The Geography of Innovative Firms

Craig A. Chikis
University of Chicago

Benny Kleinman
Stanford University

Marta Prato
Bocconi University*

October 3, 2024

[Preliminary. Please do not circulate.]

Abstract

A large body of literature has documented that knowledge spillovers between firms are local and decay with distance. In a world with multi-plant firms, such local spillovers imply that a firm's spatial expansion can impact both its own exposure to existing ideas and the spillovers it generates for others, shaping aggregate innovation and growth. We investigate this idea empirically and theoretically. We first provide evidence on the importance of geographical firm expansion for cross-firm spillovers by linking data on firms' geography and patent citations. Using an event-study design, we show that a firm's expansion into a local market increases citations to its existing patents by inventors in that market. Motivated by this evidence, we develop a model in which multi-plant firms can improve their innovation process by accessing ideas across multiple markets. The distribution of these ideas across space depends on the location decisions of all firms in the economy. A key takeaway is that when knowledge spillovers are local and ideas are sufficiently differentiated across firms, innovative firms are too geographically concentrated relative to the social optimum. Thus, policies promoting spatial expansion, like subsidies for opening new R&D branches, can enhance growth and knowledge diffusion, while reducing spatial inequality.

*This research benefited from the financial support from the Becker Friedman Institute for Economics at University of Chicago.

1. Introduction

The existence of knowledge spillovers between firms is a key idea in the literature on innovation and economic growth. Empirical measures of these spillovers show that they are highly localized and depend on geographical proximity between firms. In a world with multi-plant firms, such local spillovers imply that a firm's spatial scope can impact both its own exposure to existing ideas and the spillovers it generates for others. Consequently, a firm's geographical structure might be sub-optimal from a social perspective, and changes in it can affect the prevalence of knowledge spillovers, aggregate innovation, and growth.

In this paper, we investigate this idea, contributing new evidence and theory on the role of firm geography in shaping knowledge spillovers and growth. First, we provide new evidence on the importance of geographical firm expansion for spillovers across firms, linking data on firms' geography and the near universe of patent citations. Second, we develop a model that formalizes the role of local knowledge spillovers for firm expansion decisions and innovation. Third, we study social welfare in the model and highlight conditions under which innovative firms are too geographically concentrated in equilibrium relative to the optimum. Finally, we characterize which local markets gain and lose from policies encouraging innovative firms to increase their geographical scope.

Empirically, we provide evidence on the role of spatial expansion for knowledge spillovers across firms. We link data on changes in firms' geography from Dun & Bradstreet with cross-firm patent citations from USPTO's PatentsView, successfully covering the near-universe of firm locations and citations in the past five decades in the United States. Employing an event-study design, we establish that a firm's expansion into a local market generates a persistent increase in the number of external citations that it receives from inventors in that market. These effects are robust to excluding any citations by firms sharing inventors with the expanding firm, in the past or in the future; and to excluding any citations to any patents that the firm generates in the new market. We further show that these effects are strongest when the expanding firm opens a plant that specializes in innovative activity in the destination market. We also find evidence that knowledge flows in the opposite direction: expanding firms increase citations to inventors located in the new market.

Based on these findings, we develop a theory that links firms' geographical scope to innovation activity in a world with local knowledge spillovers. We extend the canonical "creative destruction" model from the literature on economic growth to a multi-market setting, adding two key components. First, we introduce multi-location innovative firms that employ R&D workers across local markets to increase their arrival rate of new ideas. Firms choose to operate in multiple markets due to decreasing returns to their R&D workforce in each location, and firms avoid operating in all markets since they incur a fixed cost for every new location. Second, the productivity of R&D workers in each location depends on a local stock of knowledge, which increases in the amount of local innovation activity, giving rise to agglomeration spillovers. Crucially, we assume that there are diminishing returns to knowledge spillovers from each local firm and that no local firm internalizes its contribution to this stock. Thus, local innovation efficiency depends not only on a given firm's local innovation workforce but also on the number and composition of firms, capturing the notion that there are gains from a diversity of ideas as in Jacobs (1969).

We establish a connection between economic growth and the geographical scope of innovating firms in equilibrium, and characterize social welfare in this setting. We show that the optimal allocation of firm activity across space differs from the competitive equilibrium when there are local knowledge spillovers. The sign and magnitude of the deviation depends on two additional parameters: the elasticity of new ideas to firms' local labor in the production function for ideas, and the diminishing returns to knowledge spillovers from each firm in our specification for agglomeration forces. When the latter is low enough, innovative firms are too geographically concentrated in the decentralized equilibrium.

To gain some intuition for this result, consider the spillovers that are generated when a firm opens a new innovation facility in a new market. On the one hand, it generates new knowledge spillovers in that market. On the other hand, due to labor market clearing, it must induce a reduction in the amount of labor that is allocated to some other innovation plant in some other region of the economy, since the expansion incurs a fixed cost that is paid in labor. When the diminishing returns to each plant in the agglomeration function are strong enough, the knowledge gained from opening a new plant are large relative to the knowledge lost from shrinking an existing plant. In this case, the social planner would

like to induce more expansion relative to the competitive equilibrium. In the knife-edge case that the decreasing returns in the ideas production function are the same as those in the ideas spillovers function, the decentralized equilibrium is efficient, even if knowledge spillovers are strong.

To understand whether innovating firms are too geographically concentrated in practice, and to understand the welfare implications of policies that induce them to expand, we estimate our model using data on U.S. local labor markets in 2019. We show how to identify the decreasing returns in the ideas production function and in the knowledge spillover function using data on the co-movement of patenting and patent citations within firms across space. Our estimation results are consistent with the case in which innovative firms are too geographically concentrated in equilibrium.

Finally, we investigate the implications of a simple policy that induces firms to expand more, by giving a fixed subsidy for every new market in which firms open an innovation center. We assume that this subsidy is financed by a lump sum on all agents in the economy. We find that the welfare gains are maximized when the size of such a program is around 2% of GDP. Under such a policy, growth increases by 0.08 p.p., and the present discounted value of aggregate consumption is higher by 2% relative to the benchmark equilibrium. We also find that smaller markets gain relatively more from such a policy. Our estimation suggests that a modest incentive for innovating firms to spread their activity across space can contribute to higher growth and lower spatial disparities.

Related literature. We relate to several strands of literature. First, we relate to the empirical literature that investigates the local nature of spillovers between firms, following Jaffe et al. (1993), Audretsch and Feldman (1996), Peri, 2005, Greenstone et al. (2010), Griffith et al. (2011), Carlino and Kerr, 2015, Bloom et al. (2019), Matray, 2021, and Giroud et al. (2021). Our design has two advantages relative to these papers. First, by studying changes in patent citations following firm expansion, we provide direct, well-identified evidence for knowledge spillovers that result from changes in firms' geographical scope. Second, our large-scale data and matching design allows us to investigate heterogeneity in the effect of different plant expansions. In particular, as our theory shows, understanding how knowledge spillovers relate to the size of the expanding plant is crucial to determining whether there is too much or too little geographical firm expansion in equilibrium. Finally, it is worth noting that throughout the paper, we focus on spillovers

in innovation, and not on contemporary productivity spillovers.

Second, we relate to papers that model the expansion of firms through space, including Jia (2008) and Holmes (2011) in the industrial organization literature; Argente et al. (2020), Oberfield et al. (2024), and Kleinman (2023) in the macroeconomics literature; and Helpman (1984), Ramondo and Rodríguez-Clare (2013), Tintelnot (2017), and Arkolakis et al. (2018) in the literature on trade and multinational firms. We focus on the geographical structure of innovative firms, introducing a new expansion motive based on access to ideas across multiple markets, and linking it to aggregate growth. We also show that the degree of spatial expansion in equilibrium can be sub-optimal from a social perspective.

Third, we relate to the theoretical literature on innovation, diffusion, and endogenous growth. Following models of Schumpeterian innovation, firms invest in innovation in order to improve product quality, giving rise to a quality ladder (Grossman and Helpman, 1991; Aghion and Howitt, 1992; Klette and Kortum, 2004; Akcigit and Kerr, 2018). As in models of knowledge diffusions, firms learn from others and gain insights which make their innovation process more effective (Lucas and Moll, 2014; Perla and Tonetti, 2014; Buera and Oberfield, 2020; Prato, 2022). Our contribution to this literature is to analyze the interaction between firms' innovation and spatial expansion decisions. We highlight that in the presence of local knowledge spillovers, both innovation and expansion decisions can be sub-optimal in equilibrium. There are potential gains from re-configuring firms' geography, even conditional on the amount of resources that are channeled into the innovation sector.

Fourth, we relate to recent papers that study the role of local spillovers between workers for aggregate productivity and growth, e.g. Martellini (2022), Crews (2023), and Lhuillier (2023). Relative to these papers, we focus on knowledge spillovers between firms. We highlight a new connection between the geographical scope of firms and aggregate growth and derive new policy implications for the optimal geography of firms.

Finally, we relate to the literature on the link between trade and innovation, as surveyed in Melitz and Redding (2021) and Akcigit and Melitz (2022). Rather than studying the importance of trade frictions between countries for innovation and growth, we focus on spatial frictions to the expansion of innovative firms within countries. However, the key insights from our analysis apply also to the expansion of multinationals that engage

in innovation across countries.

2. Data

We use two main data sources for the analysis. First, PatentsView provides information on the universe of U.S. patents granted after January 1976. Second, Dun & Bradstreet (D&B) provides detailed information on the spatial organization of firms between 1969 and 2021.[1] The key advantage of the Dun & Bradstreet data is its public availability and its clear delineation of firm ownership across space (i.e. headquarters-branch information); the spatial distribution of firms in Dun & Bradstreet is characterized by high accuracy when benchmarked against official, administrative datasets.[2] We merge these two datasets following the methodology of Hughes et al. (2021).[3] Table 1 shows the results of this merge. The quality is high: Over the period 1969 to 2021, we match 84% of all nonwithdrawn utility patents granted to at least a single U.S.-based assignee. Encouragingly, a large majority of these patents are matched to firms at the headquarters or ultimate owner levels, consistent with the notion that large, publicly-listed corporations account for a majority of private sector R&D activity (Bloom et al., 2013).[4] The mean and median grant and application years in the matched sample are comparable to those in the population of PatentsView data, suggesting broad balance in the match along the time dimension. Figure A.1 provides further details on the time series balance in the match as well as the performance of the match spatially. Both aspects are broadly balanced and suggest our matched sample is representative of the population along these dimensions. Lastly, the three digit CPC distribution is very similar between the matched sample and the population, reflecting technological balance in the matched sample (Figure A.2). Further information on the data and the quality of the merge between PatentsView and Dun &

[1] In our analysis, we focus on the period from 1980 to 2018. We limit to these years to maximize sample size subject to truncation considerations. Although our patent data covers patents granted after 1976, those patents in the early part of the sample were often *applied* for prior to 1976—in some cases, as far back as 1950. Since there are patents applied for prior to 1950 that were granted before 1976, our sample is not truly universal in the earliest years. Similarly, there may be patents applied for after 2018 that are not yet granted in our data. We also wish to exclude the Covid-19 pandemic from our analysis, a period in which firm location decisions differed from the historical norm.

[2] See Barnatchez et al. (2017) for further description of the representativeness and accuracy of the National Establishment Time Series (NETS) data, which is constructed from the Dun & Bradstreet data.

[3] Details of the merge can be found in Appendix A.1.

[4] In particular, Bloom et al. (2013) report Compustat firms in their sample account for 62% of the total private sector R&D spending in the economy.

Match quality				Firm resolution		
Statistic	Matched	PatentsView	Match rate	Organization	N	Cum. share
Number of assignees	150730	205655	73%	Headquarters	2375886	0.86
Number of commuting zones	679	707	96%	Single establishment	339269	0.98
Number of patents	2759380	3269141	84%	Ultimate owner	44225	1.00
Grant year						
Mean	2005	2006				
Median	2009	2009				
Application year						
Mean	2003	2003				
Median	2005	2005				

Table 1: **Merge quality between PatentsView and Dun & Bradstreet.** This table conditions on nonwithdrawn utility patents granted between 1969 and 2021 with a single U.S.-based assignee.

Overall		Firm			Firm × commuting zone		Expansion types		
			Mean	Median	Mean	Median	SIC	N	Cum. share
Number of firms	172075	Markets	4.15	1.00	—	—	50	104472	0.12
Number of commuting zones	727	Establishments	11.04	1.00	2.66	1.00	73	77508	0.21
Number of firm × commuting zones	879239	Patents	2.61	0.00	0.63	0.00	87	42223	0.26
		Inventors	2.22	0.00	0.53	0.00	99	41521	0.30
		Log D&B emp.	3.48	3.26	3.23	3.09	59	39975	0.35
		External citations	1.11	0.00	0.27	0.00	51	36961	0.39
							35	36579	0.43
							36	28087	0.46
							28	24733	0.49
							56	20083	0.52

Table 2: **Summary statistics.** All firm-level statistics condition on firm activity (measured by having at least a single active establishment). All firm × commuting zone-level statistics condition on having at least one establishment active in the commuting zone. Both pool across years. Numbers of patents and citations are flow measures; number of citations is based on the same refinement of external citations used in Section 3.1. Expansion type is measured as the modal two-digit SIC of establishments in focal commuting zone in year of expansion.

Bradstreet can be found in Appendix A.

Table 2 presents summary statistics of the merged data. The number of markets and establishments per firm is slightly higher than that reported in Hsieh and Rossi-Hansberg (2023) and Kleinman (2023), but this deviation is reasonable given our focus on patenting firms, which tend to be larger than the services firms studied in the aforementioned papers.[5]

Our analysis focuses on spillovers that emerge from firm expansion at the commuting zone level. We obtain county-level information on economic activity and aggregate it to the commuting zone level from various publicly available, tabulated micro-datasets, including the Business Dynamics Statistics and County Business Patterns data of the U.S. Census Bureau.

The merged dataset allows us to conduct a comprehensive analysis of the relationship between firm expansion and knowledge spillovers. In Section 3, we employ an event-

[5]As noted before, we utilize Dun & Bradstreet whereas Hsieh and Rossi-Hansberg (2023) and Kleinman (2023) utilize the Longitudinal Business Database (LBD), another reason for slight deviation in firm size and scope moments.

study design to investigate how a firm's expansion into a new local market affects the citations it receives from inventors in that market. We exploit the detailed geographical information in the Dun & Bradstreet data to identify firm expansions, while using the PatentsView data to track changes in citation patterns. This approach enables us to examine both the magnitude and persistence of spillover effects following firm expansion. Additionally, we explore heterogeneity in these effects based on the type of establishment opened and investigate potential reverse spillovers from local inventors to expanding firms.

3. Empirical Evidence

In this section, we empirically investigate the relationship between firm expansion and knowledge spillovers. We begin by detailing our identification strategy and econometric specification. We then present our main results on the impact of firm expansion on local citation patterns. Subsequently, we explore heterogeneity in these effects across different types of establishments and examine potential reverse spillovers. Finally, we conduct a series of robustness checks to validate our findings.

3.1 Measuring spillovers

To estimate the presence, strength, and persistence of spillovers resulting from firm expansion, we employ a matched difference-in-differences (MDiD) specification. Our unit of observation, i, is a firm $\times$ commuting zone pair (interchangeably referred to at times as an establishment). We denote by $g(i)$ the first year unit i records activity in commuting zone $CZ(i)$. A unit is considered treated when $t \geq g(i)$, with $g(i) = \infty$ for never-treated units.[6]

As in any DiD framework with staggered treatment adoption, identification relies on a particular form of the parallel trends assumption (Callaway and Sant'Anna, 2021). This argues for the careful construction of a matched control sample. For example, comparing Google in 1999, when it was founded, to a never treated unit of IBM's is unlikely to satisfy the parallel trends assumption, since these firms—and their attendant markets—are very

[6]We assume treatment is absorbing; thus, if a firm opens an establishment in a given commuting zone but then exits a commuting zone, for the purposes of analysis, the unit remains treated. In Appendix B.1 we show our main results are robust to excluding these temporarily treated commuting zones from our analysis.

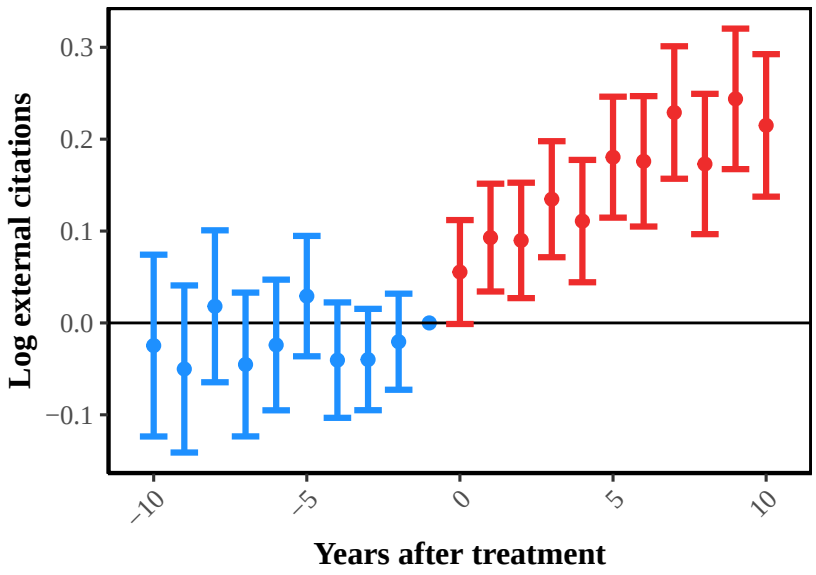


Figure 1: **Response of citations received by firm patents in expansion commuting zone preceding and following expansion.** The outcome variable is the log number of citations made to stock of patents of firm $Firm(i)$ at time t from inventors residing in commuting zone $CZ(i)$. We remove all internal citations and all external citations made from patent documents from firms $f \neq Firm(i)$ that ever shared inventors with firm $Firm(i)$; we further only consider citations made to patents originated outside of $CZ(i)$. Each point represents a loading on the relative treatment time dummies specified in (1). Standard errors are clustered at the firm $\times$ commuting zone level. Error bars represent 5% confidence intervals. The Wald test of joint nullity of the pre-treatment dummies, $\{\beta_\tau\}_{\tau=-10}^{-2}$, returns a p-value of 0.30.

likely to be on different trajectories. Therefore, we match treated commuting zones to never-treated commuting zones, *within firm*, by minimizing a distance criterion defined on pre-treatment outcomes and pre-treatment covariates.[7] To ensure that both treated units and never treated units are relevant locations for studying the innovative activity of the focal firm, we further require that both treated and control units receive positive citations in two out of ten years of the pre-treatment period $\tau \equiv t - g(i) \in \{-10, \ldots, -1\}$.

Our main specification is:

$$y_{it} = \alpha_i + \delta_t + \kappa_{Firm(i)t} + \sum_{\tau \in \{-10,\ldots,10\} \setminus \{-1\}} \beta_\tau \mathbb{I}\{t - g(i) = \tau\} + \varepsilon_{it}. \tag{1}$$

where $Firm(i)$ is the firm identifier of unit i. The outcome variable, y_{it}, is the log number of citations received by firm $Firm(i)$'s patents in commuting zone $CZ(i)$ at time t. To ensure that we are capturing spillovers due to the spatial proximity of distinct firms operating in a given commuting zone — and not, for example, within-firm spillovers or spillovers mediated by shared employees — we restrict our analysis to "external citations".

[7] See Appendix C for details of the matching procedure. For a related approach, see Arkhangelsky et al. (2021).

These are citations received from other firms to patents of firm $Firm(i)$. We also consider only citations to $Firm(i)$'s patents that were authored by inventors reporting a location other than commuting zone $CZ(i)$, and made by patent documents whose originating firms never share inventors with $Firm(i)$.[8,9]

Our main results are presented in Figure 1. The results show that firm expansion into a market results in a significant and persistent increase in the number of citations received by the expanding firm from inventors in the new market. Ten years following expansion, citations are, on average, roughly 20% higher in treated commuting zones compared to control commuting zones. Encouragingly, there is no discernible pre-trend, and the Wald test of joint nullity of the pre-treatment dummies, $\{\beta_\tau\}_{\tau=-10}^{-2}$, returns a p-value of 0.30, meaning we fail to reject the joint nullity of the pre-treatment dummies under conventional levels of statistical significance, providing further evidence in support of the parallel trends assumption.

Figure 2 shows the results when studying alternative specifications of the dependent variable. Moving from left to right, the specification becomes increasingly stringent. First, we use the log of all external citations received by the expanding firm from inventors in the focal commuting zone, $CZ(i)$; then we throw out any citations originating from patent documents associated with firms that have *ever* shared inventors with the expanding firm; and finally we throw out any citations made to patents invented in the expansion commuting zone, $CZ(i)$ (our baseline specification in Figure 1). Clearly, the three specifications return very similar results; Wald tests lead us to fail to reject joint nullity in all three specifications.

Lastly, in order to further strengthen our identification strategy and isolate the narrow effects of expansion, we extend our baseline specification to include commuting zone by year fixed effects. We estimate the following regression:

$$y_{it} = \alpha_i + \delta_t + \kappa_{Firm(i)t} + \zeta_{CZ(i)t} + \sum_{\tau \in \{-10,\ldots,10\} \setminus \{-1\}} \beta_\tau \mathbb{I}\{t - g(i) = \tau\} + \varepsilon_{it}. \tag{2}$$

[8]These latter conditions are to ensure that we are not picking up on spillovers that occur within the firm due to, for example, anticipated or mismeasured merger activity. Similarly, these restrictions rule out a mechanical increase in citations received due to an increase in patenting upon the opening of a new R&D facility in location $CZ(i)$. Purging the firm-year trend, $\kappa_{Firm(i)t}$ further rules out any mechanical increase in citations whereby, firm-wide, there is a rise in citable patents that coincides with firm expansion in the focal market.

[9]Appendix A.2 discusses how we allocate inventors to firms.

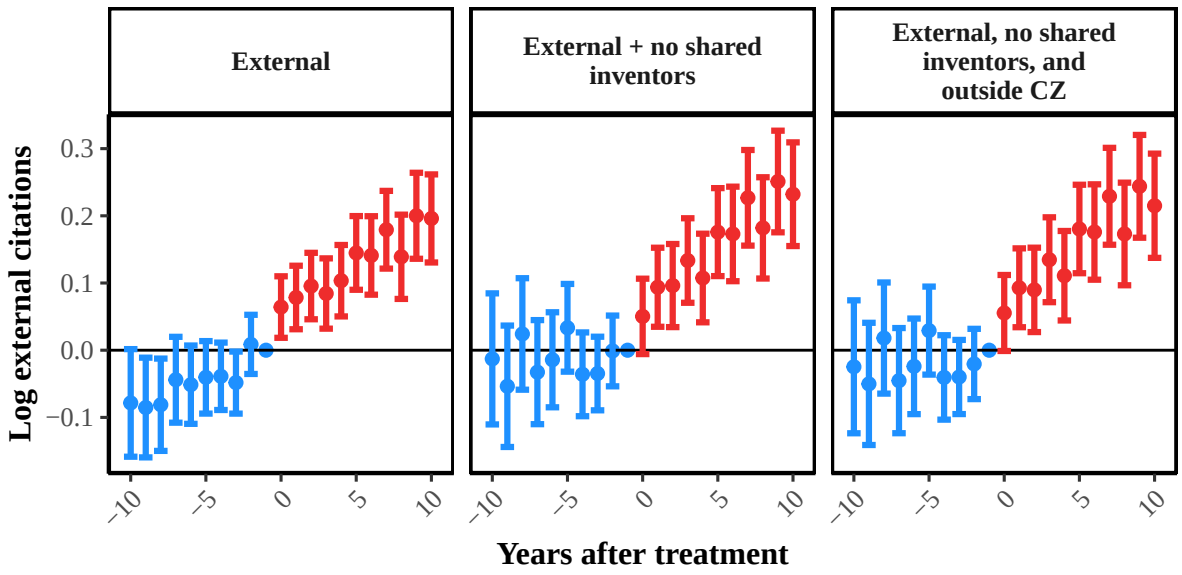


Figure 2: **Event study under alternative specifications of the dependent variable.** The outcome variable is the log number of citations made to stock of patents of firm $Firm(i)$ at time t from inventors residing in commuting zone $CZ(i)$. In all specifications, we remove all internal citations. Each point represents a loading on the relative treatment time dummies specified in (1). Standard errors are clustered at the firm $\times$ commuting zone level. Error bars represent 5% confidence intervals. The Wald test of joint nullity of the pre-treatment dummies, $\{\beta_\tau\}_{\tau=-10}^{-2}$, returns p-values of 0.18, 0.30, and 0.30, respectively.

The inclusion of these granular fixed effects allows us to control for any time-varying, location-specific factors. Such factors could include local economic booms, policy changes, or the emergence of industry clusters that might attract firms while also fostering increased innovation activity. As is evident from Figure 3, our main finding of increased citations following firm expansion remains robust to this more demanding specification. However, we observe that some of the treatment effect is absorbed by the commuting zone-year trends, resulting in a slightly attenuated effect size. Specifically, while the baseline model shows an increase of about 20% relative to control 10 years following expansion, the model with commuting zone-year fixed effects shows an increase of approximately 10% over the same period.

This attenuation suggests that part of our effect is increased citation activity at the regional level. Of course, this could still be an equilibrium effect of the expansion of the focal firm, as the new innovating plant contributes to fostering the local ecosystem. However, to the extent that it reflects a regional trend that is orthogonal to the expansion event – and that by coincidence begins in parallel to the expansion – this result provides evidence for the robustness of our finding. Importantly, the persistence of a significant treatment effect even after accounting for these local trends provides strong evidence

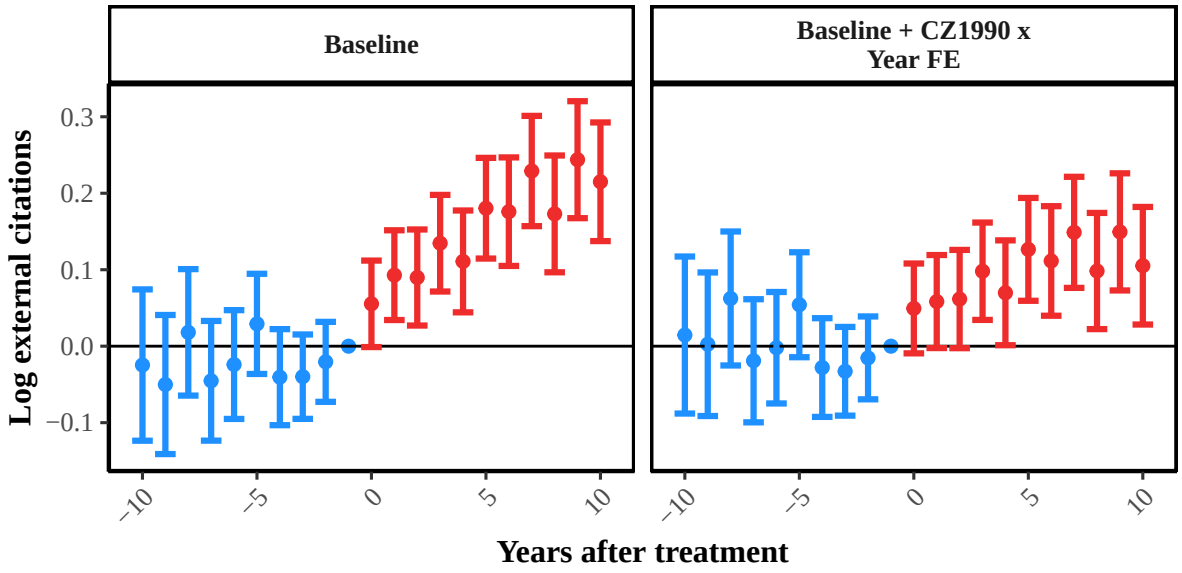


Figure 3: **Event study after purging commuting zone trends.** The outcome variable is the log number of citations made to stock of patents of firm $Firm(i)$ at time t from inventors residing in commuting zone $CZ(i)$. We remove all internal citations and all external citations made from patent documents from firms $f \neq Firm(i)$ that ever shared inventors with firm $Firm(i)$; we further only consider citations made to patents originated outside of $CZ(i)$. The right-most panel also includes a commuting zone-year FE. Each point represents a loading on the relative treatment time dummies specified in equations (1) and (2), respectively. Standard errors are clustered at the firm × commuting zone level. Error bars represent 5% confidence intervals. The Wald test of joint nullity of the pre-treatment dummies, $\{\beta_\tau\}_{\tau=-10}^{-2}$, returns p-values of 0.30 and 0.22, respectively.

that firm expansion itself leads to increased knowledge flows. This suggests that the act of establishing a local presence allows firms to tap into knowledge sources in ways that go beyond simply being present in a thriving innovation ecosystem.

3.2 Heterogeneity in treatment effect by establishment type

Figure 4 explores heterogeneity in the treatment effect based on whether the expanding firm engages in local patenting activity. We find that when firms establish a new presence in a location and also conduct patenting activity there—a proxy for the opening of a R&D facility—local firms much more intensively cite the existing patents of the expanding firm. On average, treated units receive 40% greater citations than control units after 10 years. In contrast, firms that expand but do not engage in local patenting still generate positive spillovers, but the magnitude is smaller at around 10-15% after the same time period.

This heterogeneity suggests that the physical presence of R&D activities amplifies knowledge spillovers to the local economy. The result aligns with the literature emphasizing the importance of geographic proximity for knowledge diffusion (e.g. Jaffe et al.,

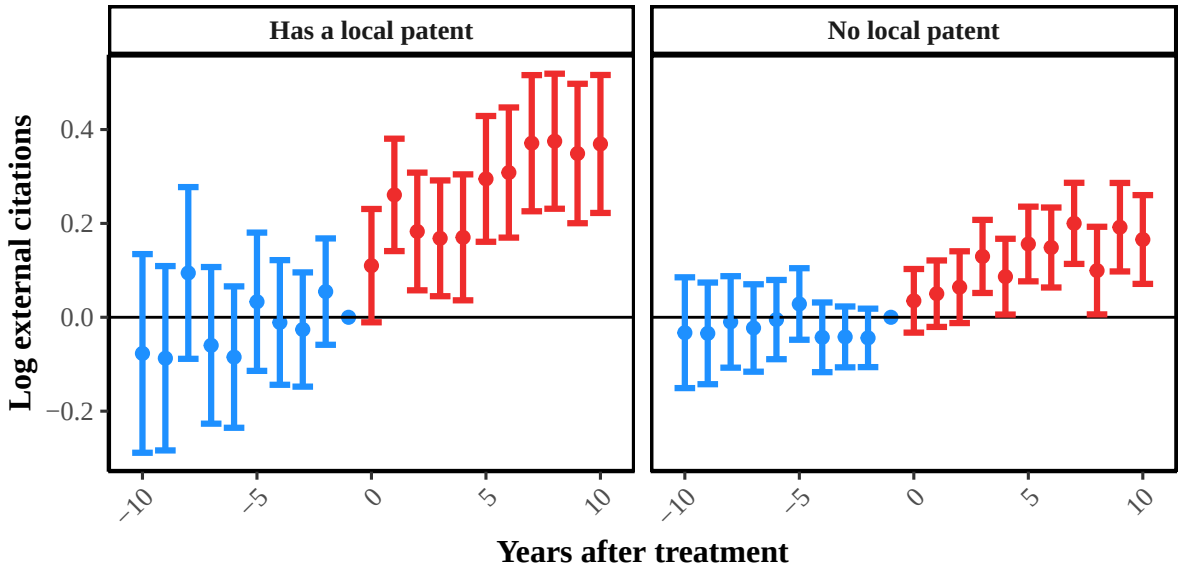


Figure 4: **Heterogeneity by plant type, local patents.** The outcome variable is the log number of citations made to stock of patents of firm $Firm(i)$ at time t from inventors residing in commuting zone $CZ(i)$. We remove all internal citations and all external citations made from patent documents from firms $f \neq Firm(i)$ that ever shared inventors with firm $Firm(i)$; we further only consider citations made to patents originated outside of $CZ(i)$. Each point represents a loading on the relative treatment time dummies specified in (1). Regressions are run separately for firms that patent at least once in $\tau \in \{0, \ldots, 10\}$ in the expansion market versus those that do not. Standard errors are clustered at the firm $\times$ commuting zone level. Error bars represent 5% confidence intervals. The Wald test of joint nullity of the pre-treatment dummies, $\{\beta_\tau\}_{\tau=-10}^{-2}$, returns a p-value of 0.36 and 0.65, respectively.

1993; Audretsch and Feldman, 1996). It extends these findings by demonstrating that not just proximity, but the nature of local activities matters for the intensity of knowledge spillovers.

Figure 5 illustrates the heterogeneity in spillover effects across expansions in the top 10 SIC's listed in Table 2. The results reveal substantial variation in the intensity of knowledge spillovers—as proxied by patent citations—generated by firm expansions across sectors. Notably, firms in knowledge-intensive industries such as Chemicals and Allied Products (SIC 28) and Industrial and Commercial Machinery and Computer Equipment (SIC 35) exhibit some of the largest spillovers. These sectors are characterized by high R&D intensity and complex technological processes, which may facilitate greater knowledge diffusion when firms expand into new locations. However, large treatment effects are also observed in expansions in Miscellaneous Retail (SIC 59); Nonclassifiable Establishments (SIC 99); and Wholesale Trade of Durable and Nondurable goods (50, 51, respectively). Table A.3 shows, however, that many of these expansions seem to occur in hard-to-classify, innovation-adjacent sectors.

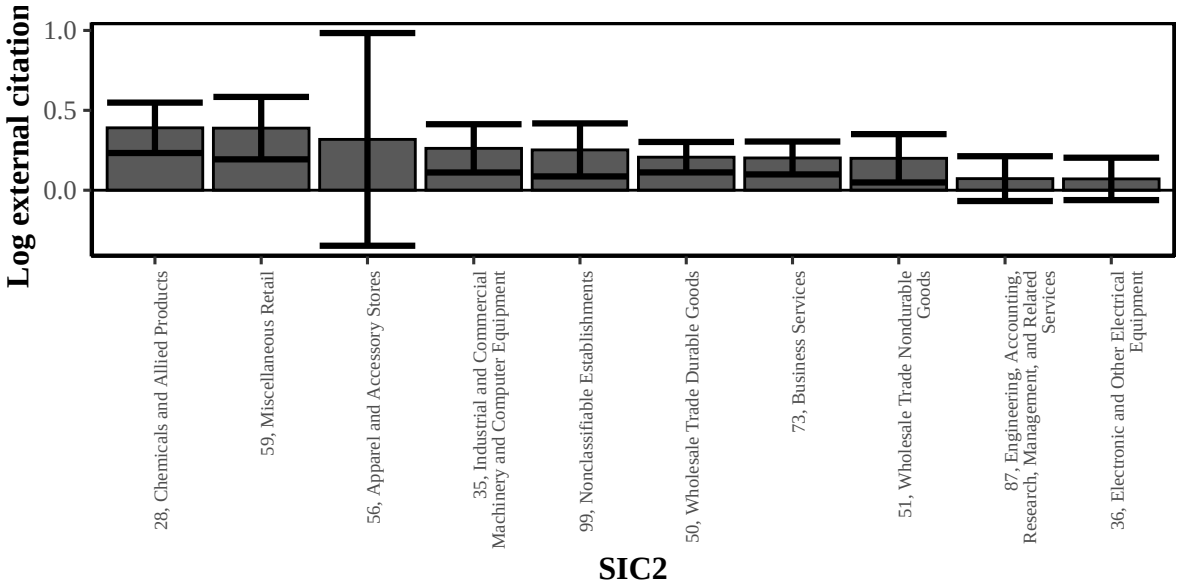


Figure 5: **Heterogeneity by SIC expansion type.** The outcome variable is the log number of citations made to stock of patents of firm $Firm(i)$ at time t from inventors residing in commuting zone $CZ(i)$. We remove all internal citations and all external citations made from patent documents from firms $f \neq Firm(i)$ that ever shared inventors with firm $Firm(i)$; we further only consider citations made to patents originated outside of $CZ(i)$. Each point represents a loading on the relative treatment time dummies specified in (3). Regression specification is static, "stacked", DiD estimator. Regressions are run separately for each SIC expansion type. Standard errors are clustered at the firm $\times$ commuting zone level. Error bars represent 5% confidence intervals.

3.3 Spillovers to the expanding firm

While our previous analysis focused on how firm expansion generates knowledge spillovers to local firms, we now turn to examining a potential driving force behind firms' decisions to expand geographically: access to local knowledge pools. Specifically, we investigate whether expanding firms increase their utilization of knowledge from their new locations following entry.

Figure 6 presents the results of our event study analysis, which examines how a firm's citations to patents originating in focal commuting zone $CZ(i)$ change after the firm establishes a new presence in that location.[10] The observed pattern provides compelling evidence that geographical expansion indeed facilitates access to local knowledge.

Following entry into a new commuting zone, firms exhibit a substantial and persistent increase in citations to patents from that area. For external citations, this effect becomes statistically significant within 3-4 years of expansion and continues to grow over the subsequent decade, reaching an increase of approximately 10% relative to control

[10] We simply replace the dependent variable in equation (1) with citations made by unit i at time t to commuting zone $CZ(i)$.

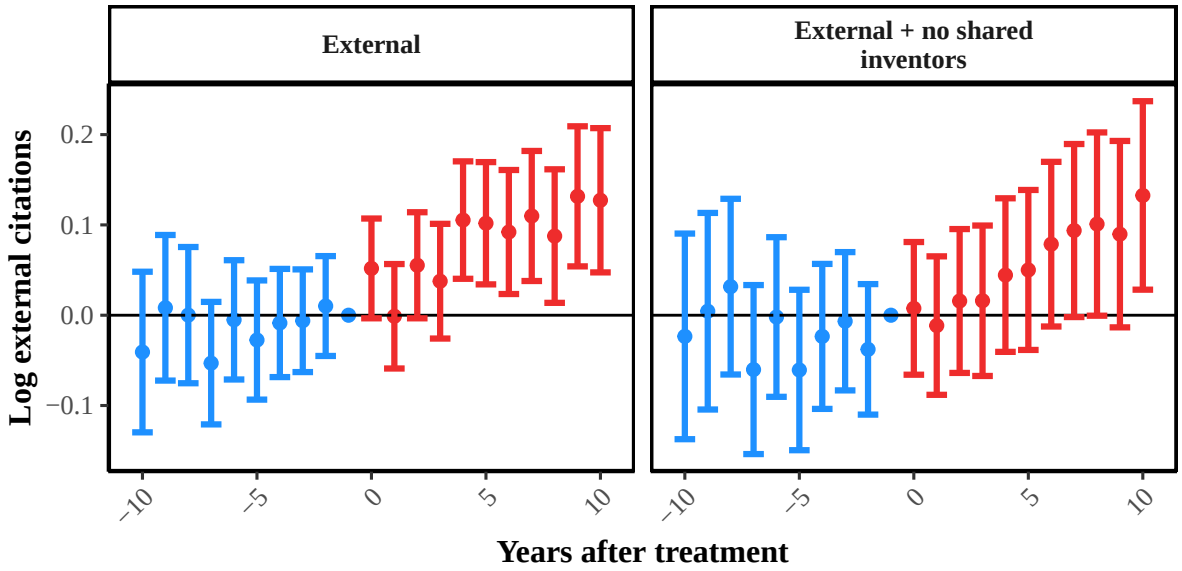


Figure 6: **Response of citation made by firm patents in the expansion commuting zone preceding and following expansion.** The outcome variable is the log number of citations made to stock of patents of firms $f \neq Firm(i)$ at time t to patents invented in $CZ(i)$. The right-most panel further excludes citations made to patents of firms that ever shared inventors with the $Firm(i)$. Each point represents a loading on the relative treatment time dummies specified in (1). Standard errors are clustered at the firm × commuting zone level. Error bars represent 5% confidence intervals. The Wald test of joint nullity of the pre-treatment dummies, $\{\beta_\tau\}_{\tau=-10}^{-2}$, returns p-values of 0.64 and 0.55, respectively.

commuting zones.[11]

These results shed light on an important strategic consideration in firms' location decisions. Beyond factors such as market access, labor costs, or tax incentives, our findings suggest that the potential to acquire new knowledge plays a significant role in driving geographical expansion. Crucially, this motive would be fully internalized in the firm, whereas the previous analysis likely reveals an externality of expansion, of interest to a social planner and not necessarily the expanding firm.

3.4 Alternative estimators

Our baseline specification utilizes the familiar dynamic two-way fixed effects (TWFE) "event-study" estimator. A voluminous recent literature has raised potential issues with this estimator, in particular in settings with staggered treatments and treatment effect heterogeneity, both of which are likely to be present in our setting (Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021; Sun and Abraham, 2021; Borusyak et al.,

[11] For the specification excluding external citations made to firms with ever-shared inventors, statistical power is clearly an issue. However, the magnitude of the estimated effect 10 years following expansion is similar to the looser specification. Regressing the same outcome variable on a dummy indicating $\tau \equiv t - g(i) \geq 0$ (i.e. "static" DiD) returns a regression loading of 0.060 with a standard error of 0.026.

2024). While our matching procedure should resolve many of the underlying problems raised in this literature, we revisit the robustness of our results to using alternative estimators.

First, we consider the "stacking" estimator employed by Cengiz et al. (2019).[12] This specification rules out invalid comparisons by making all comparisons within group-year:

$$y_{it} = \alpha_{ig(Treated(i))} + \delta_{g(Treated(i))t} + \sum_{\tau \in \{-10,\ldots,10\} \setminus \{-1\}} \beta_\tau \mathbb{I}\{t - g(i) = \tau\} + \varepsilon_{it} \tag{3}$$

where if $g(i) < \infty$, $Treated(i) = i$, and, if $g(i) = \infty$, $Treated(i)$ equals the matched treated unit associated with the given never-treated control unit.[13] The downside of this approach is its inability to purge firm or commuting zone level trends as in equations (1) and (2), and that the treatment coefficients, $\{\beta_\tau\}_\tau$, are potentially unintuitive variance-weighted averages of heterogeneous treatment effects.

Second, we employ the estimators proposed by Sun and Abraham (2021) and Dube et al. (2023). These approaches are particularly well-suited for our setting as they allow for heterogeneous treatment effects across cohorts and over time as well ruling out contamination from already-treated units and have nice, intuitive aggregation properties. The downside is that they do not straightforwardly accommodate high-dimensional fixed effect controls at the $Firm(i) \times t$ level.[14]

Lastly, we implement a Poisson regression estimator (Bergé, 2018; Correia et al., 2020). This approach is advantageous for several reasons. Most importantly, it allows us to include observations with zero citations, which are excluded when using the log-transformed outcomes of our baseline results. This inclusion can provide a more complete picture of spillovers, especially for firms or locations with initially low levels of innovation. Additionally, the Poisson model naturally accounts for the count nature of our citation data and can handle over-dispersion, which is common in innovation

[12] See also Cunningham (2021) and Baker et al. (2022).

[13] Appendix D shows how comparisons made within-group year rule out the "negative weight" problem that arises from comparing later treated units to already treated units.

[14] For Dube et al. (2023), we first project out $Firm(i) \times t$-level means using the `reghdfe` function of Correia (2016). For Sun and Abraham (2021), we are able to include $Firm(i) \times t$ fixed effects when estimating the modified version of (1), but identification of treatment parameters, $\{\beta_\tau\}_\tau$, based on parallel trends conditional on covariates is not proven in Sun and Abraham (2021).

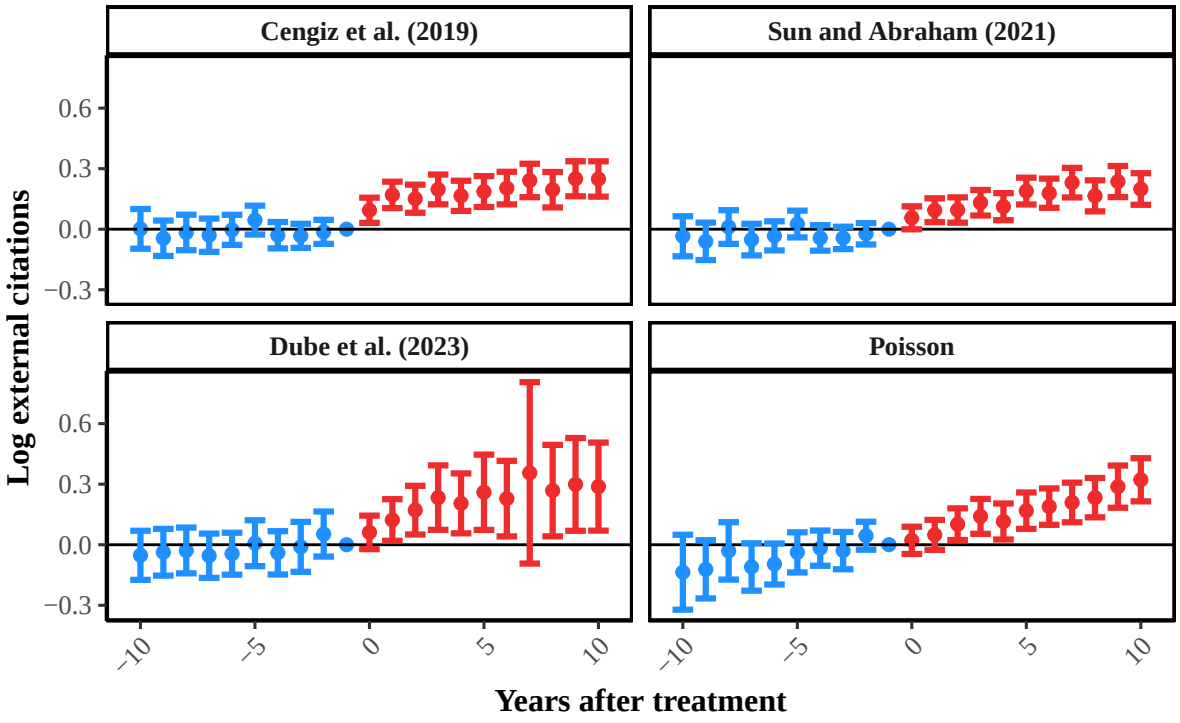


Figure 7: **Alternative estimators for the baseline event study.** The outcome variable is the log number of citations made to stock of patents of firm $Firm(i)$ at time t from inventors residing in commuting zone $CZ(i)$. We remove all internal citations and all external citations made from patent documents from firms $f \neq Firm(i)$ that ever shared inventors with firm $Firm(i)$; we further only consider citations made to patents originated outside of $CZ(i)$—for Poisson regression, the outcome variable is the count of such citations. Each point represents a loading on the relative treatment time dummies specified in (1) or (3), respectively. Standard errors are clustered at the firm × commuting zone level, except for the panel of Cengiz et al. (2019) which clusters at the firm × commuting zone × group level. Error bars represent 5% confidence intervals. The Wald test of joint nullity of the pre-treatment dummies, $\{\beta_\tau\}_{\tau=-10}^{-2}$, returns a p-value of 0.49 for Cengiz et al. (2019) and 0.17 for Poisson regression—Wald tests for the other two estimators are not straightforward to calculate, although visual inspection seemingly rules out meaningful pre-trend.

measures.

Figure 7 presents the results from these alternative estimators. Reassuringly, we find very similar patterns across all approaches. The estimated treatment effects are nearly identical in magnitude and timing, with no changes to our key conclusions. The pre-trends remain flat and close to zero, supporting the parallel trends assumption.

These consistent results across multiple estimators provide strong evidence for the robustness of our findings. They suggest that the potential biases in TWFE highlighted in the recent literature are not a major concern in our setting, likely due to our careful matching procedure. Furthermore, the Poisson results demonstrate that our findings are robust to alternative functional forms and the inclusion of zero-citation observations. Collectively, these alternative specifications increase our confidence in the validity of our empirical strategy and the reliability of our estimates of the spillover effects from firm

expansion.

4. Theory: a model of innovative firms with spatial spillovers

In this section, we introduce a model of innovation-based endogenous growth with multi-location innovative firms that generate local knowledge spillovers. Growth is driven by quality-improving innovations produced by firms that hire R&D workers, with decreasing returns to R&D labor in each location. The productivity of R&D workers depends on endogenous local knowledge spillovers, which depends on the equilibrium composition of innovative firms that open an establishment in a given location. Firms internalize the benefit of local knowledge when opening an additional establishment, but they do not internalize the spillovers they generate for other local firms. As a result, firms' spatial scope can be sub-optimal in the decentralized equilibrium, where innovative firms could be too concentrated or too dispersed.

4.1 Environment

4.1.1 Setting

Time is continuous and indexed by t. The economy consists of N locations, indexed by n.

Each location $n \in \{1, ..., N\}$ is populated by a measure $\bar{L}_n^p$ of households that can supply labor to the production sector and a measure $\bar{L}_n^i$ of households that can supply labor to the innovation sector.[15] All households discount the future at rate ρ and are endowed with one unit of labor. Households consume a freely tradable final good. The preferences of a household in location n, sector $s \in \{p, i\}$ are represented by the utility function

$$U_n^s = \int_0^\infty e^{-\rho t} u\left(Q_{nt}^s\right) \, dt,$$

where Q_{nt}^s is the amount of final-good consumption in period t. Each household owns a balanced portfolio of all the firms in the economy, B_{nt}^s. The household's budget constraint implies that $Q_{nt}^s + \dot{B}_{nt}^s = w_{nt}^s + r_t B_{nt}^s$, where w_{nt}^s is the market wage in sector s, location n, in period t, and r_t is the interest rate at time t.

[15] In the quantitative model in Section 5, we allow for labor mobility.

4.1.2 Production Technology

The final good Y_t is produced using a continuum of intermediate goods y_{jt} according to the production function:

$$\ln Y_t = \int_0^1 \ln(y_{jt})dj.$$

The final good price is normalized to 1.

Each intermediate good is produced by a monopolist that owns the rights to the frontier technology with the following production function:

$$y_{jt} = A_{jt}\left(\sum_{n=1}^{N}(Z_n L_{jnt}^p)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}.$$

where A_{jt} is the quality of intermediate j at time t, Z_n is a time-invariant productivity term specific to workers in location n, and σ is the elasticity of substitution across workers from different locations in the production of the final good. We define aggregate productivity in the economy as $\ln \mathcal{A}_t = \int_0^1 \ln A_{jt}dj$.

In each product line j a single monopolist with the rights to produce with the frontier technology A_{jt} competes against a fringe of competitive firms, that can also produce intermediate j with an obsolete technology with quality A_{jt}/γ, where γ is the step-size of technological progress in the economy. The monopolist charges a price p_{jt}, obtaining revenues equal to $y_{jt}p_{jt}$ and profits π_{jt}.

4.1.3 Innovation and Spillovers

A mass of innovative firms, indexed by ω, hires R&D labor across different markets to produce innovations. The mass of innovative firms is endogenous and denoted by M_t. A successful innovation improves the quality of a random intermediate product j by a step size γ, so that the quality of intermediate j increases to $A_{jt}(1+\gamma)$. A firm ω's employment of inventors across space influences the arrival rate of innovations, $\lambda_{\omega t}$ is given by:

$$\lambda_{\omega t} = \lambda_0 \sum_{n=1}^{N} K_{nt} z_\omega^{1-\eta} \ell_{n\omega t}^{\eta}, \tag{4}$$

where z_ω is the innovation productivity of firm ω; $\ell_{n\omega t}$ is the mass of inventors that firm ω hires in location n at time t; K_{nt} is the relative productivity of inventors, or local knowledge spillover, in location n at time t; η captures decreasing returns from employing more labor in location n; and λ_0 is a scaling constant. The assumption of decreasing returns can capture various forces, such as the challenge of hiring and supervising large research teams in the same location, and is the reason why innovative firms operate multiple plants. We also assume that setting up an innovation plant in location n is subject to a firm-location-specific fixed cost $f_{n\omega t}$ denoted in terms of the local innovation labor. We assume that $f_{n\omega t}$ is drawn from an i.i.d. distribution with cumulative distribution function $F_f(\cdot)$, which is inverse Pareto with shape dispersion θ and scale f_{max}.

The regional productivity of inventors, K_{nt}, is an increasing function of the mass of firms and inventors that are active in location n, given by

$$K_{nt} = \bar{K}_n \left(\sum_{i=1}^{N} \int_{\omega \in \Omega_i} z_\omega^{1-\beta} \ell_{i\omega t}^{\beta} d\omega \right)^{\alpha}, \tag{5}$$

where Ω_i is the set of firms that open a plant in location n; $\bar{K}_n$ is an exogenous determinant of the relative innovation efficiency of location n; α captures the overall importance of insights from other firms for idea creation; β captures the degree to which spillovers from a particular plant depend on the plant's employment.

Upon a successful innovation, a firm obtains monopoly rights over the new frontier technology in sector j, displacing the existing incumbent. As a result, an incumbent is subject to being displaced by innovation at the equilibrium arrival rate of new technologies in the economy, or the endogenous rate of creative destruction, $\bar{\lambda}$, which is equal to the sum of innovation efforts across all firms in the economy:

$$\bar{\lambda}_t = M_t \int_{\omega \in \Omega} \lambda_{\omega t} d\omega,$$

where Ω is the set of innovative firms in the economy.

The value of holding the rights to state-of-the-art technology in sector j, denoted as V_{jt}, satisfies the following equation:

$$r_t V_{jt} - \dot{V}_t = \pi_{jt} - \bar{\lambda} V_t. \tag{6}$$

where r_t is the interest rate in the economy and π_{jt} denotes the profits of the intermediate monopolist that produces good j with the frontier technology. The interpretation of this equation is that the value of holding the rights to state-of-the-art technology is equal to the discounted flow profits, π_{jt}, minus the risk of displacement due to creative destruction, captured by the last term.

Given the value of an innovation, we can express the problem of an innovative firm, which maximizes the expected value from new innovations net of labor costs in all of its locations:

$$\max \lambda_{\omega t} V_t - \sum_{n=1}^{N} \mathbb{I}_{n\omega t} \times \left[w_{nt}^{i} \left(\ell_{n\omega t} + f_{n\omega t} \right) \right],$$

where $\mathbb{I}_{n\omega t}$ is an indicator function for positive activity in location n at time t.

The mass of innovating firms in the economy is endogenous. To create a firm, a potential entrepreneur pays an entry cost f_e in terms of the final good.

4.2 Equilibrium

We focus on a balanced growth path (BGP) along which aggregate productivity $\mathcal{A}_t$, wages, consumption, profits, and the value of new innovations all grow at a constant rate g. The spatial distribution of these objects is invariant along the BGP.

4.2.1 Production

The standard household maximization problem yields the Euler equation $g = r - \rho$.

The maximization problem of the final goods producer generates the following demand for intermediate j:

$$y_{jt} = \frac{Y_t}{p_{jt}}.$$

The cost minimization problem for an intermediate monopolist implies that the marginal cost of producing an additional unit of the intermediate j when using the state-of-the-art technology is given by

$$q_{jt} = \frac{1}{A_{jt}} \left(\sum_{n=1}^{N} Z_n^{\sigma} \left(w_{nt}^{p} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}.$$

The monopolist with state-of-the-art technology engages in limit pricing and charges

a price equal to the marginal cost of the competitive fringe, such that the price of the intermediate j is $p_{jt} = \gamma q_{jt}$. As a result, the monopolist's profits are given by

$$\pi_{jt} = \bar{\pi} Y_t,$$

where we define $\bar{\pi} \equiv \left(\frac{\gamma-1}{\gamma}\right)$. We note that the profits are equal across all intermediate sectors j, so we denote $\pi_{jt} = \pi_t$.

4.2.2 Innovation

Combining equation 6 with the Euler equation, we obtain that the value of an innovation along a BGP is equal to :

$$V_t = \frac{\bar{\pi}}{\rho + \lambda} Y_t.$$

An innovative firm ω, conditional on activity in market n, chooses the following mass of inventors :

$$\ell_{n\omega} = \left(\frac{\eta K_n \lambda_0 V}{w_n^i}\right)^{\frac{1}{1-\eta}}.$$

The expected value from opening a plant in location n is then given by

$$V \lambda_0 K_n z_\omega^{1-\eta} \ell_{n\omega}^{\eta} - w_n^i \left(\ell_{n\omega} + f_{n\omega}\right)$$

and the firm chooses to open that plant if the fixed cost $f_{n\omega}$ is low enough:

$$f_{n\omega} < \left(\frac{1-\eta}{\eta}\right) \ell_{n\omega}\left(z_\omega\right).$$

The probability that a firm of type z opens a plant in location n, denoted as $\chi_n(z)$ is:

$$\chi_n(z) = F_f\left(\left(\frac{1-\eta}{\eta}\right) \ell_{n\omega}\left(z_\omega\right)\right) = f_{max}^{-\theta}\left(\left(\frac{1-\eta}{\eta}\right) \ell_{n\omega}\left(z_\omega\right)\right)^{\theta},$$

where the second equality follows from the assumption that F_f is an inverse Pareto. We note that all firms of type z hire the same number of workers for their innovation activity in location n, so from now on we omit the index ω. The expected amount of labor that a

firm of type z hires in location n to pay the fixed costs, conditional on being active in n, is:

$$\bar{f}_n(z) = \mathbb{E}\left[f | f < \left(\frac{1-\eta}{\eta}\right)\ell_n(z)\right] = \frac{\theta}{\theta+1}\left(\frac{1-\eta}{\eta}\ell_n(z)\right).$$

4.2.3 Equilibrium Definition

We characterize the equilibrium by detrending all variables, in which case we omit the time subscript. An equilibrium is an allocation of production and innovation workers for all markets, production and innovation wages in all markets, quantity and price of the final and intermediate goods, a mass of innovating firms, and a value of new innovations, that satisfy:

1. Labor market clearing in the production sector

$$w_n^p = \frac{1}{\gamma} Z_n^{\frac{\sigma-1}{\sigma}} \left(\frac{\bar{L}_n^p}{Y}\right)^{-\frac{1}{\sigma}}.$$

2. Labor market clearing in the innovation sector

$$\bar{L}_n^i = M \int_z \chi_n(z)\left(\bar{f}_n(z) + \ell_n(z)\right) dG(z).$$

where

$$\ell_n(z) = z\left(\frac{\eta K_n V}{w_n^i}\right)^{\frac{1}{1-\eta}}, \quad \bar{f}_n(z) = \frac{\theta}{\theta+1}\frac{1-\eta}{\eta}\ell_n(z), \quad \chi_n(z) = f_{max}^{-\theta}\left(\left(\frac{1-\eta}{\eta}\right)\ell_n(z)\right)^{\theta}.$$

3. Regional innovation productivity is given by

$$K_n = \bar{K}_n\left(M\int_z \chi_i(z)\, z^{1-\beta}\ell_i(z)^{\beta}\, dG(z)\right)^{\alpha}.$$

4. Market clearing in the final good market

$$Y = M f_e + \rho V + \sum_{n=1}^{N} w_{nt}^i \bar{L}_{nt}^i + \sum_{n=1}^{N} w_{nt}^p \bar{L}_{nt}^p.$$

5. The value of new innovations is given by

$$V = \frac{\left(\frac{\gamma-1}{\gamma}\right) Y}{\rho + \bar{\lambda}}.$$

where the rate of creative destruction is given by

$$\bar{\lambda} = M \int_z \sum_{n=1}^{N} \chi_n(z) K_n z^{1-\eta} \ell_n^{\eta}(z) \, dG(z).$$

6. Free entry

$$f_e = \int_z \sum_{n=1}^{N} \chi_n(z) \left[\lambda_0 K_n z^{1-\eta} \ell_n(z)^{\eta} V - w_n^i \left(\ell_n(z) + \bar{f}_n(z)\right)\right] dG(z)$$

7. Aggregate productivity $\mathcal{A}_t$, wages, consumption, profits, and the value of new innovations grow at a constant rate g, which is given by :

$$g = \bar{\lambda} \log \gamma.$$

4.3 Social Welfare

The model is inefficient due to different forces. As in Schumpeterian models, innovation could be sub-optimal due to two opposing forces. On the one hand, innovative firms do not internalize their effect on knowledge creation due to raising the level of quality upon which future innovations will be built. On the other hand, they do not internalize that their innovation efforts increase the creative destruction rate, lowering the equilibrium value of holding the frontier technology in a product line and thus discouraging innovation by others, a force known as the "business stealing" effect. The model also introduces a new spatial innovation externality. Innovative firms do not internalize that opening an establishment in a given location generates knowledge spillovers for other local firms, so that the geographic scope of innovative firms could be suboptimal. In addition, in the quantitative model in Section 5 we allow for labor mobility, introducing potential spatial misallocation.

To study the interaction between innovation and expansion inefficiency, in this sec-

tion, we consider the problem a constrained planner that cannot reallocate labor across sectors and space, but can reallocate resources between R&D and expansion, and across firm types. While the Schumpeterian growth literature typically examines the optimal allocation of resources between production and innovation, here we fix the amount of resources devoted to each sector, and we focus on the allocation of resources within the innovation sector, between R&D and geographic expansion. Thus, the objective of the constrained planner is to maximize aggregate innovation output, $\bar{\lambda}$, subject to innovation labor market clearing in each location, the spillover structure K_n, and the distribution of fixed costs and is given by:

$$\begin{aligned}
\max \quad & \bar{\lambda} \\
\text{s.t.} \quad & K_n = \bar{K}_n \left(M \int_z \chi_i(z) z^{1-\beta} \ell_i(z)^{\beta} dG(z) \right)^{\alpha} \\
& L_n^i = M \int_z \chi_n(z) \left(\bar{f}_n(z) + \ell_n(z) \right) dG(z) \\
& \bar{f}_n(z) = \int_0^{F^{-1}(\chi_n(z))} t \frac{dF_f(t)}{\chi_n(z)} = \frac{\theta}{\theta+1} f_{max} \chi_n(z)^{\frac{1}{\theta}}.
\end{aligned}$$

The following proposition characterizes the optimal allocation of the contained planner.

Proposition 1 *Constrained Planner's Allocation. Let $x_n(z)$ and $y_n(z)$ denote the amount of region-n innovation labor allocated to expansion and R&D activities of type-z firms, respectively. Their ratio in the planner's allocation is given by:*

$$\frac{x_n(z)}{y_n(z)} = \frac{\alpha \frac{s_n^{\beta}(z)}{s_n^{\eta}(z)} \left(\frac{1-\beta}{1-\eta} \right) + 1}{\alpha \frac{s_n^{\beta}(z)}{s_n^{\eta}(z)} \frac{\beta}{\eta} + 1} \frac{1-\eta}{\eta} \frac{\theta}{\theta+1}$$

where $s_n^{\eta}(z)$ and $s_n^{\beta}(z)$ are the weights of type z firms in local spillovers and in local innovation output, respectively:

$$s_n^{\eta}(z) \equiv \frac{g(z) z^{1-\eta} x_n(z)^{\frac{\theta}{\theta+1}(1-\eta)} y_n(z)^{\eta}}{\int_{z'} g(z') z'^{1-\eta} x_n(z')^{\frac{\theta}{\theta+1}(1-\eta)} y_n(z')^{\eta} dz}, \quad s_n^{\beta}(z) \equiv \frac{g(z) z^{1-\beta} x_n(z)^{\frac{\theta}{\theta+1}(1-\beta)} y_n(z)^{\beta}}{\int_{z'} g(z') z'^{1-\beta} x_n(z')^{\frac{\theta}{\theta+1}(1-\beta)} y_n(z')^{\beta} dz'}$$

We note that, when there are no local knowledge spillovers ($\alpha = 0$) or local spillovers are a function of regional innovation output ($\eta = \beta$), this solution coincides with the decentralized equilibrium. In addition, when ideas from different firms are sufficiently differentiated ($\alpha > 0$ and $\beta < \eta$), the planner allocates more resources to expansion,

meaning that innovating firms are too geographically concentrated in the decentralized equilibrium. The planner induces more expansion for firms whose contribution to spillovers is greater than their contribution to patent creation.

The next proposition characterizes the optimal allocation across firm types.

Proposition 2 *Optimal allocation across firm types. The mass of local innovation labor that the planner allocates to firms of type z is given by:*

$$\frac{\alpha s_n^{\beta}(z)\left(\beta+(1-\beta)\frac{\theta}{\theta+1}\right)+s_n^{\eta}(z)\left(\eta+(1-\eta)\frac{\theta}{\theta+1}\right)}{\int_{z'}\left(\alpha s_n^{\beta}(z')\left(\beta+(1-\beta)\frac{\theta}{\theta+1}\right)+s_n^{\eta}(z')\left(\eta+(1-\eta)\frac{\theta}{\theta+1}\right)\right)dG(z')}L_n^i$$

When $\alpha=0$ or $\eta=\beta$, this solution coincides with the decentralized equilibrium.

4.4 Extensions

4.4.1 Spillovers from all regions

The model outlined so far assumed that spillovers operate only at the local level, i.e. innovative firms gain insights only from firms that share their locations. This framework can easily be extended to allow for spillovers from all regions, where frictions for knowledge to travel across space from location i to location n are captured by a parameter κ_{in}. Let the local knowledge spillover K_n now be given by

$$K_n=\bar{K}_n\left(\sum_{i=1}^{N}\int_{\omega\in\Omega_i}\kappa_{in}z_{\omega}^{1-\beta}\ell_{i\omega}^{\beta}d\omega\right)^{\alpha}.$$

Under this assumption, the optimal allocation is the same, with

$$s_n^{\beta}(z)\equiv\sum_{k=1}^{N}\int_{z'}s_k^{\eta}(z')\,s_{nk}^{\beta}(z)\,dG(z')$$

where $s_{nk}^{\beta}(z)$ is the weight of plants from region n, type z in delivering insights to k.

4.4.2 Span of control

In our model, innovative firms need to pay an idiosyncratic fixed cost to open an establishment in a given location. An alternative assumption commonly used in the spatial literature is that, instead of random fixed costs, geographic expansion is costly due to a span of control problem. In this section, we explore the properties of our model under

this alternative assumption.

In particular, suppose that space is continuous and symmetric. Instead of random fixed costs, an innovative firm operating in $\mathcal{N}$ locations incurs a cost of expansion given by $(\mathcal{N})^{1+\frac{1}{\theta}}$. Under these assumptions, the optimal allocation between R&D and expansion follows the same expression as in Proposition 1.

5. Quantitative Analysis

In this section, we introduce a quantitative version of the model to examine the following questions. Are innovative firms too geographically concentrated (i.e., is β low enough relative to η)? What are the potential welfare gains from incentivizing them to expand in space? What are the spatial implications of such policies?

5.1 Estimation

We estimate the model for 167 U.S. local labor markets in 2019. The quantitative model includes the following additional elements relative to the baseline model. First, we introduce labor mobility, as workers sort across space and sectors subject to idiosyncratic preferences. Second, we assume that each location consists of a unit of land, with a constant expenditure share. Third, we introduce amenities.

The estimation procedure follows four steps. First, we estimated directly from the data the parameters for the decreasing returns to R&D labor in idea-creation (η) and for the spillovers (α and β). Second, we use typical values from the literature the calibrate the parameters for the discount rate (ρ), the elasticity of substitution in production (σ), and the housing and labor supply elasticities. Third, we implement a Simulated Method of Moments to obtain the values of the step-size in innovation (γ) and the distribution of fixed costs . Fourth, we implement a model inversion and we set regional productivity in production (Z_i) and innovation ($\bar{K}_i$), and amenities, to match the distribution of employment and wages from BLS-QCEW data.

Table 3 summarized the parameter values that are used for the quantified model.

5.2 Quantitative Results

The optimal allocation discussed in Section 4.3 is hard to implement in practice because it is location- and type-specific and it is non-linear. In this Section, we investigate

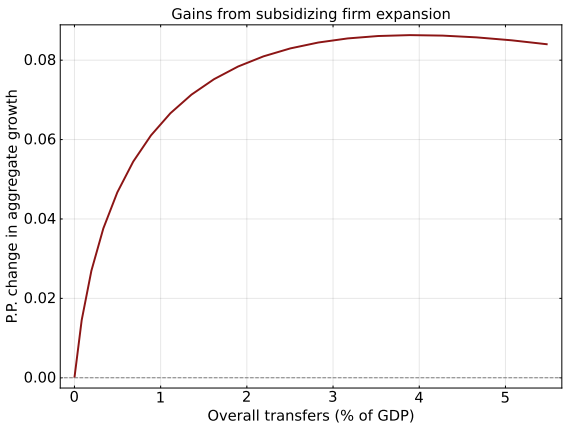
Parameter	Description	Value	Notes
Externally set			
ρ	Discount rate	0.05	
σ	EoS across locations in production	5	Costinot and Rodríguez-Clare (2014)
ψ	Housing share of consumption	0.25	Davis and Ortalo-Magne (2011)
ϵ	Location and sector preference dispersion	2	Galle et al. (2023)
Direct estimation			
β	DRS in local spillovers	0.2	
η	DRS in ideas production	0.7	
α	Elasticity of innovation to local insights	0.2	
Internal calibration			
γ	Step size for innovations	1.35	Target: innovation wage premium
λ_0	Economy-wide innovation efficiency	0.14	Target: growth rate
f_{max}	Average # of regions per firm	0.02	Target: average # of regions per firm
θ	Dispersion of fixed costs	1.5	Target: dispersion of # of regions per firm

Table 3: **Parameter values in the quantified model.**

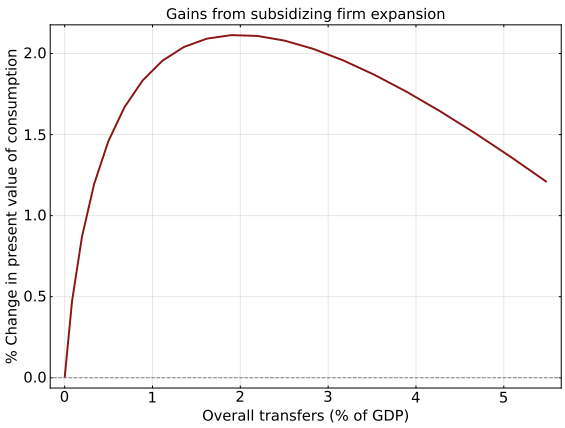
the implications of a simple subsidy scheme that could be easily implemented. We implement a subsidy at rate τ for innovative firms for every market in which they have innovation activity. The subsidy is financed by a lump sum tax levied on all households. We also introduce a tax on entry to keep the mass of firms unchanged, which allows us to isolate the gains due to spatial expansion separately from the gains due to a larger innovation sector.

Figure 8 shows the aggregate implications of subsidizing expansion for innovating firms. The left panel shows the change in the growth rate (in percentage points) relative to the benchmark equilibrium, against a measure of the scope of the subsidy program (measured in terms of total % of GDP). The right panel shows the welfare gains from such a program in terms of the present discounted value of aggregate consumption (in %-changes relative to the benchmark equilibrium) against the scope of the subsidy program. The effect of the program on the PDV of aggregate consumption is maximized when the size of the program is around 2% of GDP. In this case, aggregate growth is higher by 0.08 p.p., and the PDV of consumption is higher by around 2%.

Figure 9 shows the regional implications of subsidizing expansion for innovating firms.



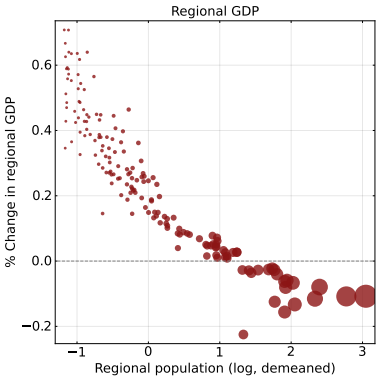
Change in growth (p.p.)



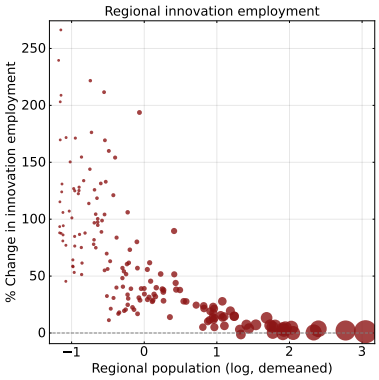
%-Change in PDV of Consumption

Figure 8: The aggregate effects of a subsidy for the spatial expansion of innovating firms.

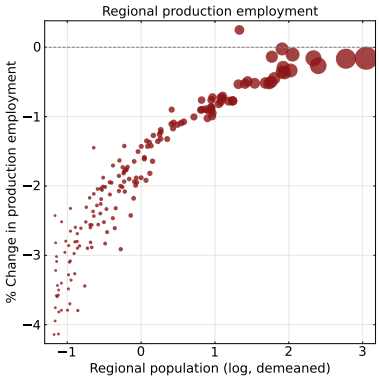
Each dot in each of the panels represents a local labor market, and they are weighted by the size of these labor markets. In all three panels, the horizontal axis shows the size of a local labor market, measured in log regional population. The left panel shows the percentage change in regional GDP relative to the baseline equilibrium when the above subsidy is implemented. The middle and right panels show the percentage changes in regional innovation and production employment, respectively. Evidently, the subsidy program induces a reallocation of economic activity from larger to smaller markets. In addition, it induces a reallocation of economic activity in the smaller markets from production to innovation. We conclude that a modest incentive for innovating firms to expand in space can lead to higher growth and lower spatial disparities.



% change in regional GDP



% change in innovation labor



% change in production labor

Figure 9: The regional effects of a subsidy for the spatial expansion of innovating firms.

6. Conclusion

When knowledge spillovers are local, firms' spatial scope affects aggregate innovation and productivity growth, and could be sub-optimal. In this paper, we provide evidence that firms' geographic expansion facilitates knowledge diffusion across innovating firms. This observation motivates a novel theoretical framework, which describes the process of creative destruction in a setup with multi-location innovative firms that generate endogenous local knowledge spillovers. We estimate the model using data from the geographic distribution of US innovative firms and we find that these firms are too geographically-concentrated compared to the welfare-maximizing allocation. We show that a simple policy such as a moderate subsidy to expansion can raise growth and reduce spatial inequality.

References

Aghion, Philippe and Peter Howitt (1992). "A Model of Growth Through Creative Destruction". In: *Econometrica* 60.2, pp. 323–351. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/2951599.

Akcigit, Ufuk and William R. Kerr (2018). "Growth through Heterogeneous Innovations". In: *Journal of Political Economy* 126.4, pp. 1374–1443. DOI: 10.1086/697901. eprint: https://doi.org/10.1086/697901. URL: https://doi.org/10.1086/697901.

Akcigit, Ufuk and Marc Melitz (2022). "International trade and innovation". In: *Handbook of International Economics*. Vol. 5. Elsevier, pp. 377–404.

Argente, David, Sara Moreira, Ezra Oberfield, and Venky Venkateswaran (2020). *Scalable Expertise*.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager (2021). "Synthetic Difference-in-Differences". In: *American Economic Review* 111.12, pp. 4088–4118. ISSN: 0002-8282. DOI: 10.1257/aer.20190159. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20190159.

Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple (2018). "Innovation and production in the global economy". In: *American Economic Review* 108.8, pp. 2128–73.

Audretsch, David B and Maryann P Feldman (1996). "R&D spillovers and the geography of innovation and production". In: *The American economic review* 86.3, pp. 630–640.

Baker, Andrew C., David F. Larcker, and Charles C. Y. Wang (2022). "How much should we trust staggered difference-in-differences estimates?" In: *Journal of Financial Economics* 144.2, pp. 370–395. ISSN: 0304-405X. DOI: 10.1016/j.jfineco.2022.01.004. URL: https://www.sciencedirect.com/science/article/pii/S0304405X22000204.

Barnatchez, Keith, Leland D. Crane, and Ryan A. Decker (2017). "An Assessment of the National Establishment Time Series (NETS) Database". In: URL: https://www.federalreserve.gov/econres/feds/an-assessment-of-the-national-establishment-time-series-nets-database.htm.

Bergé, Laurent (2018). "Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm". In: *CREA Discussion Papers* 13.

Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen (2019). "What drives differences in management practices?" In: *American Economic Review* 109.5, pp. 1648–1683.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen (2013). "Identifying Technology Spillovers and Product Market Rivalry". In: *Econometrica* 81.4. _eprint: https://onlinelibrary.wiley. pp. 1347–1393. ISSN: 1468-0262. DOI: 10.3982/ECTA9466. URL: https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA9466.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2024). *Revisiting Event Study Designs: Robust and Efficient Estimation*. DOI: 10.48550/arXiv.2108.12419. arXiv: 2108.12419[econ]. URL: http://arxiv.org/abs/2108.12419.

Buera, Francisco J. and Ezra Oberfield (2020). "The Global Diffusion of Ideas". In: *Econometrica* 88.1, pp. 83–114.

Callaway, Brantly and Pedro H. C. Sant'Anna (2021). "Difference-in-Differences with multiple time periods". In: *Journal of Econometrics*. Themed Issue: Treatment Effect 1 225.2, pp. 200–230. ISSN: 0304-4076. DOI: 10.1016/j.jeconom.2020.12.001. URL: https://www.sciencedirect.com/science/article/pii/S0304407620303948.

Carlino, Gerald and William R. Kerr (2015). "Chapter 6 - Agglomeration and Innovation". In: *Handbook of Regional and Urban Economics*. Ed. by Gilles Duranton, J. Vernon Henderson, and William C. Strange. Vol. 5. Handbook of Regional and Urban Economics. Elsevier, pp. 349–404. DOI: https://doi.org/10.1016/B978-0-444-59517-1.00006-4. URL: https://www.sciencedirect.com/science/article/pii/B9780444595171000064.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer (2019). "The Effect of Minimum Wages on Low-Wage Jobs*". In: *The Quarterly Journal of Economics* 134.3, pp. 1405–1454. ISSN: 0033-5533. DOI: 10.1093/qje/qjz014. URL: https://doi.org/10.1093/qje/qjz014.

Chaisemartin, Clément de and Xavier D'Haultfœuille (2020). "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects". In: *American Economic Review* 110.9, pp. 2964–2996. ISSN: 0002-8282. DOI: 10.1257/aer.20181169. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20181169.

Correia, Sergio (2016). *Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator*. Tech. rep. Working Paper.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin (2020). "ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects". In: *The Stata Journal: Promoting communications on statistics and Stata* 20.1, pp. 95–115. ISSN: 1536-867X, 1536-8734. DOI: 10.1177/1536867X20909691. arXiv: 1903.01690[econ]. URL: http://arxiv.org/abs/1903.01690.

Crews, Levi Garrett (2023). "A dynamic spatial knowledge economy". In.

Cunningham, Scott (2021). *Causal Inference: The Mixtape*. New Haven London: Yale University Press. 584 pp. ISBN: 978-0-300-25168-5.

Dube, Arindrajit, Daniele Girardi, Òscar Jordà, and Alan M. Taylor (2023). *A Local Projections Approach to Difference-in-Differences*. DOI: 10.3386/w31184. URL: https://www.nber.org/papers/w31184.

Giroud, Xavier, Simone Lenzu, Quinn Maingi, and Holger M Mueller (2021). *Propagation and amplification of local productivity spillovers*. Tech. rep. National Bureau of Economic Research.

Goodman-Bacon, Andrew (2021). "Difference-in-differences with variation in treatment timing". In: *Journal of Econometrics*. Themed Issue: Treatment Effect 1 225.2, pp. 254–277. ISSN: 0304-4076. DOI: 10.1016/j.jeconom.2021.03.014. URL: https://www.sciencedirect.com/science/article/pii/S0304407621001445.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti (2010). "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings". In: *Journal of political economy* 118.3, pp. 536–598.

Griffith, Rachel, Sokbae Lee, and John Van Reenen (2011). "Is distance dying at last? Falling home bias in fixed-effects models of patent citations". In: *Quantitative Economics* 2.2, pp. 211–249.

Grossman, Gene M. and Elhanan Helpman (1991). "Quality Ladders in the Theory of Growth". In: *The Review of Economic Studies* 58.1, pp. 43–61. ISSN: 0034-6527. DOI: 10.2307/2298044. eprint: https://academic.oup.com/restud/article-pdf/58/1/43/4361322/58-1-43.pdf. URL: https://doi.org/10.2307/2298044.

Helpman, Elhanan (1984). "A simple theory of international trade with multinational corporations". In: *Journal of political economy* 92.3, pp. 451–471.

Holmes, Thomas J (2011). "The diffusion of Wal-Mart and economies of density". In: *Econometrica* 79.1, pp. 253–302.

Hsieh, Chang-Tai and Esteban Rossi-Hansberg (2023). "The Industrial Revolution in Services". In: *Journal of Political Economy Macroeconomics* 1.1. Publisher: The University of Chicago Press, pp. 3–42. ISSN: 2832-9341. DOI: 10.1086/723009. URL: https://www.journals.uchicago.edu/doi/full/10.1086/723009.

Hughes, Ryan, Charles deGrazia, and Julian Kolev (2021). *Technical Documentation for Matching Patents and Trademarks to the 2017 National Establishment Time Series Database*. URL: https://www.uspto.gov/sites/default/files/documents/oce-wp-ip-to-nets.pdf.

Jacobs, Jane (1969). *The economy of cities*. Vintage.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson (1993). "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations". In: *The Quarterly Journal of Economics* 108.3. Publisher: Oxford University Press, pp. 577–598. ISSN: 0033-5533. DOI: 10.2307/2118401. URL: https://www.jstor.org/stable/2118401.

Jia, Panle (2008). "What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry". In: *Econometrica* 76.6, pp. 1263–1316.

Kleinman, Benny (2023). "Wage Inequality and the Spatial Expansion of Firms". In.

Klette, Tor Jakob and Samuel Kortum (2004). "Innovating Firms and Aggregate Innovation". In: *Journal of Political Economy* 112.5, pp. 986–1018.

Lhuillier, Hugo (2023). "Should I Stay or Should I Grow?" In.

Lucas, Robert E. and Benjamin Moll (2014). "Knowledge Growth and the Allocation of Time". In: *Journal of Political Economy* 122.1, pp. 1–51.

Martellini, Paolo (2022). *Local labor markets and aggregate productivity*. Tech. rep. Working paper.

Matray, Adrien (2021). "The local innovation spillovers of listed firms". In: *Journal of Financial Economics* 141.2, pp. 395–412. ISSN: 0304-405X. DOI: https://doi.org/10.1016/j.jfineco.2021.04.009. URL: https://www.sciencedirect.com/science/article/pii/S0304405X21001367.

Melitz, Marc J and Stephen J Redding (2021). *Trade and innovation*. Tech. rep. National bureau of economic research.

Oberfield, Ezra, Esteban Rossi-Hansberg, Pierre-Daniel Sarte, and Nicholas Trachter (2024). "Plants in space". In: *Journal of Political Economy* 132.3, pp. 000–000.

Peri, Giovanni (2005). “Determinants of Knowledge Flows and Their Effect on Innovation”. In: *The Review of Economics and Statistics* 87.2, pp. 308–322. ISSN: 0034-6535. DOI: 10.1162/0034653053970258. eprint: https://direct.mit.edu/rest/article-pdf/87/2/308/1613967/0034653053970258.pdf. URL: https://doi.org/10.1162/0034653053970258.

Perla, Jesse and Christopher Tonetti (2014). “Equilibrium Imitation and Growth”. In: *Journal of Political Economy* 122.1, pp. 52–76.

Prato, Marta (2022). “The Global Race for Talent: Brain Drain, Knowledge Transfer, and Growth”. In: Working Paper.

Ramondo, Natalia and Andrés Rodríguez-Clare (2013). “Trade, multinational production, and the gains from openness”. In: *Journal of Political Economy* 121.2, pp. 273–322.

Sun, Liyang and Sarah Abraham (2021). “Estimating dynamic treatment effects in event studies with heterogeneous treatment effects”. In: *Journal of Econometrics*. Themed Issue: Treatment Effect 1 225.2, pp. 175–199. ISSN: 0304-4076. DOI: 10.1016/j.jeconom.2020.09.006. URL: https://www.sciencedirect.com/science/article/pii/S030440762030378X.

Tintelnot, Felix (2017). “Global production with export platforms”. In: *The Quarterly Journal of Economics* 132.1, pp. 157–209.

Wasi, Nada and Aaron Flaaen (2015). “Record Linkage Using Stata: Preprocessing, Linking, and Reviewing Utilities”. In: *The Stata Journal* 15.3. Publisher: SAGE Publications, pp. 672–697. ISSN: 1536-867X. DOI: 10.1177/1536867X1501500304. URL: https://doi.org/10.1177/1536867X1501500304.

Patent ID	Raw Assignee Organization	CZ1990	Grant Date
4310440	Union Carbide Corporation	19400	1982-01-12

Table A.1: **Example patent record.** This table shows an example patent record.

dunsno	hqdunsno	parentdunsno	ultdunsno	status	companyname	nationalcode	cz1990	firstyear	lastyear
1289008	1289008	—	—	1	UNION CARBIDE CORP	0	19400	$-\infty$	1970
1289008	1289008	—	1289008	1	UNION CARBIDE CORP	0	19400	1971	1973
1289008	1289008	—	1289008	1	UNION CARBIDE CORPORATION	0	19400	1974	1976
1289008	1289008	—	1289008	1	UNION CARBIDE CORPORATION	0	19400	1978	1981
1289008	1289008	—	1289008	1	UNION CARBIDE CORPORATION	0	20901	1982	1988
1289008	1289008	—	1289008	1	UNION CARBIDE CORPORATION	0	20901	1995	2000
1289008	1289008	1381581	1381581	1	UNION CARBIDE CORPORATION	0	20901	2001	2004
1289008	1289008	1381581	1381581	1	UNION CARBIDE CORPORATION	0	32000	2005	2016
1289008	1289008	1381581	80386615	1	UNION CARBIDE CORPORATION	0	32000	2017	2018
1289008	1289008	1381581	117063457	1	UNION CARBIDE CORPORATION	0	32000	2019	∞
1289008	1289008	602844128	602844128	1	UNION CARBIDE CHEM & PLAS CO	0	20901	1989	1993
1289008	1289008	602844128	602844128	1	UNION CARBIDE CORPORATION	0	20901	1994	1994
1289008	1289008	<	1289008	1	UNION CARBIDE CORPORATION	0	19400	1977	1977

Table A.2: **Example Dun & Bradstreet record.** This table shows an example Dun & Bradstreet record. The construction of this entry follows the procedure described in Appendix A.1. Note that we are missing Dun & Bradstreet files for 1981 and 1984. We account for this by filling forward activity data in those years.

A. Data

A.1 Details of merging PatentsView and Dun & Bradstreet

Our main analysis uses a dataset that merges the PatentsView citations data with Dun & Bradstreet establishment data. Since Dun & Bradstreet establishment data is used to construct the National Establishment Time Series (NETS) data, we follow Hughes et al. (2021) and perform an "exact" merge with commuting zone level blocking.

To give a sense of our data structure, Tables A.1 and A.2 show example records that we are trying to match. Table **??** presents summary statistics on the quality of the merge.

1. Using the raw Dun & Bradstreet data, we construct a dataset with the first and last year a given Dun & Bradstreet observation appears. Dun & Bradstreet observations are indexed by their `dunsno-{companyname/secondaryname}-CZ1990`. An example for `dunsno` 1289008's `companyname` is shown in Table A.2.
2. Using Stata packages developed by Wasi and Flaaen (2015), we standardize the firm names in the Dun & Bradstreet data and in the patent data. For the patent data, we use the file `g_assignee_not_disambiguated.tsv.zip` from PatentsView, downloaded on December 15, 2023. The assignee file has geographic information on the assignee FIPS code, which we match to 1990 commuting zones using a crosswalk developed by the U.S. Department of Agriculture (USDA).

3. We perform an exact merge on name and commuting zone between the two datasets

   (a) Due to systematic asymmetric naming in the two data sets, we hand-merge all IBM records to `dunsno` `001368083`.
   (b) Using `companyname`, we match on standardized name and commuting zone
   (c) Next, we remove whatever matched in step 3b and match the residual on the compressed name (where we remove all white spaces from the standardized names) and commuting zone
   (d) Lastly, we remove whatever matched in step 3c and repeat steps 3b-3c for the `secondaryname`'s in Dun & Bradstreet.

4. Step 3 can produce many potential matches for a given patent record. Moreover, the match returns establishment-level matches, whereas we wish to aggregate patenting activity to the firm level. We do this as follows.

   (a) We assign codes to each match. In descending order of quality:
      i. "A"—If a patent's grant date falls within a `companyname-dunsno-CZ1990`'s first and last years of recorded activity in Dun & Bradstreet.
      ii. "B"'—If a patent's grant date falls within a `dunsno-CZ1990`'s first and last years of recorded activity in Dun & Bradstreet.
      iii. "C"—If a patent's grant date falls within a `dunsno`'s first and last years of recorded activity in Dun & Bradstreet.
      iv. "D"—If a patent does not satisfy any of the above.
      v. Note that for the example records in Tables A.1 and A.2, there would be a single type "A" match.
   (b) We associate each match with its "nearest" record in Dun & Bradstreet. For type "A" matches, this nearest record is an exact match. For all other types, we find the record nearest to the given patent's grant date.
   (c) Then we assign the patent to a unique Dun & Bradstreet identifier
      - If the patent is associated with no `hqdunsno`'s (i.e. it is only associated with `dunsno`'s), then we assign equal ownership to all unique `dunsno`'s
      - If the patent is associated with a single unique `hqdunsno`, then we assign the patent to the unique `hqdunsno`
      - If there are multiple unique `hqdunsno`, we refine in the following order:

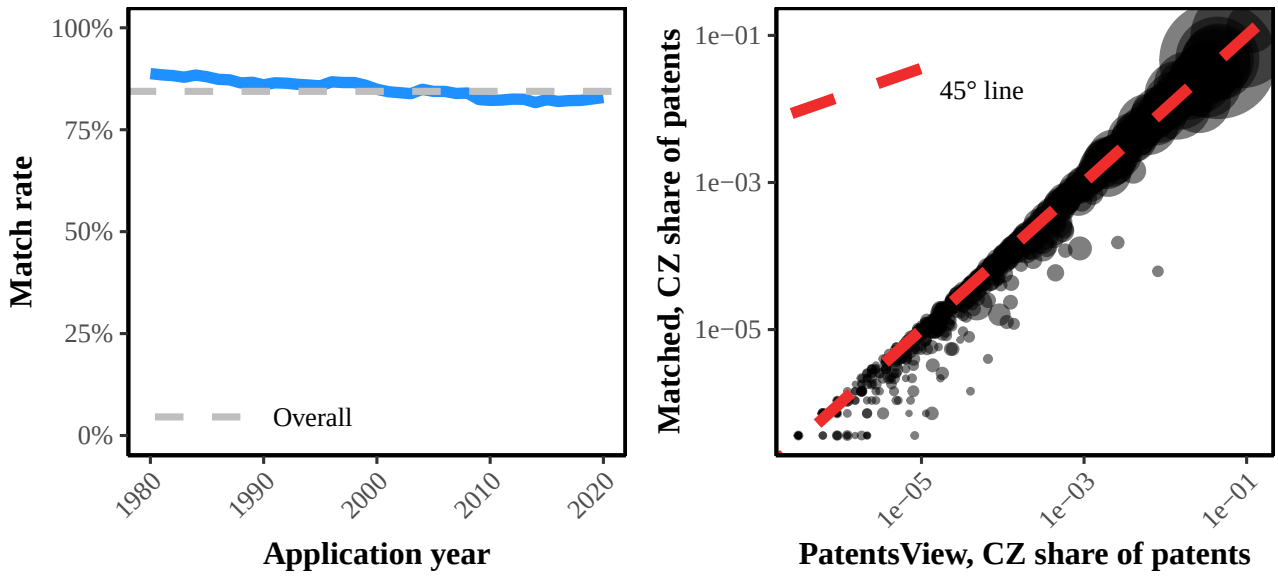


Figure A.1: **Match quality across time and space.** The left panel shows the match rate as a function of the application year of the patents. The grey dashed line is the overall match rate. The right panel shows the spatial distribution of patenting in the matched sample and the population, PatentsView. Each dot represents a commuting zone, and the sizes of the dots correspond to employment in each commuting zone. The right panel has $\log_{10}$ scaling along the x- and y-axes

SIC	Modal SIC4
50	5084, Industrial Machinery and Equipment
73	7389, Business Services, Not Elsewhere Classified
87	8711, Engineering Services
99	9999, Nonclassifiable Establishments
59	5999, Miscellaneous Retail Stores, Not Elsewhere Classified
51	5191, Farm Supplies

Table A.3: **Retail, wholesale trade, and services SIC codes in further detail.** For each non-manufacturing SIC2 in Table 2, this table shows the modal SIC4 associated with expansions in our data.

i. Multiple `dunsno` report HQ activity. Take all unique `ultdunsno`'s.

ii. All `dunsno` report single location or branch activity. Take all unique `hqdunsno`'s.

iii. There is a single `ultdunsno`. Take the unique `ultdunsno`.

iv. Otherwise, take all unique `hqdunsno`.

A.2 Allocating inventors to firms

In order to calculate our baseline citation series, we require a panel of inventor-firm links. The patent data do not provide this directly, but we are able to construct one as follows.

1. `g_inventor_disambiguated.tsv.zip` provides the disambiguated inventors for each of our patents as well as their self-reported locations.
2. We merge the patent ownership dataset (Appendix A.1) onto the Dun & Bradstreet

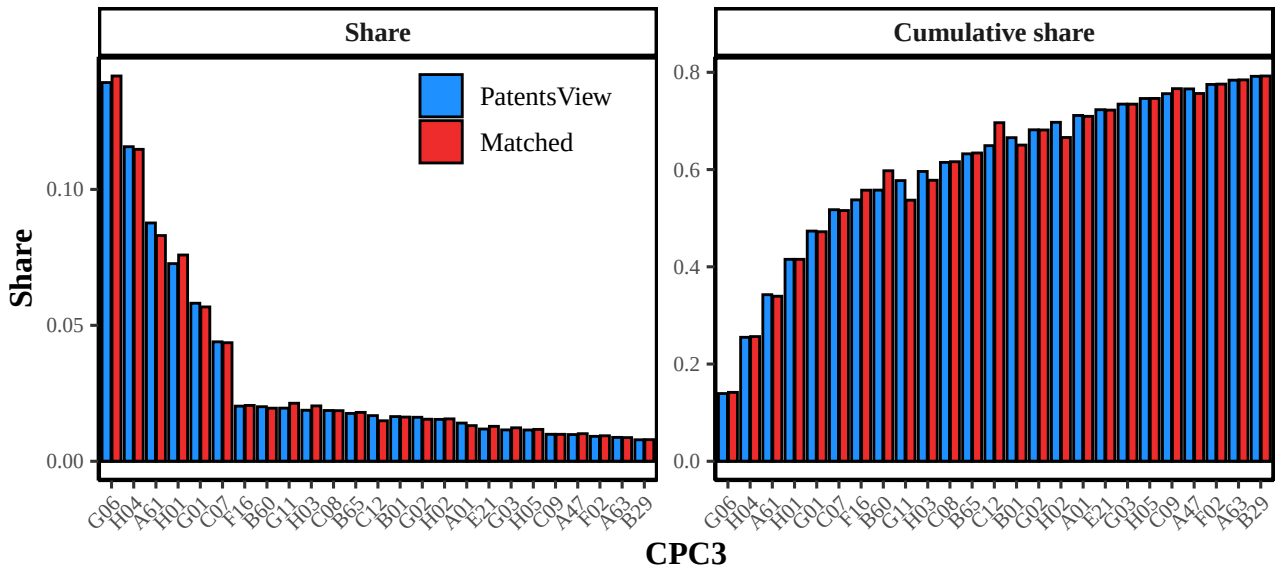


Figure A.2: **Distribution of three digit CPC in PatentsView and in the matched sample.**

data by `dunsno-year` (where `year` for the patent ownership data is the application year). Establishment activity is considered within a $+2$ year window around the report date (i.e. if a patent is applied for in 2008 and Dun & Bradstreet reports establishment activity any time between 2008-2010, this would be included in our merge). This is a many-to-many merge.

3. By `patent_id` and `cz1990` we merge the inventor file onto the dataset created in step 2. We take all unique `inventor_id-year-firm_id-cz1990` matches.
   - If an inventor reports multiple commuting zones in a year, it may be assigned to multiple firm branches in a given year.
   - We attempt to refine this match by looking at the inventor's self-reported FIPS (i.e. state-county pair) and checking for establishment activity in those counties. If this procedure enables us to eliminate a multiplicitous commuting zone, we do so; if multiple counties have establishment activity, we keep multiple matches.
4. We residualize any remaining `inventor_id-year`'s who could not be exact matched on commuting zone and merge this residual, by `patent_id`, on to the ownership file. Within an `inventor_id`-year, we take the inventor-firm pair with the smallest geographical distance between the `inventor_id-FIPS` and the `firm_id- FIPS` as defined using the geographical centroids.

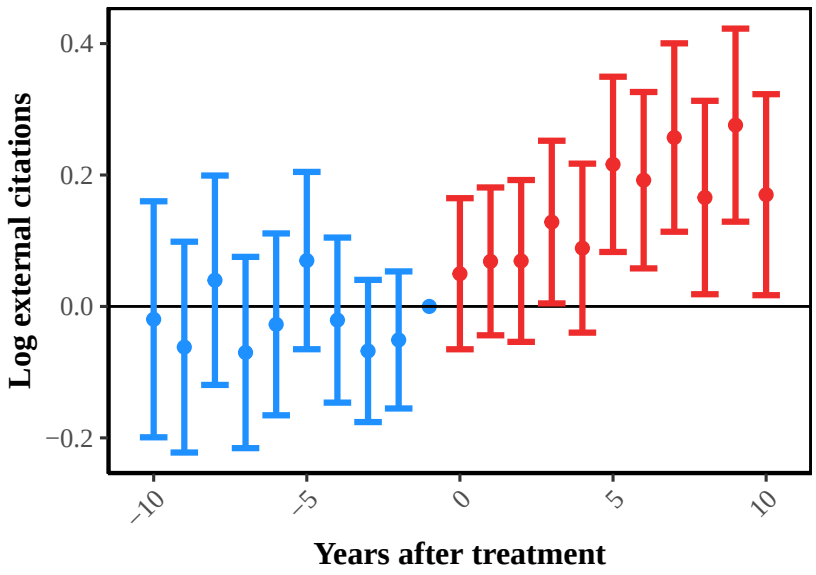


Figure B.1: **Response of citations received by firm patents in expansion commuting zone preceding and following expansion, excluding temporarily treated units.** The outcome variable is the log number of citations made to stock of patents of firm $Firm(i)$ at time t from inventors residing in commuting zone $CZ(i)$. We remove all internal citations and all external citations made from patent documents from firms $f \neq Firm(i)$ that ever shared inventors with firm $Firm(i)$; we further only consider citations made to patents originated outside of $CZ(i)$. Each point represents a loading on the relative treatment time dummies specified in (1). Standard errors are clustered at the firm $\times$ commuting zone level. Error bars represent 5% confidence intervals. The Wald test of joint nullity of the pre-treatment dummies, $\{\beta_\tau\}_{\tau=-10}^{-2}$, returns a p-value of 0.30.

5. Any patent that remains is either associated with an inventor outside the U.S. or is unmatched in our sample (i.e. missing an `inventor_id` `cz1990` or a `dunsno-cz1990`). Thus, it cannot be matched.

B. Robustness

B.1 Excluding nonabsorbing commuting zones

Figure B.1 shows the results after running the regression in (1) but drops treated units that eventually exit the expansion market. This constricts the sample size, resulting in a slight loss of statistical power, but the main result is preserved. Ten years following expansion, the number of citations received, relative to control units, is roughly 20% higher.

B.2 Million dollar plant design

C. Details of the matching procedure for DiD analysis

C.1 Exact match on firm

- For a treated unit i (i.e. where $g(i) < \infty$), consider all units i' where $g(i') = \infty$ and $Firm(i') = Firm(i)$. These i' are potential control units. We allow for potential control units to be repeated within a multi-establishment firm, across treated units.
- Treated and potential control units must have received a citation within 10 years of the treated unit's expansion event.
  - For matching on Log's, we additionally require that both units have positive citations in at least 2 out of 10 years prior to expansion (these need not occur in the same years).
- Let τ denote the relative time for the treated units ($\tau = t - g(i)$, where t is year and $g(i)$ is the first year of expansion).
- Let $y_{i\tau}$ denote an outcome of interest for firm $\times$ CZ1990 i at relative time τ.
- **Matching on pre-treatment outcomes, *changes***

$$\texttt{SLOPE}_i = \frac{\sum_{\tau=-10}^{-6} y_{i\tau}}{5} - \frac{\sum_{\tau=-5}^{-1} y_{i\tau}}{5}$$
$$\texttt{CRIT1}_{ii'} = \frac{\texttt{SLOPE}_i - \texttt{SLOPE}_{i'}}{0.5\left(|\texttt{SLOPE}_i| + |\texttt{SLOPE}_{i'}|\right)}$$

  - If $\texttt{SLOPE}_i$ and $\texttt{SLOPE}_{i'}$ have different **signs**, then remove these potential matches.
- **Matching on pre-treatment outcomes, *levels***
  - Outcome in $\log, \log(1+\cdot)$:

$$\texttt{CRIT2}_{ii'} = \frac{1}{10}\sum_{\tau=-10}^{-1}\left(y_{i\tau} - y_{i'\tau}\right)$$

  - Outcome in level or indicator function:

$$\texttt{CRIT2}_{ii'} = \frac{1}{10}\sum_{\tau=-10}^{-1}\frac{y_{i\tau} - y_{i'\tau}}{0.5\left(y_{i\tau} + y_{i'\tau}\right)}$$

  - If $y_{i\tau} = y_{i'\tau} = 0$, set this DHS difference to 0 (i.e. don't consider it missing when taking the average over $[-10, -1]$).
- **Matching on pre-treatment covariates**
  1. **Employment shares**—2 digit NAICS supersectors in commuting zones $CZ(i)$

and $CZ(i')$

2. **Business Dynamics**—Employment to firm and establishment to firm ratios
   (a) Large firms
   (b) Overall
3. **Demographic**
   (a) Log income per capita
   (b) Log employment
   (c) Ratio of college-educated to noncollege-educated workforces

$\texttt{CRIT3}_{ii'}$ is equal to the mean of the DHS differences between $CZ(i)$ and $CZ(i')$ for all of the listed covariates over $\tau \in [-10, -1]$.

- **Select control unit for each treated unit**
  - Take $i'(i) = \arg\min_{i'} \texttt{CRIT}_{ii'}$ (use random numbers to break ties)

$$\texttt{CRIT}_{ii'} = \frac{1}{3}\left(\texttt{CRIT1}_{ii'} + \texttt{CRIT2}_{ii'} + \texttt{CRIT3}_{ii'}\right).$$

C.2 Inexact match on firm

- Repeat above **but** no longer require $\text{Firm}(i) = \text{Firm}(i')$.
- This massively expands the number of potential pairwise comparisons.
- Therefore, to make computationally tractable, we do the following:
  - We bin all treated units into 500 bins ($\approx$ 30 firm $\times$ CZ1990 per bin).
  - For each of these treated units, we allow them to be potentially matched to a 10% random sample (without replacement) of potential control units.
  - This sample is drawn independently for each of the 500 chunks.

D. Estimator of Cengiz et al. (2019)

$$y_{it} = \alpha_i + \delta_t + \gamma_{g(Treated(i))t} + \kappa_{Firm(i)t} + \beta_i \mathbb{I}\{t \geq g(i)\} + \varepsilon_{it}$$

Note that due to the presence of a Firm $\times$ year FE, this model makes **within firm-year** comparisons—it also makes within **group-year** comparisons.

We aim to show how invalid comparisons can arise even in the presence of a cohort $\times$ year fixed effect. Consider two units i and i' such that $Firm(i) = Firm(i')$ and $g(i) < g(i') < \infty$. Note that, in practice, unit fixed effects are actually unit $\times$ stack fixed effects

(i.e. $\alpha_i = \alpha_{ig(Treated(i))}$ where the second subindex is redundant for treated units but not for potentially repeated control units) as in equation (3).

Consider the difference between outcomes for units i and i' at time t

$$y_{it} - y_{i't} = (\alpha_i - \alpha_{i'}) + \left(\gamma_{g_i t} - \gamma_{g_{i'} t}\right) + \beta_i \mathbb{I}\{t \geq g_i\} - \beta_{i'} \mathbb{I}\{t \geq g_{i'}\} + (\varepsilon_{it} - \varepsilon_{i't})$$

Three periods:

1. Both units not yet treated—t_0:

$$y_{it_0} - y_{i't_0} = (\alpha_i - \alpha_{i'}) + \left(\gamma_{g_i t_0} - \gamma_{g_{i'} t_0}\right) + (\varepsilon_{it_0} - \varepsilon_{i't_0})$$

2. Unit i treated; unit i' not yet treated (i.e. $g_i \leq t_1 < g_{i'}$)

$$y_{it_1} - y_{i't_1} = (\alpha_i - \alpha_{i'}) + \left(\gamma_{g_i t_1} - \gamma_{g_{i'} t_1}\right) + \beta_i + (\varepsilon_{it_1} - \varepsilon_{i't_1})$$

3. Both units treated—$t_2 \geq g_i, g_{i'}$

$$y_{it_2} - y_{i't_2} = (\alpha_i - \alpha_{i'}) + \left(\gamma_{g_i t_2} - \gamma_{g_{i'} t_2}\right) + (\beta_i - \beta_{i'}) + (\varepsilon_{it_2} - \varepsilon_{i't_2})$$

Now, take differences in differences:

$$(y_{it_1} - y_{i't_1}) - (y_{it_0} - y_{i't_0}) = \left(\gamma_{g_i t_1} - \gamma_{g_{i'} t_1}\right) - \left(\gamma_{g_i t_0} - \gamma_{g_{i'} t_0}\right) + \beta_i + \mathcal{E}_{10}$$

$$(y_{it_2} - y_{i't_2}) - (y_{it_1} - y_{i't_1}) = \left(\gamma_{g_i t_2} - \gamma_{g_{i'} t_2}\right) - \left(\gamma_{g_i t_1} - \gamma_{g_{i'} t_1}\right) \underbrace{-\beta_{i'}}_{\text{Negative weight problem}} + \mathcal{E}_{21}$$

In contrast, suppose you only had a Stack x Year FE. Then, there are no restricted firm $\times$ year comparisons being made, only group $\times$ year.

Two periods:

1. Both units not yet treated—t_0:

$$y_{it_0} - y_{i't_0} = (\alpha_i - \alpha_{i'}) + (\varepsilon_{it_0} - \varepsilon_{i't_0})$$

2. Both units treated—$t_2 \geq g_i = g_{i'}$

$$y_{it_2} - y_{i't_2} = (\alpha_i - \alpha_{i'}) + (\varepsilon_{it_2} - \varepsilon_{i't_2})$$

Clearly we cannot identify $\{\beta_i\}_i$ by comparing i and i'. But suppose we compare i to any control unit c with $g(Treated(i)) = g_i = g_{i'}$:

1. Both units not yet treated—t_0:

$$y_{it_0} - y_{ct_0} = (\alpha_i - \alpha_c) + (\varepsilon_{it_0} - \varepsilon_{ct_0})$$

2. Unit i treated; unit c not yet treated

$$y_{it_1} - y_{ct_1} = (\alpha_i - \alpha_c) + \beta_i + (\varepsilon_{it_1} - \varepsilon_{ct_1})$$

Then we can make only a valid comparison:

$$(y_{it_1} - y_{ct_1}) - (y_{it_0} - y_{ct_0}) = \beta_i + \mathcal{E}_{10}$$

One caveat: **Why doesn't a year FE cause invalid comparisons with the presence of a group $\times$ year FE?**

- Within a year, after all, there are all sort of invalid comparisons that could arise
  - For example, one of them is given in the example with a Firm $\times$ Year FE
- In reality, the year FE is redundant (whenever it is switched on for a particular year, it is also switched on for the relevant Group $\times$ Year FE which, as we showed, rules out bad comparisons)

References

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer (2019). "The Effect of Minimum Wages on Low-Wage Jobs*". In: *The Quarterly Journal of Economics* 134.3, pp. 1405–1454. ISSN: 0033-5533. DOI: 10.1093/qje/qjz014. URL: https://doi.org/10.1093/qje/qjz014.

Hughes, Ryan, Charles deGrazia, and Julian Kolev (2021). *Technical Documentation for Matching Patents and Trademarks to the 2017 National Establishment Time Series Database*. URL: https://www.uspto.gov/sites/default/files/documents/oce-wp-ip-to-nets.pdf.

Wasi, Nada and Aaron Flaaen (2015). "Record Linkage Using Stata: Preprocessing, Linking, and Reviewing Utilities". In: *The Stata Journal* 15.3. Publisher: SAGE Publications, pp. 672–697. ISSN: 1536-867X. DOI: 10.1177/1536867X1501500304. URL: https://doi.org/10.1177/1536867X1501500304.