The Geography of Innovative Firms

Craig A. Chikis
University of Chicago

Benny Kleinman
Stanford University

Marta Prato
Bocconi University*

June 15, 2025

Abstract

Most U.S. innovation output originates from firms that operate R&D facilities across multiple local markets. We study how this geographic structure influences aggregate innovation and growth, and whether it is socially optimal. First, we develop an endogenous growth model featuring multi-market innovative firms that generate knowledge spillovers to geographically proximate firms. In equilibrium, firms may operate in too few or too many local markets, depending on how sensitive are the local spillovers they generate to their local size. Second, to quantify these effects, we link the model to data on firms' R&D locations, patents, and citation networks. Using an event-study design, we show that firms' spatial expansion increases spillovers to other firms and estimate how these spillovers depend on a firm's local footprint. Our estimates imply that U.S. innovative firms operate in too few markets relative to the social optimum. Third, using quantitative counterfactuals, we find that policies promoting broader spatial scope yield larger welfare gains than standard R&D subsidies. Moreover, unlike R&D subsidies, such policies can also reduce regional inequality.

Keywords: Innovation, Firm Geography, Knowledge Spillovers, Economic Growth, R&D Policy, Regional Disparities

JEL Classification: O30, O40, R10, L10

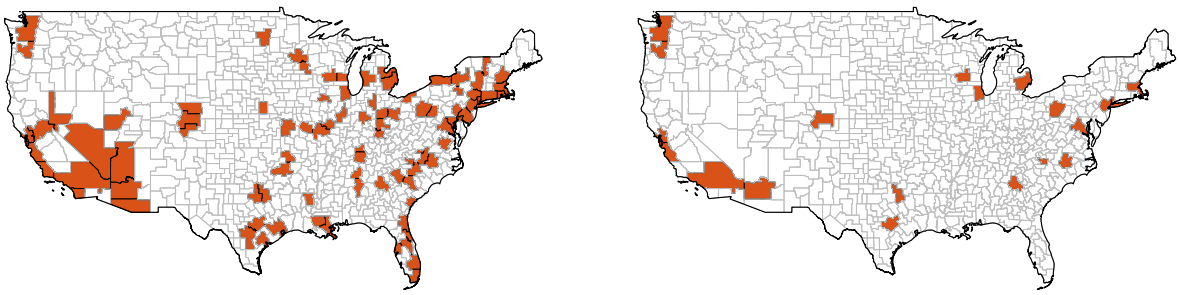
*Chikis: cachikis@uchicago.edu. Kleinman: bennyk@stanford.edu. Prato: marta.prato@unibocconi.it. We are grateful to University of Chicago, including the Becker Friedman Institute for Economics, Stanford University, and Bocconi University, for research support. We would like to thank Adrien Bilal, Pablo D. Fajgelbaum, Chad Jones, Pete Klenow, Esteban Rossi-Hansberg, Christopher Tonetti, and conference and seminar participants at the Chicago-Princeton Conference on Spatial Economics, West Coast Spatial Workshop, Junior Spatial Economics Conference, SED, USC, UCLA, UC Berkeley, Penn State, Stanford, Stanford GSB, EUI, LUISS, Helsinki, for helpful comments and suggestions.

1 Introduction

The majority of U.S. innovation output originates from firms that operate R&D facilities across multiple local markets. Figure 1 illustrates two examples: IBM conducts innovation activities in over 50 local markets, while Google operates in around 20. If such facilities generate local knowledge spillovers to surrounding firms and inventors, firms' location decisions affect not only their own innovation output but also that of others. In this context, a firm's decision about how many markets to operate in can influence aggregate innovation and economic growth—and may be socially suboptimal.

Figure 1: Innovation facility locations of IBM and Google

(a) IBM (b) Google



Note: Panel (a) displays a map of IBM's innovation facility locations across U.S. commuting zones. We define an innovation facility as a location where the firm has an establishment (based on Dun & Bradstreet data) and has filed at least one patent with inventors from that location (based on PatentView data). Panel (b) presents a corresponding map for Google.

In this paper, we study how the geographic structure of innovative firms affects aggregate innovation and growth, whether it is socially optimal, and the welfare implications of policies that alter firms' geographical scope. To address these questions, we develop an endogenous growth model with multiple local markets, multi-location innovating firms, and knowledge spillovers from local firm activity. We characterize both the competitive equilibrium and the social planner's solution. We show that even when the allocation of labor across space is efficient, the geographical scope of firms may not be. Depending on how sensitive spillovers are to a firm's local activity, firms in equilibrium may operate in too few or too many local markets. To quantify these forces, we link the model to data on firms' spatial organization, innovation activity, and citation networks. We provide empirical evidence for the model's key assumptions, including the presence of local knowledge spillovers from firms' R&D facilities. Estimating the model, we find parameter values suggesting that U.S. innovating firms operate in too few markets relative to the social optimum. Using the quantified model,

we evaluate the effectiveness of policies that promote geographical expansion and show that they outperform conventional R&D subsidies.

Our theoretical framework extends the canonical "creative destruction" model from the economic growth literature to a multi-market setting with two key components. First, we introduce multi-location innovative firms that employ R&D workers across multiple local markets. Firms incur a fixed cost for each location, generating increasing returns that incentivize the concentration of economic activity. At the same time, firms face diminishing labor efficiency within each location,[1] generating incentives for firm-level geographical diversification.

Second, following the economic geography literature, we allow the productivity of R&D workers in each market to depend on the mass of local R&D employment and innovation firms. Importantly, we assume that these spillovers exhibit diminishing learning from each individual firm. Our assumptions imply that, all else equal, workers in more efficient or larger firms contribute more to local agglomeration spillovers. At the same time, the marginal worker in a given firm may generate smaller spillovers when the firm already employs many local workers. Intuitively, learning from the marginal Google employee in a given market may be limited if Google already has a large local footprint. One way to interpret this spillover structure is that there are gains from a diversity of ideas, as emphasized by Jacobs (1969).

Solving the social planner's problem, we establish that firms in the decentralized equilibrium may operate in either too few or too many markets. On the one hand, when a firm opens an R&D plant in a new market, it generates knowledge spillovers in that location. On the other hand, due to labor-market clearing, new R&D plants reallocate labor from existing ones. If local spillovers from a given plant increase with its size, this reallocation reduces spillovers from the shrinking plants. The optimal number of locations per firm thus depends on how sensitive these spillovers are to the firm's local employment. Intuitively, if knowledge spillovers from a Google office depend only weakly on its size—so that even a single employee generates substantial benefits for nearby firms—a policymaker would prefer Google to open many small offices nationwide. At the opposite extreme, when nearby firms benefit only if Google operates a large local team, the planner might prefer fewer, larger facilities. In the knife-edge case where the spillover elasticity matches the firm's own innovation elasticity, expansion decisions are efficient even in the presence of knowledge spillovers.[2]

[1] A potential micro-foundation for this assumption is that workers are differentiated either by the amenity value they derive from each firm or by the efficiency units they contribute. Firms then face diminishing marginal efficiency of labor in each location.

[2] These results resemble optimal entry results in the spirit of Dixit and Stiglitz (1977), although here, the extent of sub-optimal plant entry is determined by the structure of knowledge spillovers across firms rather

We also examine the spatial allocation of innovation labor, which raises a distinct question from that of firms' optimal spatial scope. To clarify, in our model, the planner may wish to reallocate resources between idea creation and the entry of additional innovative plants within each market, even when the spatial distribution of labor is held fixed. When we allow for free mobility of innovation labor, the allocation of innovation labor across space in the baseline model is efficient, despite the presence of knowledge spillovers.[3] More generally, in our quantified model—where we allow for limited labor mobility and compensating differentials across space—the spatial allocation of labor is inefficient. However, we find that the welfare gains from reallocating labor across space are small relative to those from changing the spatial scope of firms. Notably, a key difference between optimal spatial scope and the spatial allocation of labor lies in the presence of non-rivalries: while labor can be in only one market at a time, multiple markets can simultaneously benefit from the knowledge generated by a single firm. In our model, this takes the form of a firm-level entry cost that is paid only once—regardless of the number of markets in which the firm operates—combined with a local spillover structure that generates uninternalized gains from multi-market activity by the same firm.

To understand whether innovating firms operate in too few or too many locations in practice, and to understand the welfare implications of policies that induce them to expand, we estimate our model for the U.S. economy. To this end, we link data on 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.

We demonstrate that the data aligns with the key assumptions of our model. First, innovating firms that operate in multiple markets account for most of the patenting output in the U.S. economy. Second, when these firms expand in space, they generate local knowledge spillovers, as measured by external patent citations received from other firms in their new location. To establish this, we employ an event-study design, and show 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, as well as an increase in patenting activity by other firms in that market. Third, we show that these spillovers indeed increase with the size of the new R&D facility, but less than one-to-one, in line with our assumption of diminishing returns to size in the spillover function.

Estimating the key model elasticities, we find evidence that U.S. innovative firms operate

than the tradeoff between love-of-variety and business stealing.

[3]This result follows from the assumption of constant-elasticity spillovers and the absence of compensating differentials across regions, such as differences in amenities. See Glaeser and Gottlieb (2008); Kline and Moretti (2014); Fajgelbaum and Gaubert (2020).

in too few markets relative to the social optimum. Specifically, we identify the elasticity of local knowledge spillovers with respect to a firm's local footprint by examining how external patent citations in a given market respond to the cited firm's local employment. We identify firms' innovation elasticity from the responsiveness of their local patent output to their local employment. We find the spillover elasticity to be particularly low, which—through the lens of our model—implies that firms expand too little across space in equilibrium.

We utilize our estimated model to assess the potential benefits of incentivizing innovative firms to increase their spatial scope. Specifically, we explore a series of subsidies aimed at promoting geographic expansion and compare them to a standard R&D subsidy featured in the endogenous growth literature. This comparison allows us to evaluate the additional gains from geographical expansion beyond the typical effect of allocating more resources to innovation. We calibrate the size of all programs to 0.1% of GDP, which is approximately the cost of R&D support in the United States. A standard R&D subsidy of this magnitude results in a 0.15% increase in consumption-equivalent (CE) welfare relative to the baseline equilibrium. In contrast, subsidizing spatial expansion at the same cost yields higher gains of 0.19%-0.29% increase in CE welfare. We also find that depending on implementation, incentivizing spatial expansion might result in greater misallocation of labor across space. However, we find this effect to be quantitatively very small relative to the gains from increasing firms' spatial scope.

Incentivizing spatial expansion has distinct distributional implications compared to standard R&D subsidies. To illustrate this, we decompose our preferred welfare metric into two components: aggregate real income gains (in present discounted value) and a residual distributional effect that captures inequality across space, as well as inequality between innovation and production workers. All the policies we consider have a negative distributional effect, as they increase income more for innovation workers. However, the negative distributional effect is less pronounced when incentivizing spatial expansion. Specifically, while R&D subsidies tend to boost income in high-wage regions like Silicon Valley, exacerbating spatial disparities, certain variants of the expansion subsidies we examine can help reduce regional inequality.

Related literature. This papers relates to several strands of literature.

First, we contribute to the theoretical literature on endogenous growth, combining elements of both innovation-based and diffusion-based growth models. On the innovation side, we present a quality-ladder framework, building on creative destruction models following Aghion and Howitt (1992); Grossman and Helpman (1991); Klette and Kortum (2004); Akcigit and Kerr (2018). Relative to this literature, we allow innovative firms to operate across multiple local markets, and highlight the importance of these decisions for aggregate growth

in the presence of local spillovers. On the knowledge diffusion side, we relate to models of idea flows, in which agents improve their innovation process by learning from others, such as Perla and Tonetti (2014); Lucas and Moll (2014); Akcigit et al. (2018); Buera and Oberfield (2020); Prato (2024). This literature has also examined the role of knowledge diffusion in space as a source of long-run growth (Eckert and Peters, 2022; Cai et al., 2022b; Berkes et al., 2025), and the role of trade in diffusing knowledge across locations (Sampson, 2015; Santacreu, 2015; Cai et al., 2022a; Perla et al., 2021; Hsieh et al., 2023; Lind and Ramondo, 2023; Ayerst et al., 2023). Relative to this literature, we connect knowledge diffusion to location decisions by multi-plant firms, in an environment in which spillovers are local in nature. We highlight that in such a setting, the amount of knowledge diffusion in the economy depends on firms' geographical scope, and both innovation and expansion decisions can be sub-optimal in equilibrium.

Second, we relate to the literature on multi-location firms in spatial economics, including Argente et al. (2020), Oberfield et al. (2024), and Kleinman (2023); and the related literature on multinational firms in international trade, including Helpman (1984), Ramondo and Rodríguez-Clare (2013), Garetto (2013), Keller and Yeaple (2013), Tintelnot (2017), and Arkolakis et al. (2018). Relative to these two strands of literature, we emphasize that firms' geographic scope may be suboptimal from a planner's perspective when local plants generate spatial knowledge spillovers. In addition, our analysis focuses on the spatial distribution of firms' innovation activities, rather than on their production structure.[4]

Third, we relate to the literature that studies the optimal allocation of labor across space, e.g. Glaeser and Gottlieb (2008); Kline and Moretti (2014); Fajgelbaum and Gaubert (2020, 2025). In contrast to this literature, we focus on the optimal allocation of resources to firm expansion, which in our model takes the form of the allocation of labor between fixed and variable costs in each location. We highlight that for a given allocation of labor across space, firms can be too spatially dispersed or concentrated. We also 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, highlighting a new connection between the geographical scope of firms and aggregate growth and deriving new policy implications for the optimal geography of firms.

Empirically, we relate to the literature that investigates the local nature of spillovers between firms, documenting the role of technology clusters (Jaffe et al., 1993; Audretsch

[4]Although we do not model firms' production location decisions explicitly, such considerations are indirectly reflected in the idiosyncratic costs of establishing innovation facilities across locations. Interestingly, empirically, we find that a substantial share of innovation facilities in our data are not colocated with production sites.

and Feldman, 1996; Peri, 2005), and in particular the existence of spillovers on firms' productivity (Greenstone et al., 2010; Bloom et al., 2019; Giroud et al., 2024), and spillovers on innovation activity (Griffith et al., 2011; Carlino and Kerr, 2015; Berkes and Gaetani, 2020; Moretti, 2021; Matray, 2021); see Chatterji et al. (2014) and Kerr and Robert-Nicoud (2020) for a review. Our main empirical 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.

2 A theory of innovative firms with local spillovers

In this section, we introduce a model of innovation-driven endogenous growth with multi-location innovative firms and local knowledge spillovers. Growth is fueled by quality-enhancing innovations produced by firms that hire R&D workers, subject to decreasing returns to R&D labor within each local market. The productivity of R&D workers depends on local agglomeration spillovers, which are shaped by the equilibrium composition of innovative firms that establish a presence in each location. Firms internalize the private benefits of local knowledge when opening additional establishments but do not internalize the spillovers they generate for other local firms.

2.1 Environment

2.1.1 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 $\bar{L}_n$ of households that can supply labor to the production sector and a measure $\bar{H}_n$ of households that can supply labor to the innovation sector. For now, we abstract from labor mobility across space and across sectors, to highlight more clearly the new insights from our framework, but we relax this assumption later, including in the quantitative model in Section 4. 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 $s \in \{p, i\}$ (where p indicates

production and i innovation), are represented by the following utility function:

$$U_{ns0} = \int_0^\infty e^{-\rho t} \frac{(C_{nst})^{1-\phi} - 1}{1-\phi} \, dt,$$

where C_{nst} is the amount of final-good consumption in period t and ϕ is a parameter that governs the intertemporal elasticity of substitution.

2.1.2 Production Technology

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

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

Intermediate goods are produced by production labor using the following production function:

$$y_{jt} = A_{jt} \left(\sum_{n=1}^{N} (Z_n L_{jnt})^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{2}$$

where A_{jt} is the state-of-the-art technology in the production of good j, L_{jnt} is the mass of production workers from location n that participate in the production of j, Z_n is the relative productivity of production workers in n, and σ is the elasticity of substitution between production workers from different locations.

Growth in this economy occurs due to improvements in A_{jt}, keeping relative productivity across locations constant. The time-invariant local productivity component Z_n and the imperfect substitution across production workers from different locations allow for a realistic degree of heterogeneity in local employment and wages, but are not a key component of the analysis. In this framework, we intentionally keep the structure of the production sector simple and focus primarily on the organization of firms within the innovation sector. Specifically, we abstract from the question of firm boundaries and firms' spatial organization in the production sector.

2.1.3 Innovation

Growth results from improvements in intermediate technology A_{jt} due to investment in innovation, which produces new ideas. Innovation is undirected, and each new idea increases A_{jt} for some intermediate $j \in [0, 1]$ by a factor of γ. Ideas are produced by innovative

firms, with endogenous mass M_t and indexed by ω, which hire innovation-sector labor across different markets to produce ideas.

The arrival rate of ideas for a firm ω, $\lambda_{\omega t}$, depends on its employment of R&D labor, and is given by

$$\lambda_{\omega t} = \lambda_0 \sum_{n=1}^{N} K_{nt}(z_\omega)^{1-\eta}(\ell_{n\omega t})^{\eta}, \tag{3}$$

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 endogenous relative productivity of inventors in market n, which depends also on the activity of other firms in that area through local knowledge spillovers; η captures decreasing returns from employing more labor in location n; and λ_0 is a scaling constant. The assumption of decreasing returns at the market level can capture various forces, such as the challenge of hiring and supervising large research teams in the same location. Another possible micro-foundation is the differentiation of workers in terms of how many effective units of labor can they supply to each firm, or in terms of how they value firm-specific amenities. In these formulations, firms face an upward-sloping supply of labor in each location, generating diminishing labor efficiency in each local market. The assumption of $\eta < 1$, which we later verify in the data, is the reason that firms choose to spread their activity across space, despite their output (ideas) being completely tradable and non-rival.

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. These fixed costs discourage firms from fragmenting innovation activity across space, giving rise to increasing returns to scale when local innovation teams are small: innovation output per unit of local labor is non-monotonic—initially increasing with team size and eventually decreasing once local employment becomes sufficiently large. We impose a distributional assumption that makes aggregation particularly tractable. We assume that the inverse of $f_{n\omega t}$ is drawn from an i.i.d. Pareto distribution with shape parameter θ and scale f_{max}. This assumption also implies that a firm's marginal cost of expansion increases with the number of markets in which it operates. In the case of a continuous spatial environment, the distributional assumption is isomorphic to imposing an isoelastic "span-of-control" cost function in the mass of locations.

A new innovative firm can enter the market subject to the the payment of a fixed cost of $f_{e,t}$ units of the final good.[5] Upon entry, firms draw their firm-level productivity z_ω from

[5]Denominating the cost of entry in terms of the final good enables us to isolate the allocation of resources between firm expansion and R&D within the innovation sector, independent of the issue concerning the optimal size of the innovation sector. However, this approach necessitates the assumption that the cost of entry grows at the same rate as output, in order to maintain a well-behaved balanced-growth-path. In any

some distribution $\Psi(z)$.

The aggregate innovation rate in the economy, $\bar{\lambda}$, is equal to the total innovation efforts across all firms in the economy:

$$\bar{\lambda}_t = \int_{\omega \in \Omega} \lambda_{\omega t} d\omega, \tag{4}$$

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

2.1.4 Spillovers

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

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

where Ω_n 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; and β captures the degree to which spillovers from a particular firm depend on its presence and employment in market n. One natural interpretation for this structure is that regional innovation productivity scales with the mass of insights that a firm can gain from its peers. In this case, $z_\omega^{1-\beta} \ell_{n\omega t}^{\beta}$ is the arrival rate of insights from each firm ω, and α is the sensitivity of innovation productivity to insights from others. The arrival rate of insights from each firm ω is increasing with its local employment and with its overall efficiency, with their relative importance determined by β.

To get intuition for the implications of this structure, it is worth considering a few special cases. First, when $\beta = 1$, local innovation productivity is a power-function of total inventors in location n, which resembles standard specifications in the spatial economic literature, with α capturing the strength of agglomeration spillovers. Second, when $\beta = 0$, local innovation productivity is a function of the mass of locally active firms, weighted by their efficiency z_ω. In this case, it is sufficient for a firm ω to open a tiny local plant in n for all other firms in n to fully absorb all available insights from ω. When $\beta = \eta$, local innovation productivity is a function of local innovation output, and the contribution of each firm to local knowledge is proportional to its own innovation output. Finally, when firms are homogeneous in their innovation efficiency z_ω, local knowledge is a Cobb-Douglas function of the mass of locally-active firms and the mass of local inventors.

case, the choice of how to denominate the cost of entry is not consequential for the key new insights from our framework.

This completes the set-up of the physical environment as viewed from the perspective of a social planner. Table 1 provides a summary of the preferences, technology, and resource constraints in the economy.

Table 1: **Summary of the economic environment**

Preferences	
Utility function	$U_{nst} = \int_{h=t}^{\infty} e^{-\rho(h-t)} u\left(C_{nsh}\right) dh$
Production	
Production of the final good	$Y_t = \exp\left(\int_0^1 \ln y_{jt} dj\right)$
Production of intermediate goods	$y_{jt} = A_{jt}\left(\sum_{n=1}^{N}\left(Z_n L_{jnt}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$
Innovation	
Growth of A_{jt} upon successful idea	γ
Aggregate arrival rate of ideas	$\bar{\lambda}_t = \int_{\omega\in\Omega_t} \lambda_{\omega t} d\omega$
Production of ideas	$\lambda_{\omega t} = \sum_{n=1}^{N} K_{nt} z_\omega^{1-\eta} \ell_{\omega nt}^{\eta}$
Knowledge spillovers	$K_{nt} = \bar{K}_n\left(\int_{\omega\in\Omega_t} z_\omega^{1-\beta} \ell_{\omega nt}^{\beta} d\omega\right)^{\alpha}$
Innovation efficiency and fixed costs	$z_\omega \sim \Psi\left(z\right) \quad f_{\omega n}^{-1} \sim \text{Pareto}\left(\theta, f_{max}^{-1}\right)$
Resource constraints	
Production labor in location n	$\bar{L}_n = \int_0^1 L_{jnt} dj$
Innovation labor in location n	$\bar{H}_n = \int_{\omega\in\Omega_t}\left(\ell_{\omega nt} + f_{\omega n} \times \mathbb{I}\left\{\ell_{\omega nt} > 0\right\}\right) d\omega$
Output	$Y_t = \sum_{n=1}^{N}\sum_{s\in\{L,H\}} C_{nst} + \int_{\omega\in\Omega_t} f_{et} d\omega$

2.2 Equilibrium

In this section, we analyze the equilibrium solution of the model.

We examine a market structure characterized by perfect competition in final good production and labor markets, and monopoly in the intermediate goods sector, as commonly assumed in the growth literature. In each intermediate product line j, the producer is a monopolist that owns the rights to the frontier technology A_{jt}. The monopolist competes against a fringe of competitive firms, that can also produce intermediate j with an obsolete technology with quality A_{jt}/γ, where recall that γ is the step-size of technological progress in the economy.

We normalize the price of the aggregate good to 1. We denote the price of production

labor in market n by w_{npt}, the price of innovation labor in market n by w_{nit}, and the price of intermediate j by p_{jt}. We denote the profits from owning the production rights to a particular product line by π_{jt}. We denote the value of a new idea by V_{jt}.

We focus on a balanced growth path (BGP) along which aggregate productivity, defined by $\log \mathcal{A}_t = \int_0^1 \log A_{jt} dj$, consumption, all components of income, and the value of new innovations, all grow at a constant rate g. The spatial distribution of these objects is invariant along the BGP. We represent the equilibrium conditions for the economy after de-trending all variables, omitting time subscripts.

To determine how profits are allocated, we assume that innovation rents are distributed by a national fund to workers in proportion to their wages. In addition, households can save in a bond in zero net supply. The standard household maximization problem yields $\phi g = r - \rho$ along the balanced growth path, where r is the interest rate, which is constant along a BGP. [6]

2.2.1 Production sector solutions

The maximization problem of the final goods producer generates the demand for a particular intermediate j, given by $y_{jt} = \frac{Y_t}{p_{jt}}$. 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

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

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 profits are equal across all intermediate product lines j, so we denote $\pi_{jt} = \pi_t$.

The value of holding the rights to state-of-the-art technology in product-line j, denoted

[6] This assumption is motivated by similarity to the distribution of housing rents in the quantitative model. Alternatively, we could assume that households can save by investing in a share of a balanced portfolio of all firms in the economy, B_{nst}. The household's budget constraint would imply that $C_{nst} + \dot{B}_{nst} = w_{nst} + r_t B_{nst}$. This assumption would also yield a standard Euler equation along a BGP, $\phi g = r - \rho$. As another possibility, we could assume that all agents are hand-to-mouth, and allocate profits to an absentee capitalist that spends all income on the final consumption good. In this case, the interest rate r would be exogenous, but all the main results in the paper would still hold.

as V_{jt}, satisfies the following equation:

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

The interpretation of this equation is that the value of holding the rights to state-of-the-art technology is equal to the flow profits, π_{jt}, minus the risk of displacement due to a new innovation at rate $\bar{\lambda}$, captured by the last term. Along a balanced growth path, the value V_{jt} satisfies the following equation

$$V_{jt} = V_t = \frac{\bar{\pi}}{r - g + \bar{\lambda}} Y_t. \tag{6}$$

2.2.2 Innovation sector solutions

Upon a successful innovation, an innovative firm develops a new frontier technology on a random product line j which displaces the existing incumbent monopolist. The innovative firms sells the technology to an intermediate good producer extracting all the surplus. The price of a new idea when sold to a producer is equal in equilibrium to the value from holding the rights for an idea, V_t.[7]

Given the value of a new idea, 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:

$$\begin{aligned} \max_{\{\ell_{n\omega t}, \mathbb{I}_{n\omega t}\}_{n=1}^{N}} \quad & \lambda_{\omega t} V_t - \sum_{n=1}^{N} \mathbb{I}_{n\omega t} \times \left[w_{nit} \left(\ell_{n\omega t} + f_{n\omega t} \right) \right] \\ \text{s.t.} \quad & \lambda_{\omega t} = \lambda_0 \sum_{n=1}^{N} K_{nt} (z_\omega)^{1-\eta} (\ell_{n\omega t})^{\eta}, \end{aligned} \tag{7}$$

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

Solving this problem yields an expression for the mass of inventors chosen by firm ω with

[7] We assume that an innovative firm sells the technology to a random intermediate good producer. The identity of the buyer and the exact market structure in the market for ideas are irrelevant for the main results of the paper. Given that the value of an idea in equation (6) does not depend on A_{jt}, the return to an innovation is the same for any intermediate producer who would gain the market and displace the existing incumbent, independently of the current level of factor productivity in the intermediate. Only the existing incumbent has a lower valuation for the innovation because it cannibalizes the existing profits. However, a random match with the incumbent is a zero-probability event, given that there is an infinite number of intermediates.

productivity z in market n in period t, conditional on choosing to be active there[8]:

$$\ell_{n\omega t} = \ell_{nt}(z) = z\left(\frac{\eta K_{nt}\lambda_0 V_t}{w_{nit}}\right)^{\frac{1}{1-\eta}}.$$

We note that the optimal innovative employment conditional on activity is the same for all firms with productivity z, so we suppress the index ω and write $\ell_{nt}(z) = \ell_{n\omega t}$ for all ω such that $z_\omega = z$. We denote as $\psi_t(z)$ the CDF of innovative firms' productivity at time t. Given the optimal innovative employment conditional on activity in market n, the probability that a firm of type z chooses to be active in market n, denoted as $\chi_n(z)$, taking into account the distribution of fixed costs across firms, is the following:

$$\chi_{nt}(z) = f_{max}^{-\theta}\left(\left(\frac{1-\eta}{\eta}\right)\ell_{nt}(z_\omega)\right)^{\theta}.$$

It is also useful to define 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, which we denote $\bar{f}_n(z)$:

$$\bar{f}_{nt}(z) = \frac{\theta}{\theta+1}\left(\frac{1-\eta}{\eta}\ell_{nt}(z)\right).$$

We observe that the optimal innovative employment conditional on activity, the probability of entering a market and the expected fixed costs are higher in markets with higher regional inventors productivity, K_{nt}, and with lower wages, w_{nt}, and is increasing in firm's productivity, z.

Finally, we assume that fixed entry cost grows at the same rate of output, that is: $f_{e,t} = f_e e^{gt}$. As a result, we obtain the following free entry condition:

$$f_e = \int_z \sum_{n=1}^{N} \chi_n(z)\left[\lambda_0 K_n z^{1-\eta}\ell_n(z)^{\eta} V - w_{ni}\left(\ell_n(z) + \bar{f}_n(z)\right)\right] d\Psi(z). \tag{8}$$

This expressions states that, in equilibrium, the cost of entry for a new innovative firm is equal to the expected profits, averaged across the productivity type z, which is drawn after entry. In Appendix A we present a summary of all the equilibrium conditions.

2.2.3 Firm expansion, R&D, and aggregate growth in equilibrium

We now characterize the relationship between firm expansion and R&D in equilibrium. We omit the time subscripts, since allocations are time-invariant along the BGP. We begin by

[8] See Appendix A for derivations.

analyzing the forces that shape the allocation of innovative labor between R&D (variable costs, captured by $\ell_n(z)$ above) and firm expansion (fixed costs), described in the following proposition.

Proposition 1. ***Allocation to spatial expansion in equilibrium.***
Let $x_n(z) \equiv \chi_n(z)\bar{f}_n(z)$ and $y_n(z) \equiv \chi_n(z)\ell_n(z)$ denote the expected amount of region-n innovation labor allocated to expansion (paying fixed plant costs) and R&D activities (hiring inventors) for type-z firms, respectively. In equilibrium, their ratio is given by:

$$\frac{x_n(z)}{y_n(z)} = \frac{1-\eta}{\eta}\frac{\theta}{\theta+1}.$$

Proof of Proposition 1. See Appendix A.

Proposition 1 states that the relative allocation of innovation labor between expansion and R&D activities for firms of type z in market n is decreasing in η, decreasing in θ, and it is independent of firm type. When innovative firms choose whether to expand into a new market, they balance out two forces. On the one hand, the benefit of spreading out their R&D workers is higher when η is lower, indicating stronger decreasing returns to labor in each location in the R&D production function. On the other, spreading out R&D workers is associated with the fixed cost of running an additional establishment. The cost of expansion is summarized by θ, the shape parameter of the fixed costs distribution.

Next, we examine how the allocation of innovative activity in space affects the rate of economic growth. To do this, we first present the equilibrium growth rate along a BGP.

Proposition 2. ***Equilibrium growth.***
Along a BGP equilibrium, aggregate output grows at a constant rate g, given by

$$g = \bar{\lambda}\ln(\gamma) \tag{9}$$

where $\bar{\lambda}$ is the aggregate rate of creative destruction:

$$\bar{\lambda} = M\lambda_0 \int_z \sum_{n=1}^{N} K_n(\{\chi,\ell\})\,\chi_n(z)\,z^{1-\eta}\ell_n(z)^{\eta}\,d\Psi(z).$$

Proof of Proposition 2. See Appendix A.

The model preserves the standard result of Schumpeterian models that the aggregate growth rate depends on the aggregate rate of creative destruction, $\bar{\lambda}$, and the step size of innovation, γ. The novelty introduced in this framework is that the geography of innovative

firms affects the rate of creative destruction and aggregate growth, with a two-way feedback effect. On the one hand, the growth rate is affected by firms' propensity to locate in regions with higher R&D productivity, K_{nt}. On the other hand, local R&D productivity is endogenously determined in equilibrium as a result of firm's geographic expansion decisions.

How does an improvement in the ability of innovative firms to expand in space affect growth in equilibrium? We answer this question in the following proposition, by considering a shift in the scale of fixed costs, f_{max}, that lowers them across the board.

Proposition 3. ***Firm expansion and growth.***
The elasticity of the growth rate to the scale of the cost of expansion is given by:

$$\frac{\partial \ln g}{\partial \ln f_{max}} = \frac{(1-\eta)+\alpha(1-\beta)}{1+\theta}\left(\frac{\partial \ln M}{\partial \ln f_{max}} - \theta\right).$$

A decline in f_{max} raises growth if there are decreasing returns at the plant level ($\eta < 1$) or spillovers with diminishing learning ($\alpha > 0$ and $\beta < 1$). The effect holds for a given mass of firms (M), and is amplified under free entry $\left(\frac{\partial \ln M}{\partial \ln f_{max}} < 0\right)$.

Proof of Proposition 3. See Appendix A.

The proposition suggests that a technological reduction in the fixed costs of establishing innovation plants fosters growth through three channels. First, when $\eta < 1$, lower expansion costs enhance innovation efficiency by enabling firms to distribute their activities across more locations, which is beneficial under decreasing returns. Second, in the presence of knowledge spillovers ($\alpha > 0$) and diminishing learning from each firm ($\beta < 1$), expanding the number of locations per firm amplifies knowledge spillovers, thereby increasing local innovation productivity K_n. Third, these effects are further reinforced under free entry: greater profitability of innovation encourages more firms to enter the innovation sector, which in turn enhances innovation efficiency when there are decreasing returns in both idea production and the spillover function.

2.3 Social Welfare

We now turn to characterize the planner's allocation. The model is characterized by a few different externalities. First, the model features the externalities typical of Schumpeterian models, where 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.

Conditional on the allocation of resources to innovation, potential inefficiencies arise due to uninternalized knowledge spillovers between innovating firms. Firms fail to account for the fact that establishing a new facility in a given location generates knowledge spillovers that benefit other local firms. Additionally, they do not internalize that by drawing labor away from existing firms when opening a new plant, they diminish the knowledge spillovers from these firms, as spillovers depend positively on a firm's local size ($\beta > 0$).

To study the implications of these forces, we consider a planner that maximizes the aggregate present discounted value (PDV) of aggregate welfare. We allow the planner to freely redistribute ex-post through transfers of the final good, abstracting from any distributional concerns in our analysis.[9] We also restrict the planner to choose solutions that satisfy balanced growth.

The planner's problem aggregates the of each household in location n and sector s at time t with some non-negative weights $\{\{\nu_{nt}^s\}_{s\in\{p,i\}}\}_{n=1}^N$ and is given by:

$$\begin{aligned}
\max_{\left\{[\ell_n(z),\chi_n(z)]_z\right\}_{n=1}^N, M} \quad & \int_0^\infty e^{-\rho t} \sum_{n=1}^N \sum_{s=\{p,i\}} \nu_{nt}^s \frac{\left(C_{ns0} e^{gt}\right)^{1-\phi} - 1}{1-\phi}\, dt \qquad (10) \\
\text{s.t.} \quad & g = \ln(\gamma) M \int_z \sum_{n=1}^N \chi_n(z)\, K_n z^{1-\eta} \ell_n^\eta(z)\, d\Psi(z) \\
& K_n = \bar{K}_n \left(M \int_z \chi_n(z)\, z^{1-\beta} \ell_n(z)^\beta\, d\Psi(z) \right)^\alpha \\
& \bar{H}_n = M \int_z \chi_n(z) \left(\bar{f}_n(z) + \ell_n(z) \right) d\Psi(z) \\
& \bar{f}_n(z) = \frac{\theta}{\theta+1} f_{max} \left(\chi_n(z)\right)^{\frac{1}{\theta}}.
\end{aligned}$$

The Schumpeterian growth literature typically examines the optimal allocation of resources between production and innovation. Our model maintains the property of this class of models where innovation in equilibrium could be either too large or too small relative to the planner's solution, due to the opposing forces of the knowledge externality and the business stealing externality. However, for large enough values of the innovation step size γ, innovation is too low in equilibrium. Here we focus on the novel spatial expansion externality that emerges in our framework which may induce misallocation of resources within

[9] In appendix A, we show that the planner's objective can be expressed as a function of aggregate consumption.

the innovation sector between R&D and geographic expansion. The following proposition characterizes the planner's optimal allocation of R&D human capital between expansion and R&D.

Proposition 4. ***Optimal allocation to spatial expansion.***
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:

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

Proof of Proposition 4. See Appendix A.

Proposition 4 illustrates how the planner's relative expenditure on fixed plant costs diverges from that in equilibrium (Proposition 1). First, when there are no knowledge spillovers ($\alpha = 0$), both solutions are identical. Second, the two solutions may still coincide even when spillovers are present ($\alpha > 0$), but only if the elasticity of innovation output with respect to local labor (η) exactly matches the elasticity of spillovers with respect to local labor (β). When β is sufficiently low relative to η, firms in equilibrium under-invest in expansion compared to the planner's optimal solution. Conversely, when β is relatively high, firms over-invest in expansion. Intuitively, a low β suggests that, given a level of z, smaller plants contribute to spillovers as much as larger ones. In such cases, the planner would favor more smaller plants to boost innovation efficiency across the economy. Conversely, when β is high, spillovers scale strongly with plant size, and the planner would prefer fewer, larger plants. In the special case where $\beta = \eta$, firms' equilibrium decisions align exactly with the planner's optimal strategy.

The next proposition characterizes the optimal allocation across firm types.

Proposition 5. ***Optimal allocation across firm types.***
The mass of local innovation labor that the planner allocates to firms of type z is given by:

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

Proof of Proposition 5. See Appendix A.

This proposition states that the proportion of local innovation labor allocated to firms of type z is given by a weighted sum of its contributions to spillovers, $s_n^{\beta}(z)$, and direct innovation output, $s_n^{\eta}(z)$. When spillovers are stronger (α is large), the first term in the numerator gets larger, and the planner places more weight on firms that contribute strongly to spillovers. If β is low relative to η, the planner encourages more expansion of smaller firms, since smaller plants still contribute to spillovers.

Interestingly, whether firms in equilibrium over- or under-expand does not depend on the magnitude of spillovers, as captured by α. Instead, the direction of inefficiency depends solely on the relative magnitude of β compared to η. However, the extent of the inefficiency is naturally influenced by the importance of knowledge spillovers for innovating firms, i.e., the size of α. In what follows, we aim to estimate these parameters for the case of the innovation sector in the U.S. economy.

2.4 Labor mobility

Up until now we have considered the case in which labor is immobile. We now characterize the allocation of labor across space in the simple case of free labor mobility. We focus on the mobility of innovation labor.[10] In our full quantified model below, we consider a more realistic environment with limited mobility, as well as mobility between the production and innovation sectors.

We assume that there is an exogenous economy-wide endowment of $\bar{H}$ innovation workers. In equilibrium, free mobility implies the equalization of wages across space. For the central planner, this implies solving also for H_n in the optimization problem 10, subject to the constraint $\sum_{n=1}^{N} H_n = \bar{H}$. The next proposition characterizes the allocation of innovation labor across space.

Proposition 6. ***Allocation of innovation labor across space.***
Under free labor mobility, the share of innovation labor in market n, given by $H_n/\bar{H}$, is efficient. In both the decentralized equilibrium and the planner's solution, this share is given by

$$\frac{H_n}{\sum_{n'=1}^{N} H_{n'}} = \frac{\left(\bar{K}_n\right)^{\frac{1}{\frac{1-\eta}{1+\theta}-\alpha\frac{\beta+\theta}{1+\theta}}}}{\sum_{n'=1}^{N}\left(\bar{K}_{n'}\right)^{\frac{1}{\frac{1-\eta}{1+\theta}-\alpha\frac{\beta+\theta}{1+\theta}}}}.$$

Proof of Proposition 6. See Appendix A.

[10] The production sector features no spillovers, and the optimal allocation of labor across space in that sector is efficient.

The allocation of labor across space is efficient, despite the presence of knowledge spillovers. This result follows from two key assumptions: constant-elasticity spillovers and the absence of compensating differentials across regions (such as differences in amenities)—see Glaeser and Gottlieb (2008); Kline and Moretti (2014); Fajgelbaum and Gaubert (2020) for further discussion. In our quantified model, we relax this assumption by allowing for compensating differentials, in which case the spatial allocation of labor is no longer efficient. However, we find that the potential gains from reallocating labor across space are quantitatively small, consistent with findings elsewhere in the literature.

The results outlined in Section 2.3 hold also under free labor mobility. Notably, in this setting, the planner would like to alter firms' spatial scope without changing the allocation of labor across space. This implies to increase firms' spending on plant-level fixed costs relative to their spending on variable costs across all markets.

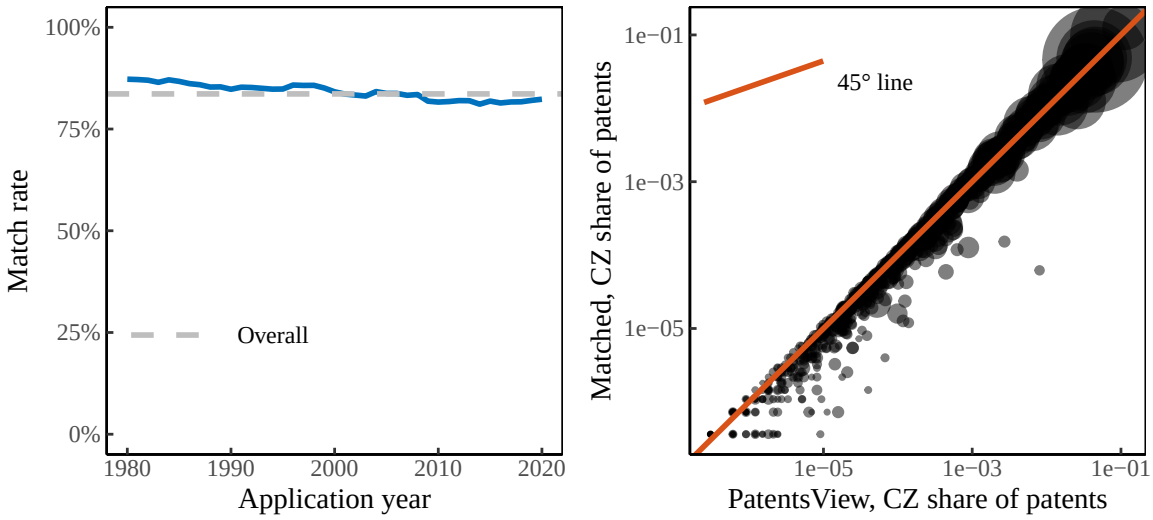
2.5 Extensions

In Appendix Section A.8, we explore alternative specifications for the spillover function and the cost of expansion. First, we demonstrate that our main results remain robust when accounting for decreasing returns to the number of firm locations, conditional on the labor employed, by introducing a span-of-control cost for the number of locations, as in Oberfield et al. (2024). In this setup, the curvature of the span-of-control cost replaces the role of the fixed cost dispersion parameter, θ. Second, we extend our analysis to a more general spillover function, allowing firms to learn not only from plants in the same location but also from those in other locations, while incorporating spatial frictions that hinder cross-location knowledge transfer. The main takeaways from our theoretical analysis remain the same with this specification.

3 Empirical Evidence

In this section, we present evidence supporting the main assumptions of our model and examine key moments that inform its estimation. First, we demonstrate that firms engaged in R&D activities across multiple markets contribute the majority of U.S. innovation output, underscoring the significance of analyzing their geographical structure. Second, we provide evidence of local spillovers from firm expansion, consistent with our assumption that a firm's local innovation efficiency (K_n) depends on the activity of other firms in the same market. Third, we show that these spillovers grow with the size of a firm's local employment, though less than one-to-one, aligning with our assumption that $\beta \in (0, 1)$.

Figure 2: Match quality across time and space.



Note: 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.

3.1 Data

Firm-level 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. Appendix B describes further details of our data.

Matching patents and firms' geography. We merge these two datasets following the methodology of Hughes et al. (2021). Table B.3 shows the results of this merge. The quality of the merge 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. Figure 2 shows this match rate is stable across time and space. In Appendix B.1 we provide additional details on the matching process and match quality. Encouragingly, a large majority of 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).[11]

Table 2 presents summary statistics of our merged data, pooled across years. The average firm in our data operates in 2 markets and generates 7 patents per year, while employing 11 inventors. We classify the industry segment of an expansion event by the looking at the

[11]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.

Table 2: **Summary statistics**

Firm-level statistics			Industry of expansion		
	Mean	Median	SIC	Description	Share (%)
Market	2.09	1.00	73	Business Services	13%
Establishments	4.62	1.00	50	Wholesale Trade Durable Goods	10%
Patents	6.93	1.00	35	Industrial and Commercial Machinery and Computer Equipment	8%
Inventors	10.74	2.00	87	Engineering, Accounting, Research, Management, and Related Services	8%
D&B employment	507.69	26.00	99	Nonclassifiable Establishments	7%
External citations	63.30	2.00	—	Other	55%

Note: Firm-level statistics from merged PatentsView-Dun & Bradstreet data, for firms with at least one patent and positive inventor employment. Data pooled across years. Industry codes reflect the two-digit SIC of largest establishments in expansion year.

two-digit SIC code of the largest establishment in the expansion year. We find our data selects largely for industries well-known for their innovation intensity.[12]

Regional data. Our analysis focuses on spillovers that emerge from firm expansion at the commuting zone level (our notion of a "market"). 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.

3.2 The centrality of multi-market innovation firms

We begin by documenting the dominant role of multi-market firms in driving innovation in the United States. Figure 3 displays the distribution of U.S. patents across firms, categorized by the number of R&D markets in which they operate. The figure shows both the aggregate pattern and snapshots from 1980 and 2015. We define an R&D market of a firm as a commuting zone in which a firm maintains both patenting activity and establishments in R&D-intensive sectors. Specifically, we identify patenting activity using inventor locations from the PatentsView database, and we classify R&D-intensive sectors according to the taxonomy in Table B.4 using establishment data from Dun & Bradstreet (see Appendix B.2

[12]We also track the extent to which manufacturing activity is co-located with innovation activity, an issue studied by Fort et al. (2020). Consistent with Fort et al. (2020), although we find manufacturing firms account for a large share of overall innovation in our sample, their decisions to co-locate production and innovation has atrophied over time, tracking the broader decline of manufacturing activity in the U.S. In 2015, only 45% of markets in which firms conducted innovation also recorded manufacturing activity.

Figure 3: Share of patents by firm scope

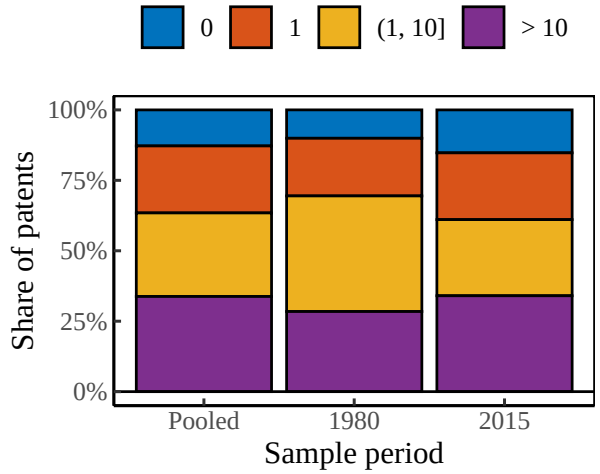


Table 3: Top companies by R&D markets

Company	Number of R&D markets
INTERNATIONAL BUS MCHS CORP	75
HONEYWELL INTERNATIONAL INC	71
ORACLE CORPORATION	59
LOCKHEED MARTIN CORPORATION	49
BOEING COMPANY	46
UNITED TECHNOLOGIES CORP	45
CISCO SYSTEMS INC	41
E I DU PONT DE NEMOURS & CO	41
ILLINOIS TOOL WORKS INC	41
PARKER-HANNIFIN CORPORATION	41
INTEL CORPORATION	38
AT&T INC	36
EATON CORPORATION	36
RAYTHEON COMPANY	35
MEDTRONIC INC	34
BECTON DICKINSON AND COMPANY	31
JOHNSON & JOHNSON	31
TEXAS INSTRUMENTS INCORPORATED	30
CATERPILLAR INC	29
HUBBELL INCORPORATED	28

Figure 3. Note: The figure displays the distribution of patents across firms categorized by their number of R&D markets (commuting zones), shown for pooled data and specific years (1980 and 2015). Firms are grouped into four categories based on their R&D market count: those with no markets (0), single markets (1), multiple markets (2-10), and extensive presence (> 10). The definition of an R&D market considers both the physical presence of inventors and establishments being classified under R&D-specific SIC codes (Table B.4).
Table 3. Note: The table presents the leading companies ranked by their number of R&D markets. For each company, the table displays the SIC code associated with their headquarters branch and total number of markets where they maintain research facilities.

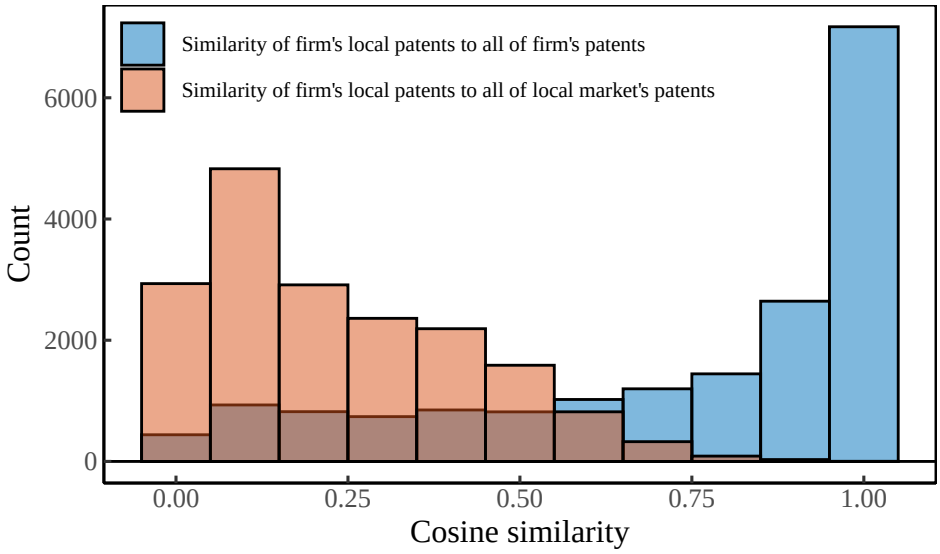
for further details on how we allocate inventors to firm establishments).

The data reveal that firms with multi-market R&D activity consistently account for approximately two-thirds of all patenting activity throughout the period. This pattern has remained remarkably stable over time, suggesting a persistent structural feature of the U.S. innovation landscape. A small share of patenting firms have zero R&D markets under our definition, as we do not observe establishments in R&D-intensive sectors for these firms in the D&B data. Relaxing this requirement would yield an even higher share of U.S. innovative output attributable to firms with multi-market R&D activity.

The significance of multi-market innovation becomes even more apparent when examining the top innovating firms. Table 3 lists the twenty firms with the largest R&D market footprints, revealing that industry leaders across various sectors maintain extensive networks of research facilities. IBM, for instance, maintains R&D operations across 75 distinct markets, while Honeywell and Johnson & Johnson operate in 71 and 31 markets, respectively.

We also examine the degree of differentiation in firms' innovation activities across their multiple markets. One possibility is that firms conduct distinct types of innovation in each location, potentially reflecting the comparative advantage of local markets in different innovation areas. Alternatively, firms may pursue similar innovation activities across all their locations. The data support the latter. To assess this, we compare the technological simi-

Figure 4: Similarity of Firms' Local Patents to Their National Portfolio vs. the Local Avg.



Note: This figure shows the distribution of cosine similarity between a firm's local innovation portfolio and (i) its national portfolio in other markets (red) and (ii) the local market's average across other firms (gray). Innovation portfolios are measured as vectors of patent counts across CPC classes. The sample includes firms with patenting activity in multiple markets.

larity of patenting in each firm-by-market unit to both the firm's national portfolio and the local market's average. We measure technological similarity using the cosine distance between the firm's local distribution of innovations across knowledge classes and (i) the firm's overall distribution and (ii) the local market's distribution. Figure 4 plots the histogram of these similarity measures across all firm-market pairs. We find that firms tend to pursue technologically similar innovations across their multiple markets, while differing more from other firms operating in the same locations.

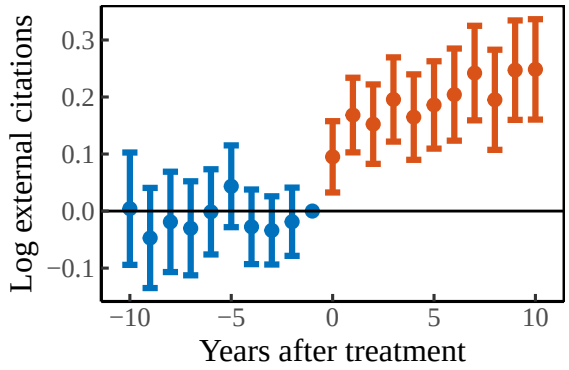
3.3 Evidence for spillovers from firm expansion

We now provide evidence that changes in firms' spatial scope generate knowledge spillovers in their new expansion markets, in line with our assumption on the structure of K_n. Specifically, we investigate: (a) does the expansion of some firm ω into some market n affect the citations of ω's patents by other firms in market n? (b) does it result in higher patenting efficiency for firms in market n, excluding ω?

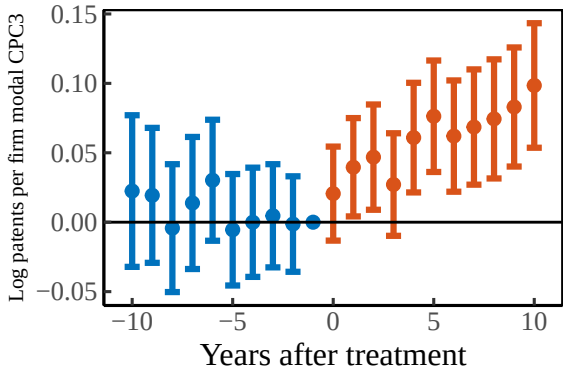
We adopt an event study specification, leveraging the staggered expansion of firms into new markets. Identification relies on a parallel trends assumption. The key challenge is that comparing expansions across different types of firms would likely violate the parallel trends assumption due to fundamentally different firm trajectories and market conditions. We

Figure 5: The effect of firm expansion on local citations and patenting efficiency.

(a) Response of citations received by firm patents in expansion commuting zone preceding and following expansion.



(b) Response of patenting efficiency in expansion commuting zone × expanding firm modal CPC3 preceding and following expansion.



Note: Each point represents a loading on the relative treatment time dummies specified in (11). We fail to reject the Wald test of jointy nullity of pre-treatment dummies for both panels (p-values of 0.49 and 0.22, respectively). Standard errors are clustered at the firm × commuting zone ($\omega \times n$) level.

therefore implement a within-firm matching approach. For example, we compare citations of IBM in a market to which it expanded to a market in which it did not expand, and not to citations of other firms. We match each treated commuting zone to never-treated commuting zones from the *same* firm, based on a distance criterion incorporating both pre-treatment outcomes and pre-treatment covariates. Appendix E discusses further details of the matching procedure.

Within-firm matching helps ensure we compare expansion events across similar market conditions and firm characteristics. 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 patents originating in both the treated and control units receive citations in two out of ten years of the pre-treatment period. To demonstrate robustness beyond our baseline matching approach, we also implement the synthetic difference-in-differences methodology of Arkhangelsky et al. (2021) and Ben-Michael et al. (2022), which provides a more systematic framework for control group construction—those results are reported in Appendix D.

We focus on two main outcome variables: (1) the log number of external citations received by firm ω in commuting zone n in year t and (2) the log of patenting efficiency in

the expanding firm's modal three-digit CPC classification,[13,14,15] and employ the following specification:

$$\log y_{\omega nt} = \alpha_{\tilde{g}_{\omega n}\omega n} + \delta_{\tilde{g}_{\omega n}t} + \sum_{\tau \in \{-10,\ldots,10\}\setminus\{-1\}} \beta_\tau \mathbb{I}\{t - g_{\omega n} = \tau\} + \varepsilon_{\omega nt}. \tag{11}$$

Recall the term $g_{\omega n}$ represents the year firm ω first enters commuting zone n. For treated units, $\tilde{g}_{\omega n}$ equals their own treatment year, $g_{\omega n}$. For control units, $\tilde{g}_{\omega n}$ equals the treatment year of their matched treated unit. The matching structure allows us to estimate relative time effects $\{\beta_\tau\}_\tau$ while controlling for cohort-unit fixed effects $\alpha_{\tilde{g}_{\omega n}\omega n}$ and cohort-time effects $\delta_{\tilde{g}_{\omega n}t}$.[16] In robustness exercises, we show that our results survive inclusion of a commuting zone-time fixed effect, ϑ_{nt}, as well as a cohort-firm-time fixed effect $\delta_{\tilde{g}_{\omega n}\omega t}$ (Appendix D.2).

Figure 5 presents our results. The left panel shows 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 25% higher in treated commuting zones compared to control commuting zones. The right panel shows that patenting efficiency in the expansion firm's modal three digit CPC is 10% higher, on average. Encouragingly, in both exercises there is no discernible pre-trend, and the Wald test of joint nullity of the pre-treatment dummies fails to be rejected. Appendix D shows our results are robust to numerous perturbations of the baseline specification.

[13]For firm ω in commuting zone n patenting efficiency at time t is calculated as the ratio of patents applied for (and subsequently granted) by firms in the expansion firm's primary technology class to the count of those same firms. Specifically, we divide the number of patents applied for by other firms in the expansion firm's modal three-digit CPC by the count of those firms in the commuting zone. We define a firm's modal three-digit CPC as the technology class in which it has accumulated the most patents. In Appendix D, we show results of our event study when using both the numerator (patenting) and the denominator (log firm count) as dependent variables.

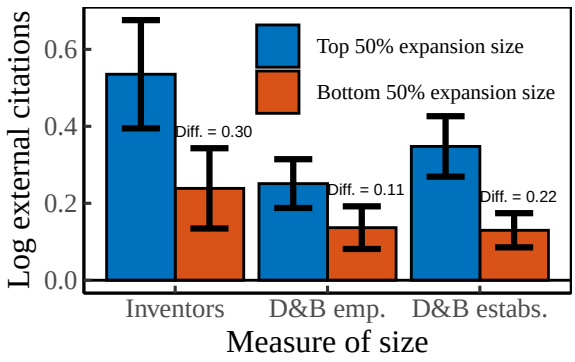
[14]In the synthetic control specifications (Figure D.4a), we will instead use a binary, indicating any positive citation activity as the dependent variable: $\mathbb{I}\{y_{\omega nt} > 0\}$. Figure D.4b shows our results when using this binary and the matched control sample.

[15]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." These are citations received from other firms $\omega' \neq \omega$ to patents of firm ω. We also consider only citations to ω's patents that were authored by inventors reporting a location $n' \neq n$, and made by patent documents whose originating firms never share inventors with ω. Appendix B.2 discusses how we allocate inventors to firms. 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 n. Figure D.1 shows the results are very similar when using looser definitions of external citations.

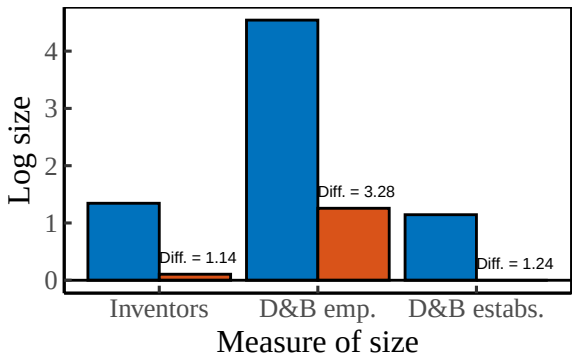
[16]This approach is a variant of the "stacking" estimator adopted in Cengiz et al. (2019). As discussed in Roth et al. (2023), Gardner (2022) shows that this framework estimates a convex weighted average of cohort-specific treatment effects under parallel trends and no anticipation.

Figure 6: Heterogeneity by size of local expansion.

(a) Estimation by expansion size splits



(b) Size of splits used in Panel A.



Note: The figure splits expansions based on their size as measured by inventor count, Dun & Bradstreet employment count, and establishment count in the first five years of the expansion and estimates loadings on partition-specific treatment dummies as well as giving the mean log size of each bucket. The difference between values of the blue and red bars is printed above the red bars. Standard errors for Panel A are clustered at the firm × commuting zone level.

3.4 Heterogeneity by the type of expanding establishment

In the model, the contribution of a firm to local knowledge K_n increases with its productivity and with its presence in market n, governed by the elasticity β. To see that this dimension of heterogeneity is meaningful in the data, we repeat specification (11), splitting the data by the size of the expanding firm in the new market, between the largest 50% of expansions and the smallest 50% of expansions, where recall that we still consider only within-firm comparisons.

Results from this estimation can be seen in Panel (A) of Figure 6, which summarizes the estimated effect of expansion on external citations by the size of the expansion. We split expansions by their size in the first five years of the expansion. We define size using three alternative metrics, using inventor counts, employment measures from D&B, and number of establishments that a firm opens in market n. Larger expansions systematically generate higher spillovers as measured by post-expansion citations of other firms to the expanding firm. This finding supports our model's prediction that spillovers increase in plant size, though with possibly diminishing returns. Indeed, in section 4, we will utilize these results to measure β, the elasticity regulating how strong spillovers depend on a firm's local presence.

Appendix D shows additional results for other dimensions of effect heterogeneity, by the industry of the expanding plant and by local patenting activity. We find that spillovers are higher, on average, in expansions that engage in local patenting in the first five years of firm expansion and are highest for industries known for high R&D intensity and complex technological processes.

4 Estimation

In this section, we introduce a quantitative version of the model to examine the following questions. First, do innovative firms operate in too few or too many markets—that is, is β low enough relative to η? Second, what are the potential welfare gains from policies that influence firms' spatial expansion, particularly in comparison to more conventional interventions such as R&D subsidies? Third, given the spatial heterogeneity in the incidence of such policies, what are their implications for regional inequality?

We estimate the model for the U.S. economy in 2019. We map regions in the model to U.S. labor market areas (LMAs - aggregations of commuting zones) and aggregate small LMAs to a single region, resulting in 300 local markets.

4.1 Quantified model

First, we enrich the model with additional components that help to discipline the baseline equilibrium, and which are standard in the economic geography literature.

Labor Mobility. Households choose both where to live and which sector to work in—either innovation or production. To preserve tractability, we assume logarithmic flow utility, $u(C_{nst}) = \ln(C_{nst})$. We focus on a balanced growth path equilibrium, and assume that agents choose a location and sector at time 0, taking into account that the economy is on a balanced growth path. Choices are subject to additive idiosyncratic preference shocks for locations and sectors, drawn from a Type-1 Extreme Value distribution. Under these assumptions, agents choose location and sector to maximize

$$\max_{n\in\{1,\ldots,N\},\, s\in\{p,i\}} \frac{1}{\rho}\log C_{ns} + \frac{g}{\rho^2} + \varepsilon_{ns}, \quad \varepsilon_{ns} \sim \text{Gumbel}\left(\frac{1}{\rho\epsilon}, \log B_{ns}^{\frac{1}{\rho\epsilon}}\right),$$

where $\frac{1}{\rho\epsilon}$ and $\log B_{ns}^{\frac{1}{\rho\epsilon}}$ are the scale and location parameters governing the distribution of preference shocks for sector s in location n, and C_{ns} is the detrended relative consumption if choosing sector s and location n.

Local housing. We now assume that agents consume both a national traded good and a local good in the form of housing, introducing an additional congestion force and enabling realistic variation in local prices. The consumption bundle is a Cobb-Douglas aggregator of tradable goods and housing, with expenditure shares ϖ and $1-\varpi$, respectively. Each location is endowed with one unit of housing, and the local rental rate is denoted by R_{nt}. We assume that all housing rents, like monopolistic profits, accrue to a national fund and are redistributed to households in proportion to their labor income. Letting ς denote the

aggregate ratio of total income to the aggregate wage bill, households' detrended relative consumption is given by

$$C_{ns} = \varsigma R_n^{-\varpi} w_{ns},$$

where R_n and w_{ns} are the detrended rental rate and wage, respectively.

Under the above assumptions, the share of agents choosing location n and sector s takes the standard choice probability form:

$$\frac{B_{ns}\left(R_n^{-\varpi} w_{ns}\right)^{\epsilon}}{\sum_{n'=1}^{N}\sum_{s'\in\{i,p\}} B_{n's'}\left(R_{n'}^{-\varpi} w_{n's'}\right)^{\epsilon}}.$$

Ex-ante welfare—i.e., expected utility before idiosyncratic preference shocks are realized—is given by

$$U = \frac{1}{\rho}\log\left(\sum_{n'=1}^{N}\sum_{s'\in\{i,p\}} B_{n's'}\left(R_{n'}^{-\varpi}\varsigma w_{n's'}\right)^{\epsilon}\right)^{\frac{1}{\epsilon}} + \frac{g}{\rho^2}. \tag{12}$$

4.2 Direct estimation of key parameters

We now describe how we estimate the key parameters that govern the novel aspects of the model: α, β, η, and θ.

4.2.1 Estimation of β

Recall that β measures the sensitivity of the spillovers generated by a firm ω to other firms in market n as a function of its size in that market. To discipline this parameter, we utilize data on citations across firms, assuming that bilateral citations serve as a reasonable proxy for knowledge transfers between firms.

We estimate β in two ways. First, we rely on an OLS regression implied by the model's structure. Recall our specification for the structure of spillovers, $K_{nt} = \bar{K}_n\left(\int_{\omega\in\Omega_{nt}} \iota_{\omega nt} d\omega\right)^{\alpha}$, where we denote $\iota_{\omega nt} \equiv z_{\omega}^{1-\beta}\ell_{\omega nt}^{\beta}$. $\iota_{\omega nt}$ is the contribution of firm ω to local knowledge or to local innovation efficiency, and it can be thought of as the arrival rate of insights from some locally active firm ω, such that $\int_{\omega\in\Omega_{nt}} \iota_{\omega nt} d\omega$ is the total expected number of insights that a firm can get from other firms active in market n. Integrating over the mass of active firms and adopting a discretized, log-linear approximation yields the following estimating equation:

$$\log \iota_{\omega nt} = \underbrace{\log|\Omega_{nt}|}_{n\times t \text{ FE}} + \underbrace{(1-\beta)\log z_{\omega}}_{\omega\times t \text{ FE}} + \beta \log \ell_{\omega nt} + \varepsilon_{\omega nt} \tag{13}$$

where we map model-based insights $\iota_{\omega nt}$ to realized external citations. We also include a firm-time fixed effect to account for firm-specific productivity, z_ω.

Results are shown in Panel A of Table 4. Our estimates for β are estimated with high precision and are broadly similar when using inventor counts or broader Dun & Bradstreet employment counts to measure $\ell_{\omega nt}$. We find that β is in the range of $(0.2, 0.3)$.

The above specification is valid under the assumptions of the model. However, it could be that a firm opens a large plant in a market precisely because it already receives more citations from that market, e.g. because it anticipates high synergies with existing firms. Therefore, we employ an alternative approach to estimate β that leverages our event-study design from section 3, relying on a parallel trends justification for identification. We use our estimates for the effects of firm expansion split by the size of the expansion in Figure 6. We estimate β from the difference in the change in external citations by size (left panel of Figure 6) relative to the gap in the size of the expansion (right panel of Figure 6), relying on the following approximation:

$$\beta = \frac{\partial \log \iota_{\omega nt}}{\partial \log \ell_{\omega nt}} \approx \frac{\bar{\beta}_\tau^{\text{Large}} - \bar{\beta}_\tau^{\text{Small}}}{\bar{\ell}^{\text{Large}} - \bar{\ell}^{\text{Small}}} \tag{14}$$

We find values for β ranging between close-to-zero and 0.3, reasssuring us of the general bounds of our estimate of β.

4.2.2 Estimation of α and η

Using the model structure, we are able to estimate a composite of, α, the overall importance of spillovers for innovation, and η, the decreasing returns in the production of ideas. We combine the production function for ideas, the structure of spillovers, and the firm's solution for local labor, and obtain an equation that connects a firm's local innovation output, $\lambda_{\omega nt}$, with the expected mass of insights that a firm gets from others in location n, $\mathcal{L}_{nt} \equiv \int_{\omega' \in \Omega_{nt}} z_{\omega' t}^{1-\beta} \ell_{\omega' nt}^{\beta} d\omega'$:

$$\log \lambda_{\omega nt} = \frac{\alpha}{1-\eta} \log \mathcal{L}_{nt} + \mathbf{X}'_{\omega nt} \boldsymbol{\gamma} + \varsigma_{\omega nt}. \tag{15}$$

We proxy the outcome variable – a firm's innovation output – using its patents. We construct the treatment variable, $\mathcal{L}_{nt}$, using data on firms' local employment and our estimate for β. $\mathbf{X}_{\omega nt}$ captures additional variables that we are able to control for, including firm productivity which is captured by firm-year fixed effects, and regional wages.[17] The error term, $\varsigma_{\omega nt}$, captures unobserved regional characteristics, $\bar{K}_{nt}$. Since we proxy for a

[17] We use the employment-weighted average wage in each market, drawing from the industries listed in Table B.4. Results are robust to using college-educated worker wages in these sectors from the Current Population Survey (CPS) and estimating (15) at the state-level.

firm's innovation output using only its patents, the error term also captures errors in the measurement of innovation.

We use different approaches to construct $\mathcal{L}_{nt}$. First, we use two distinct measures of employment: inventor counts from the patents data, and employment measures from the D&B data. Second, we need to take a stance on the weight of each firm, which depends in the model on its productivity z_ω. We thus employ two alternative approaches. First, we estimate the regression under an assumption of homogeneous firms, in which case z_ω integrates out as a constant. Second, we estimate the regression under the assumption that a firm's geographical scope scales with its productivity, as in the model. In practice, results from these different approaches are rather similar. In all cases, the error term $\varsigma_{\omega nt}$ might also include measurement error in $\mathcal{L}_{nt}$.

Equation (15) resembles a standard specification in the literature that estimates agglomeration spillovers, by connecting firms' local innovation efficiency to a composite measure of overall local activity, $\mathcal{L}_{nt}$. Diverging from this literature, the measure of local activity depends critically on the number and composition of local innovative firms, and not solely on local employment, and therefore on the value of β. In addition, the coefficient on local activity depends not only on the agglomeration spillover α, but also on the decreasing returns in the production of ideas, η, since improvements in regional productivity translate only partially to higher output in the presence of such decreasing returns.

Similar to the agglomeration literature, a key challenge is the identification of α and η in the presence of confounding, unobserved regional factors in the error term. An exogenous improvement in a location's patenting efficiency, captured by an increase in $\bar{K}_n$, would also lead to higher local innovation activity, $\mathcal{L}_{nt}$, generating simultaneity bias. An additional challenge is the above-mentioned potential measurement error in $\mathcal{L}_{nt}$.

We therefore adopt an instrumental variable strategy to address these concerns. We utilize the immigration instrument developed in Burchardi et al. (2020), which exploits quasi-random variation in local immigration flows based on the interaction of historical push-pull factors with recent immigration patterns. Intuitively, an exogenous influx of labor should increase the number of firms, shifting $\mathcal{L}_{nt}$ independently from $\bar{K}_n$. We report the results of this estimation in Panel B of Table 4. As robustness, we also estimate (15) at the state-level using state-level variation in the tax price of R&D. This approach follows Babina and Howell (2024), Bloom et al. (2013), and König et al. (2019) with data from Wilson (2009). These estimates are reported in Table F.3.

We show four different specifications that vary with our different measures of $\mathcal{L}_{nt}$ as described above. For all of them, the first-stage results indicate a strong relationship between predicted immigration flows and $\mathcal{L}_{nt}$, with F-statistics well above conventional thresholds

Table 4: Estimation results for β and $\alpha/(1-\eta)$.

	Panel A: Estimating β				Panel B: Estimating $\alpha/(1-\eta)$			
	Inventors		D&B employment		Inventors		D&B employment	
	Pooled	Bootstrap	Pooled	Bootstrap	Pooled	Bootstrap	Pooled	Bootstrap
No z-adjustment								
OLS	0.31	0.31	0.19	0.18	0.14	0.16	0.11	0.12
	(0.013)	(0.006)	(0.009)	(0.005)	(0.024)	(0.023)	(0.024)	(0.023)
TSLS	—	—	—	—	0.14	0.15	0.14	0.14
	—	—	—	—	(0.030)	(0.031)	(0.030)	(0.031)
N	98039	—	92972	—	98039	—	98028	—
KP F-stat.	—	—	—	—	10839	—	12099	—
z adjustment								
OLS	—	—	—	—	0.09	0.11	0.03	0.04
	—	—	—	—	(0.029)	(0.027)	(0.024)	(0.022)
TSLS	—	—	—	—	0.22	0.22	0.20	0.19
	—	—	—	—	(0.047)	(0.047)	(0.042)	(0.045)
N	—	—	—	—	98039	—	98028	—
KP F-stat.	—	—	—	—	7577	—	8274	—

Note: The table shows the results of estimating β and $\alpha/(1-\eta)$ using equations (13) and (15). See footnote 18 for details of the bootstrap. Standard errors are clustered at the firm-level. See the main text for discussion on z-adjustment variants of (15). We report first-stage estimates in Table F.1.

(Stock and Yogo, 2005). In all specifications, the TSLS estimates are in a similar range of around $(0.14, 0.22)$.[18]

4.2.3 Separating α and η

Table 4 yields joint estimates for $\alpha/(1-\eta)$, and we employ a few alternative approaches to separate them:

1. Draw η from the innovation literature: η is the elasticity of R&D output to R&D inputs, a key object in the innovation literature. Some attempts to estimate this elasticity include Kortum (1993); Terry et al. (2023); Burchardi et al. (2020); Acemoglu et al. (2016). The range of estimates from these studies is $\eta \in (0.35, 0.80)$, which implies $\alpha \in (0.04, 0.12)$.

2. OLS estimation of η: a projection of a firm's local R&D output (in terms of patents) on local labor—controlling for firm-year and market-year fixed effects—yields estimates in the range of $\eta \in (0.40, 0.95)$.

[18] To account for estimation uncertainty of β when estimating $\alpha/(1-\eta)$ we bootstrap equation (13) with 1,000 replicates, drawing with replacement from unique firms to maintain clustering. Each iteration uses the estimated $\beta^{(k)}$ and resampled data to run the IV regression (15). Appendix F reports the empirical distributions of $\{\beta^{(k)}\}_{k=1}^{1000}$ and $\{(\alpha/(1-\eta))^{(k)}\}_{k=1}^{1000}$. For pooled estimates, we use the full sample for $\{\beta^{(k)}\}_{k=1}^{1000}$ and $\bar{\beta} = \frac{1}{1000}\sum_{k=1}^{1000}\beta^{(k)}$ to calculate $\mathcal{L}_{nt}$ when estimating $\alpha/(1-\eta)$.

3. Draw α from the "spillovers" literature. A similar innovation efficiency spillover elasticity is estimated in Moretti (2021) using a combination of DiD and a movers-design. He reports $\alpha \in (0.07, 0.09)$. Relatedly, Greenstone et al. (2010) find a spillover elasticity of approximately 0.05 when assessing how local manufacturing TFP responds to the arrival of a million dollar plant. These values of α imply a range of $\eta \in (0.49, 0.72)$.

While the exact results depend on the specific approach taken, a clear takeaway is that the results align with existing estimates from the literature.

Comparing β and η. Despite the sensitivity of the estimates for β and η to the exact specification, we consistently find that $\beta << \eta$ across all estimations. Figure 7 shows various estimates for (β, η), fixing α. The shaded-red area captures our range of estimates for β. The shaded-blue area captures the range of estimates for η in the literature. The shaded-orange area captures their overlap. The individual points capture our attempts at estimating both, with the circle and the triangle capturing simple OLS regressions for both η and β, imposing the same sample and definition of local firm employment. Although there is wide dispersion in terms of the levels of β and η, one can have a high degree of confidence that the estimates for η exceed the estimates for β (i.e. all points and their error bands lie in the region above the 45° line) and that most estimates for (β, η) concord with the literature (i.e. fall within the shaded-orange region).

In our baseline specification, we impose $\eta = 0.50$ – a value common to the quantitative growth literature (Acemoglu et al., 2016, 2018; Akcigit and Ates, 2023) – and take a simple mean of our estimates for β and $\alpha/(1-\eta)$ – implying $\beta = 0.20$ and $\alpha = 0.09$.

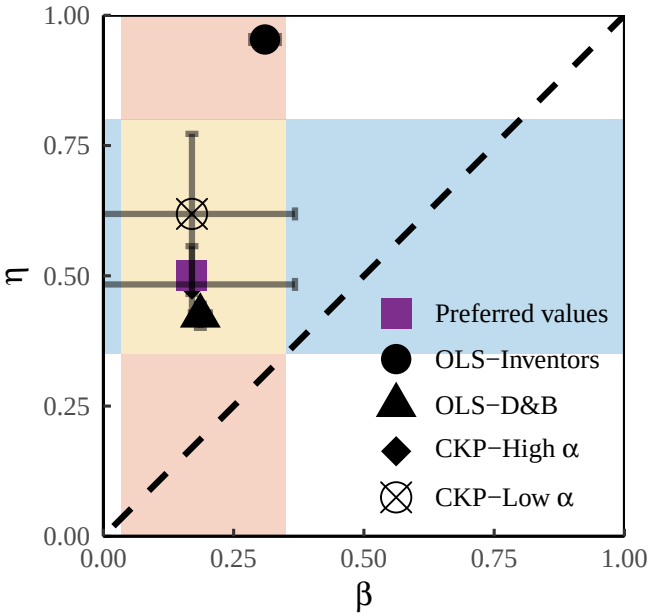
4.2.4 Estimating θ

The distribution for fixed costs admits the following log-linear relationship between the conditional probability of a type-z firm choosing to operate in market n and, conditional on operating, its optimal size, $\ell_n(z)$:

$$\log \chi_{nt}(z) = \theta \log \ell_{nt}(z) + \mathbf{X}'_{nt}(z)\boldsymbol{\gamma} + \varepsilon_{nt}(z). \tag{16}$$

We estimate this equation directly from the data. We bin firms into ten deciles based on their firm-wide employment as measured by inventor count and D&B employment to proxy z. We then regress the log of the share of firms in that decile that operate in a market n on the average log employment maintained by active establishments in that decile in market n along with market-year, market-decile, and decile-year fixed effects—$\mathbf{X}_{nt}(z)$. We estimate this regression, as we did those for β and $\alpha/(1-\eta)$, at the semi-decadal, LMA-level. Table F.2 shows the results. We find $\theta \approx 0.50$.

Figure 7: Parameter uncertainty for β and η.



Note: The width of the red box is the range of estimated β reported in Table 4 and Figure 6. The width of the blue box is the range of reported η in the literature—Acemoglu and Akcigit (2012) report $\eta = 0.35$ and Burchardi et al. (2020); Terry et al. (2023) report $\eta \approx 0.80$. The square purple marker corresponds to the quantitative value of η we use, which equals 0.50, as in Acemoglu et al. (2018). The "high" value of α comes from Table 2 of Moretti (2021), and the "low" value comes from Table 3 of Moretti (2021)

4.3 Remaining parameters

Externally Set Parameters. We calibrate the time discounting rate, ρ, to 0.05. The housing expenditure share is set to 0.25, consistent with evidence from Davis and Ortalo-Magné (2011). The dispersion of workers' preference shocks, ϵ, is fixed at 2.0, aligning with the estimated range for labor supply elasticity across locations and occupations, as documented in Galle et al. (2023). Similarly, the elasticity of substitution across locations in the production sector, σ, is set to 5.0, following standard values in trade and economic geography literature, such as those in Costinot and Rodríguez-Clare (2014).

Internally calibrated parameters. The remaining parameters we need to determine are the productivity step-size following new innovations, γ; the eoconomy-wide efficiency of innovation, λ_0; the scale of the fixed costs distribution, f_{max}; the cost of entry, f_e; and the efficiency advantage of large firms, $\bar{z}$. To determine these parameters, we match the following moments: the share of innovation labor in total wage bill (13% in the data), aggregate growth rate (2%), the average number of markets per innovating firm (2.09), the share of the top 10% of innovating firms in total patenting (70%) and the average employment in patenting

firms (70). We develop an efficient iterative algorithm to compute these parameters using a set of closed-form equations that link these 5 parameters and 5 moments, exactly matching all targeted moments. We provide further information on the computation of these moments and the estimation algorithm in Appendix G.

Table 5 summarizes the parameter values that are used for the quantified model.

Table 5: **Parameter values in the quantified model**

Parameter	Description	Value	Notes
	Externally set		
ρ	Discount rate	0.05	
σ	EoS across locations in production	5	Costinot and Rodríguez-Clare (2014)
ϖ	Housing share of consumption	0.25	Davis and Ortalo-Magné (2011)
ϵ	Dispersion of preference shocks	2	Galle et al. (2023)
	Direct estimation		
η	Elasticity of innovation to R&D inputs	0.5	See text
α	Elasticity of innovation to local spillovers	0.08	
β	Sensitivity of spillovers to local firm activity	0.17	
θ	Dispersion of fixed costs	0.5	
	Internal calibration		
γ	Step size for innovations	1.32	Target: share of innovation in total wage bill
λ_0	Economy-wide innovation efficiency	5.59	Target: growth rate
f_{max}	Scale of fixed costs	15×10^4	Target: average # of regions per firm
$\bar{z}$	Efficiency advantage of large firms	8.1	Target: share of top 10% firms in total patents
f_e	Cost of entry into innovation	275.9	Target: average size of patenting firms

5 Gains from incentivizing spatial expansion

We use our estimated model to assess the potential welfare gains of incentivizing innovative firms to expand into more markets. Specifically, we analyze different variants of a spatial expansion subsidy and compare them to a traditional R&D subsidy. A well-known feature of the endogenous growth framework used in our analysis is the generically inefficient allocation of resources to innovation in equilibrium. Since spatial expansion subsidies also stimulate innovation activity, they are expected to generate positive welfare gains, even in the absence of local spillovers across innovating firms. By benchmarking these policies against the traditional R&D subsidy, we can evaluate their additional benefits in the presence of local spillovers.

5.1 Aggregate gains

In our main analysis, we begin with the baseline equilibrium that we estimate and compare the impact of four different policy tools. First, we consider a standard R&D subsidy, as described above, which lowers firms' total wage bill, both for variable and fixed costs. Second, we explore a "spatial expansion subsidy" that is proportional to firms' expenditure on fixed plant costs, $w_{ni}f_{\omega n}$. This subsidy addresses the inefficiency identified in Proposition 4 by correcting the wedge between total spending on variable costs and spending on fixed costs. A uniform version of this subsidy would be first-best in the case of homogeneous firms and locations. However, two important aspects of this subsidy should be noted. First, by construction, it lowers the cost of expansion more significantly in locations with high innovation wages. Second, it is challenging to implement, as fixed costs (and expansion costs more generally) are often unobserved, especially when they vary randomly across firms.

The third policy tool we consider is a per-market expansion subsidy, which grants firms a transfer τ for each additional market in which they open an innovative plant. While this subsidy is relatively simple to implement, it falls short of being first-best. Although it encourages greater expansion, it also reallocates resources to less productive firms and markets. Finally, we evaluate a variant of this subsidy aimed exclusively at large firms, providing the top 10% of the largest firms with a subsidy for each new market in which they open an innovative plant.

To compare these policies, we choose the scale of each policy such that its total fiscal cost equals 0.1% of GDP, aligning with the estimated cost of U.S. innovation incentives such as the R&D tax credit. All subsidies are financed by a lump-sum tax on households. We then compute the new steady-state balanced growth path under each policy.

Our benchmark welfare metric is the expression for ex-ante expected utility from Equation 12.[19] This metric captures both the aggregate and distributional effects of each policy, reflecting our specification of agents' sectoral and locational preferences. For example, holding other factors constant, welfare may decline if a policy increases the dispersion of real income across space. To clarify these channels, we decompose the welfare metric into two components: one capturing aggregate real income gains, and the other reflecting distributional implications:

$$U = \underbrace{\frac{1}{\rho}\log\frac{Y}{R^{\varpi}} + \frac{g}{\rho^2}}_{\text{PDV of aggregate real income}} + \underbrace{\frac{1}{\rho}\log\left(\sum_{n=1}^{N}\sum_{s\in\{i,p\}} B_{ns}\left(S_{ns}^{Y}/S_{ns}^{R}\right)^{\epsilon}\right)^{\frac{1}{\epsilon}}}_{\text{Distributional effects}}, \tag{17}$$

[19] We adjust this equation by deducting the lump-sum tax from each agent's income.

where Y is total de-trended aggregate household income, and $R \equiv \left(\sum_{n',s'} L_{n's'} R_{n'}^{\varpi}\right)^{\frac{1}{\varpi}}$ is the aggregate price index of local housing services (so that R^{ϖ} serves as the aggregate consumer price index). The term $S_{ns}^{Y} \equiv \frac{L_{ns}Y_{ns}}{\sum_{n',s'} L_{n's'}Y_{n's'}}$ denotes the share of agents in location n, sector s in aggregate income, while $S_{ns}^{R} \equiv \frac{L_{ns}R_{n}^{\varpi}}{\sum_{n',s'} L_{n's'}R_{n'}^{\varpi}}$ captures the relative cost of housing faced by agents in location n, sector s.

The first term in the welfare expression corresponds to the present discounted value of aggregate real income. The second term captures distributional considerations: it is higher when the income share S_{ns}^{Y} is relatively large or the cost of living S_{ns}^{R} is relatively low in location-sector pairs that agents tend to prefer, as indicated by high amenity values B_{ns}. It is also higher when, other things equal, the dispersion in amenities-weighted real income across space and sectors is low.

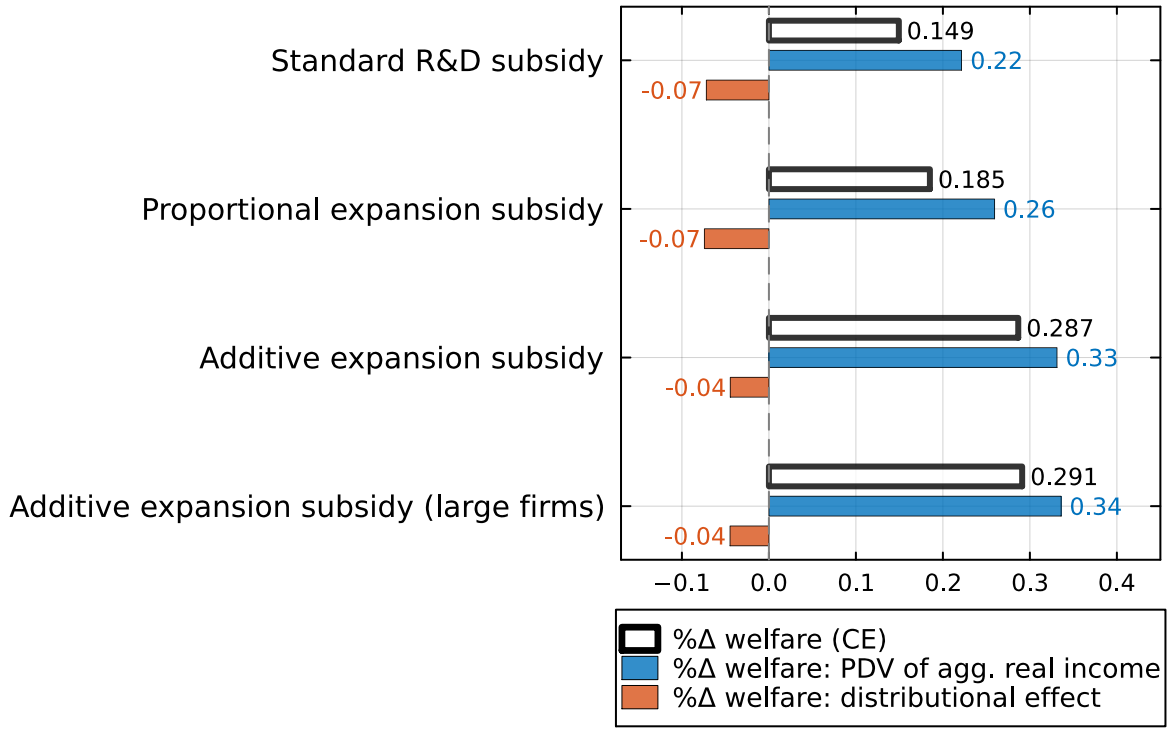
Results for the aggregate gains from each of our considered policies are shown in Figure 8. An RD subsidy equivalent to 0.1% of GDP raises consumption-equivalent welfare by 0.15%. This increase is driven by a 0.22% rise in the present discounted value (PDV) of aggregate real income. However, the distributional effect is negative, equivalent to a 0.07% welfare loss. This reflects the regressive nature of R&D subsidies, which tend to raise real income in the innovation sector and in relatively richer markets ex ante.

The gains from the proportional expansion subsidy are higher—by approximately 0.04 percentage points, or about a quarter more than the gains from the textbook R&D subsidy. Like the R&D subsidy, it reallocates more resources toward innovation. In addition, it expands the spatial scope of innovating firms. These larger gains are consistent with the results from Proposition 4 and our estimated values of η and β, and are quantitatively meaningful.

Turning to the additive expansion subsidies, these generate even higher aggregate gains and smaller negative distributional effects than the proportional expansion subsidy. The smaller distributional effect is to be expected, as these subsidies incentivize relatively more expansion—and thus increase demand for innovation labor—in ex-ante smaller and poorer regions. Perhaps more puzzling are the larger aggregate gains from the per-market (additive) subsidy compared to the proportional subsidy, given that the former may induce greater spatial misallocation. In our benchmark model, these gains arise because the per-market subsidy has a stronger effect on entry into innovation. Specifically, it acts as a form of insurance for low-productivity entrants, ensuring some profits even when the drawn productivity z is very low. Since entry into innovation is insufficient in equilibrium, this subsidy is particularly effective at stimulating additional entry for a given cost. Indeed, as discussed below, these extra gains from the per-market subsidy disappear when entry is not allowed.

To further highlight the economic forces at play in the model, we repeat the policy

Figure 8: Subsidizing R&D vs spatial expansion - aggregate implications



Note: This figure reports the percentage change in consumption-equivalent welfare and its components, as defined in Equation 17, under each of the four policy interventions discussed in Section 5. The white-outlined bars indicate the total change in consumption-equivalent welfare relative to the benchmark equilibrium. The blue bars represent the change in the present discounted value of aggregate real income, corresponding to the first term in Equation 17. The orange bars capture the distributional welfare effect, associated with the second term in Equation 17.

counterfactuals in restricted versions of the model that shut down various adjustment mechanisms. Results are shown in Figure 9. Each panel replicates the analysis from Figure 8 under a different restriction. The left panel considers a version of the model in which we shut down free entry, holding the mass of firms fixed at its baseline equilibrium level. The middle panel shuts down spatial labor reallocation, keeping total employment in each market fixed at its baseline level. The right panel shuts down both entry and labor reallocation, across sectors and across space.

Shutting down the entry of new firms (left panel of Figure 9) and fixing the mass of firms at its baseline equilibrium level leads to smaller increases in the present discounted value of aggregate real income across all policies. This result is intuitive, as the restriction limits the reallocation of resources into innovation. Consistent with the earlier discussion, the reduction in gains is particularly pronounced under the additive expansion subsidies,

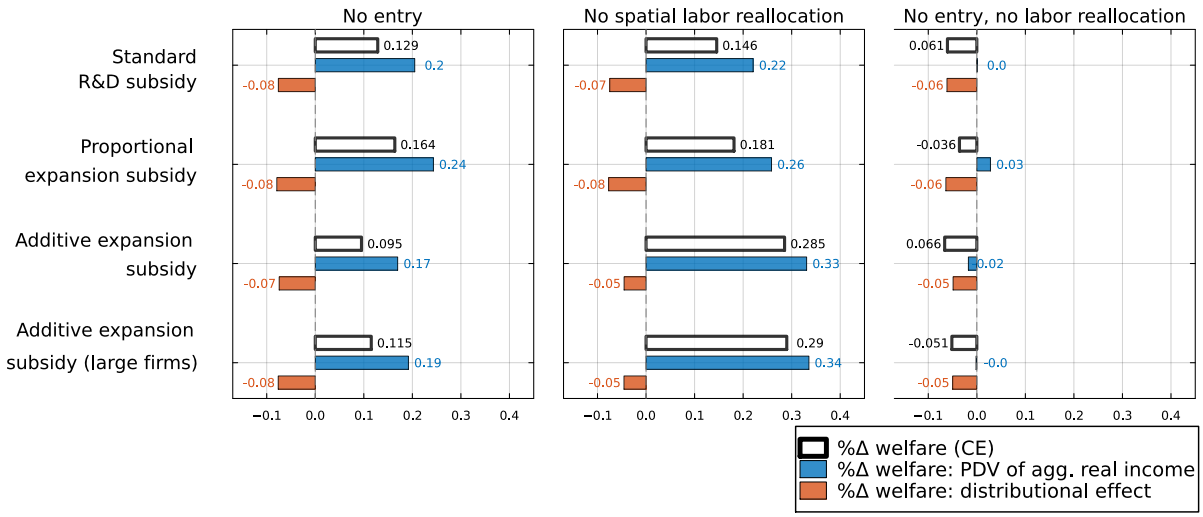
which are characterized by stimulating firm entry into innovative activities. In this version of the model, the proportional expansion subsidy emerges as the superior policy, generating approximately 0.04% higher welfare gains relative to the standard R&D subsidy. This advantage arises because the gains from broader spatial expansion—via enhanced knowledge spillovers across innovating firms—remain in place even when firm entry is restricted.

Restricting labor reallocation across space (middle panel of Figure 9) has minimal impact on our baseline results, and the ranking of the different policy interventions remains unchanged. Recall that in our quantified model, labor may be misallocated across space in the baseline equilibrium, and any policy could potentially improve or worsen this margin of efficiency. Nevertheless, we find that the welfare effects of all policies through changes in the spatial allocation of labor are quantitatively small. In particular, these effects are marginal compared to the welfare gains from reallocating resources toward innovation, and from expanding the spatial scope of innovating firms.

Finally, in the version of the model in which we shut down firm entry and shut down any labor reallocation across space or across sectors (right panel of Figure 9), there are no aggregate gains from the standard R&D subsidy, since resources in the innovation sector are held fixed. Nevertheless, the proportional expansion subsidy still raises the present discounted value of aggregate real income by 0.03%, due to its positive effect on local knowledge spillovers across innovating firms. This gain is approximately the same as the difference between this subsidy and the standard R&D subsidy in our baseline model, providing further evidence of the importance of this channel. Note that under all policies, the negative distributional effects remain similar to those in the baseline equilibrium, so that the total effect on welfare is negative. However, a planner could achieve a Pareto improvement by combining the proportional expansion subsidy with a redistribution policy. Also note that in this restricted version of the model, the additive expansion subsidy does not generate positive aggregate gains, even when targeted only at large firms (to avoid misallocation across firm types). This is due to overshooting: the subsidy leads to excessive spatial expansion relative to the optimum. Indeed, we find that a more modest version of the subsidy yields a positive effect on the present discounted value of aggregate real income.

To conclude, subsidies designed to stimulate spatial expansion generally outperform simple R&D subsidies. The optimal form of expansion subsidy depends on how elastic is firm entry and the extent to which labor can be reallocated toward innovation. While more complex to implement, a proportional expansion subsidy consistently proves more effective than a standard R&D subsidy in the scenarios considered. Additionally, a moderate per-market subsidy can generate positive outcomes, even when firm entry is inelastic, particularly when targeted at larger firms.

Figure 9: Welfare implications in different models



Note: This figure repeats the results from Figure 8 for restricted versions of the model. Each panel reports the percentage change in consumption-equivalent welfare and its components, as defined in Equation 17, under each of the four policy interventions discussed in Section 5. The white-outlined bars indicate the total change in consumption-equivalent welfare relative to the benchmark equilibrium. The blue bars represent the change in the present discounted value of aggregate real income, corresponding to the first term in Equation 17. The orange bars capture the distributional welfare effect, associated with the second term in Equation 17. The left panel considers a version of our baseline model in which we shut down free entry, holding the mass of firms fixed at its baseline equilibrium level. The middle panel shuts down spatial labor reallocation, keeping total employment in each market fixed at its baseline level. The right panel shuts down both entry and labor reallocation, across sectors and across space.

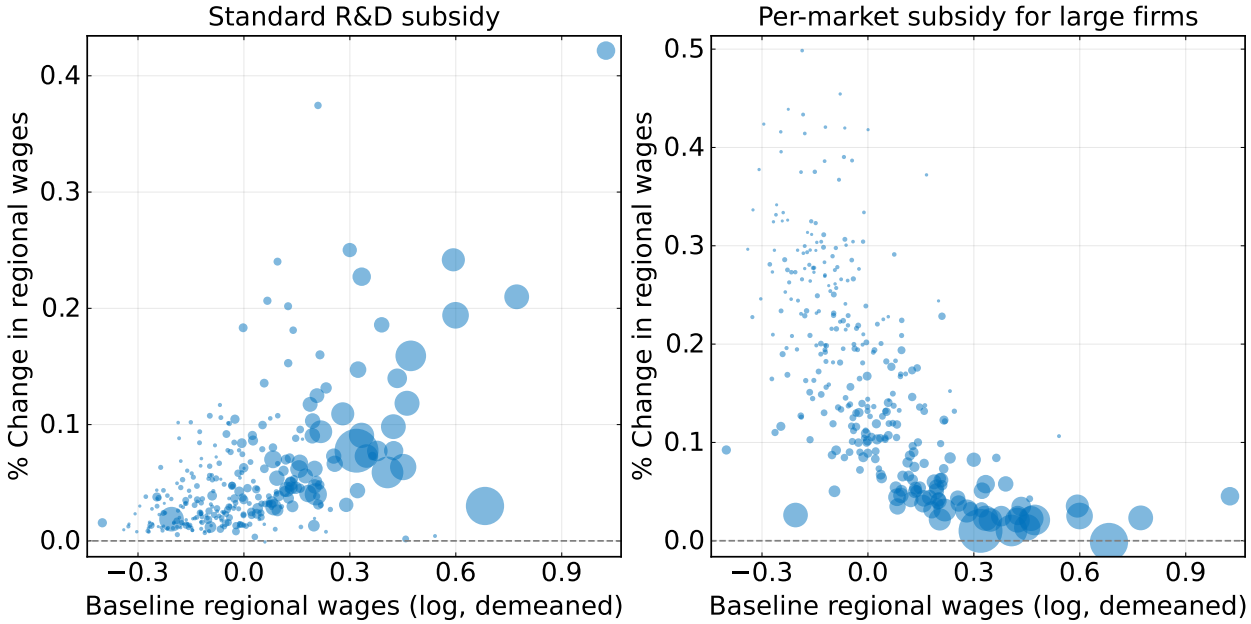
5.2 Regional implications

Beyond the aggregate gains, the model also predicts significant spatial heterogeneity in response to the above policies. A useful benchmark for comparison is again the standard R&D subsidy, which tends to exacerbate spatial inequality. This occurs because the subsidy primarily benefits locations that already specialize in innovation, which are characterized by higher wages in the baseline equilibrium. To illustrate this, the left panel of Figure 10 shows the percent change in regional wages between the baseline equilibrium and the scenario with an R&D subsidy, plotted against the baseline (demeaned) log of regional wages. Each circle represents a U.S. labor market area, with the circle size reflecting the area's population. The results show that wage gains are higher in initially high-wage regions. A similar pattern emerges with the proportional expansion subsidy: although it generates positive aggregate gains, these benefits are concentrated in higher-wage markets, which experience the greatest reduction in the cost of expansion.

In contrast, the per-market expansion subsidy generates a greater incentive to expand

into less profitable markets. The right panel of Figure 10 demonstrates this for the case of a per-market subsidy that is aimed at large firms. Locations with ex-ante lower wages see greater wage growth following the imposition of the subsidy, which results from higher growth of local innovation activity. Therefore, this variant of subsidy is able to both raise growth and reduce spatial disparities in the baseline model.

Figure 10: Subsidizing R&D vs spatial expansion - regional implications



Note: This figure shows the regional implications of the subsidies considered in Section 5. The left panel plots the percentage change in regional wages from implementing the standard R&D subsidy, relative to the baseline equilibrium, against the log of average regional wages in the baseline. Each circle represents a local labor market area, weighted by population. The right panel replicates the left panel for the case of the additive expansion subsidy targeted at the top 10% of firms.

6 Conclusion

When knowledge spillovers are local, the geographic reach of firms influences aggregate innovation and productivity growth and may be suboptimal from a social perspective. We examine this idea by developing an endogenous growth model with multi-location innovative firms and local knowledge spillovers. Our analysis shows that innovative firms may not expand into enough markets in equilibrium when the sensitivity of local spillovers to firm employment is sufficiently low. Using data on firms' R&D locations, patents, and citation networks, we estimate model parameters consistent with the under-expansion of innovative firms across locations. Model counterfactuals suggest that modest financial incentives

to encourage further geographical expansion of R&D activity could increase growth and, depending on implementation, reduce spatial inequality.

References

Acemoglu, Daron and Ufuk Akcigit, "Intellectual Property Rights Policy, Competition and Innovation," *Journal of the European Economic Association*, 2012, *10* (1), 1–42.

_ , _ , Douglas Hanley, and William Kerr, "Transition to Clean Technology," *Journal of Political Economy*, 2016, *124* (1), 52–104.

_ , _ , Harun Alp, Nicholas Bloom, and William Kerr, "Innovation, reallocation, and growth," *American Economic Review*, 2018, *108* (11), 3450–3491.

Aghion, Philippe and Peter Howitt, "A Model of Growth Through Creative Destruction," *Econometrica*, 1992, *60* (2), 323–351.

Akcigit, Ufuk and Sina T. Ates, "What Happened to US Business Dynamism?," *Journal of Political Economy*, 2023, *131* (8), 2059–2124.

_ and William R. Kerr, "Growth through Heterogeneous Innovations," *Journal of Political Economy*, 2018, *126* (4), 1374–1443.

_ , Santiago Caicedo, Ernest Miguelez, Stefanie Stantcheva, and Valerio Sterzi, "Dancing with the Stars: Innovation Through Interactions," Working Paper 24466, National Bureau of Economic Research March 2018.

Argente, David, Sara Moreira, Ezra Oberfield, and Venky Venkateswaran, "Scalable Expertise," Manuscript 2020.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager, "Synthetic Difference-in-Differences," *American Economic Review*, 2021, *111* (12), 4088–4118.

Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple, "Innovation and production in the global economy," *American Economic Review*, 2018, *108* (8), 2128–73.

Audretsch, David B and Maryann P Feldman, "R&D spillovers and the geography of innovation and production," *The American economic review*, 1996, *86* (3), 630–640.

Ayerst, Stephen, Faisal Ibrahim, Gaelan MacKenzie, and Swapnika Rachapalli, "Trade and diffusion of embodied technology: An empirical analysis," *Journal of Monetary Economics*, 2023, *137*, 128–145.

Babina, Tania and Sabrina T. Howell, "Entrepreneurial Spillovers from Corporate R&D," *Journal of Labor Economics*, 2024, *42* (2), 469–509.

Barnatchez, Keith, Leland D. Crane, and Ryan A. Decker, "An Assessment of the National Establishment Time Series (NETS) Database," *Federal Reserve Board of Governors Working Paper*, 2017.

Ben-Michael, Eli, Avi Feller, and Jesse Rothstein, "Synthetic Controls with Staggered Adoption," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 2022, *84* (2), 351–381.

Bergé, Laurent, "Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm," Technical Report 13 2018.

Berkes, Enrico and Ruben Gaetani, "The Geography of Unconventional Innovation," *The Economic Journal*, 09 2020, *131* (636), 1466–1514.

_ , _ , and Martí Mestieri, "Technological Waves, Knowledge Diffusion, and Local Growth," *Journal of Political Economy Macroeconomics*, 2025, *3* (1), 75–121.

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

_ , Mark Schankerman, and John Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

Buera, Francisco J. and Ezra Oberfield, "The Global Diffusion of Ideas," *Econometrica*, 2020, *88* (1), 83–114.

Burchardi, Konrad B., Thomas Chaney, Tarek Alexander Hassan, Lisa Tarquinio, and Stephen J. Terry, "Immigration, Innovation, and Growth," Technical Report 27075, National Bureau of Economic Research May 2020.

Cai, Jie, Nan Li, and Ana Maria Santacreu, "Knowledge Diffusion, Trade, and Innovation across Countries and Sectors," *American Economic Journal: Macroeconomics*, 2022, *14* (1), 104–45.

Cai, Sheng, Lorenzo Caliendo, Fernando Parro, and Wei Xiang, "Mechanics of Spatial Growth," Working Paper 30579, National Bureau of Economic Research October 2022.

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

Carlino, Gerald and William R. Kerr, "Chapter 6 - Agglomeration and Innovation," in "Handbook of Regional and Urban Economics," Vol. 5, Elsevier, 2015, pp. 349–404.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The Effect of Minimum Wages on Low-Wage Jobs," *The Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

Chatterji, Aaron, Edward Glaeser, and William Kerr, "Clusters of Entrepreneurship and Innovation," *Innovation Policy and the Economy*, 2014, *14*, 129–166.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin, "ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects," *The Stata Journal*, 2020, *20* (1), 95–115.

Costinot, Arnaud and Andrés Rodríguez-Clare, "Trade theory with numbers: Quantifying the consequences of globalization," in "Handbook of international economics," Vol. 4, Elsevier, 2014, pp. 197–261.

Crews, Levi Garrett, "A dynamic spatial knowledge economy." PhD dissertation, The University of Chicago 2023.

Davis, Morris A and François Ortalo-Magné, "Household expenditures, wages, rents," *Review of Economic Dynamics*, 2011, *14* (2), 248–261.

Dixit, Avinash K and Joseph E Stiglitz, "Monopolistic competition and optimum product diversity," *The American economic review*, 1977, *67* (3), 297–308.

Eckert, Fabian and Michael Peters, "Spatial Structural Change," Working Paper 30489, National Bureau of Economic Research September 2022.

Fajgelbaum, Pablo D and Cecile Gaubert, "Optimal spatial policies, geography, and sorting," *The Quarterly Journal of Economics*, 2020, *135* (2), 959–1036.

_ **and** _ , "Place-Based Policies: Lessons from Theory," Working Paper 33517, National Bureau of Economic Research February 2025.

Fort, Teresa C, Wolfgang Keller, Peter K Schott, Stephen Yeaple, and Nikolas Zolas, "Colocation of Production and Innovation: Evidence from the United States," 2020.

Galle, Simon, Andrés Rodríguez-Clare, and Moises Yi, "Slicing the pie: Quantifying the aggregate and distributional effects of trade," *The Review of Economic Studies*, 2023, *90* (1), 331–375.

Gardner, John, "Two-stage differences in differences," Technical Report, arXiv July 2022. arXiv:2207.05943.

Garetto, Stefania, "Input Sourcing and Multinational Production," *American Economic Journal: Macroeconomics*, April 2013, *5* (2), 118–51.

Giroud, Xavier, Ernest Liu, and Holger Mueller, "Innovation Spillovers across US Tech Clusters," Technical Report, National Bureau of Economic Research 2024.

Glaeser, Edward L and Joshua D Gottlieb, "The economics of place-making policies," Technical Report, National Bureau of Economic Research 2008.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings," *Journal of political economy*, 2010, *118* (3), 536–598.

Griffith, Rachel, Sokbae Lee, and John Van Reenen, "Is distance dying at last? Falling home bias in fixed-effects models of patent citations," *Quantitative Economics*, 2011, *2* (2), 211–249.

Grossman, Gene M. and Elhanan Helpman, "Quality Ladders in the Theory of Growth," *The Review of Economic Studies*, 1991, *58* (1), 43–61.

Helpman, Elhanan, "A simple theory of international trade with multinational corporations," *Journal of political economy*, 1984, *92* (3), 451–471.

Hsieh, Chang-Tai, Peter J. Klenow, and Ishan Nath, "A Global View of Creative Destruction," *Journal of Political Economy Macroeconomics*, 2023, *1* (2), 243–275.

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

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

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *The Quarterly Journal of Economics*, 1993, *108* (3), 577–598. Publisher: Oxford University Press.

Keller, Wolfgang and Stephen Ross Yeaple, "The Gravity of Knowledge," *American Economic Review*, June 2013, *103* (4), 1414–44.

Kerr, William R. and Frederic Robert-Nicoud, "Tech Clusters," *Journal of Economic Perspectives*, 2020, *34* (3), 50–76.

Kleinman, Benny, "Wage Inequality and the Spatial Expansion of Firms," *Working Paper*, 2023.

Klette, Tor Jakob and Samuel Kortum, "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 2004, *112* (5), 986–1018.

Kline, Patrick and Enrico Moretti, "Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority," *The Quarterly journal of economics*, 2014, *129* (1), 275–331.

Kortum, Samuel, "Equilibrium R&D and the Patent–R&D Ratio: U.S. Evidence," *The American Economic Review*, 1993, *83* (2), 450–457.

König, Michael D., Xiaodong Liu, and Yves Zenou, "R&D Networks: Theory, Empirics, and Policy Implications," *The Review of Economics and Statistics*, 2019, *101* (3), 476–491.

Lhuillier, Hugo, "Should I Stay or Should I Grow?," 2023. Working Paper.

Lind, Nelson and Natalia Ramondo, "Global Innovation and Knowledge Diffusion," *American Economic Review: Insights*, 2023, *5* (4), 494–510.

Lucas, Robert E. and Benjamin Moll, "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 2014, *122* (1), 1–51.

Martellini, Paolo, "Local labor markets and aggregate productivity," Technical Report, Working paper 2022.

Matray, Adrien, "The local innovation spillovers of listed firms," *Journal of Financial Economics*, 2021, *141* (2), 395–412.

Moretti, Enrico, "The Effect of High-Tech Clusters on the Productivity of Top Inventors," *American Economic Review*, 2021, *111* (10), 3328–3375.

Oberfield, Ezra, Esteban Rossi-Hansberg, Pierre-Daniel Sarte, and Nicholas Trachter, "Plants in space," *Journal of Political Economy*, 2024, *132* (3), 000–000.

Peri, Giovanni, "Determinants of Knowledge Flows and Their Effect on Innovation," *The Review of Economics and Statistics*, 2005, *87* (2), 308–322.

Perla, Jesse and Christopher Tonetti, "Equilibrium Imitation and Growth," *Journal of Political Economy*, 2014, *122* (1), 52–76.

_ , _ , and Michael E. Waugh, "Equilibrium Technology Diffusion, Trade, and Growth," *American Economic Review*, 2021, *111* (1), 73–128.

Prato, Marta, "The Global Race for Talent: Brain Drain, Knowledge Transfer, and Growth," *The Quarterly Journal of Economics*, 2024, *140* (1), 165–238.

Ramondo, Natalia and Andrés Rodríguez-Clare, "Trade, multinational production, and the gains from openness," *Journal of Political Economy*, 2013, *121* (2), 273–322.

Roth, Jonathan, Pedro H. C. Sant'Anna, Alyssa Bilinski, and John Poe, "What's trending in difference-in-differences? A synthesis of the recent econometrics literature," *Journal of Econometrics*, 2023, *235* (2), 2218–2244.

Sampson, Thomas, "Dynamic Selection: An Idea Flows Theory of Entry, Trade, and Growth," *The Quarterly Journal of Economics*, 2015, *131* (1), 315–380.

Santacreu, Ana Maria, "Innovation, diffusion, and trade: Theory and measurement," *Journal of Monetary Economics*, 2015, *75*, 1–20.

Scherer, F. M and Dietmar Harhoff, "Technology policy for a world of skew-distributed outcomes," *Research Policy*, April 2000, *29* (4), 559–566.

Stock, James H. and Motohiro Yogo, "Testing for Weak Instruments in Linear IV Regression," *Identification and Inference for Econometric Models*, 2005.

Terry, Stephen J, Toni M Whited, and Anastasia A Zakolyukina, "Information versus Investment," *The Review of Financial Studies*, 2023, *36* (3), 1148–1191.

Tintelnot, Felix, "Global production with export platforms," *The Quarterly Journal of Economics*, 2017, *132* (1), 157–209.

Wasi, Nada and Aaron Flaaen, "Record Linkage Using Stata: Preprocessing, Linking, and Reviewing Utilities," *The Stata Journal*, 2015, *15* (3), 672–697. Publisher: SAGE Publications.

Wilson, Daniel J., "Beggar Thy Neighbor? The in-State, Out-of-State, and Aggregate Effects of R&D Tax Credits," *The Review of Economics and Statistics*, 2009, *91* (2), 431–436.

A Proofs and Derivations

A.1 Equilibrium conditions

We summarize the conditions for a balanced growth path equilibrium, which consists of 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 the value from new innovations, that satisfy:

1. Labor market clearing in the production sector

$$w_{np} = \frac{1}{\gamma} Z_n^{\frac{\sigma-1}{\sigma}} \left(\frac{\bar{L}_n}{Y} \right)^{-\frac{1}{\sigma}}. \tag{18}$$

2. Labor market clearing in the innovation sector

$$\bar{H}_n = M \int_z \chi_n(z) \left(\bar{f}_n(z) + \ell_n(z) \right) d\Psi(z), \tag{19}$$

where we solved for $\ell_n(z)$, $\chi_n(z)$, and $\bar{f}_n(z)$ above in terms of equilibrium objects in Section 2.2.2.

3. Regional innovation productivity is given by

$$K_n = \bar{K}_n \left(M \int_z \chi_n(z) z^{1-\beta} \ell_n(z)^{\beta} d\Psi(z) \right)^{\alpha}. \tag{20}$$

4. Market clearing in the final good market

$$Y = M f_e + (r-g)V + \sum_{n=1}^{N} w_{ni} \bar{H}_n + \sum_{n=1}^{N} w_{np} \bar{L}_n. \tag{21}$$

5. The value of new innovations is given by equation (6).

6. Free entry into innovation satisfies equation (8).

7. Aggregate productivity $\mathcal{A}_t$, wages, consumption, profits, and the value of new innovations grow at a constant rate g, given by equation (9).

A.2 Proof of Proposition 1.

Consider the problem of an innovative firm ω described in equation (7).

Conditional on being active in market n, the firm's innovative employment decision for such market solves the following problem:

$$\max_{\ell_{n\omega t}} \lambda_{\omega t} V_t - \left[w_{nit}\left(\ell_{n\omega t} + f_{n\omega t}\right)\right]$$
$$\text{s.t.} \quad \lambda_{\omega t} = \lambda_0 \sum_{n=1}^{N} K_{nt}(z_\omega)^{1-\eta}(\ell_{n\omega t})^{\eta}$$

The first order condition of this problem is the following:

$$\eta \lambda_0 K_{nt}(z_\omega)^{1-\eta}(\ell_{n\omega t})^{\eta-1} V_t = w_{nit}.$$

Rearranging this expression, we obtain:

$$\ell_{n\omega t} = z_\omega \left(\frac{\eta K_{nt} \lambda_0 V_t}{w_{nit}}\right)^{\frac{1}{1-\eta}}.$$

Given the optimal innovative employment conditional on activity in market n, a firm will choose to enter such market if the additional idea arrival rate obtained from an R&D establishment in that location is higher than the total fixed and variable labor costs, as described by the following problem:

$$\max\left\{\lambda_0 K_{nt}(z_\omega)^{1-\eta}(\ell_{n\omega t})^{\eta} V_t - \left[w_{nit}\left(\ell_{n\omega t} + f_{n\omega t}\right)\right], 0\right\}$$
$$\text{s.t.} \quad \ell_{n\omega t} = z_\omega \left(\frac{\eta K_{nt} \lambda_0 V_t}{w_{nit}}\right)^{\frac{1}{1-\eta}}.$$

The solution to this problem is a threshold rule, which indicates that the firm will enter market n if the draw for $f_{n\omega t}$ is low enough :

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

As a result, the probability that a firm ω chooses to be active in market n, before observing the fixed cost draw, is:

$$\chi_{n\omega t} = Pr\left(f_{n\omega t} < \left(\frac{1-\eta}{\eta}\right) \ell_{n\omega t}\right) = f_{max}^{-\theta}\left(\left(\frac{1-\eta}{\eta}\right) \ell_{n\omega t}\right)^{\theta},$$

where the second equality follows from the assumption that the fixed cost distribution is inverse Pareto.

The expected amount of labor that a firm ω hires in location n to pay the fixed costs, conditional on being active in n, is:

$$\bar{f}_{n\omega t} = \mathbb{E}\left[f | f < \left(\frac{1-\eta}{\eta}\right)\ell_{n\omega t} f\right] = \int_0^{\left(\frac{1-\eta}{\eta}\right)\ell_{n\omega t}} f \frac{\theta f_{max}^{-\theta} f^{\theta-1}}{f_{max}^{-\theta}\left(\frac{1-\eta}{\eta}\ell_{n\omega t}\right)^{\theta}} df = \frac{\theta}{\theta+1}\left(\frac{1-\eta}{\eta}\ell_{n\omega t}\right). \tag{22}$$

We note that the optimal innovative employment conditional on activity, the probability of entering a market, and the expected fixed cost are the same for all firms with productivity z, so we suppress the index ω and write $\ell_{nt}(z) = \ell_{n\omega t}$, $\chi_{nt}(z) = \chi_{n\omega t}$, and $\bar{f}_{nt}(z) = \bar{f}_{n\omega t}$ for all ω such that $z_\omega = z$.

Finally, from equation (22), we obtain that the ratio of the expected amount of region-n innovation labor allocated to expansion relative to R&D activities for type-z firms is:

$$\frac{x_n(z)}{y_n(z)} = \frac{\chi_n(z)\bar{f}_n(z)}{\chi_n(z)\ell_n(z)} = \frac{1-\eta}{\eta}\frac{\theta}{\theta+1}.$$

A.3 Proof of Proposition 2.

Consider the expression for aggregate output in equation (1). Plugging in the expression for intermediate goods production in equation (2), we can express aggregate output as:

$$\ln Y_t = \int_0^1 \ln A_{jt} dj + \ln\left(\sum_{n=1}^N (Z_n L_n^p)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} = \ln(\mathcal{A}_t) + \ln(\mathcal{L}^p), \tag{23}$$

where we define aggregate productivity, $\mathcal{A}_t$, such that $\ln(\mathcal{A}_t) \equiv \int_0^1 \ln(A_{jt})dj$ and $\mathcal{L}^p \equiv \left(\sum_{n=1}^N (Z_n L_n^p)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$. Note that aggregate output and aggregate productivity grow at the same rate g.

We can write the change in aggregate productivity $\mathcal{A}_t$ in a small time interval, Δt, as $\frac{\ln(\mathcal{A}_{t+\Delta t}) - \ln(\mathcal{A}_t)}{\Delta t}$ and obtain the growth rate on the balanced growth path by letting Δt go to

zero, i.e.:

$$\begin{aligned} g &= \lim_{\Delta t \to 0} \frac{\ln(\mathcal{A}_{t+\Delta t}) - \ln(\mathcal{A}_t)}{\Delta t} \\ &= \lim_{\Delta t \to 0} \frac{\int_0^1 \left(\bar{\lambda}_t \Delta t \ln\left(\gamma A_{jt}\right) + \left(1 - \bar{\lambda}_t \Delta t\right) \ln\left(A_{jt}\right)\right) dj - \ln(\mathcal{A}_t)}{\Delta t} \\ &= \lim_{\Delta t \to 0} \frac{\bar{\lambda}_t \Delta t \ln(\gamma) + \ln(\mathcal{A}_t) - \ln(\mathcal{A}_t)}{\Delta t} \\ &= \bar{\lambda}_t \ln(\gamma). \end{aligned}$$

The second line is derived by noting that, in an interval of time Δt, each intermediate j is subject to being innovated upon with probability $\bar{\lambda}_t \Delta t$. As a consequence, in the aggregate, a fraction $\bar{\lambda}_t \Delta t$ of the intermediates is hit by an innovation in the time interval Δt, and the productivity of those intermediates increase by a factor γ. Notice that a constant aggregate growth rate along a BGP requires a constant rate of creative destruction, $\bar{\lambda}$.

The aggregate rate of creative destruction is derived by combining equations (3) and (4), which delivers:

$$\bar{\lambda} = M \lambda_0 \int_z \sum_{n=1}^{N} K_n\left(\{\chi, \ell\}\right) \chi_n\left(z\right) z^{1-\eta} \ell_n\left(z\right)^{\eta} dG\left(z\right).$$

A.4 Proof of Proposition 3.

Consider the growth rate along a BGP presented in equation (9). Plugging in equation (20) and the expressions for the innovative firms optimal decisions, we obtain:

$$\begin{aligned} g &= \ln(\gamma) M \lambda_0 \sum_{n=1}^{N} \bar{K}_n \left(M \int_{z'} \chi_n(z') z'^{1-\beta} \ell_n(z')^{\beta} \, d\Psi(z') \right)^{\alpha} \int_z \chi_n(z) z^{1-\eta} \ell_n^{\eta}(z) \, d\Psi(z) \\ &= \ln(\gamma) M \lambda_0 \sum_{n=1}^{N} \bar{K}_n \left(M \int_{z'} z'^{1-\beta} f_{max}^{-\theta} \left(\frac{1-\eta}{\eta}\right)^{\theta} \ell_n(z')^{\theta+\beta} \, d\Psi(z') \right)^{\alpha} \times \\ &\qquad \times \int_z z^{1-\eta} f_{max}^{-\theta} \left(\frac{1-\eta}{\eta}\right)^{\theta} \ell_n^{\theta+\eta}(z) \, d\Psi(z) \end{aligned}$$

Plugging the firm's optimal innovative labor hiring decision we get:

$$g = \ln(\gamma) \left(M f_{max}^{-\theta} \left(\frac{1-\eta}{\eta}\right)^{\theta} \mathbb{E}(z^{1+\theta}) \right)^{\alpha+1} \lambda_0 \sum_{n=1}^{N} \bar{K}_n \left(\frac{\eta K_n \lambda_0 V}{w_{ni}} \right)^{\frac{\theta+\eta+\alpha(\theta+\beta)}{1-\eta}}.$$

The innovative labor market clearing delivers the following expression:

$$\begin{aligned}\bar{H}_n &= M\int_z \chi_n(z)\left(\ell_n(z)+\bar{f}_n(z)\right)\psi(z)dz \\ &= M\int_z f_{\max}^{-\theta}\left(\frac{1-\eta}{\eta}\right)^{\theta}\left(1+\frac{\theta}{\theta+1}\frac{1-\eta}{\eta}\right)(\ell_n(z))^{1+\theta}\psi(z)dz.\end{aligned}$$

Combining this with the firm's innovation labor decision, we get:

$$\frac{\eta K_n V}{w_{ni}} = \left(\frac{\bar{H}_n}{M\mathbb{E}[z^{1+\theta}]f_{\max}^{-\theta}\left(\frac{1-\eta}{\eta}\right)^{\theta}\left(1+\frac{\theta}{\theta+1}\frac{1-\eta}{\eta}\right)}\right)^{\frac{1-\eta}{1+\theta}}.$$

Plugging this expression into the equation for g we obtain:

$$g = \ln(\gamma)\lambda_0\left(Mf_{max}^{-\theta}\left(\frac{1-\eta}{\eta}\right)^{\theta}\mathbb{E}(z^{1+\theta})\right)^{\alpha+1-\frac{\theta+\eta+\alpha(\theta+\beta)}{1+\theta}}\left(1+\frac{\theta}{\theta+1}\frac{1-\eta}{\eta}\right)^{-\frac{\theta+\eta+\alpha(\theta+\beta)}{1+\theta}}\sum_{n=1}^{N}\bar{K}_n\bar{H}_n^{\frac{\theta+\eta+\alpha(\theta+\beta)}{1+\theta}}$$

From this expression we obtain that the elasticity of the growth rate relative to f_{max} is given by:

$$\frac{\partial \ln g}{\partial \ln f_{max}} = \frac{(1-\eta)+\alpha(1-\beta)}{1+\theta}\left(\frac{\partial \ln M}{\partial \ln f_{max}}-\theta\right).$$

A.5 Proof of Proposition 4.

A.5.1 Restricted problem

For ease of exposition, we consider first the problem of a restricted planner that takes the mass of innovating firms in the economy as given. In addition, recall that the labor endowments are fixed. Thus, the tool available to the planner is to determine the allocation of resources within the innovation sector between R&D and firm expansion. In such a scenario, the planner wishes to maximize the aggregate arrival rate of ideas, since it raises growth

without reallocating resources away from production. The planner's problem is therefore

$$\begin{aligned}
\max_{\{\ell_n(z),\chi_n(z)\}_{n=1}^N} & M\int_z \sum_{n=1}^N \chi_n(z) K_n z^{1-\eta}\ell_n^\eta(z)\, d\Psi(z) \\
\text{subject to:} K_n &= \bar{K}_n\left(M\int_z \chi_n(z) z^{1-\beta}\ell_n(z)^\beta\, d\Psi(z)\right)^\alpha \\
\bar{H}_n &= M\int_z \chi_n(z)\left(\bar{f}_n(z)+\ell_n(z)\right) d\Psi(z) \\
\bar{f}_n(z) &= \frac{\theta}{\theta+1} f_{max}\left(\chi_n(z)\right)^{\frac{1}{\theta}}
\end{aligned}$$

Let $x_n(z) \equiv \chi_n(z)\bar{f}_n(z)$ and $y_n(z) \equiv \chi_n(z)\ell_n(z)$ denote the amount of region-n innovation labor allocated to expansion and R&D activities of type-z firms, respectively. Let $\psi(z)$ be the pdf of firms of type z in the economy, and define $H_n(z) = \psi(z)(x_n(z)+y_n(z))$ and $\tilde{H}_n(z) = H_n(z)/\psi(z)$. Note that the probability that a firm of type z opens a plant in location n, $\chi_n(z)$, and the expected amount of labor that a firm of type z hires in location n to pay the fixed costs conditional on being active there, $\bar{f}_n(z)$, are linked by the following equation:

$$\chi_n(z) = \left(f_{max}\frac{\theta}{\theta+1}\right)^{-\theta}\bar{f}_n(z)^\theta.$$

Combining this expression with the definitions of $x_n(z)$ and $y_n(z)$ we obtain:

$$\bar{f}_n(z) = \left(\frac{\theta}{\theta+1}f_{max}\right)^{\frac{\theta}{\theta+1}} x_n(z)^{\frac{1}{\theta+1}} \tag{24}$$

$$\chi_n(z) = \left(\frac{\theta}{\theta+1}f_{max}\right)^{-\frac{\theta}{\theta+1}} x_n(z)^{\frac{\theta}{\theta+1}} \tag{25}$$

$$y_n(z) = \frac{H_n(z)}{\psi(z)} - x_n(z). \tag{26}$$

Using these definitions, we note that, for any constant $a > 0$, we obtain:

$$\chi_n(z)z^{1-a}\ell_n(z)^a = \chi_n(z)z^{1-a}\left(\frac{y_n(z)}{\chi_n(z)}\right)^a = \left(\frac{\theta}{\theta+1}f_{max}\right)^{-\frac{\theta}{\theta+1}(1-a)} x_n(z)^{\frac{\theta}{\theta+1}(1-a)} z^{1-a} y_n(z)^a.$$

We then define the following objects:

$$\begin{aligned}
D_n(z) &\equiv \psi(z)z^{1-\eta}x_n(z)^{\frac{\theta}{\theta+1}(1-\eta)}y_n(z)^\eta \\
B_n(z) &\equiv \psi(z)z^{1-\beta}x_n(z)^{\frac{\theta}{\theta+1}(1-\beta)}y_n(z)^\beta.
\end{aligned}$$

Combining the expressions above, we define the following variables:

$$s_n^{\beta}(z) \equiv \frac{\psi_n(z)\chi_n(z)z^{1-\beta}(\ell_n(z))^{\beta}}{\int_{z'}\psi_n(z')\chi_n(z')(\ell_n(z'))^{\beta}dz'} = \frac{B_n(z)}{\int_{z'}B_n(z')dz'}$$
$$s_n^{\eta}(z) \equiv \frac{\psi_n(z)\chi_n(z)z^{1-\eta}(\ell_n(z))^{\eta}}{\int_{z'}\psi_n(z')\chi_n(z')(\ell_n(z'))^{\eta}dz'} = \frac{D_n(z)}{\int_{z'}D_n(z')dz'}.$$

Using these expressions, we can write the following maximization problem, which yields a solution equivalent to the planner's allocation:

$$\max_{x_n(z),y_n(z)} M^{1+\alpha}\sum_{n=1}^{N}\bar{K_n}\left(\int_{z'}B_n(z')dz'\right)^{\alpha}\int_z D_n(z)dz + \\ +\sum_{n=1}^{N}\mathcal{L}_n\left(\bar{H}_n - M\int_z\psi(z)(x_n(z)+y_n(z))dz\right).$$

The first order conditions of this problem with respect to x_n and y_n are:

$$M^{1+\alpha}\bar{K_n}\left[\left(\int_{z'}B_n(z')dz'\right)^{\alpha}\frac{\partial D_n(z)}{\partial x_n(z)} + \alpha\left(\int_{z'}B_n(z')dz'\right)^{\alpha-1}\frac{\partial B_n(z)}{\partial x_n(z)}\int_z D_n(z)dz\right] = \mathcal{L}_n\psi(z)$$
$$M^{1+\alpha}\bar{K_n}\left[\left(\int_{z'}B_n(z')dz'\right)^{\alpha}\frac{\partial D_n(z)}{\partial y_n(z)} + \alpha\left(\int_{z'}B_n(z')dz'\right)^{\alpha-1}\frac{\partial B_n(z)}{\partial y_n(z)}\int_z D_n(z)dz\right] = \mathcal{L}_n\psi(z).$$

Note that

$$\frac{\partial D_n(z)}{\partial x_n(z)} = \frac{\theta}{\theta+1}(1-\eta)\frac{D_n(z)}{x_n(z)} \qquad \text{and} \qquad \frac{\partial D_n(z)}{\partial y_n(z)} = \eta\frac{D_n(z)}{y_n(z)}$$
$$\frac{\partial B_n(z)}{\partial x_n(z)} = \frac{\theta}{\theta+1}(1-\beta)\frac{B_n(z)}{x_n(z)} \qquad \text{and} \qquad \frac{\partial B_n(z)}{\partial y_n(z)} = \beta\frac{B_n(z)}{y_n(z)}$$

Thus equation the first order conditions can be rewritten as:

$$M^{1+\alpha}\frac{\theta}{\theta+1}\frac{\bar{K_n}\left(\int_{z'}B_n(z')dz'\right)^{\alpha}\int_z D_n(z)dz}{x_n(z)}\left[(1-\eta)s_n^{\eta}(z) + \alpha(1-\beta)s_n^{\beta}(z)\right] = \mathcal{L}_n\psi(z) \quad (27)$$
$$M^{1+\alpha}\frac{\bar{K_n}\left(\int_{z'}B_n(z')dz'\right)^{\alpha}\int_z D_n(z)dz}{y_n(z)}\left[\eta s_n^{\eta}(z) + \alpha\beta s_n^{\beta}(z)\right] = \mathcal{L}_n\psi(z) \quad (28)$$

Taking the ratio of these two equations we find that:

$$\begin{aligned}\frac{x_n(z)}{y_n(z)} =& \frac{\theta}{\theta+1}\frac{\left[(1-\eta)s_n^{\eta}(z)+\alpha(1-\beta)s_n^{\beta}(z)\right]}{\left[\eta s_n^{\eta}(z)+\alpha\beta s_n^{\beta}(z)\right]} \\ =& \frac{\left(\frac{s_n^{\beta}(z)}{s_n^{\eta}(z)}\right)\alpha\frac{1-\beta}{1-\eta}+1}{\left(\frac{s_n^{\beta}(z)}{s_n^{\eta}(z)}\right)\alpha\frac{\beta}{\eta}+1}\left(\frac{1-\eta}{\eta}\frac{\theta}{\theta+1}\right). \end{aligned} \tag{29}$$

A.5.2 Generalized Welfare Function

Next we turn to the general case where the planner can choose the number of innovative firms M. The planner takes into account the fixed cost of opening an innovative firm, which is $f_e Y_t$ in terms of the final good. Increasing innovative firms' entry thus reduces aggregate consumption, C_t, which equal to final good output minus the resources spent on fixed costs: $C_t = Y_t - M f_e Y_t$.

Consider the unrestricted problem of a planner along a balanced growth path. The planner maximizes the following welfare function, which aggregates the utility of each household in location n and sector s at time t with some non-negative weights $\left\{\{\nu_{nt}^s\}_{s\in\{p,i\}}\right\}_{n=1}^N$:

$$W_0 = \int_0^{\infty} e^{-\rho t}\sum_{n=1}^{N}\sum_{s=\{p,i\}}\nu_{nt}^s u\left(C_{ns0}e^{gt}\right)\,dt,$$

where $C_{nst} = C_{ns0}e^{gt}$ is the amount of final-good consumption in period t for each household, summing up to total aggregate consumption in each period: $C_t = \sum_{n=1}^{N}\sum_{s=\{p,i\}} C_{nst}$. Note that for a given amount of final good dedicated to consumption, C_t, the planner could choose transfers to redistribute resources across households to maximize welfare. We are not interested in solving for those transfers, but in understanding the optimal allocation of resources between R&D and expansion. As a consequence, note that we can equivalently express aggregate welfare as a function of aggregate consumption:

$$W_0 = \int_0^{\infty} e^{-\rho t} v\left(C_t\right)\,dt,$$

where the function $v(\cdot)$ is strictly increasing and concave as long as $u(\cdot)$ is strictly increasing and concave and the weights $\left\{\{\nu_{nt}^s\}_{s\in\{p,i\}}\right\}_{n=1}^N$ are non-negative.

As a result, using the expression from equation (23), the problem of the planner becomes:

$$\begin{aligned}
&\max_{\{\ell_n(z),\chi_n(z)\}_{n=1}^N,M} \int_0^\infty e^{-\rho t} v\left(\mathcal{A}_0 \mathcal{L}^p (1 - M f_e) e^{gt}\right) dt \\
\text{s.t.} \quad & g = \ln(\gamma) M \int_z \sum_{n=1}^N \chi_n(z) K_n z^{1-\eta} \ell_n^\eta(z) \, d\Psi(z) \\
& K_n = \bar{K}_n \left(M \int_z \chi_n(z) z^{1-\beta} \ell_n(z)^\beta \, d\Psi(z) \right)^\alpha \\
& \bar{H}_n = M \int_z \chi_n(z) \left(\bar{f}_n(z) + \ell_n(z) \right) d\Psi(z) \\
& \bar{f}_n(z) = \frac{\theta}{\theta+1} f_{max} \left(\chi_n(z)\right)^{\frac{1}{\theta}}.
\end{aligned}$$

Using the expressions for $B_n(z)$ and $D_n(z)$ introduced in the previous section, we can write

$$K_n = \left(\frac{\theta}{\theta+1} f_{max} \right)^{-\frac{\theta}{\theta+1}(1-\beta)} \bar{K}_n M^\alpha \left(\int_{z'} B_n(z') dz' \right)^\alpha$$

and we can express the growth rate as

$$g = \left(\frac{\theta}{\theta+1} f_{max} \right)^{-\frac{\theta}{\theta+1}(1-\beta+1-\eta)} \ln(\gamma) M^{1+\alpha} \sum_{n=1}^N \bar{K}_n \left(\int_{z'} B_n(z') dz' \right)^\alpha \int_z D_n(z) dz.$$

Using these expressions, we can write the following maximization problem, which yields a solution equivalent to the planner's allocation:

$$\begin{aligned}
\max_{x_n(z),y_n(z),M} & \int_0^\infty e^{-\rho t} v\left(\mathcal{A}_0 \mathcal{L}^p (1 - M f_e) e^{gt}\right) dt + \\
& + \mu \left(g - \left(\frac{\theta}{\theta+1} f_{max} \right)^{-\frac{\theta}{\theta+1}(1-\beta+1-\eta)} \ln(\gamma) M^{1+\alpha} \sum_{n=1}^N \bar{K}_n \left(\int_{z'} B_n(z') dz' \right)^\alpha \int_z D_n(z) dz \right) \\
& + \sum_{n=1}^N \mu_n \left(H_n - M \int_z \psi(z)(x_n(z) + y_n(z)) dz \right).
\end{aligned}$$

The first order conditions of this problem are:

$$-\mu\frac{\partial g}{\partial x_n(z)} - \mu_n M\psi(z) = 0 \quad (30)$$

$$-\mu\frac{\partial g}{\partial y_n(z)} - \mu_n M\psi(z) = 0 \quad (31)$$

$$-\int_0^\infty e^{(g-\rho t)}v'(C_t)\left(\mathcal{A}_0\mathcal{L}^p f_e\right)dt - \mu(1+\alpha)\frac{g}{M} - \mu_n\left(\int_z \psi(z)(x_n(z)+y_n(z))dz\right) = 0 \quad (32)$$

Note that equations (30) and (31) have the same solution as the first order conditions in the restricted planner problem in equations (27) and (28). As a result, the relative allocation of resources between R&D and expansion, for a given level of firm entry M, is the same in the restricted and unrestricted problem, and it satisfies equation (29).

A.6 Proof of Proposition 5.

To derive the planner's allocation across firm types, we continue with the analysis presented in Section A.5.1. Solving the first order conditions in equations (27) and (28) for $x_n(z)$ and $y_n(z)$ we can solve for the mass of labor hired by firms of type z in market n, $H_n(z)$:

$$\begin{aligned} H_n(z) &= M\psi(z)(x_n(z)+y_n(z)) = \\ &= M^{1+\alpha}\frac{\bar{K}_n\left(\int_{z'}B_n(z')dz'\right)^\alpha\int_z D_n(z)dz}{\mathcal{L}_n}\times \\ &\qquad\times\left(\frac{\theta}{\theta+1}\left[(1-\eta)s_n^\eta(z)+\alpha(1-\beta)s_n^\beta(z)\right]+\left[\eta s_n^\eta(z)+\alpha\beta s_n^\beta(z)\right]\right) \end{aligned}$$

As a result, the share of local R&D labor allocated to type-z firms in market n is:

$$\begin{aligned} \frac{H_n(z)}{H_n} &= \frac{\frac{\theta}{\theta+1}\left[(1-\eta)s_n^\eta(z)+\alpha(1-\beta)s_n^\beta(z)\right]+\left[\eta s_n^\eta(z)+\alpha\beta s_n^\beta(z)\right]}{\int_{z'}\left(\frac{\theta}{\theta+1}\left[(1-\eta)s_n^\eta(z)+\alpha(1-\beta)s_n^\beta(z)\right]+\left[\eta s_n^\eta(z)+\alpha\beta s_n^\beta(z)\right]\right)d(z')} \\ &= \frac{\alpha s_n^\beta(z)\left(\beta+(1-\beta)\frac{\theta}{\theta+1}\right)+s_n^\eta(z)\left(\eta+(1-\eta)\frac{\theta}{\theta+1}\right)}{\int_{z'}\left(\alpha s_n^\beta(z')\left(\beta+(1-\beta)\frac{\theta}{\theta+1}\right)+s_n^\eta(z')\left(\eta+(1-\eta)\frac{\theta}{\theta+1}\right)\right)d(z')}. \end{aligned}$$

A.7 Proof of Proposition 6.

In this section, we examine the planner's allocation under free labor mobility in the innovation sector. The planner's problem is

$$\begin{aligned}
\max_{H_n,\{\ell_n(z),\chi_n(z)\}_{n=1}^N} \; & M\int_z \sum_{n=1}^N \chi_n(z)\, K_n z^{1-\eta} \ell_n^\eta(z)\, d\Psi(z) \\
\text{subject to:} \; & K_n = \bar{K}_n \left(M\int_z \chi_n(z)\, z^{1-\beta} \ell_n(z)^\beta\, d\Psi(z) \right)^\alpha \\
& H_n = M\int_z \chi_n(z) \left(\bar{f}_n(z) + \ell_n(z) \right) d\Psi(z) \\
& \bar{f}_n(z) = \frac{\theta}{\theta+1} f_{max} \left(\chi_n(z)\right)^{\frac{1}{\theta}} \\
& \sum_{n=1}^N H_n = \bar{H}.
\end{aligned}$$

To solve the planner's problem, we define $x_n(z) \equiv \chi_n(z)\,\bar{f}_n(z)$ and $y_n(z) \equiv \chi_n(z)\,\ell_n(z)$ as the amount of region-n innovation labor allocated to expansion and R&D activities of type-z firms, respectively, as in Section A.5. Using these definitions, we note that, for any constant $a > 0$, we obtain:

$$\chi_n(z) z^{1-a} \ell_n(z)^a = \chi_n(z) z^{1-a} \left(\frac{y_n(z)}{\chi_n(z)}\right)^a = \left(\frac{\theta}{\theta+1} f_{max}\right)^{-\frac{\theta}{\theta+1}(1-a)} x_n(z)^{\frac{\theta}{\theta+1}(1-a)} z^{1-a} y_n(z)^a.$$

We then define the following objects:

$$\begin{aligned}
D_n(z) &\equiv \psi(z) z^{1-\eta} x_n(z)^{\frac{\theta}{\theta+1}(1-\eta)} y_n(z)^\eta \\
B_n(z) &\equiv \psi(z) z^{1-\beta} x_n(z)^{\frac{\theta}{\theta+1}(1-\beta)} y_n(z)^\beta.
\end{aligned}$$

Combining the expressions above, we define the following variables, as in Section A.5:

$$\begin{aligned}
s_n^\beta(z) &\equiv \frac{\psi_n(z)\,\chi_n(z)\, z^{1-\beta} (\ell_n(z))^\beta}{\int_{z'} \psi_n(z')\,\chi_n(z')\,(\ell_n(z'))^\beta dz'} = \frac{B_n(z)}{\int_{z'} B_n(z') dz'} \\
s_n^\eta(z) &\equiv \frac{\psi_n(z)\,\chi_n(z)\, z^{1-\eta} (\ell_n(z))^\eta}{\int_{z'} \psi_n(z')\,\chi_n(z')\,(\ell_n(z'))^\eta dz'} = \frac{D_n(z)}{\int_{z'} D_n(z') dz'}.
\end{aligned}$$

Using these expressions, we can write the following maximization problem, which yields a

solution equivalent to the planner's allocation:

$$\max_{x_n(z),y_n(z),H_n} M^{1+\alpha}\sum_{n=1}^{N}\bar{K_n}\left(\int_{z'}B_n(z')dz'\right)^{\alpha}\int_z D_n(z)dz+ \\ +\sum_{n=1}^{N}\mathcal{L}_n\left(H_n - M\int_z \psi(z)(x_n(z)+y_n(z))dz\right)+ \\ +\bar{\mathcal{L}}\left(\bar{H}-\sum_{n=1}^{N}H_n\right).$$

The first order conditions of this problem with respect to x_n, y_n, and H_n are:

$$M^{1+\alpha}\bar{K_n}\left[\left(\int_{z'}B_n(z')dz'\right)^{\alpha}\frac{\partial D_n(z)}{\partial x_n(z)}+\alpha\left(\int_{z'}B_n(z')dz'\right)^{\alpha-1}\frac{\partial B_n(z)}{\partial x_n(z)}\int_z D_n(z)dz\right]=\mathcal{L}_n\psi(z)$$

$$M^{1+\alpha}\bar{K_n}\left[\left(\int_{z'}B_n(z')dz'\right)^{\alpha}\frac{\partial D_n(z)}{\partial y_n(z)}+\alpha\left(\int_{z'}B_n(z')dz'\right)^{\alpha-1}\frac{\partial B_n(z)}{\partial y_n(z)}\int_z D_n(z)dz\right]=\mathcal{L}_n\psi(z)$$

$$\mathcal{L}_n=\bar{\mathcal{L}}.$$

Note that

$$\frac{\partial D_n(z)}{\partial x_n(z)}=\frac{\theta}{\theta+1}(1-\eta)\frac{D_n(z)}{x_n(z)} \qquad \text{and} \qquad \frac{\partial D_n(z)}{\partial y_n(z)}=\eta\frac{D_n(z)}{y_n(z)}$$

$$\frac{\partial B_n(z)}{\partial x_n(z)}=\frac{\theta}{\theta+1}(1-\beta)\frac{B_n(z)}{x_n(z)} \qquad \text{and} \qquad \frac{\partial B_n(z)}{\partial y_n(z)}=\beta\frac{B_n(z)}{y_n(z)}$$

Thus equation the first order conditions can be rewritten as:

$$M^{1+\alpha}\frac{\theta}{\theta+1}\frac{\bar{K_n}\left(\int_{z'}B_n(z')dz'\right)^{\alpha}\int_z D_n(z)dz}{x_n(z)}\left[(1-\eta)s_n^{\eta}(z)+\alpha(1-\beta)s_n^{\beta}(z)\right]=\mathcal{L}_n\psi(z)$$

$$M^{1+\alpha}\frac{\bar{K_n}\left(\int_{z'}B_n(z')dz'\right)^{\alpha}\int_z D_n(z)dz}{y_n(z)}\left[\eta s_n^{\eta}(z)+\alpha\beta s_n^{\beta}(z)\right]=\mathcal{L}_n\psi(z)$$

Taking the ratio of these two equations we find that:

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

Solving these equations for $x_n(z)$ and $y_n(z)$ and plugging them into the local labor market

clearing equation, we get:

$$H_n = M \int_z \psi(z)(x_n(z) + y_n(z))dz =$$
$$= M^{1+\alpha} \int_z \frac{\bar{K}_n \left(\int_{z'} B_n(z')dz'\right)^{\alpha} \int_z D_n(z)dz}{\mathcal{L}_n} \times$$
$$\times \left(\frac{\theta}{\theta+1} \left[(1-\eta)s_n^{\eta}(z) + \alpha(1-\beta)s_n^{\beta}(z)\right] + \left[\eta s_n^{\eta}(z) + \alpha\beta s_n^{\beta}(z)\right] \right) dz$$

Define the following variable:

$$\Upsilon_n \equiv \bar{K}_n \left(\int_{z'} B_n(z')dz' \right)^{\alpha} \int_z D_n(z)dz.$$

Note that $\int_z s_n^{\eta}(z)dz = \int_z s_n^{\beta}(z)dz = 1$. It follows that:

$$H_n = M^{1+\alpha} \frac{\Upsilon_n}{\bar{\mathcal{L}}} \left(\frac{\theta}{\theta+1} \left((1-\eta) + \alpha(1-\beta)\right) + (\eta + \alpha\beta) \right).$$

Combining the expressions for x_n, y_n, B_n, and D_n, note that

$$\Upsilon_n \propto \bar{K}_n \Upsilon_n^{\alpha\frac{\theta}{\theta+1}(1-\beta)+\beta+\frac{\theta}{\theta+1}(1-\eta)+\eta}.$$

Rearranging this expression, we obtain that:

$$H_n \propto \Upsilon_n \propto \bar{K}_n^{\frac{1}{\frac{1-\eta}{1+\theta} - \alpha\frac{\beta+\theta}{\theta+1}}}.$$

In the decentralized equilibrium, from the model inversion, we obtain that:

$$\bar{K}_n \propto w_{ni} \left(\bar{H}_n\right)^{\frac{1-\eta}{1+\theta} - \alpha\frac{\beta+\theta}{\theta+1}}$$

As a result, under labor mobility, when wages are equalized across locations, the allocations of innovative labor in the decentralized equilibrium and in the planner's solution are the same. In particular, in both the equilibrium and the planner' allocation, we obtain that:

$$\frac{H_n}{\sum_{n'=1}^{N} H_{n'}} = \frac{\bar{K}_n^{\frac{1}{\frac{1-\eta}{1+\theta} - \alpha\frac{\beta+\theta}{\theta+1}}}}{\sum_{n'=1}^{N} \bar{K}_{n'}^{\frac{1}{\frac{1-\eta}{1+\theta} - \alpha\frac{\beta+\theta}{\theta+1}}}}.$$

A.8 Additional details for Section 2.5

A.8.1 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

$$K_n = \bar{K}_n \left(\sum_{i=1}^{N} \int_{\omega \in \Omega_i} \kappa_{in} z_\omega^{1-\beta} \ell_{i\omega}^{\beta} d\omega \right)^{\alpha}. \tag{33}$$

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

$$s_n^{\beta}(z) \equiv \sum_{k=1}^{N} \int_{z'} s_k^{\eta}(z') \, s_{nk}^{\beta}(z) \, dG(z') \tag{34}$$

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

A.8.2 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 4.

B Data

B.1 Details of merging PatentsView and Dun & Bradstreet

Table B.1: Example patent record.

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

Note: This table shows an example patent record.

Table B.2: Example Dun & Bradstreet 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

Note: This table shows an example Dun & Bradstreet record. The construction of this entry follows the procedure described in Appendix B.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.

Table B.3: Merge quality between PatentsView and Dun & Bradstreet.

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				

Note: This table conditions on nonwithdrawn utility patents granted between 1969 and 2021 with a single U.S.-based assignee.

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 B.1 and B.2 show example records that we are trying to match. Table B.3 presents summary statistics on the quality of the merge. We merge records using the following procedure.

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`.

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 using the configuration of Wasi and Flaaen (2015) suggested by Hughes et al. (2021).

   (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.

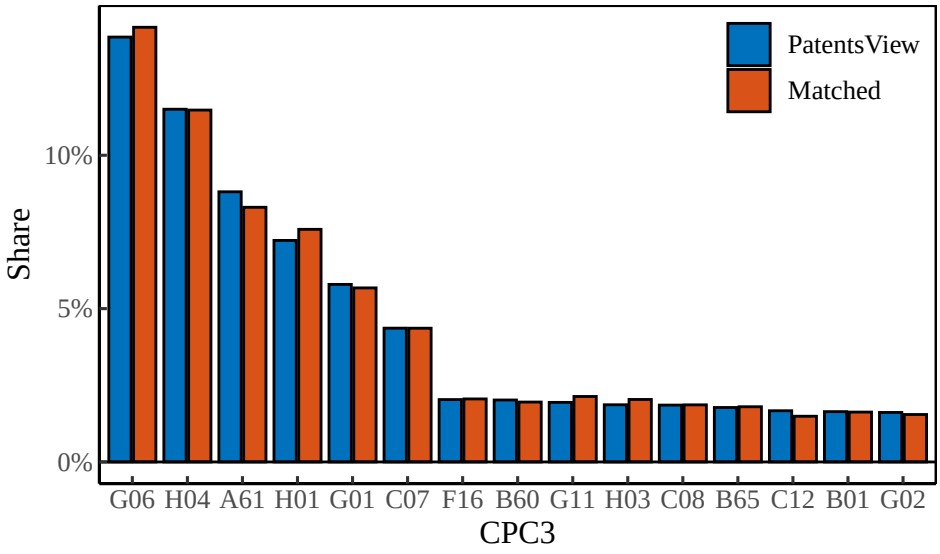
   (a) We assign codes to each match. In descending order of quality:

      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.

Figure B.1: Frequency distribution of three digit CPC in PatentsView and in the matched sample.



iv. "D"—If a patent does not satisfy any of the above.

v. Note that for the example records in Tables B.1 and B.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:
  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`.

Table B.4: R&D SIC codes.

SIC	Description
28	Chemical & Allied Products
35	Industrial Machinery & Equipment
36	Electronic & Other Electric Equipment
37	Transportation Equipment
38	Instruments & Related Products
73	Business Services
87	Engineering & Management Services

B.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 B.1) onto the Dun & Bradstreet 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.

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.3 Sample selection

Our analysis covers the period from 1980 to 2018. We select these years to maximize sample size while avoiding truncation issues at both ends of the data. While our patent data includes patents granted after 1976, many early-sample patents were applied for prior to 1976, creating incomplete coverage in the earliest years. Similarly, patents applied for after 2018 may not yet be granted in our data. We also exclude the post-2018 period to avoid complications from the Covid-19 pandemic, which disrupted the typical relationship between inventor location and workplace location.

B.4 Comparing Dun & Bradstreet to the Longitudinal Business Database

Our analysis uses Dun & Bradstreet (D&B) to identify firm locations and activities. D&B is used to construct the National Establishment Time Series (NETS) data. Barnatchez et al. (2017) study how similar NETS is to the Longitudinal Business Database (LBD). While it has known limitations compared to the LBD, these concerns are largely mitigated for our specific application. Our research design is well-suited to D&B's strengths while avoiding its main weaknesses. First, as we show, patenting activity is heavily concentrated among larger establishments, where Barnatchez et al. (2017) show that NETS correlates highly with official sources (above 95% for establishment counts and above 80% for employment levels). Second, our analysis focuses solely on the extensive margin—whether firm activity exists in a given location—rather than firm size or growth patterns, making the concerns Barnatchez et al. (2017) raise about imputation largely irrelevant to our study.

For our purposes, D&B provides comprehensive geographic coverage and industry detail that makes it well-suited for linking to patent records and identifying the spatial distribution of innovation activity.

Table C.1: Estimation results for main text event studies (Figure 5).

τ	External citations	Patenting efficiency
-10	0.00	0.02
	(0.05)	(0.03)
-9	-0.05	0.02
	(0.04)	(0.02)
-8	-0.02	-0.00
	(0.04)	(0.02)
-7	-0.03	0.01
	(0.04)	(0.02)
-6	-0.00	0.03
	(0.04)	(0.02)
-5	0.04	-0.01
	(0.04)	(0.02)
-4	-0.03	-0.00
	(0.03)	(0.02)
-3	-0.03	0.00
	(0.03)	(0.02)
-2	-0.02	-0.00
	(0.03)	(0.02)
0	0.10	0.02
	(0.03)	(0.02)
1	0.17	0.04
	(0.03)	(0.02)
2	0.15	0.05
	(0.04)	(0.02)
3	0.20	0.03
	(0.04)	(0.02)
4	0.16	0.06
	(0.04)	(0.02)
5	0.19	0.08
	(0.04)	(0.02)
6	0.20	0.06
	(0.04)	(0.02)
7	0.24	0.07
	(0.04)	(0.02)
8	0.20	0.07
	(0.04)	(0.02)
9	0.25	0.08
	(0.04)	(0.02)
10	0.25	0.10
	(0.04)	(0.02)
Adjusted R^2	0.35	0.35
Wald test pre-trend p-value	0.49	0.73
N	136721	255493
Number of treatment units	6071	5884

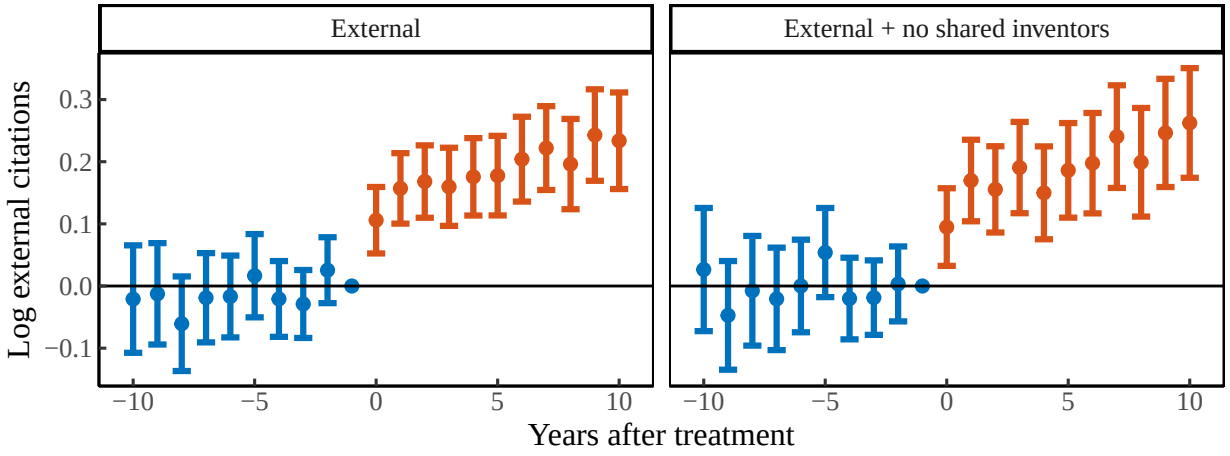
Note: This table shows the regression loadings from the event studies pictured in Figure 5.

C Regression tables

D Robustness

D.1 Different specifications of the dependent variable

Figure D.1: Estimation with looser specifications of the dependent variable.



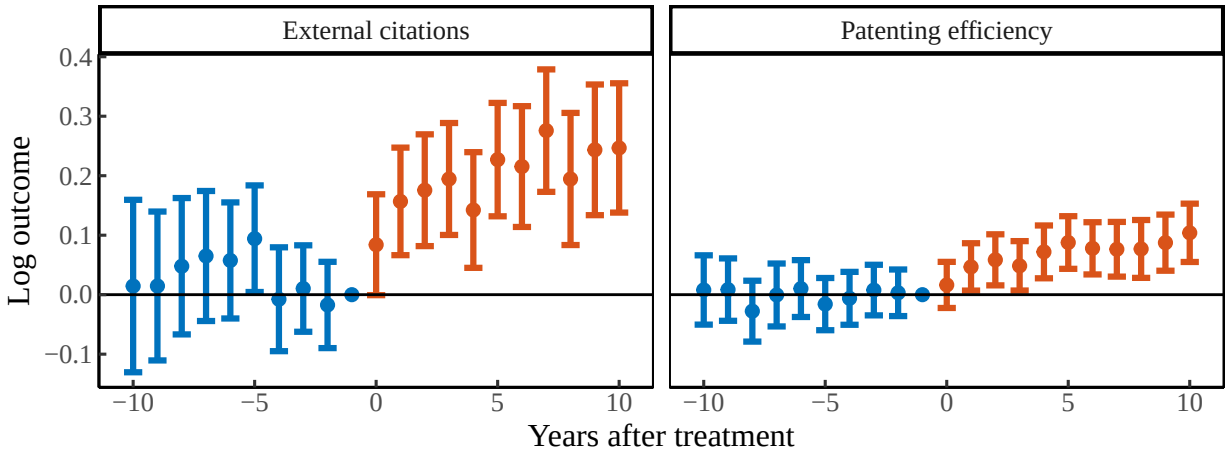
Note: Wald tests of joint nullity of pre-treatment covariates fail to be rejected with p-values of 0.50 and 0.48, respectively.

Figure D.1 shows the baseline regression for external citations when the outcome variable is defined less rigorously: the left panel shows the results when the outcome variable is the log number of external citations (without the additional restrictions discussed in footnote 15); the right panel shows the results when the outcome variable is the log number of external citations, excluding citations made between firms that ever shared inventors but not restricting to patents originated outside the expansion commuting zone. Results are highly similar to the main text.

D.2 With added controls

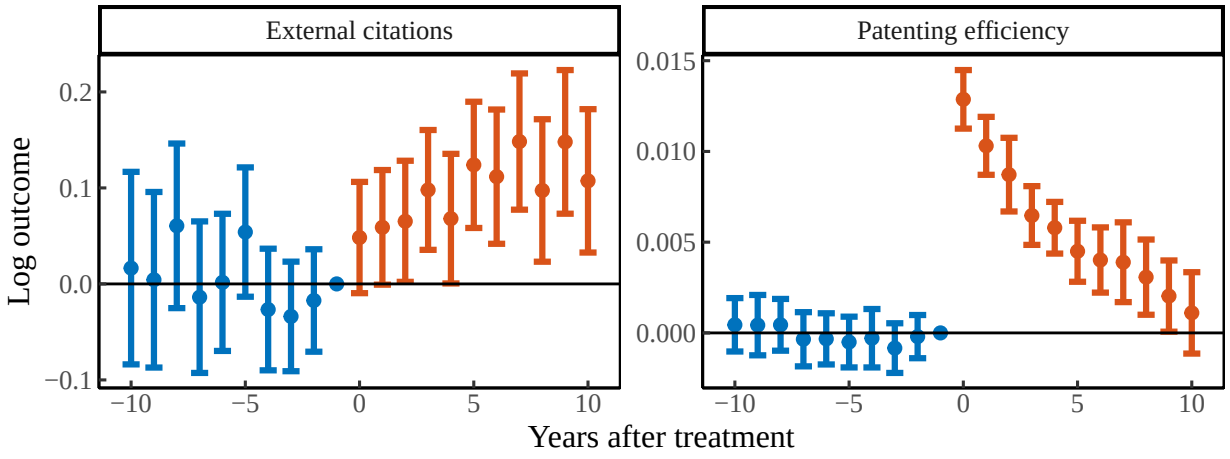
Figures D.2 and D.3 show the baseline regression when including additional controls. Figure D.2 shows results when including cohort-firm-time fixed effects, $\delta_{\tilde{g}_{\omega n}\omega t}$; Figure D.3 shows results when including commuting zone-time fixed effects, ϑ_{nt}, and firm-time fixed effects, $\delta_{\omega t}$. When the outcome variable is patenting efficiency, we further refine the fixed effect to control for any variation at the commuting zone-time-CPC3-level, $\vartheta_{n,t,CPC3}$. Results remain statistically significant, and we fail to reject joint nullity of the pre-treatment covariates at conventional levels of significance. The magnitudes of the estimated effects are lower in Figure D.3, which makes sense, as some of the variation is likely absorbed by local technological factors.

Figure D.2: Estimation with cohort-firm-time fixed effect.



Note: Wald tests of joint nullity of pre-treatment covariates fail to be rejected with p-values of 0.22 and 0.88, respectively.

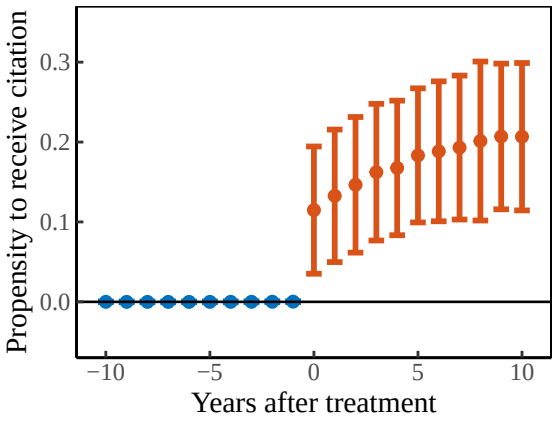
Figure D.3: Estimation with firm-time fixed effect and commuting zone-time fixed effect.



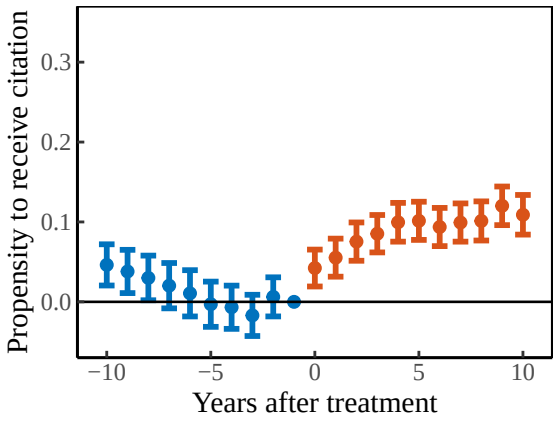
Note: Wald tests of joint nullity of pre-treatment covariates fail to be rejected with p-values of 0.49 and 0.88, respectively. Fixed effect for patenting efficiency is at the commuting zone-CPC3-time level.

Figure D.4: Robustness—synthetic controls and binary indicator.

(a) Response of citations received by firm patents in expansion commuting zone preceding and following expansion—indicator dependent variable (Ben-Michael et al., 2022).



(b) Response of citations received by firm patents in expansion commuting zone preceding and following expansion—indicator dependent variable (OLS).

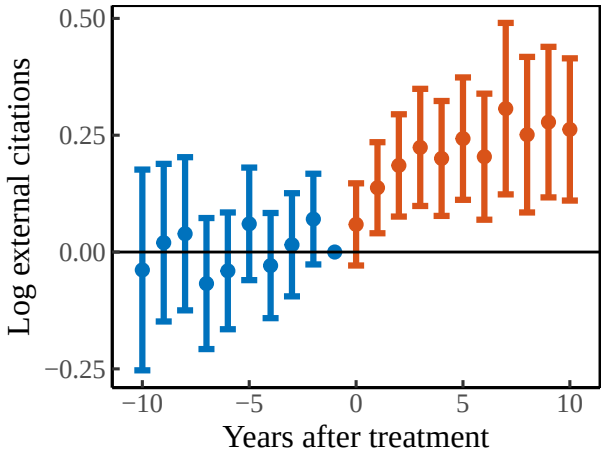


Note: We reject joint nullity of the pre-treatment dummies at the 0.1% level for Figure D.4b.

D.3 Alternative estimators

Figures D.4 and D.5 show the baseline regression when estimated using alternative estimators. Figure D.4a shows two perturbations of the main specification. First, it replaces the dependent variable with an indicator for whether positive citations are received, $\mathbb{I}\{y_{\omega nt} > 0\}$. Second, it estimates (11) using the synthetic DiD estimator of Ben-Michael et al. (2022) (code from the authors is available at `https://github.com/ebenmichael/augsynth/tree/master`). The main drawback to using the synthetic controls specification for the main analysis is twofold. First, the `R` package of Ben-Michael et al. (2022) does not allow for the use of log external citations for the outcome variable, as it cannot handle zero-citation observations. Second, the `R` package cannot handle the full frame of data due to memory limitations triggered by the construction of convex-weighted control units. To get around these limitations, we implement a bootstrap procedure and specify the outcome variable as a dummy that is equal to 1 when positive citations are received. We perform 1,000 bootstrap replicates on random 0.5% samples. Each replicate samples over $\omega \times n$ units from the full data without replacement. Figure D.6 shows the empirical densities of the bootstrap samples. The error bars in Figure D.4 are the empirical 2.5^{th} and 97.5^{th}-percentiles of the bootstrap replicates. To make error bars symmetric, for each treatment time dummy, we take the maximum absolute difference between point estimate and percentiles (Figure D.6 shows the empirical densities of the bootstrap estimates, which are approximately symmetric, which makes this

Figure D.5: Response of citations received by firm patents in expansion commuting zone—Poisson regression.



Note: Wald test of joint nullity of pre-treatment covariates fails to be rejected with a p-value of 0.30.

additional detail almost trivial). Figure D.4b shows the results of estimating (11) using the indicator, $\mathbb{I}\{y_{\omega nt} > 0\}$ as the dependent variable. Results are broadly similar to Figure D.4a, however, here, there is meaningful pre-trend, and we reject joint nullity of the pre-treatment covariates at the 0.1% level.

Figure D.5 shows the results of estimating (11) using Poisson regression (Correia et al., 2020; Bergé, 2018), which allows for the inclusion of zero's. Results are not meaningfully different from the main text.

D.4 Excluding nonabsorbing commuting zones

Figure D.7 shows the results after running the regression in (11) but drops treated units that exit the expansion market within ten years of entering. This constricts the sample size, resulting in a slight loss of statistical power, but the main result is preserved.

D.5 Additional heterogeneity analysis

Figure D.8 explores whether treatment effects are larger when firms engage in patenting in the first five years they are newly active in a market. On average, treated units that engage in local patenting 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 15-20% 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., 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 D.9 illustrates the heterogeneity in spillover effects across expansions in selected two digit SIC's. The results reveal substantial variation in estimated effects. Intuitively, expansions in industries that are known to generate many patents generate the highest spillovers, in line with the results in Figure D.8. Many of 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 effects are also observed in expansions in sectors not typically thought of as being particularly innovative: retail and wholesale trade and financial services all observe large and significant effects. Intuitively, the baseline effect discussed in Section 3 are nearest those sectors with the greatest weight in our sample.

E Details of the matching procedure for DiD analysis

E.1 Exact match on firm

- For a treated unit ωn (i.e. where $g_{\omega n} < \infty$), consider all units $\omega n'$ where $g_{\omega n'} = \infty$. These $\omega n'$ are potential control units. We allow for potential control units to be repeated within a multi-establishment firm, across treated units.

- When matching on log's, we require both treated and potential control units to have received positive citations at least twice for $\tau \equiv t - g_{\omega n} \in \{-10, \ldots, -1\}$.

- Let $y_{\omega n \tau}$ denote an outcome of interest for firm $\times$ CZ1990 ωn at relative time τ..

- **Matching on pre-treatment outcomes, *changes***

$$\texttt{SLOPE}_{\omega n} = \frac{\sum_{\tau=-10}^{-6} y_{\omega n \tau}}{5} - \frac{\sum_{\tau=-5}^{-1} y_{\omega n \tau}}{5}$$

$$\texttt{CRIT1}_{\omega n n'} = \frac{\texttt{SLOPE}_{\omega n} - \texttt{SLOPE}_{\omega n'}}{0.5\left(|\texttt{SLOPE}_{\omega n}| + |\texttt{SLOPE}_{\omega n'}|\right)}$$

  – If $\texttt{SLOPE}_{\omega n}$ and $\texttt{SLOPE}_{\omega n'}$ have different **signs**, then remove these potential matches.

- **Matching on pre-treatment outcomes, *levels***

$$\texttt{CRIT2}_{\omega nn'} = \frac{1}{10} \sum_{\tau=-10}^{-1} (\log y_{\omega n\tau} - \log y_{\omega n'\tau})$$

- **Matching on pre-treatment covariates**

  1. **Employment shares**—2 digit NAICS supersectors in commuting zones n and n'
  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

  $\texttt{CRIT3}_{\omega nn'}$ is equal to the mean of the DHS differences between n and n' for all of the listed covariates over $\tau \in \{-10, \ldots, -1\}$.

- **Select control unit for each treated unit**

  – Take $n'_\omega(n) = \arg\min_{n'} \texttt{CRIT}_{\omega nn'}$ (break ties randomly)

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

F Estimation of β, α, η, and θ

Table F.1: First-stage regression results for (15).

	Inventors		D&B employment	
	Pooled	Bootstrap	Pooled	Bootstrap
No z-adjustment	0.0102	0.0102	0.0103	0.0103
	(0.0001)	(0.0002)	(0.0001)	(0.0002)
N	98039	—	92972	—
R^2	0.87	—	0.86	—
z-adjustment	0.0066	0.0066	0.0072	0.0073
	(0.0001)	(0.0004)	(0.0001)	(0.0004)
N	98039	—	92972	—
R^2	0.82	—	0.80	—

Note: The table shows the results of the first stage regression for results of regression equation (15) in Table 4.

Table F.2: Estimating θ.

	Pooled		1990		2015	
	Inventors	D&B employment	Inventors	D&B employment	Inventors	D&B employment
$\log \ell_{nt}(z)$	0.58	0.35	0.48	0.34	0.39	0.46
	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)

Note: The table shows the results of estimating regression (16). Standard errors are clustered at the LMA level.

Table F.3: Estimation results for β and $\alpha/(1-\eta)$.

	Panel A: Estimating β				Panel B: Estimating $\alpha/(1-\eta)$			
	Inventors		D&B employment		Inventors		D&B employment	
	Pooled	Bootstrap	Pooled	Bootstrap	Pooled	Bootstrap	Pooled	Bootstrap
No z-adjustment								
OLS	0.31	0.35	0.19	0.27	0.39	0.40	0.38	0.39
	(0.013)	(0.007)	(0.009)	(0.006)	(0.022)	(0.024)	(0.023)	(0.026)
TSLS	—	—	—	—	0.62	0.62	0.76	0.75
	—	—	—	—	(0.070)	(0.062)	(0.087)	(0.084)
N	98039	—	92972	—	87129	—	87129	—
KP F-stat.	—	—	—	—	764	—	542	—
z adjustment								
OLS	—	—	—	—	0.46	0.46	0.35	0.34
	—	—	—	—	(0.031)	(0.042)	(0.027)	(0.051)
TSLS	—	—	—	—	1.86	1.87	2.96	3.10
	—	—	—	—	(0.234)	(0.328)	(0.488)	(1.373)
N	—	—	—	—	87129	—	87129	—
KP F-stat.	—	—	—	—	164	—	49	—

Note: State-level variation. Data from Bloom et al. (2013) and Wilson (2009).

Figure D.6: Bootstrap distribution of $\{\beta_\tau\}_\tau$ from synthetic DiD estimator of Ben-Michael et al. (2022).

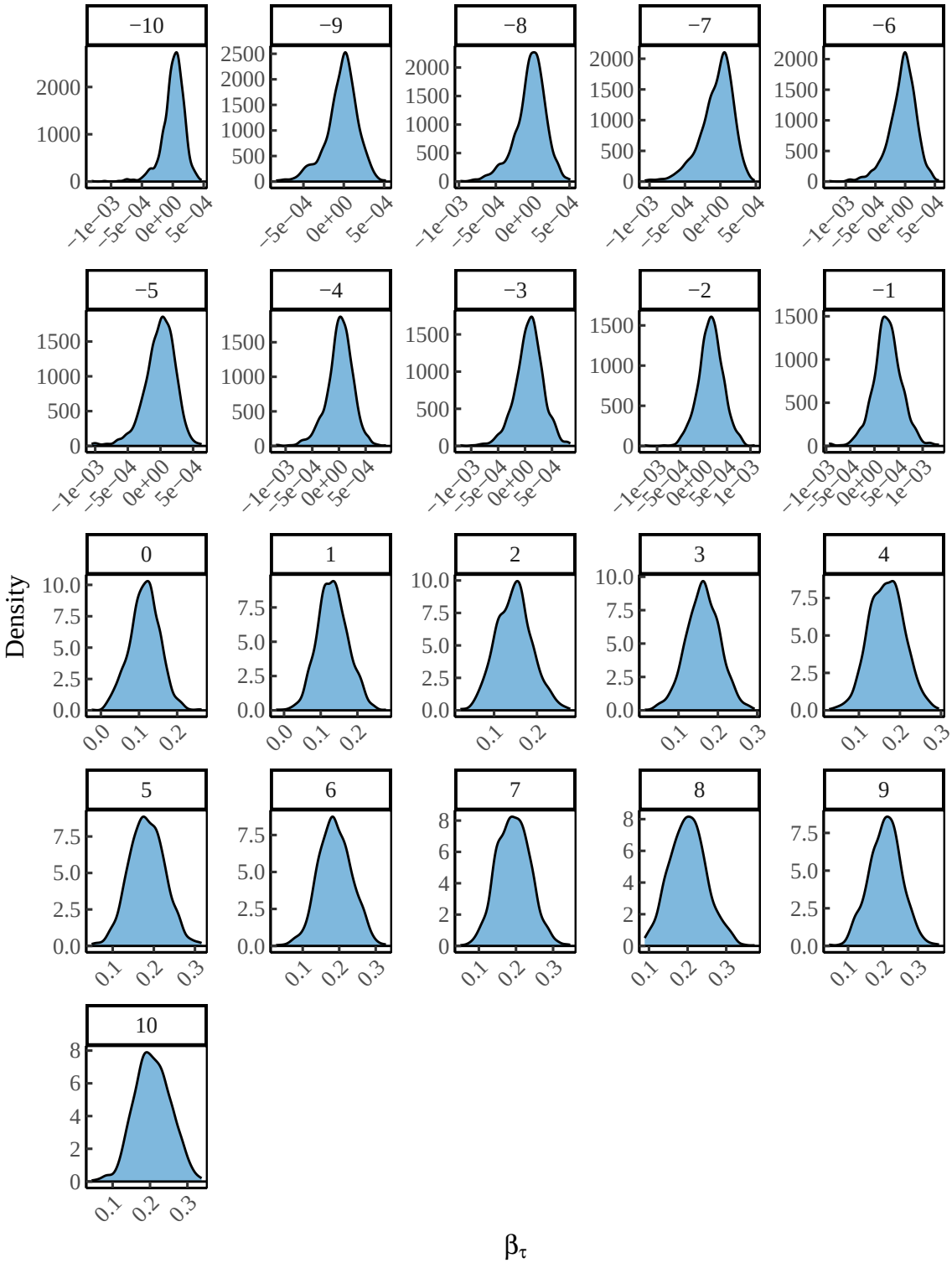
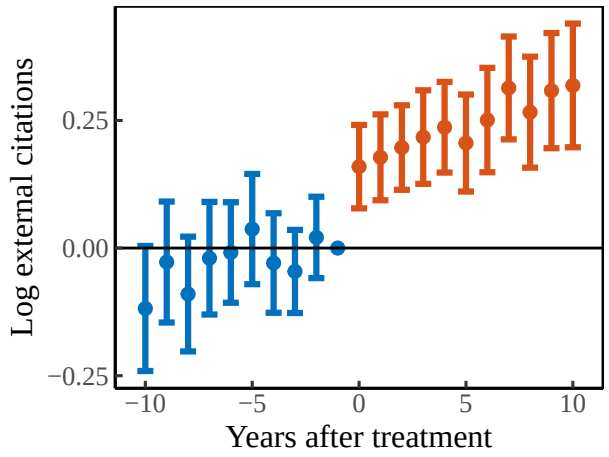
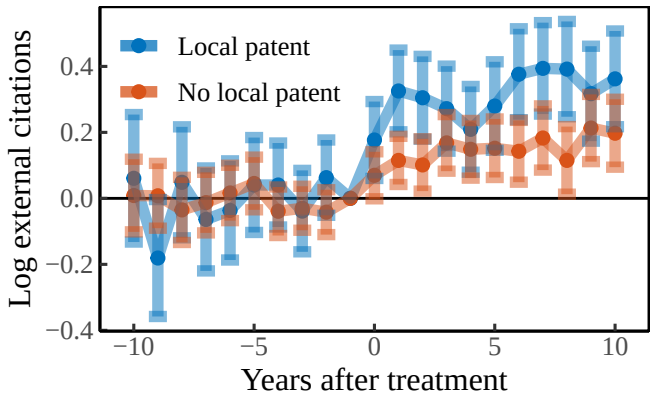


Figure D.7: Response of citations received by firm patents in expansion commuting zone preceding and following expansion, excluding temporarily treated units.



Note: Wald tests of pre-treatment covariates fail to be rejected with a p-value of 0.26.

Figure D.8: Heterogeneity by plant type, local patents.

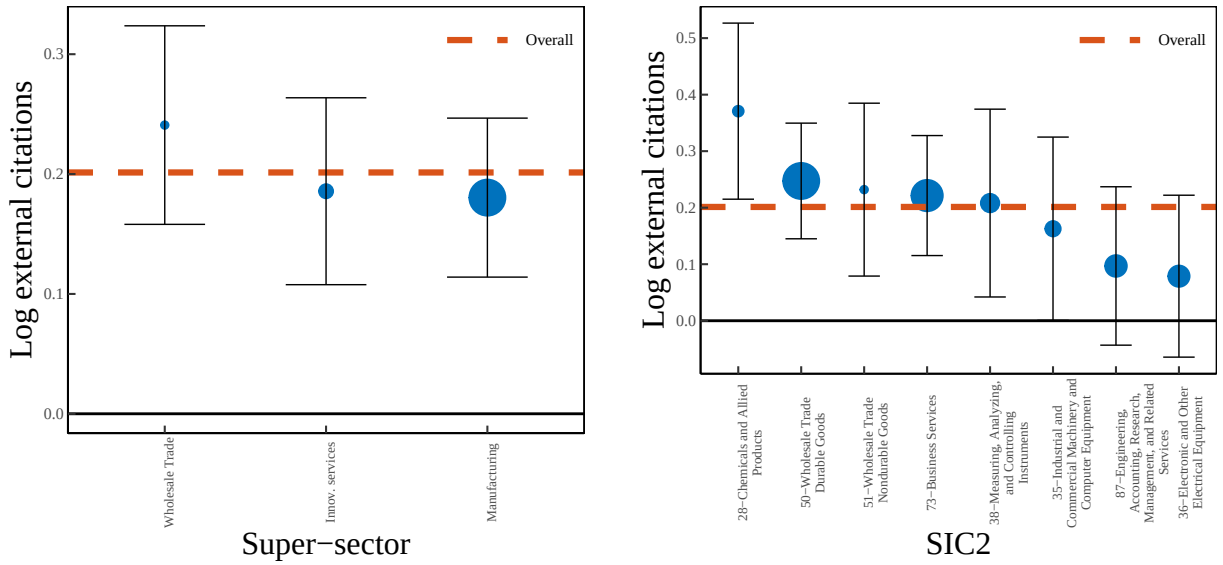


Note: Each point represents a loading on the relative, partition-interacted treatment time dummies. We fail to reject the Wald test of jointy nullity of pre-treatment dummies both separately (p-values of 0.53 and 0.16 for blue and red, respectively) and jointly (p-value of 0.27). Standard errors are clustered at the firm × commuting zone ($\omega \times n$) level.

Figure D.9: Heterogeneity by sector.

(a) Heterogeneity in effect by super-sector.

(b) Heterogeneity in effect by selected two-digit SIC.



Note: Standard errors are clustered at the firm × commuting zone ($\omega \times n$) level. Size of dots corresponds to number of observations in sector. We shows results for all sectors with greater than 10000 and 5000 observations, respectively.

Figure F.1: Empirical distribution of β estimates from bootstrap.

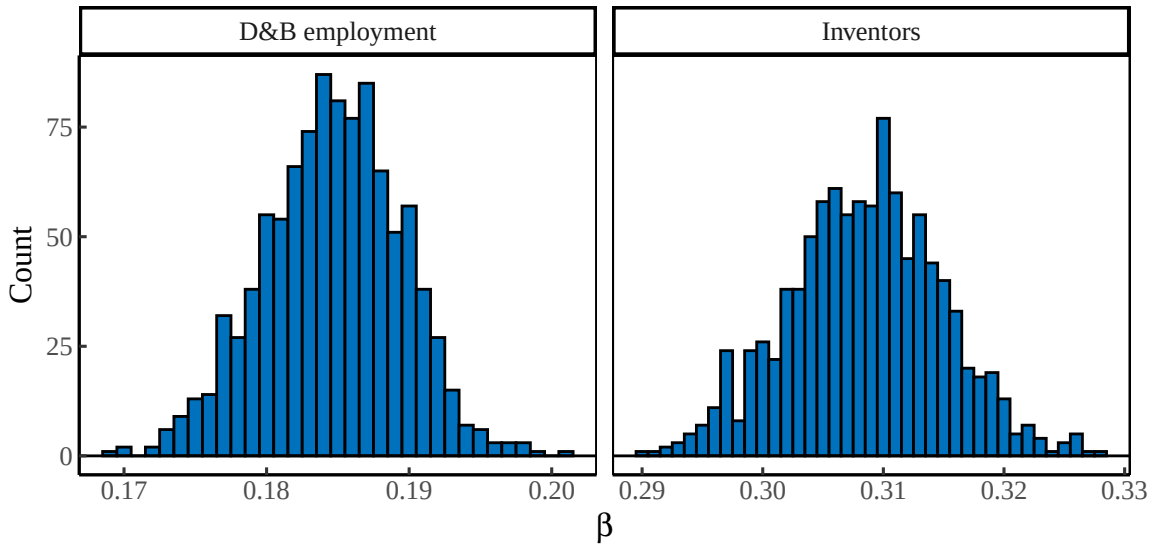


Figure F.2: Empirical distribution of $\alpha/(1-\eta)$ estimates from bootstrap.

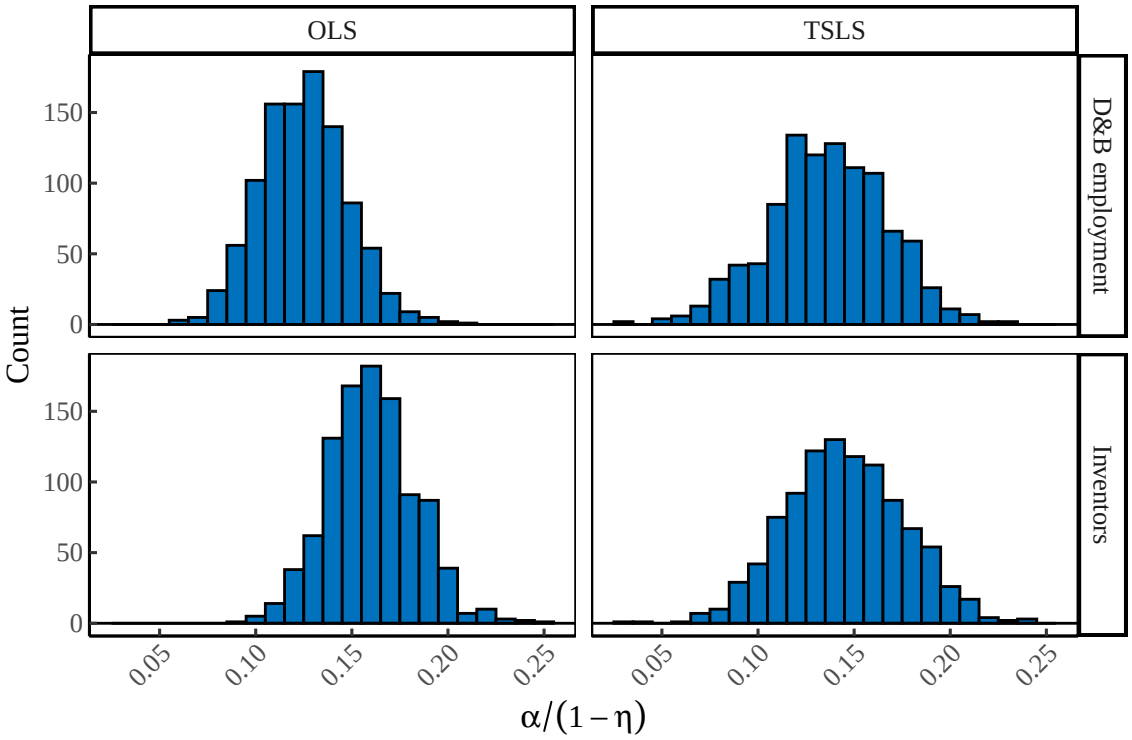
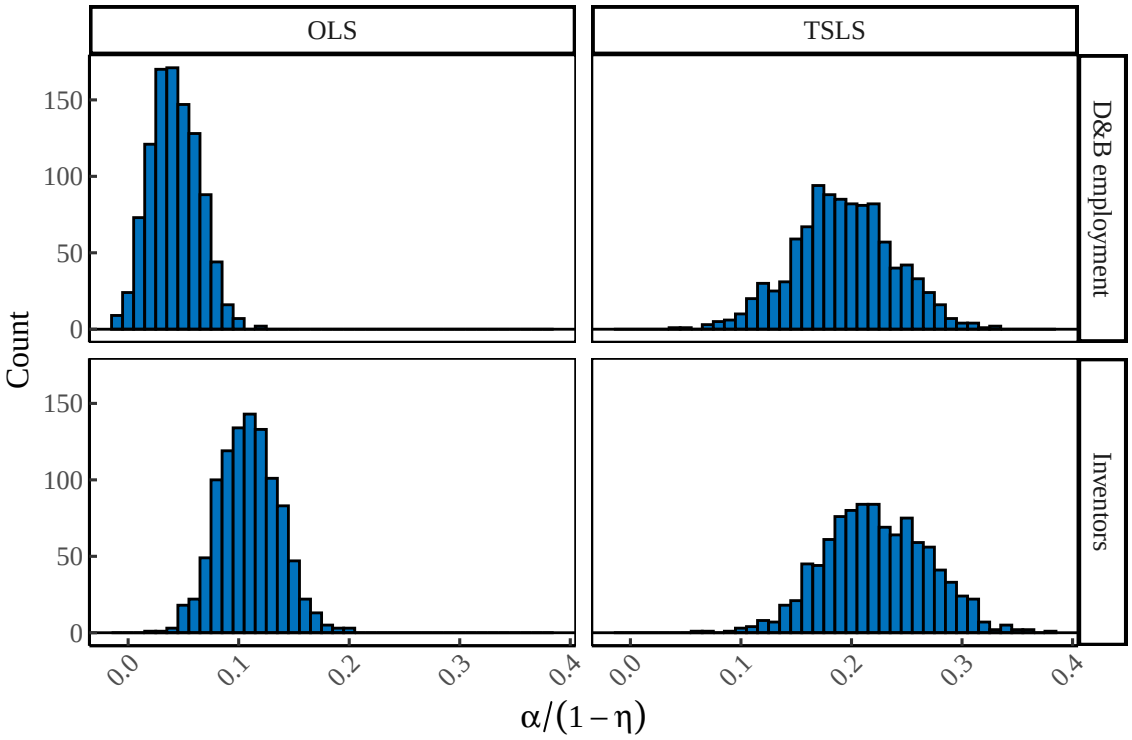


Figure F.3: Empirical distribution of $\alpha/(1-\eta)$ estimates from bootstrap with z-adjustment.



G Additional details on the estimation

G.1 Additional details on the estimation of α, β, η, and θ

For estimation of all structural parameters, we consider the period 1980-2018. Estimation of key parameters is performed at the Labor Market Area (LMA) level using semi-decadal (five-year) intervals. We first aggregate firm-level data to the LMA-year level by summing all activity within each LMA. We then consolidate these annual data into five-year periods: patents and external citations are summed across each interval, while employment figures use the maximum Dun & Bradstreet (D&B) values and inventor counts reflect the total number of unique individuals.

As described in Section 4, for estimation of the joint parameter $\alpha/(1-\eta)$ we utilize the instrument of Burchardi et al. (2020). This instrument is available at the county level in five-year intervals.[20] To apply it at the LMA level, we construct an employment-weighted mean across the constituent counties of each Labor Market Area.

For estimation of the shape parameter θ, we examine all potential markets for all firms across all years after their initial appearance in our dataset, including markets where they have no operations. In model parlance, this approach excludes only those firms that paid the entry cost f_e but chose not to pay any location-specific idiosyncratic fixed cost $f_{\omega n}$ in any market. Such firms constitute a small overall measure in our quantitative model. To proxy for firm productivity z_ω, we categorize firms into ten equal-sized bins (deciles) based on their firm-wide Dun & Bradstreet employment and inventor counts, calculated for each five-year period. We then estimate θ by regressing, within each decile, the share of firms operating in a given market against the decile's total employment in that market (using both inventor counts and D&B employment figures as measures of employment) controlling for all possible bilateral fixed effects.

G.2 Additional details on the model inversion

G.2.1 Location fundamentals

For $\{Z_n, \bar{K}_n, B_{ni}, B_{np}\}_n$, we obtain closed-form expressions from first-order conditions of the firms and households that identify these parameters up to scale.

Production employment and wages identify $\{Z_n\}$. The monopolist producer of product

[20] Data is available at https://www.immigrationshock.com/.

line j solves the following cost-minimization problem:

$$\min_{L_{jnt}} -\sum_n w_{np} L_{jnt} \quad \text{s.t.} \quad 1 = A_{jt} \left(\sum_n \left(Z_n L_{jnt} \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

FOC's imply that:

$$L_{jnt} = \left(\frac{w_{np}}{w_{n'p}} \right)^{-\sigma} \left(\frac{Z_{n'}}{Z_n} \right)^{1-\sigma} L_{jn't}$$

Plugging back into the constraint we see that:

$$Z_{n'} \propto w_{n'p}^{\frac{\sigma}{\sigma-1}} L_{jn't}^{\frac{1}{\sigma-1}}, \quad \forall n', j$$

In a like fashion, innovation employment and wages identify $\{\bar{K}_n\}$. First, aggregating firm-level innovation hiring in market n yields:

$$\begin{aligned} L_{nit} &= M \int \chi_n(z) \left(\ell_{nt}(z) + \bar{f}_{nt}(z) \right) d\Psi(z) \\ &= M \int f_{max}^{-\theta} \left(\frac{1-\eta}{\eta} \right)^{\theta} \left(\ell_{nt}(z)^{\theta+1} + \frac{\theta}{\theta+1} \frac{1-\eta}{\eta} \ell_{nt}(z)^{\theta+1} \right) d\Psi(z) \\ &= M f_{max}^{-\theta} \left(\frac{1-\eta}{\eta} \right)^{\theta} \left(1 + \frac{\theta}{\theta+1} \frac{1-\eta}{\eta} \right) \left(\frac{\eta K_{nt} \lambda_0 V_t}{w_{nit}} \right)^{\frac{\theta+1}{1-\eta}} \int z^{\theta+1} d\Psi(z) \end{aligned}$$

From the above expression it is clear that $L_{nit} \propto \left(K_{nt} / w_{nit} \right)^{(\theta+1)/(1-\eta)}$

Note that:

$$\begin{aligned} K_{nt} &= \bar{K}_n \left(M \int \chi_n(z) z^{1-\beta} \ell_{nt}(z)^{\beta} d\Psi(z) \right)^{\alpha} \\ &= \bar{K}_n \left(M f_{max}^{-\theta} \left(\frac{1-\eta}{\eta} \right)^{\theta} \left(\frac{\eta K_{nt} \lambda_0 V_t}{w_{nit}} \right)^{\frac{\theta+\beta}{1-\eta}} \int z^{\theta+\beta} d\Psi(z) \right)^{\alpha} \end{aligned}$$

Thus:

$$K_{nt} \propto \bar{K}_n L_{nit}^{\frac{\alpha(\theta+\beta)}{1+\theta}}$$

Combining with above we get:

$$\bar{K}_n \propto w_{nit} L_{nit}^{\frac{1-\eta-\alpha(\theta+\beta)}{\theta+1}}$$

Amenities are inverted using the gravity structure implied by Gumbel preference shocks. Cobb-Douglas preferences over housing and consumption give rise to the following expression

for rents in market n:

$$R_n = (1 - \varpi) \sum_{s \in \{p,i\}} w_{ns} L_{ns}$$

Note that:

$$L_{ns} \propto B_{ns} \left(R_n^{-\varpi} w_{ns} \right)^{\epsilon}$$

due to Gumbel distribution for $\{\varepsilon_{ns}\}_{ns}$ (see main text). This easily admits an expression for $\{B_{ns}\}_{s \in \{p,i\}}$:

$$B_{ns} \propto L_{ns} \left(\frac{w_{ns}}{\left(\sum_{s' \in \{p,i\}} w_{ns'} L_{ns'} \right)^{\varpi}} \right)^{-\epsilon}$$

G.2.2 Model parameters

The remaining parameters $\{\gamma, \lambda_0, f_{max}, \bar{z}, f_e\}$ are common across locations and are calibrated within the solution to the balanced growth path equilibrium. We do this by first posing a guess for the equilibrium—which is characterized by de-trended wages in both sectors in all locations; de-trended rents in all locations; innovation efficiencies (given $\bar{K}_n$ from above) in all locations; the mass of active firms; and the de-trended market value of a firm under a set of initial parameters: $\{\gamma, \lambda_0, f_{max}, \bar{z}, f_e\}_0$. We embed in this equilibrium a set of moments from the data that we use to update the equilibrium, which then implies new values for parameters (described in further detail below). We iterate until the equilibrium vector converges to an arbitrary precision.

For γ note that national income accounting implies that:

$$Y = \rho V + M f_e + \sum_{s,n} w_{ns} L_{ns}$$

The quality ladder structure of the model plus Cobb-Douglas aggregation of intermediate varieties further implies that net profits of the firm equal $\frac{\gamma - 1}{\gamma} Y$. Since revenues equal Y, this leaves $\frac{1}{\gamma} Y$ units left to compensate labor. Therefore $\sum_n w_{np} L_{np} = \frac{1}{\gamma} Y$. Combining these two expressions yields:

$$\begin{aligned} Y &= M f_e + \rho V + \frac{1}{\gamma} Y \left(1 + \frac{\sum_n w_{ni} L_{ni}}{\sum_n w_{np} L_{np}} \right) \\ \gamma &= \frac{1 + \frac{\sum_n w_{ni} L_{ni}}{\sum_n w_{np} L_{np}}}{1 - \frac{M f_e + \rho V}{Y}} \end{aligned} \tag{35}$$

where $\frac{\sum_n w_{ni} L_{ni}}{\sum_n w_{np} L_{np}}$ is the ratio of labor income in the innovation sector to labor income in the

production sector and is taken from data.

We discipline λ_0 using the growth rate. Note that $g = \bar{\lambda}\log\gamma$ (Proposition 2). The expression for $\bar{\lambda}$ is messy, but λ_0 is easily separable. From Proposition 2:

$$\bar{\lambda} = M\lambda_0 \int \sum_n K_n \chi_n(z) z^{1-\eta} \ell_n(z)^\eta d\Psi(z)$$

Note that:

$$\chi_n(z) z^{1-\eta} \ell_n(z)^\eta = \underbrace{f_{max}^{-\theta}\left(\frac{1-\eta}{\eta}\right)^\theta}_{\mathcal{C}_1} \left(\frac{\eta K_{nt}\lambda_0 V_t}{w_{nit}}\right)^{\frac{\theta+\eta}{1-\eta}} \times \overbrace{\int z^{1+\theta} d\Psi(z)}^{\bar{\mathcal{Z}}}$$

So that:

$$\lambda_0 = \left[\frac{M}{\bar{\lambda}} \sum_n K_n^{\frac{\theta+1}{1-\eta}} \mathcal{C}_1 \left(\frac{\eta V_t}{w_{nit}}\right)^{\frac{\theta+\eta}{1-\eta}} \bar{\mathcal{Z}}\right]^{\frac{\eta-1}{1+\theta}} \tag{36}$$

We get g from data.

f_{max} is a scale parameter for the draws of fixed costs; it is thus intimately tied to firms scope. The measure of active firms is given by:

$$M\left(1 - \int \prod_n (1 - \chi_n(z)) d\Psi(z)\right) = M\bar{\chi}.$$

The number of establishments (markets) that are active is given by:

$$M \int \sum_n \chi_n(z) d\Psi(z)$$

Therefore the average number of markets per firm is given by:

$$\frac{\#\text{ Markets}}{\#\text{ Firms}} = \frac{\int \sum_n \chi_n(z) d\Psi(z)}{\bar{\chi}} = \frac{f_{max}^{-\theta}\left(\frac{1-\eta}{\eta}\right)^\theta (\eta\lambda_0 V_t)^{\frac{\theta}{1-\eta}} \sum_n \left(\frac{K_{nt}}{w_{ni}}\right)^{\frac{\theta}{1-\eta}} \int z^\theta d\Psi(z)}{\bar{\chi}}$$

yielding:

$$f_{max} = \left[\frac{\left(\frac{1-\eta}{\eta}\right)^\theta (\eta\lambda_0 V_t)^{\frac{\theta}{1-\eta}} \sum_n \left(\frac{K_{nt}}{w_{ni}}\right)^{\frac{\theta}{1-\eta}} \int z^\theta d\Psi(z)}{\bar{\chi}\frac{\#\text{ Markets}}{\#\text{ Firms}}}\right]^{\frac{1}{\theta}} \tag{37}$$

where $\frac{\#\text{ Markets}}{\#\text{ Firms}}$ comes from data.

We adopt a simple distribution for $\Psi(z)$. With probability 0.9, a firm draws $z = \underline{z} = 1$; with complementary probability 0.1 it draws $z = \bar{z}$. We set $\bar{z}$ such that 72% of innovations are accounted for by the $z = \bar{z}$ (i.e. the top 10%) firms, which matches our data.[21]

Lastly, one can verify that:

$$f_e = \frac{\bar{\lambda} V_t (1 - \eta)}{(\theta + 1) M} \tag{38}$$

which completes specification of our model inversion equations.

To complete the description of the model inversion, we provide a rough outline of how we jointly solve and calibrate model parameters.

1. Invert $\{Z_n, \bar{K}_n, B_{ni}, B_{np}\}_n$ according to section G.2.1.
2. Guess $\{\gamma, \lambda_0, f_{max}, \bar{z}, f_e\}$ and $\{\{w_{ni}, w_{np}, R_n, K_n\}_n, M, V\}$
3. Using the guess, solve for $\{L_{np}, L_{ni}\}_n$ and update $\{R_n\}_n$ according to:

$$R_n^{(k+1)} = \Delta R_n^k + (1 - \Delta) R_n^k$$

   where Δ is some number close to but less than 1
4. Given the guess for γ and $\{L_{np}\}_n$ update $\{w_{np}\}_n$.
5. Given the guess for $\{\{K_n\}_n, \lambda_0, V, M\}$ and $\{L_{ni}\}_n$ update $\{w_{ni}\}_n$.
6. Given the guess for $\{M, \lambda_0, V\}$ and the new values for $\{w_{ni}\}$ update $\{K_n\}_n$.
7. Given the guess for f_{max}, $\{L_{ni}\}_n$, and data on the average firm size, update M.
8. Update f_e according to (38).
9. Using guess for $\gamma, \bar{\lambda}$, and $\{L_{np}\}_n$ update V.
10. Using guess for γ and data on the growth rate, update $\bar{\lambda}$.
11. Update λ_0 using (36).
12. Update γ using (35).
13. Update f_{max} using (37) and dampener Δ.
14. Check for convergence of $\{\{w_{ni}, w_{np}, R_n, K_n\}_n, M, V\}$; if converged then stop; otherwise repeat.

[21] Our patent data roughly obeys the Pareto principle, consistent with evidence reported elsewhere that innovation is a fat-tailed phenomenon (Scherer and Harhoff, 2000).

References

Acemoglu, Daron and Ufuk Akcigit, "Intellectual Property Rights Policy, Competition and Innovation," *Journal of the European Economic Association*, 2012, *10* (1), 1–42.

_ , _ , Douglas Hanley, and William Kerr, "Transition to Clean Technology," *Journal of Political Economy*, 2016, *124* (1), 52–104.

_ , _ , Harun Alp, Nicholas Bloom, and William Kerr, "Innovation, reallocation, and growth," *American Economic Review*, 2018, *108* (11), 3450–3491.

Aghion, Philippe and Peter Howitt, "A Model of Growth Through Creative Destruction," *Econometrica*, 1992, *60* (2), 323–351.

Akcigit, Ufuk and Sina T. Ates, "What Happened to US Business Dynamism?," *Journal of Political Economy*, 2023, *131* (8), 2059–2124.

_ and William R. Kerr, "Growth through Heterogeneous Innovations," *Journal of Political Economy*, 2018, *126* (4), 1374–1443.

_ , Santiago Caicedo, Ernest Miguelez, Stefanie Stantcheva, and Valerio Sterzi, "Dancing with the Stars: Innovation Through Interactions," Working Paper 24466, National Bureau of Economic Research March 2018.

Argente, David, Sara Moreira, Ezra Oberfield, and Venky Venkateswaran, "Scalable Expertise," Manuscript 2020.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager, "Synthetic Difference-in-Differences," *American Economic Review*, 2021, *111* (12), 4088–4118.

Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple, "Innovation and production in the global economy," *American Economic Review*, 2018, *108* (8), 2128–73.

Audretsch, David B and Maryann P Feldman, "R&D spillovers and the geography of innovation and production," *The American economic review*, 1996, *86* (3), 630–640.

Ayerst, Stephen, Faisal Ibrahim, Gaelan MacKenzie, and Swapnika Rachapalli, "Trade and diffusion of embodied technology: An empirical analysis," *Journal of Monetary Economics*, 2023, *137*, 128–145.

Babina, Tania and Sabrina T. Howell, "Entrepreneurial Spillovers from Corporate R&D," *Journal of Labor Economics*, 2024, *42* (2), 469–509.

Barnatchez, Keith, Leland D. Crane, and Ryan A. Decker, "An Assessment of the National Establishment Time Series (NETS) Database," *Federal Reserve Board of Governors Working Paper*, 2017.

Ben-Michael, Eli, Avi Feller, and Jesse Rothstein, "Synthetic Controls with Staggered Adoption," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 2022, *84* (2), 351–381.

Bergé, Laurent, "Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm," Technical Report 13 2018.

Berkes, Enrico and Ruben Gaetani, "The Geography of Unconventional Innovation," *The Economic Journal*, 09 2020, *131* (636), 1466–1514.

_ , _ , and Martí Mestieri, "Technological Waves, Knowledge Diffusion, and Local Growth," *Journal of Political Economy Macroeconomics*, 2025, *3* (1), 75–121.

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

_ , Mark Schankerman, and John Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

Buera, Francisco J. and Ezra Oberfield, "The Global Diffusion of Ideas," *Econometrica*, 2020, *88* (1), 83–114.

Burchardi, Konrad B., Thomas Chaney, Tarek Alexander Hassan, Lisa Tarquinio, and Stephen J. Terry, "Immigration, Innovation, and Growth," Technical Report 27075, National Bureau of Economic Research May 2020.

Cai, Jie, Nan Li, and Ana Maria Santacreu, "Knowledge Diffusion, Trade, and Innovation across Countries and Sectors," *American Economic Journal: Macroeconomics*, 2022, *14* (1), 104–45.

Cai, Sheng, Lorenzo Caliendo, Fernando Parro, and Wei Xiang, "Mechanics of Spatial Growth," Working Paper 30579, National Bureau of Economic Research October 2022.

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

Carlino, Gerald and William R. Kerr, "Chapter 6 - Agglomeration and Innovation," in "Handbook of Regional and Urban Economics," Vol. 5, Elsevier, 2015, pp. 349–404.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The Effect of Minimum Wages on Low-Wage Jobs," *The Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

Chatterji, Aaron, Edward Glaeser, and William Kerr, "Clusters of Entrepreneurship and Innovation," *Innovation Policy and the Economy*, 2014, *14*, 129–166.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin, "ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects," *The Stata Journal*, 2020, *20* (1), 95–115.

Costinot, Arnaud and Andrés Rodríguez-Clare, "Trade theory with numbers: Quantifying the consequences of globalization," in "Handbook of international economics," Vol. 4, Elsevier, 2014, pp. 197–261.

Crews, Levi Garrett, "A dynamic spatial knowledge economy." PhD dissertation, The University of Chicago 2023.

Davis, Morris A and François Ortalo-Magné, "Household expenditures, wages, rents," *Review of Economic Dynamics*, 2011, *14* (2), 248–261.

Dixit, Avinash K and Joseph E Stiglitz, "Monopolistic competition and optimum product diversity," *The American economic review*, 1977, *67* (3), 297–308.

Eckert, Fabian and Michael Peters, "Spatial Structural Change," Working Paper 30489, National Bureau of Economic Research September 2022.

Fajgelbaum, Pablo D and Cecile Gaubert, "Optimal spatial policies, geography, and sorting," *The Quarterly Journal of Economics*, 2020, *135* (2), 959–1036.

_ **and** _ , "Place-Based Policies: Lessons from Theory," Working Paper 33517, National Bureau of Economic Research February 2025.

Fort, Teresa C, Wolfgang Keller, Peter K Schott, Stephen Yeaple, and Nikolas Zolas, "Colocation of Production and Innovation: Evidence from the United States," 2020.

Galle, Simon, Andrés Rodríguez-Clare, and Moises Yi, "Slicing the pie: Quantifying the aggregate and distributional effects of trade," *The Review of Economic Studies*, 2023, *90* (1), 331–375.

Gardner, John, "Two-stage differences in differences," Technical Report, arXiv July 2022. arXiv:2207.05943.

Garetto, Stefania, "Input Sourcing and Multinational Production," *American Economic Journal: Macroeconomics*, April 2013, *5* (2), 118–51.

Giroud, Xavier, Ernest Liu, and Holger Mueller, "Innovation Spillovers across US Tech Clusters," Technical Report, National Bureau of Economic Research 2024.

Glaeser, Edward L and Joshua D Gottlieb, "The economics of place-making policies," Technical Report, National Bureau of Economic Research 2008.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings," *Journal of political economy*, 2010, *118* (3), 536–598.

Griffith, Rachel, Sokbae Lee, and John Van Reenen, "Is distance dying at last? Falling home bias in fixed-effects models of patent citations," *Quantitative Economics*, 2011, *2* (2), 211–249.

Grossman, Gene M. and Elhanan Helpman, "Quality Ladders in the Theory of Growth," *The Review of Economic Studies*, 1991, *58* (1), 43–61.

Helpman, Elhanan, "A simple theory of international trade with multinational corporations," *Journal of political economy*, 1984, *92* (3), 451–471.

Hsieh, Chang-Tai, Peter J. Klenow, and Ishan Nath, "A Global View of Creative Destruction," *Journal of Political Economy Macroeconomics*, 2023, *1* (2), 243–275.

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

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

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *The Quarterly Journal of Economics*, 1993, *108* (3), 577–598. Publisher: Oxford University Press.

Keller, Wolfgang and Stephen Ross Yeaple, "The Gravity of Knowledge," *American Economic Review*, June 2013, *103* (4), 1414–44.

Kerr, William R. and Frederic Robert-Nicoud, "Tech Clusters," *Journal of Economic Perspectives*, 2020, *34* (3), 50–76.

Kleinman, Benny, "Wage Inequality and the Spatial Expansion of Firms," *Working Paper*, 2023.

Klette, Tor Jakob and Samuel Kortum, "Innovating Firms and Aggregate Innovation," *Journal of Political Economy*, 2004, *112* (5), 986–1018.

Kline, Patrick and Enrico Moretti, "Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority," *The Quarterly journal of economics*, 2014, *129* (1), 275–331.

Kortum, Samuel, "Equilibrium R&D and the Patent–R&D Ratio: U.S. Evidence," *The American Economic Review*, 1993, *83* (2), 450–457.

König, Michael D., Xiaodong Liu, and Yves Zenou, "R&D Networks: Theory, Empirics, and Policy Implications," *The Review of Economics and Statistics*, 2019, *101* (3), 476–491.

Lhuillier, Hugo, "Should I Stay or Should I Grow?," 2023. Working Paper.

Lind, Nelson and Natalia Ramondo, "Global Innovation and Knowledge Diffusion," *American Economic Review: Insights*, 2023, *5* (4), 494–510.

Lucas, Robert E. and Benjamin Moll, "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 2014, *122* (1), 1–51.

Martellini, Paolo, "Local labor markets and aggregate productivity," Technical Report, Working paper 2022.

Matray, Adrien, "The local innovation spillovers of listed firms," *Journal of Financial Economics*, 2021, *141* (2), 395–412.

Moretti, Enrico, "The Effect of High-Tech Clusters on the Productivity of Top Inventors," *American Economic Review*, 2021, *111* (10), 3328–3375.

Oberfield, Ezra, Esteban Rossi-Hansberg, Pierre-Daniel Sarte, and Nicholas Trachter, "Plants in space," *Journal of Political Economy*, 2024, *132* (3), 000–000.

Peri, Giovanni, "Determinants of Knowledge Flows and Their Effect on Innovation," *The Review of Economics and Statistics*, 2005, *87* (2), 308–322.

Perla, Jesse and Christopher Tonetti, "Equilibrium Imitation and Growth," *Journal of Political Economy*, 2014, *122* (1), 52–76.

_ , _ , and Michael E. Waugh, "Equilibrium Technology Diffusion, Trade, and Growth," *American Economic Review*, 2021, *111* (1), 73–128.

Prato, Marta, "The Global Race for Talent: Brain Drain, Knowledge Transfer, and Growth," *The Quarterly Journal of Economics*, 2024, *140* (1), 165–238.

Ramondo, Natalia and Andrés Rodríguez-Clare, "Trade, multinational production, and the gains from openness," *Journal of Political Economy*, 2013, *121* (2), 273–322.

Roth, Jonathan, Pedro H. C. Sant'Anna, Alyssa Bilinski, and John Poe, "What's trending in difference-in-differences? A synthesis of the recent econometrics literature," *Journal of Econometrics*, 2023, *235* (2), 2218–2244.

Sampson, Thomas, "Dynamic Selection: An Idea Flows Theory of Entry, Trade, and Growth," *The Quarterly Journal of Economics*, 2015, *131* (1), 315–380.

Santacreu, Ana Maria, "Innovation, diffusion, and trade: Theory and measurement," *Journal of Monetary Economics*, 2015, *75*, 1–20.

Scherer, F. M and Dietmar Harhoff, "Technology policy for a world of skew-distributed outcomes," *Research Policy*, April 2000, *29* (4), 559–566.

Stock, James H. and Motohiro Yogo, "Testing for Weak Instruments in Linear IV Regression," *Identification and Inference for Econometric Models*, 2005.

Terry, Stephen J, Toni M Whited, and Anastasia A Zakolyukina, "Information versus Investment," *The Review of Financial Studies*, 2023, *36* (3), 1148–1191.

Tintelnot, Felix, "Global production with export platforms," *The Quarterly Journal of Economics*, 2017, *132* (1), 157–209.

Wasi, Nada and Aaron Flaaen, "Record Linkage Using Stata: Preprocessing, Linking, and Reviewing Utilities," *The Stata Journal*, 2015, *15* (3), 672–697. Publisher: SAGE Publications.

Wilson, Daniel J., "Beggar Thy Neighbor? The in-State, Out-of-State, and Aggregate Effects of R&D Tax Credits," *The Review of Economics and Statistics*, 2009, *91* (2), 431–436.