbent firms to industries grouped by their relation to the MDP's industry.

Table 6: 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. 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." 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. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8165).

Column 1 in Table 6 shows the total number of employees who depart incumbent firms to start a new firm regardless of industry. Column 1 shows that firms in winning counties do not lose more employees to entrepreneurship than those in counterfactual counties. This fact is consistent regardless of whether the incumbent firm is in the same industry as the MDP or in its supply chain. Column 2 shows that supply chain firms in winning counties do spawn more upstream firms. This effect is particularly pronounced for incumbents operating the MDP's industry. Column 3 similarly shows that supply chain incumbents spawn more entrepreneurs in the MDP's industry, and column 4 shows the effect is similarly strong for downstream entrepreneurship. The diagonal coefficients in columns 2-4 are interpreted as *spinoffs* since these are individuals who depart supply chain firms to start a new firm in the same supply chain segment as their employer. Aside from upstream startups that are disproportionately generated by anchor industry incumbents, we see a strong effect of the MDP on supply chain spinoffs. On average, upstream incumbents spawn .594 spinouts,

[13] The excluded interaction in Equation 7 is "all other industries".

MDP industry firms generate .570, and downstream firms generate .650 spinoffs. Therefore the diagonal coefficients imply an increase of 30%, 82%, and 61% upstream, MDP industry, and downstream spinout employees. Importantly, though, the total number of entrepreneurial departures is unchanged relative to counterfactual counties (column 1), but the composition of entrepreneurship is what changes by the MDP's arrival.

Table 6 compares extant incumbents in winning counties to those in counterfactual counties. This means that the estimated coefficients may represent declines in entrepreneurship in the counterfactual counties rather than an increase in winning counties. I provide suggestive evidence that the coefficients in Table 6 are indeed driven by increases within winning counties. I use the sample of incumbent firms in winning counties in year $t-8$ and tabulate the number of incumbent employees who found a firm by year $t-2$. I then use the same treatment group as in Table 6 and re-estimate Equation 7, substituting $Post$ for Win. The estimated coefficients represent changes in the number of employee departures within winning counties. Table A.6 confirms that incumbent firms in winning counties lose more employees to entrepreneurship after the anchor firm arrives. Upstream incumbents generate .1797 more local upstream cluster entrepreneurs, an effect that represents about a 30% increase above relative to the baseline upstream spinoff rate (.1797 / .5940). Downstream incumbents in winning counties generate 60% more downstream spinoffs than prior to the MDP (.3918 / .6500). These estimates also mimic estimates in Table 6, suggesting that the effects identified are indeed driven by the MDP's arrival to the winning county and not by declines in runner-up counties.

The results so far suggest that the total number of entrepreneurs in a county does not change after the arrival of a MDP, but the composition of new ventures shifts in favor of supply chain industries. To highlight this compositional change, I investigate characteristics at the individual level to test whether having experience in a supply chain industry correlates to an eventual entrepreneur selecting into an MDP's supply chain industry. To do this, I re-estimate Equation 7 at the individual level and use as the outcome a dichotomous variable that takes the value of 1 if the individual is a firm founder in the MDP's supply chain. The sample consists of all individuals with full-time employment in either the winning or runner-up county in year $t-2$ who found a firm by

year $t+5$. Specifically, I estimate the following:

$$\begin{aligned}\mathbf{1}[SupplyChainFounder]_{pijc} = \beta_0 + \beta_1 \times Win_{pjc} + \beta_2 \times Upstream_{pijc} + \\ \beta_3 \times Anchor_{pijc} + \beta_4 \times Downstream_{pijc} + \beta_5 \times Upstream_{pijc} \times Win_{pjc} + \\ \beta_6 \times Anchor_{pijc} \times Win_{pjc} + \beta_7 \times Downstream_{pijc} \times Win_{pjc} + \\ \beta_8 \times X_{pjt} + \alpha_j + \lambda_{it} + \epsilon_{pijt},\end{aligned} \tag{8}$$

where $\mathbf{1}[SupplyChainFounder]_{pijc}$ is one if the entrepreneur eventually founds a startup in the supply chain of the MDP. X_{pjt} is a vector of individual characteristics that includes age, age-squared, sex, race, and educational attainment. Because Equation 8 explores entrepreneur experience based on industry of employers in year $t-2$, the model does not include SEIN fixed effects. Therefore, it is possible that estimates also reflect sorting of entrepreneurial individuals into firms at $t-2$. The identifying assumption is that this sorting is consistent across both winning and runners-up counties.

Table 7 shows the effect of the MDP's arrival on selection into supply chain entrepreneurship. 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) industries are 8.7% (4.4%) more likely to found an upstream (anchor) startup. The strongest effect is found in entrepreneurs from downstream incumbents who become 17.2% more likely to have their startup in a downstream industry. Table 7 also shows having experience in an anchor industry only has a significant and positive impact on founding an MDP industry startup and corresponds to a strongly negative (-17.3%) impact on founding a downstream venture.

Focusing on the worker sample facilitates uncovering distributional impacts of attracting MDPs. One key policy relevant margin of interest is whether attracting an MDP leads to entrepreneurship by individuals who are of higher income or reduces the threshold needed to select into supply chain entrepreneurship. To investigate this, I run analogous regressions to Equation 8 by adding an

Table 7: 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-squared, sex, race, and education. Regressions are unweighted and standard errors are clustered at the SEIN-county level. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8124).

additional interaction with the eventual entrepreneur's normalized wage:

$$\begin{aligned}\mathbf{1}[SupplyChainFounder]_{pijc} = {} & \beta_0 + \beta_1 \times Win_{pjc} + \beta_2 \times Win_{pjc} \times W_{pjc} + \\ & \beta_3 \times Upstream_{pijc} + \beta_4 \times Anchor_{pijc} + \beta_5 \times Downstream_{pijc} + \\ & \beta_6 \times Upstream_{pijc} \times Win_{pjc} + \beta_7 \times Anchor_{pijc} \times Win_{pjc} + \\ & \beta_8 \times Downstream_{pijc} \times Win_{pjc} + \beta_9 \times Upstream_{pijc} \times Win_{pjc} \times W_{pjc} + \\ & \beta_{10} \times Anchor_{pijc} \times Win_{pjc} \times W_{pjc} + \beta_{11} \times Downstream_{pijc} \times Win_{pjc} \times W_{pjc} + \\ & \beta_{12} \times X_{pj} + \alpha_j + \lambda_{it} + \epsilon_{fijt},\end{aligned} \tag{9}$$

where β_2, β_9, β_{10}, and β_{11} are the coefficients of interest. Table 8 estimates Equation 9 as a linear probability model. The main interaction in the first row demonstrates 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 summarize the heterogeneous effect of income and being in a winning county by supply chain industry. 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 upstream entrepreneurship. This pattern applies to entrepreneurs initially employed in the MDP's industry as well. However, wages do not appear to have a significant impact on the founding industry of entrepreneurs from downstream employers.

Table 8: Wages and Likelihood of Founding Firms

	(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. Log wages are normalized to have mean 0 and unit standard deviation and are based on the entrepreneur's most recent wage-employment spell prior to the MDP. Interactions are conducted after standardization to restore main effects. Controls include age, age-squared, sex, race, and education. Regressions are unweighted. Standard errors are clustered at the SEIN-county level. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8124).

5.3 Alternative Explanations

Thus far I have shown that entrepreneurial entry in counties winning a Million Dollar Plant is driven by job transitions of employees already employed by incumbents in the MDP's supply chain. An alternative explanation is that the arrival of a MDP represents a potential business opportunity to incumbent firms in the existing local supply chain and employees leave bureaucratic firms that are unable to take advantage of these new opportunities. Klepper (2007), Klepper and Buenstorf (2009), and Klepper and Thompson (2010) describe this *clustering through disagreements* process as leading to the prevalence of automobile firms in Detroit, MI and the tyre industry in Akron, OH.[14] To test the entrepreneurship through disagreements hypothesis, I test whether larger and

[14]Other studies such as Oettl and Agrawal (2008) describe new firms forming by entrepreneurs who pursue projects complementary to their former employer.

older firms see more employee departures to entrepreneurship.

Table 9: 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. Regressions are unweighted. Standard errors are clustered at the SEIN-county level. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8165).

Tables 9 presents estimates of Equation 9 using *age* as the firm characteristic of interest. The table shows that older firms in winning counties generate more entrepreneurs who start firms upstream and anchor industries, though this effect is not economically meaningful. The triple interactions show the effect does not vary by industry grouping and in some cases older firms spawn more entrepreneurs. Table 10 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.

Firm size and age regressions do not provide particularly strong evidence that would be consistent with an entrepreneurship through disagreements story. However, this does not mean that these explanations are necessarily incorrect in this context as the administrative data does not afford a detailed examination of internal firm bureaucracies. The microdata does facilitate testing predictions from the literature motivated in personnel economics (Lazear and Shaw, 2007; Lazear and Oyer, 2004). Mean firm wage may be closely related to employee departures (Campbell et al., 2012; Carnahan, Agarwal, and Campbell, 2012; Franco and Filson, 2006). Operating under the assumption that a firm's wages relative to the industry are reflective of firm quality, I estimate

Table 10: 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. Firm size is calculated in the LEHD sample. Size is tabulated from the number of employees who worked in the firm during the year with an average quarterly wage above minimum wage in all quarters worked. Controls include age, age-squared, sex, race, education, and SEIN fixed effects. Regressions are unweighted. Standard errors are clustered at the SEIN-county. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8124).

Equation 9 using mean wages of workers in the firm calculated in the LEHD.[15] The variable is standardized to mean zero 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 the effect of firm compensation relative to the industry on employee departure to entrepreneurship.

The main wage interaction with the winning county dummy variable in Column 1 of Table 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 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 11 show higher paying firms generate more entrepreneurs and particularly into their own supply chain category. Upstream (downstream) firms have significantly

[15]Firm mean wages are tabulated as the average quarterly earnings of workers retained in the sample by SEIN-county in the Employee History File (EHF).

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 6 and Table 7 that the number of employee departures from anchor industry incumbent is concentrated in a limited set of firms, Table 11 shows that the firms generating entrepreneurs tend to be higher paying firms.

Table 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-county employment. Standard errors are clustered at the SEIN-county. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8165).

Finally, I consider the possibility that proximity to a large buyer in the form of the MDP is not the relevant channel but rather occupational similarity. It is possible that the MDP represents a large shock the demand for local factors of production such as worker skills and land. In the short-run, wages and rents rise quickly and forces incumbent firms to perhaps shed workers. If this is the case, then I expect to find that industries that use labor similar to the MDP will be the industries that employees depart. In effect, these departures are not necessarily departures by choice but necessity. I confirm that supply chain industries are also linked through shared labor inputs using the Bureau of Labor Statistics (BLS) Occupational Employment Statistics (OES). I tabulate the percentage of industry employees in each occupation and compare the occupational share of industry across anchor and supply chain industry (Tables A.10 and A.11). MDP-supply chain 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 in the US economy (0.0598). 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 broadly speaking and not necessarily solely attributable to an entrepreneur's proximity to a large buyer.

6 Discussion

Clusters with thicker initial labor market size experience higher rates of entrepreneurship that is driven by local founders with industry experience (Figure 1). I used Million Dollar Plants as a shock to labor market size within a cluster and investigated its impact on related clusters through input-output linkages. I found that the MDP induced employment growth in supply chain industries that was facilitated by startup formation. In fact, a larger share of supply chain jobs are concentrated in startup firms after the arrival of the MDP. Interestingly, these startups were not merely founders chasing a major news event, rather these startups grow faster and have lower 5-year failure rates than those in counterfactual counties. Founders are typically well positioned to enter supply chain entrepreneurship with local cluster experience.

These findings are consistent with theories of startup performance in clusters driven by both proximity to buyers and suppliers as well as specialized human capital (Glaeser and Kerr, 2009). 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. This paper's key contribution to the entrepreneurship literature is that spinoffs can be induced by shocks outside of the focal firm or industry.

6.1 Managerial and Policy Implications

I build on the literature describing entrepreneurship through employee spinoffs (Klepper, 2007; Franco and Filson, 2006; Klepper and Sleeper, 2005). I show that the industry shocks propagate

through the supply chain and induce spinoffs. My findings imply that even place and firm-specific industrial policies affect other employers located proximate to firms receiving subsidies. I show that policies that induce large, multinational firms to open a new facility in a county leads to employee departures into entrepreneurship. Because these new entrepreneurs start firms in the same industry as their former employer, these new firms potentially represent local competition. This may relate to one rationale for utilizing non-compete agreements (Marx, 2022), though outside the scope of this study.

This paper is among the first to investigate the effect of place-based business subsidy programs on local entrepreneurship. While the public and labor economics literatures have begun to unpack the welfare implications of large subsidy offers, management and entrepreneurship scholarship have largely avoided investigating the implications of industrial policy on local firms. Policymakers use place-based subsidies to lure typically large, multinational firms to specific locations. The policy goal is often to revitalize communities that have seen large declines in employment and wages (Juhász, Lane, and Rodrik, 2023). However, policymakers are paradoxically confronted with a tradeoff of revitalizing communities through a large plant opening but through a subsidy that potentially favors large, megafirms at the expense of smaller or younger firms. I show that policymakers can induce startup formation by using Million Dollar Plants as an anchor for entrepreneurship.

6.2 Limitations and Welfare

The mechanics behind the firm formation I observe after a Million Dollar Plant is announced is left as a black box. I do not observe business transactions in the administrative data, so I cannot say that proximity to buyers or suppliers drives my result because new firms directly support the incoming plant. Unlike the Klepper studies on spin-offs through disagreements, I do not know whether employees leave their employers to market an idea the employer was reluctant to pursue or whether former employees compete directly against their employers and represents a business stealing effect. While I proxy for bureaucratic firms using age and size, I do not have detailed information on disagreements or rigidity of firm structure.

This paper stops short of discussing welfare implications or delivering a recommendation for the ideal subsidy size. It is possible that employee departures represents reorganization of incumbent firms as shed managerial layers (Caliendo, Monte, and Rossi-Hansberg, 2015) and become more

competitive (Bloom et al., 2019; Greenstone, Hornbeck, and Moretti, 2010) and these aggregate productivity gains lead to widespread improvements in local welfare. I also do not measure the net effect on jobs between employee departures and new hiring at new firms: Do the new firms merely represent a change in the firm age distribution of current employment levels? I also do not measure wage growth broadly speaking over all residents of the winning counties relative to those in counterfactual counties. Finally, while the narrow event window about the MDP announcement of $[t-5, t+7]$ is useful for identifying effects plausibly driven by the MDP increasing labor market size, it prevents examining longer term impacts of the MDP and cluster dynamics long after the MDP makes its decision.

While several notable studies (Slattery and Zidar, 2020; Slattery, 2018) obtain subsidy information on winning deals and model counterfactual subsidies as an English auction, I do not observe winning deals for all 230 cases in the *Site Selection* sample. Limited information and prior literature suggests that typical subsidy sizes for the plants included in the sample average over quarter billion dollars. I find that approximately 2,300 jobs are created in startups as a result of the MDP which amounts to approximately $110,000 per startup job, considerably more expensive than targeted startup programs such as set-asides (Chatterji, Chay, and Fairlie, 2014). Moreover, I find that the spillover effects of MDPs are limited to economically proximate industries.

References

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.

Babina, Tania. 2017. "Destructive Creation at Work: How Financial Distress Spurs Entrepreneurship." *U.S. Census Bureau Center for Economic Studies* Working Papers 17-29.

Bartik, Timothy. 2017. "A New Panel Database on Business Incentives for Economic Development Offered by State and Local Governments in the United States." *Prepared for the Pew Charitable Trusts* .

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, Xavier Jaravel, and Jann Spiess. 2022. "Revisiting Event Study Designs: Robust and Efficient Estimation.".

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

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with multiple time periods." *Journal of Econometrics* 225(2): 200–230.

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.

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

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.

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. "Clusters of entrepreneurship and innovation." *Innovation Policy and the Economy* 14(1): 129–166.

Chatterji, Aaron K, Kenneth Y Chay, and Robert W Fairlie. 2014. "The impact of city contracting set-asides on black self-employment and employment." *Journal of Labor Economics* 32(3): 507–561.

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

Chávez, Steff. 2022. "Foxconn in Wisconsin: Tech Mega-Deal Faces Reality Check." *Financial Times* (October).

Cobb, James C. 2019. "State Incentives for Attracting Business Were Controversial Long Before Amazon. Here's What That History Can Teach Us." *Time* (February).

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.

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

De Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review* 110(9): 2964–2996.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2015. "Where Has All the Skewness Gone? The Decline in High-Growth (Young) Firms in the U.S." *Center for Economic Studies Discussion Paper* Working Papers 15-43.

Delgado, Mercedes, Michael E Porter, and Scott Stern. 2010. "Clusters and Entrepreneurship." *Journal of Economic Geography* 10(4): 495–518.

Delgado, Mercedes, Michael E Porter, and Scott Stern. 2014. "Clusters, Convergence, and Economic Performance." *Research Policy* 43(10): 1785–1799.

Due, John F. 1961. "Studies of State-Local Tax Influences on Location of Industry." *National Tax Journal* 14(2): 163–173.

Eckert, Fabian, Teresa C. Fort, Peter K. Schott, and Natalie J. Yang. 2021. Imputing Missing Values in the US Census Bureau's County Business Patterns. NBER Working Paper 26632.

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

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

Feldman, Maryann P. 1999. "The new economics of innovation, spillovers and agglomeration: Areview of empirical studies." *Economics of Innovation and New Technology* 8(1-2): 5–25.

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

Fujita, Masahisa, Paul R Krugman, and Anthony Venables. 2001. *The spatial economy: Cities, regions, and international trade.* MIT press.

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., 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.

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.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of Econometrics* 225(2): 254–277.

Greenstone, Michael, and Enrico Moretti. 2003. Bidding for Industrial Plants: Does Winning a 'Million Dollar Plant' Increase Welfare? Working Paper w9844 National Bureau of Economic Research.

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. 2024. "Go West Young Firm: The Impact of Startup Migration on the Performance of Migrants." *Management Science* 70(7): 4824–4846.

Guzman, Jorge, and Scott Stern. 2015. "Where is Silicon Valley?" *Science* 347(6222): 606–609.

Jarmin, Ron, and Javier Miranda. 2002. "The Longitudinal Business Database." *Center for Economic Studies Discussion Paper* CES-WP-02-17.

Juhász, Réka, Nathan Lane, and Dani Rodrik. 2023. "The New Economics of Industrial Policy." *Annual Review of Economics* 16.

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.

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.

LeRoy, Greg. 2005. *The Great American Jobs Scam: Corporate Tax Dodging and the Myth of Job Creation.* Berrett-Koehler Publishers.

Marley, Patrick, and Jason Stein. 2017. "Foxconn Announces $10 billion Investment in Wisconsin and Up to 13,000 Jobs." *Milwaukee Journal Sentinel* (July).

Marx, Matt. 2022. "Employee non-compete agreements, gender, and entrepreneurship." *Organization Science* 33(5): 1756–1772.

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

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.

Slattery, Cailin. 2018. "Bidding for firms: Subsidy competition in the U.S." *SSRN* 3250256.

Slattery, Cailin, and Owen Zidar. 2020. "Evaluating state and local business incentives." *Journal of Economic Perspectives* 34(2): 90–118.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of econometrics* 225(2): 175–199.

Tate, Geoffrey, and Liu Yang. 2013. "The Bright Side Of Corporate Diversification: Evidence From Internal Labor Markets." *U.S. Census Bureau Center for Economic Studies* Working Papers 13-24.

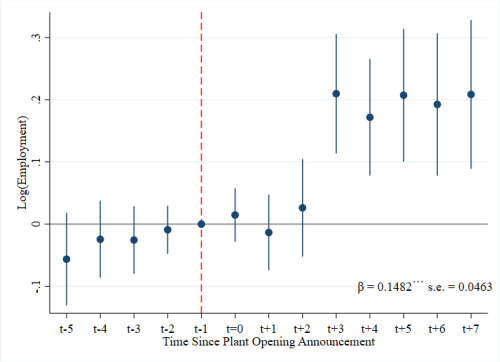
Vilhuber, Lars, and Kevin McKinney. 2014. "LEHD Data Documentation Lehd-Overview-S2011: LEHD Infrastructure Files in the Census RDC Äì Overview." *U.S. Census Bureau Center for Economic Studies* CES-WP-14-26.

Williams, Mark. 2022. "Biden Signs CHIPS Act, Paving Way for Intel Buildout of New Albany Site." *The Columbus Dispatch* (August).

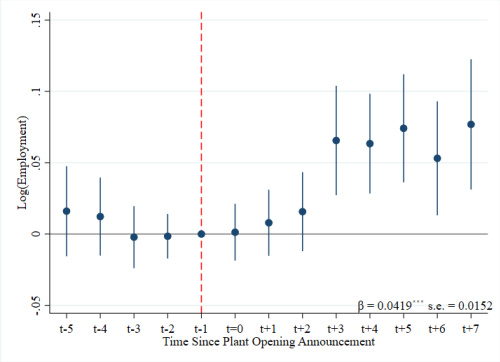
A Appendix

Figure A.1: Weighting by Initial Cluster Size

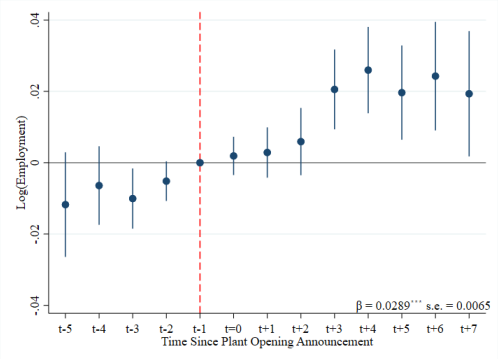
(a) MDP Industry



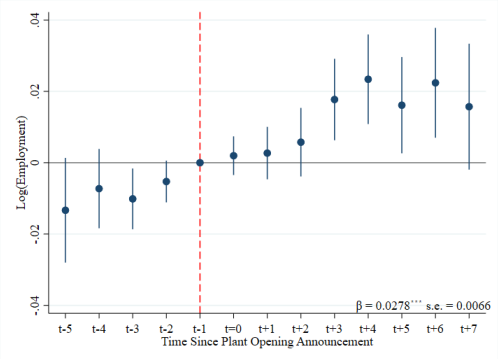
(b) Supply Chain



(c) All Industries



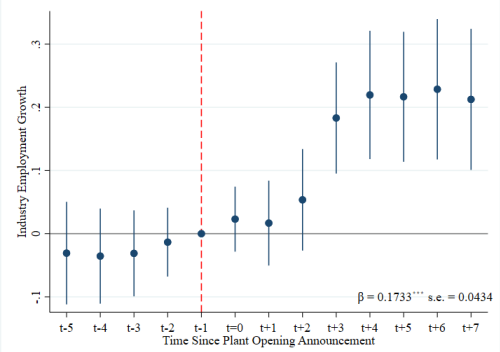
(d) Unrelated Industries Only



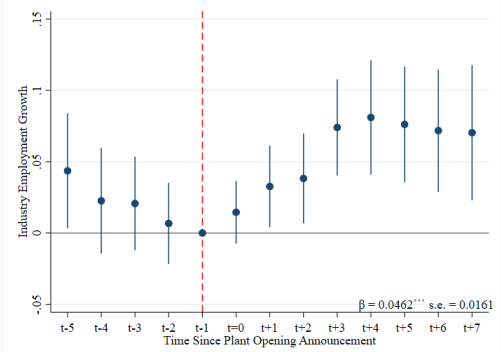
Notes: See Figure 6. Regressions are weighted by the log of total county-industry employment in year $t-1$.

Figure A.2: Alternative Cluster Growth Calculation

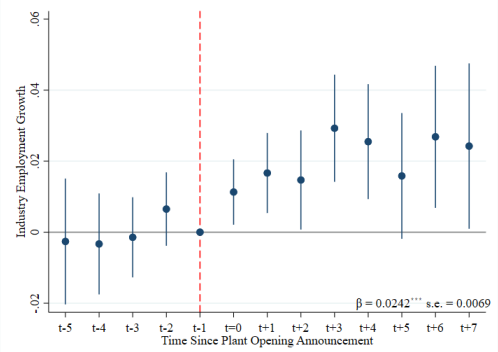
(a) MDP Industry



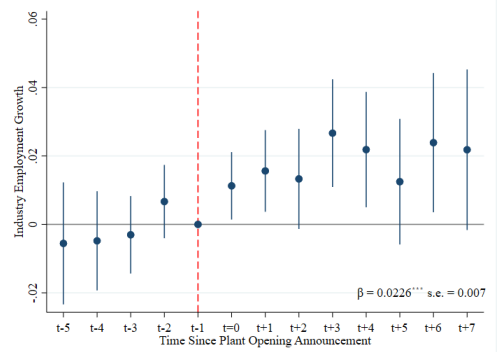
(b) Supply Chain



(c) All Industries



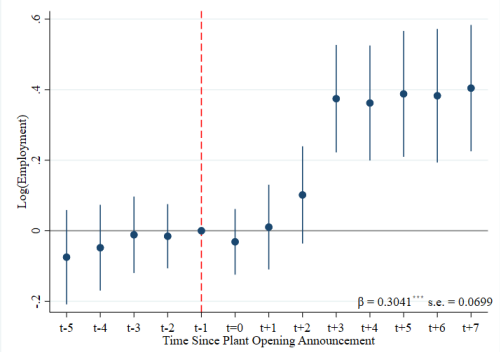
(d) Unrelated Industries Only



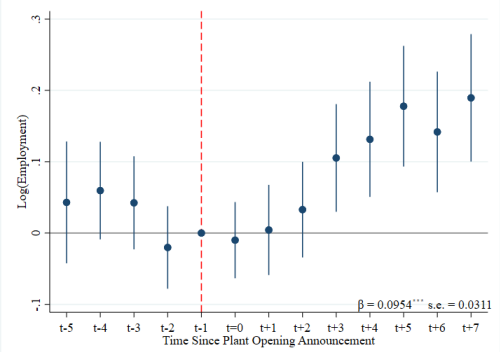
Notes: See Figure 6. County-industry growth is calculated as the Davis, Haltiwanger, and Schuh (1998) symmetrical growth measure.

Figure A.3: Excluding Already Treated Counties

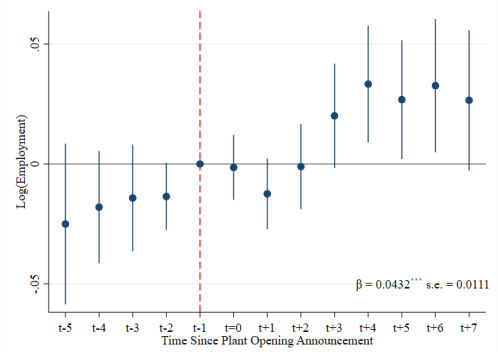
(a) MDP Industry



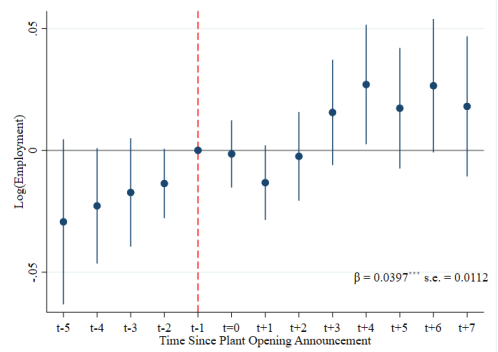
(b) Supply Chain



(c) All Industries

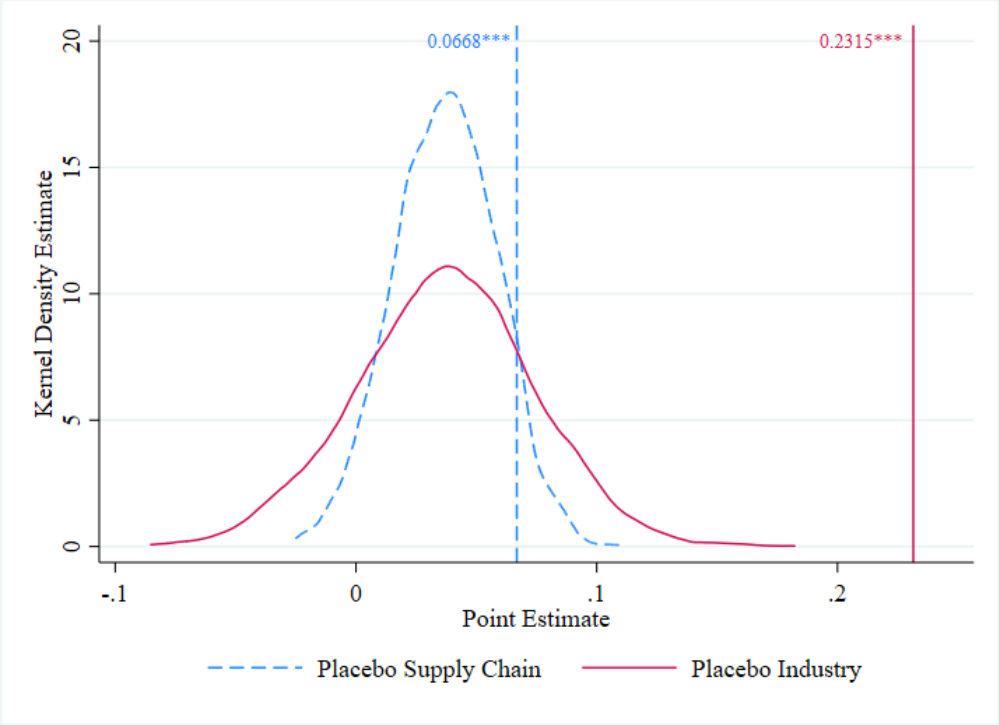


(d) Unrelated Industries Only



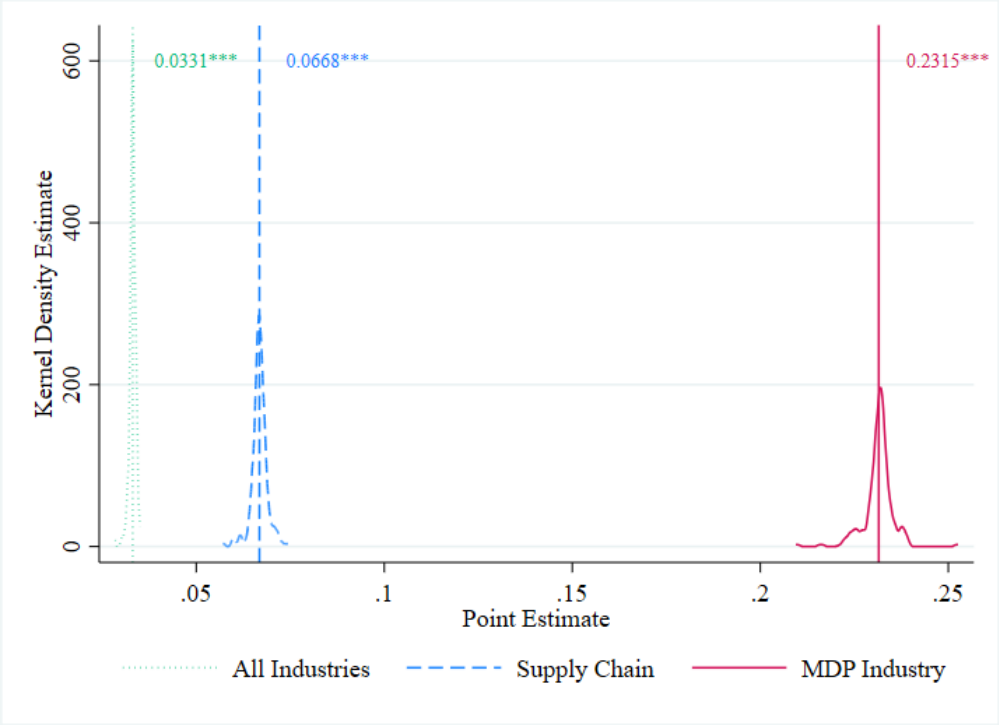
Notes: See Figure 6. Regressions exclude counties that have ever been treated before regardless of industry.

Figure A.4: Placebo Treatment



Notes: The figure plots the kernel density of estimating Equation 5 when randomly assigning a treated industry to the winning county. This figure shows that the treatment effect in the MDP industry in the actual data is in the far right tail of the placebo test distribution. The supply chain effects in simulated data is closer to the estimates because the MDP leads to some spillovers in unrelated industries. Moreover, industries that are large intermediate goods producers to the MDP are likely to in the supply chain of the placebo industry.

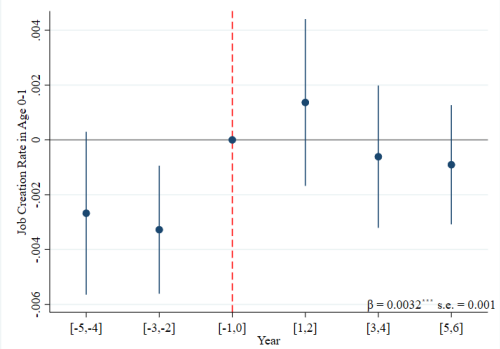
Figure A.5: Individual Case Sensitivity



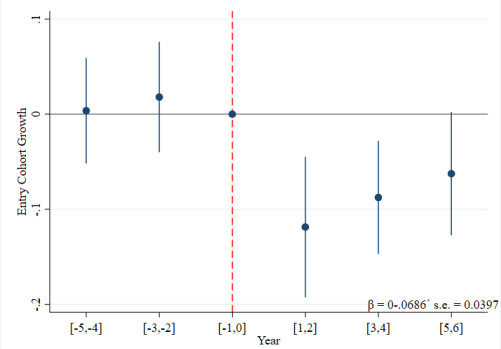
Notes: This figure shows kernel density estimates of Equation 5 when dropping each case iteratively. The tightness of the density functions about the actual coefficient estimate suggests that estimates are not sensitive to individual large cases.

Figure A.6: Industry Heterogeneity

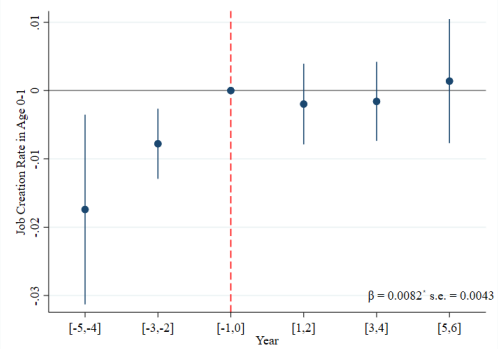
(a) Age 0-1 Job Creation (Mfg MDPs)



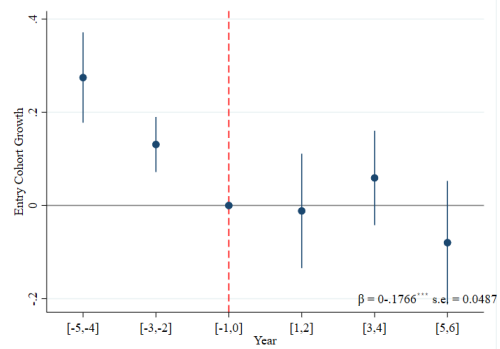
(b) Cohort Growth (Mfg MDPs)



(c) Age 0-1 Job Creation (R&D MDPs)



(d) Cohort Growth (R&D MDPs)



Notes: See Figure 7.

Table A.1: Upstream Entry and Firm 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 (1998) 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 facility's 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. Standard errors are clustered by county. All regressions are weighted by total county-industry employment. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8165).

Table A.2: Anchor Industry Entry and Firm 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 (1998) 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. Standard errors are clustered by county. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8165).

Table A.3: Downstream Entry and Firm 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 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 (1998) 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 facility's 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. Standard errors are clustered by county. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8165).

Table A.4: Relative Wages and Likelihood of Entrepreneurship

	(1)	(2)	(3)
	Pr(Found upstream firm)	Pr(Found anchor industry firm)	Pr(Found downstream firm)
Win * Upstream * Relative Wage	-0.0011	0.0057	0.0063
	(0.0110)	(0.0049)	(0.0047)
Win * Anchor Industry * Relative Wage	0.0405***	0.0464***	0.0390***
	(0.0134)	(0.0137)	(0.0126)
Win * Downstream * Relative Wage	-0.0056	-0.0048	-0.0045
	(0.0069)	(0.0065)	(0.0101)
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.	0.1060	0.0602	0.1493
N	2124000	2124000	2124000

Notes: * 0.1 ** 0.05 *** 0.01. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8124).

Table A.5: Change in Winning County Founders by Wage

	(1)	(2)	(3)
	Pr(Found upstream firm)	Pr(Found anchor industry firm)	Pr(Found downstream firm)
Post * Wage	0.0002	0.0003**	0.0008***
	(0.0079)	(0.0002)	(0.0002)
Post * Upstream * Wage	0.0644***	0.0047	0.0024
	(0.0111)	(0.0047)	(0.0052)
Post * Anchor Industry * Wage	0.0051	0.0220	0.0071
	(0.0103)	(0.0199)	(0.0123)
Post * Downstream * Wage	-0.0229***	-0.0124	-0.0225
	(0.0084)	(0.0104)	(0.0144)
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.	0.3136	0.2266	0.4067
N	4007000	4007000	4007000

Notes: * 0.1 ** 0.05 *** 0.01. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8124).

Table A.6: 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 7 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 7, 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. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8165).

Table A.7: 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-county. 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-county employment. Standard errors are clustered at the SEIN-county. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8165).

Table A.8: 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 9 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 9, 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. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8165).

Table A.9: 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 9 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 9, 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. Disclosure conducted under FSRDC Project 1841. (CBDRB-FY19-P1841-R8165).

Table A.10: 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 occupational 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.

Table A.11: 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 A.10. The table displays the Top 10 correlations between the anchor industry and any one of its supply chain pairs.